



US 20250255940A1

(19) **United States**

(12) **Patent Application Publication**

Finn et al.

(10) **Pub. No.: US 2025/0255940 A1**

(43) **Pub. Date:** Aug. 14, 2025

(54) **COMPOSITIONS AND METHODS FOR
EXPRESSING FACTOR IX**

62/829,009, filed on Apr. 3, 2019, provisional application No. 62/747,509, filed on Oct. 18, 2018.

(71) Applicants: **Intellia Therapeutics, Inc.**, Cambridge, MA (US); **Regeneron Pharmaceuticals, Inc.**, Tarrytown, NY (US)

(72) Inventors: **Jonathan Douglas Finn**, Cambridge, MA (US); **Hon-Ren Huang**, Cambridge, MA (US); **Moitri Roy**, Cambridge, MA (US); **KehDih Lai**, Yardley, PA (US); **Rachel Sattler**, New York, NY (US); **Christos Kyratsous**, Irvington, NY (US); **Cheng Wang**, Beijing (CN)

(21) Appl. No.: **18/999,448**

(22) Filed: **Dec. 23, 2024**

Related U.S. Application Data

(63) Continuation of application No. 16/657,961, filed on Oct. 18, 2019, now Pat. No. 12,214,023.

(60) Provisional application No. 62/840,352, filed on Apr. 29, 2019, provisional application No. 62/829,621, filed on Apr. 4, 2019, provisional application No.

Publication Classification

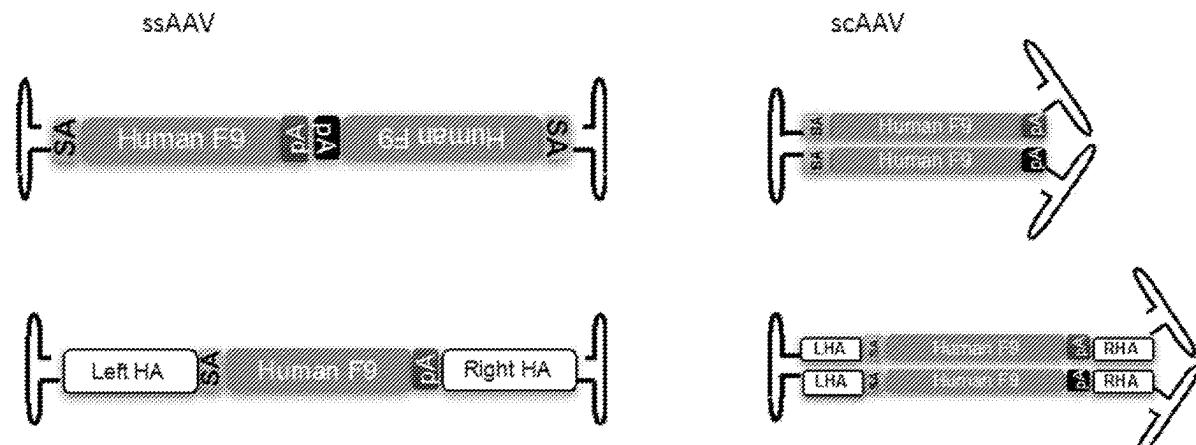
(51) Int. Cl.	
<i>A61K 38/48</i>	(2006.01)
<i>A61K 38/46</i>	(2006.01)
<i>A61K 48/00</i>	(2006.01)
<i>A61P 7/00</i>	(2006.01)
<i>C12N 15/113</i>	(2010.01)
<i>C12N 15/86</i>	(2006.01)
<i>C12N 15/90</i>	(2006.01)
(52) U.S. Cl.	
CPC	<i>A61K 38/4846</i> (2013.01); <i>A61K 38/465</i> (2013.01); <i>A61K 48/005</i> (2013.01); <i>A61P 7/00</i> (2018.01); <i>C12N 15/113</i> (2013.01); <i>C12N 15/86</i> (2013.01); <i>C12N 15/90</i> (2013.01); <i>C12Y 304/21022</i> (2013.01)

(57)

ABSTRACT

Compositions and methods for expressing Factor IX in a host cell or a population of host cells are provided. Also provided are engineered host cells expressing Factor IX.

Specification includes a Sequence Listing.



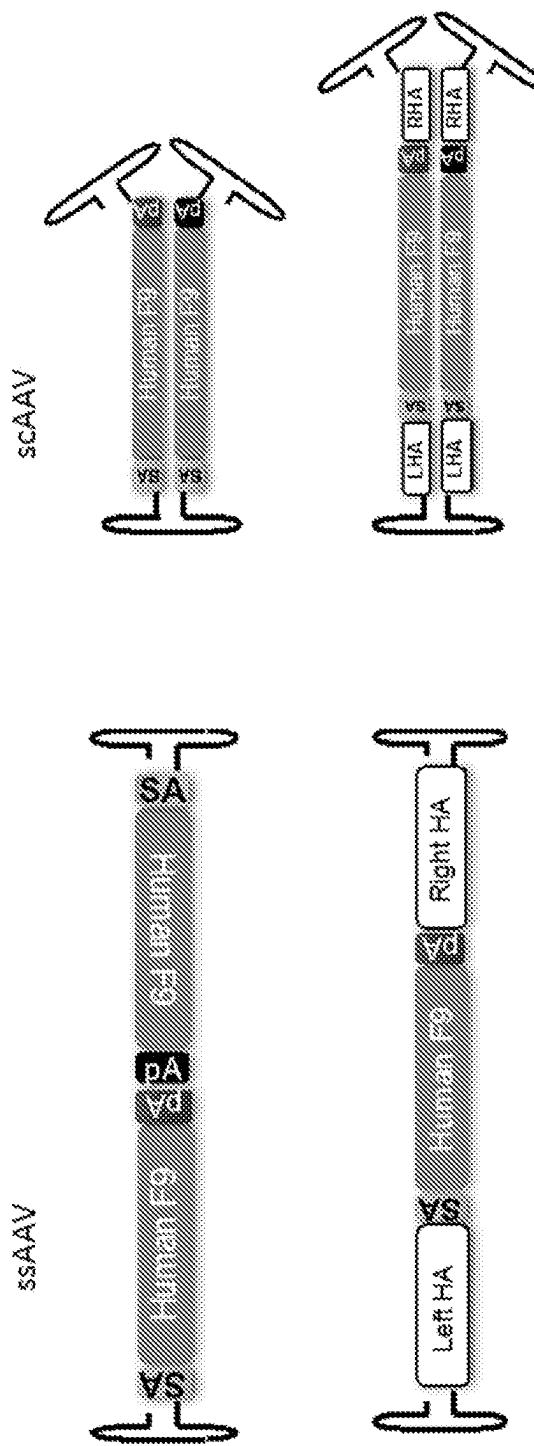


FIG. 1

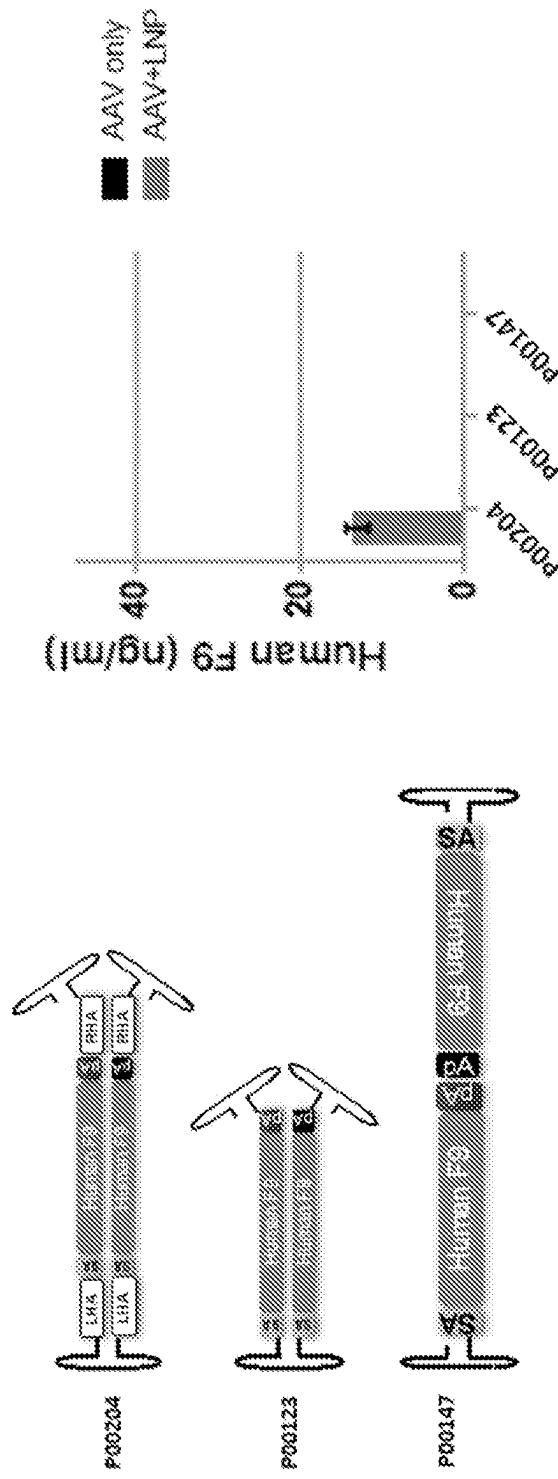


FIG. 2

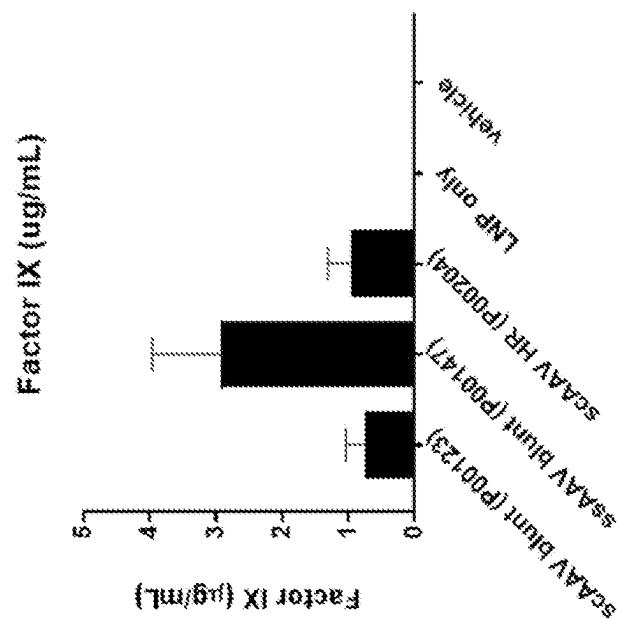


FIG. 3B

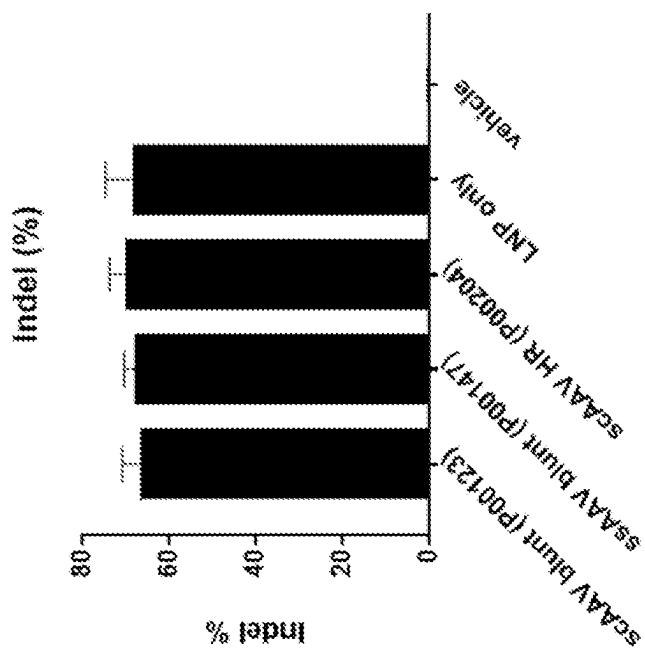
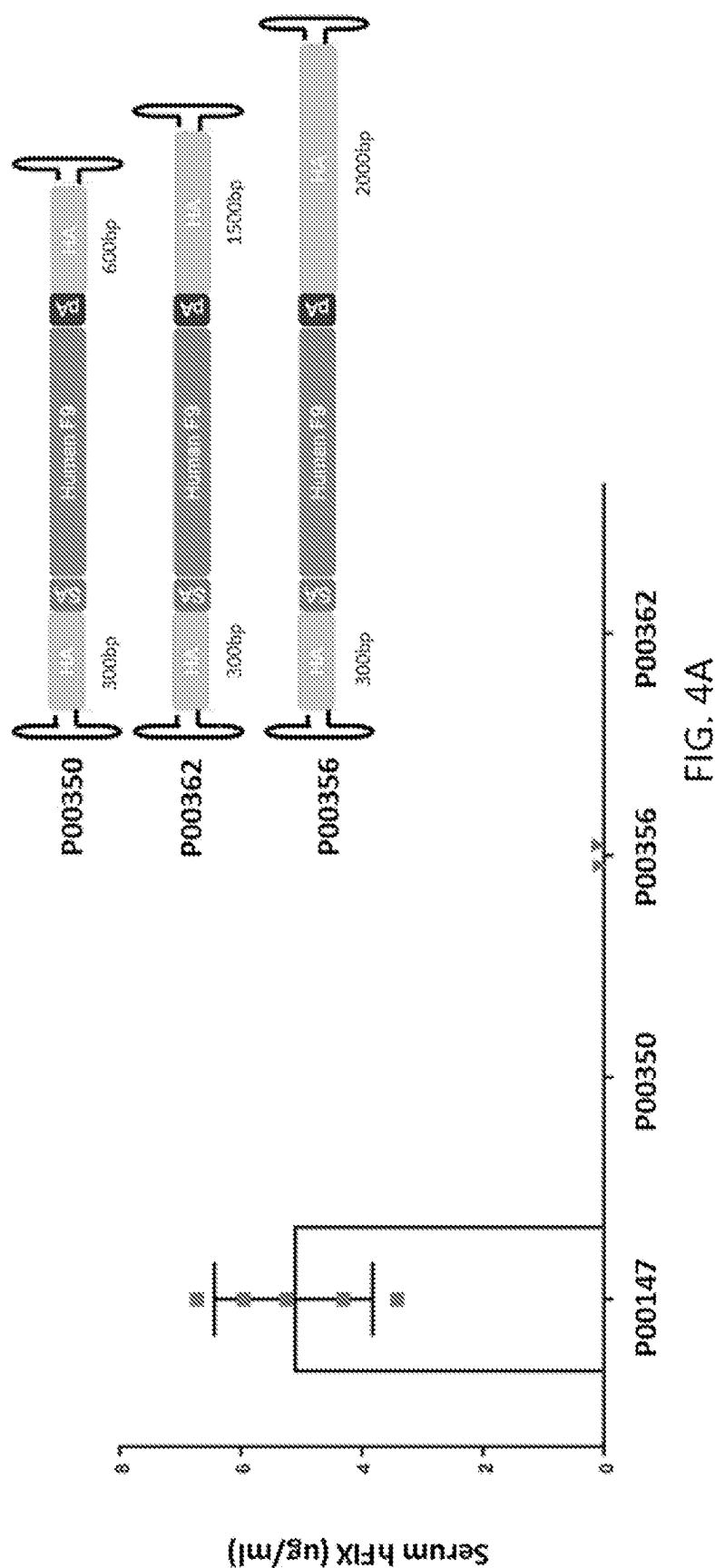


FIG. 3A



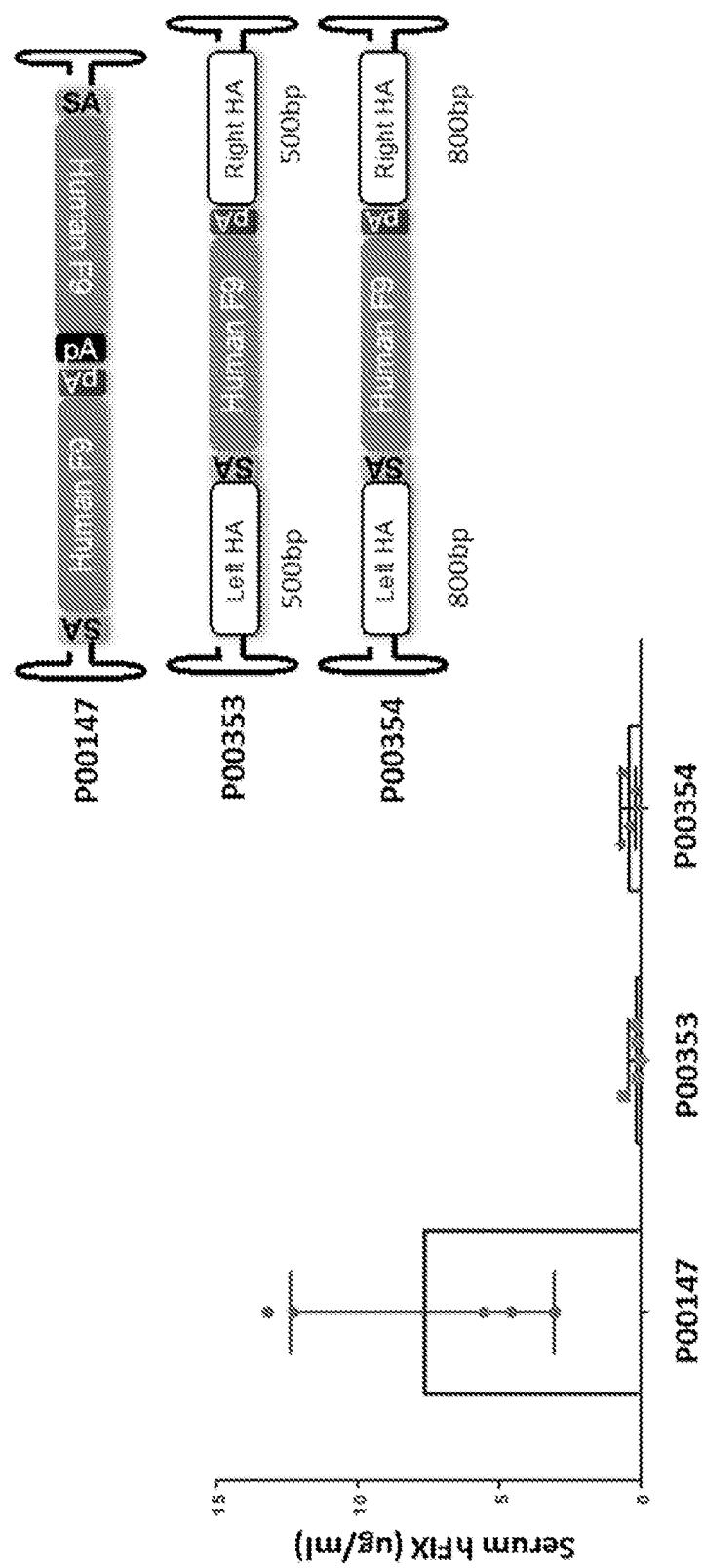


FIG. 4B

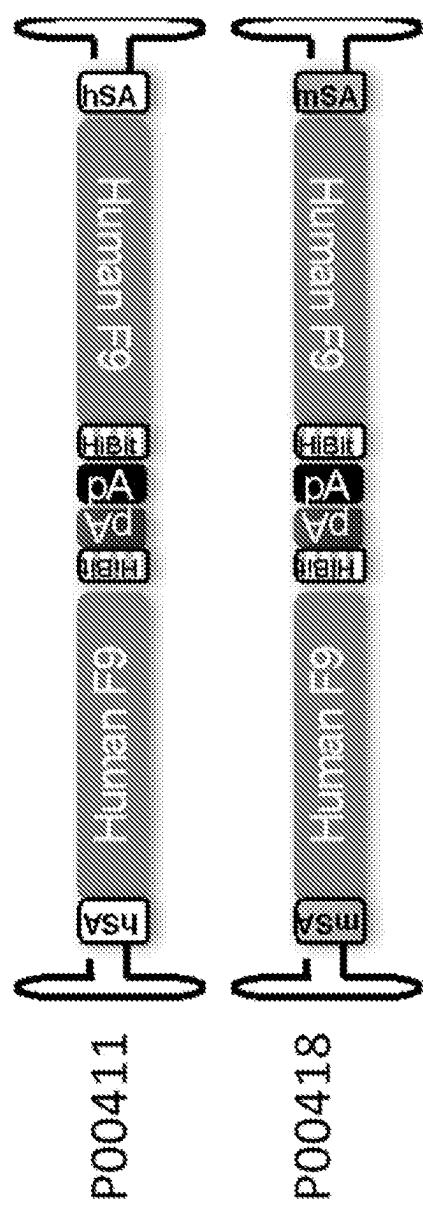
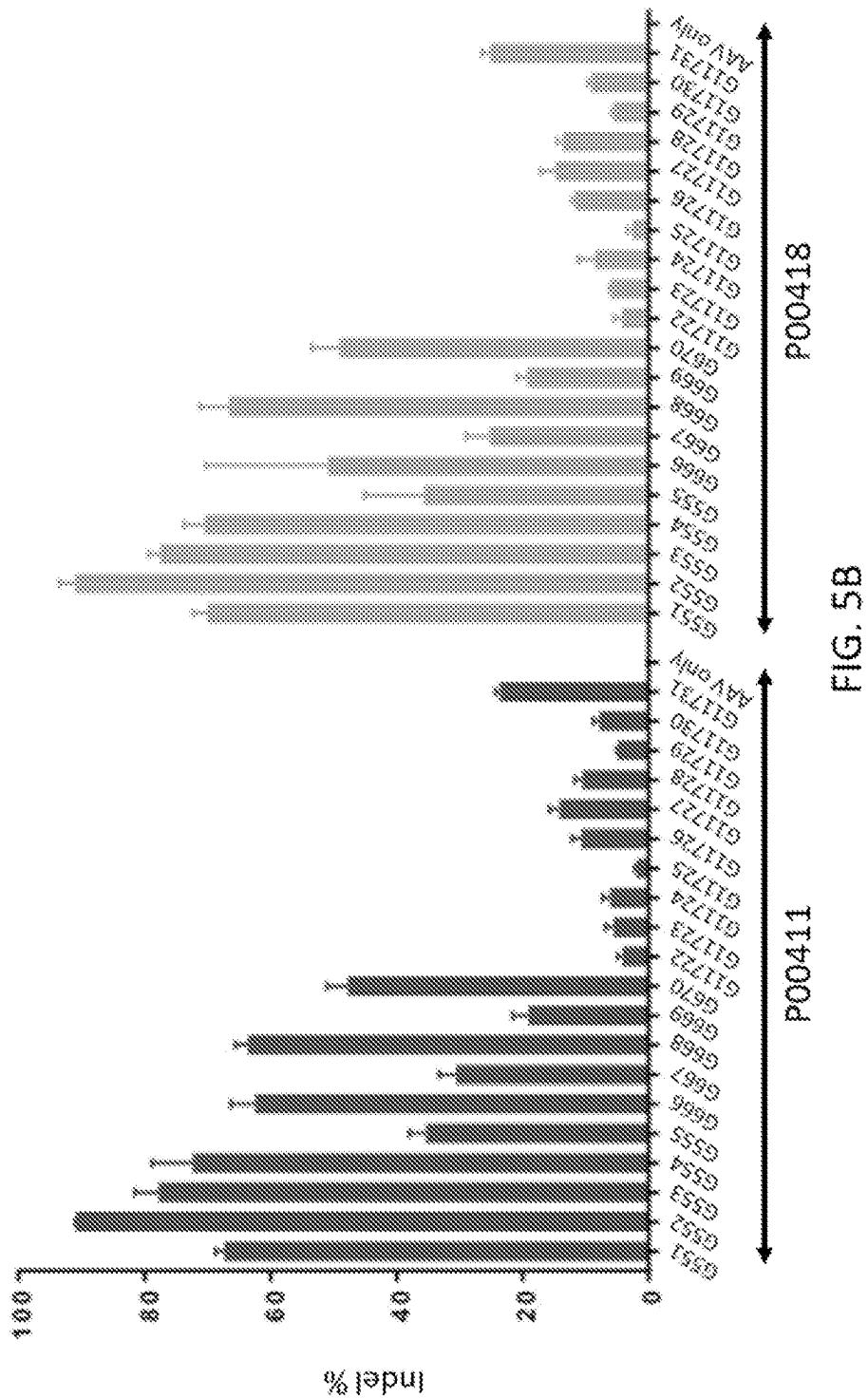
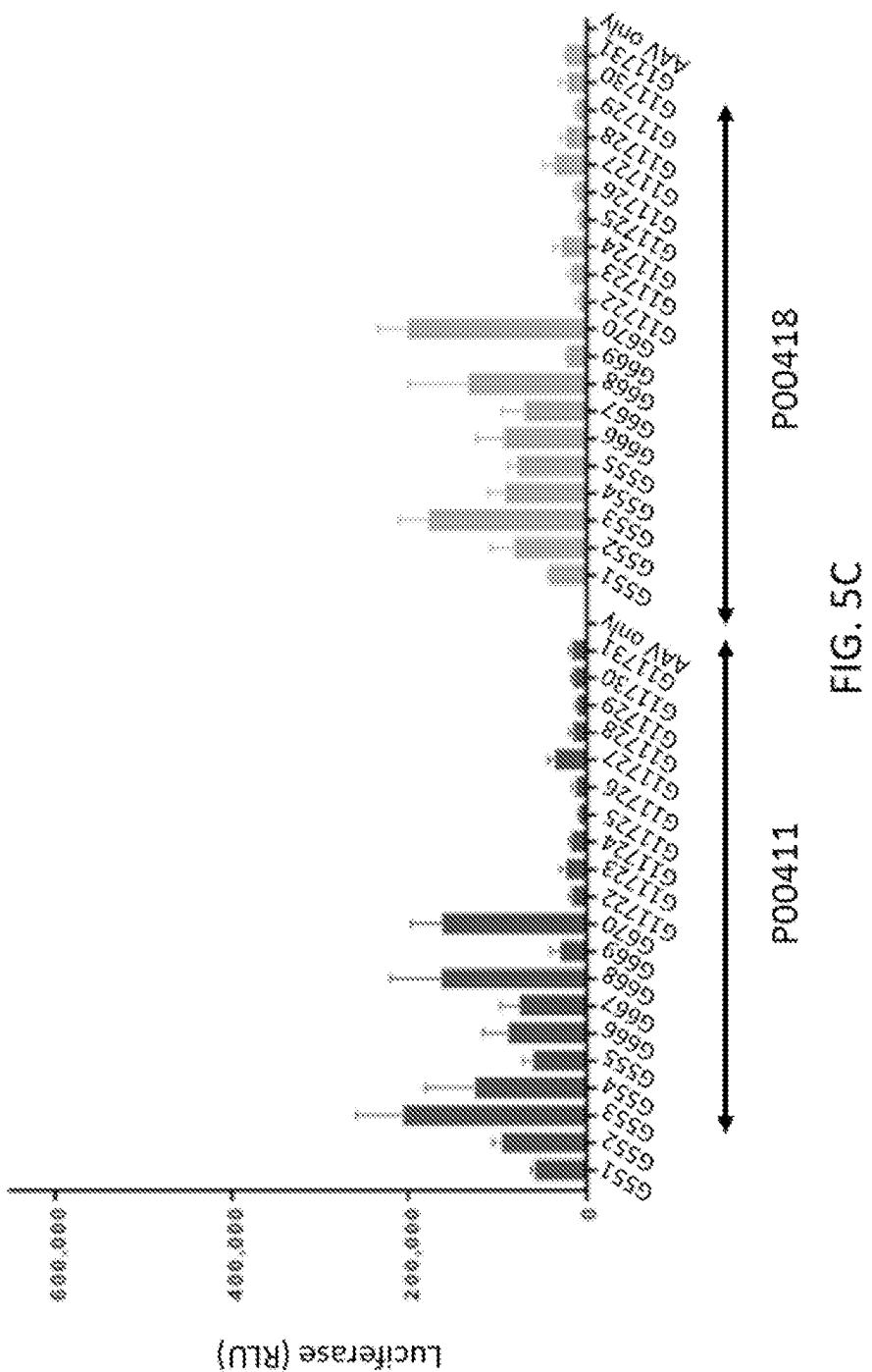


FIG. 5A





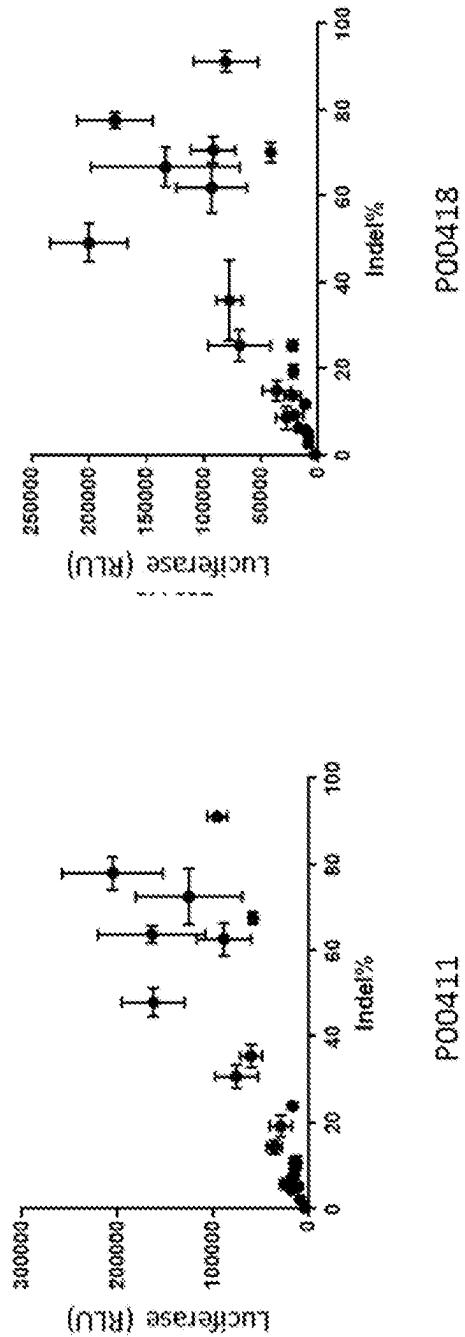
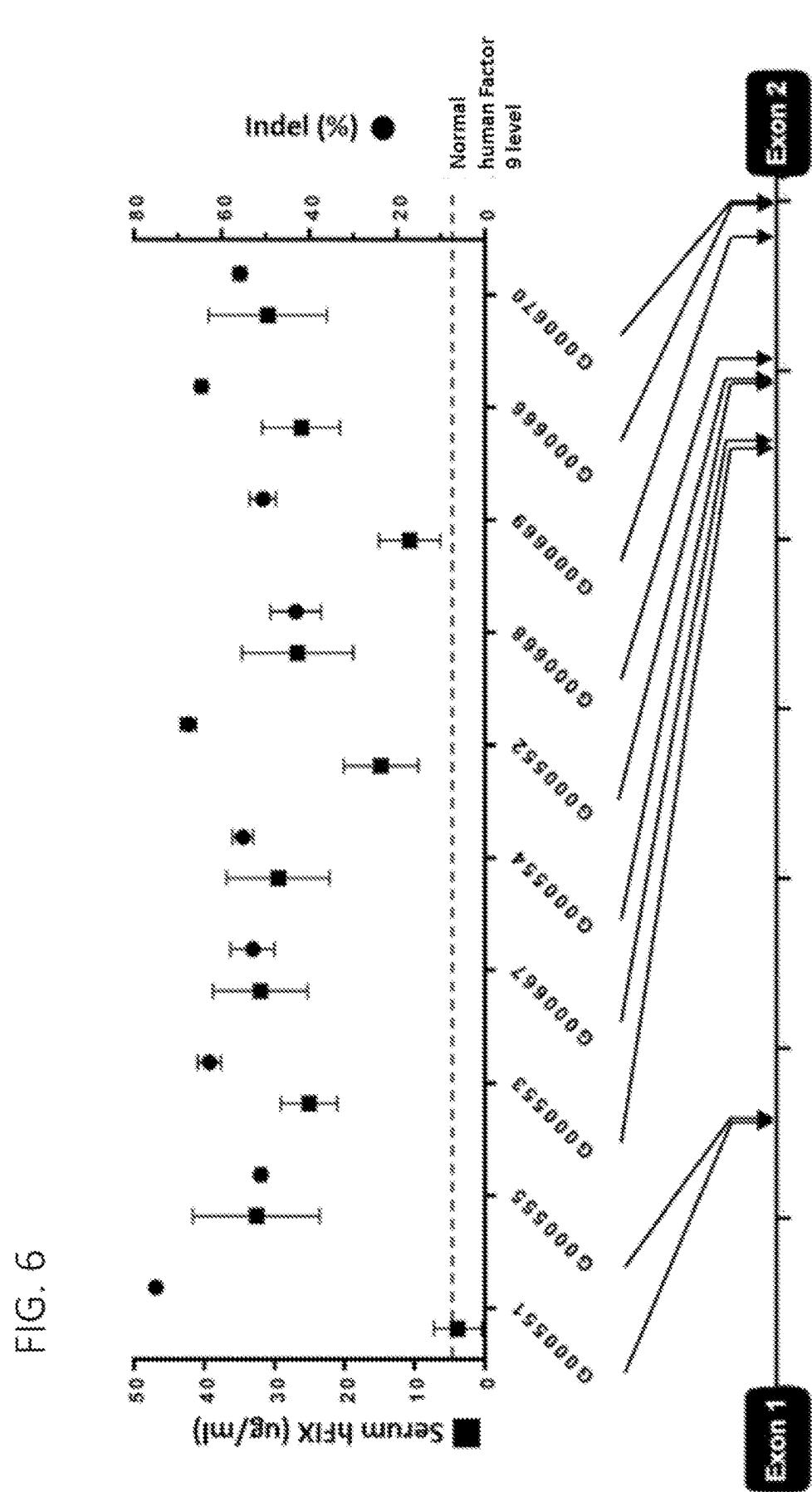
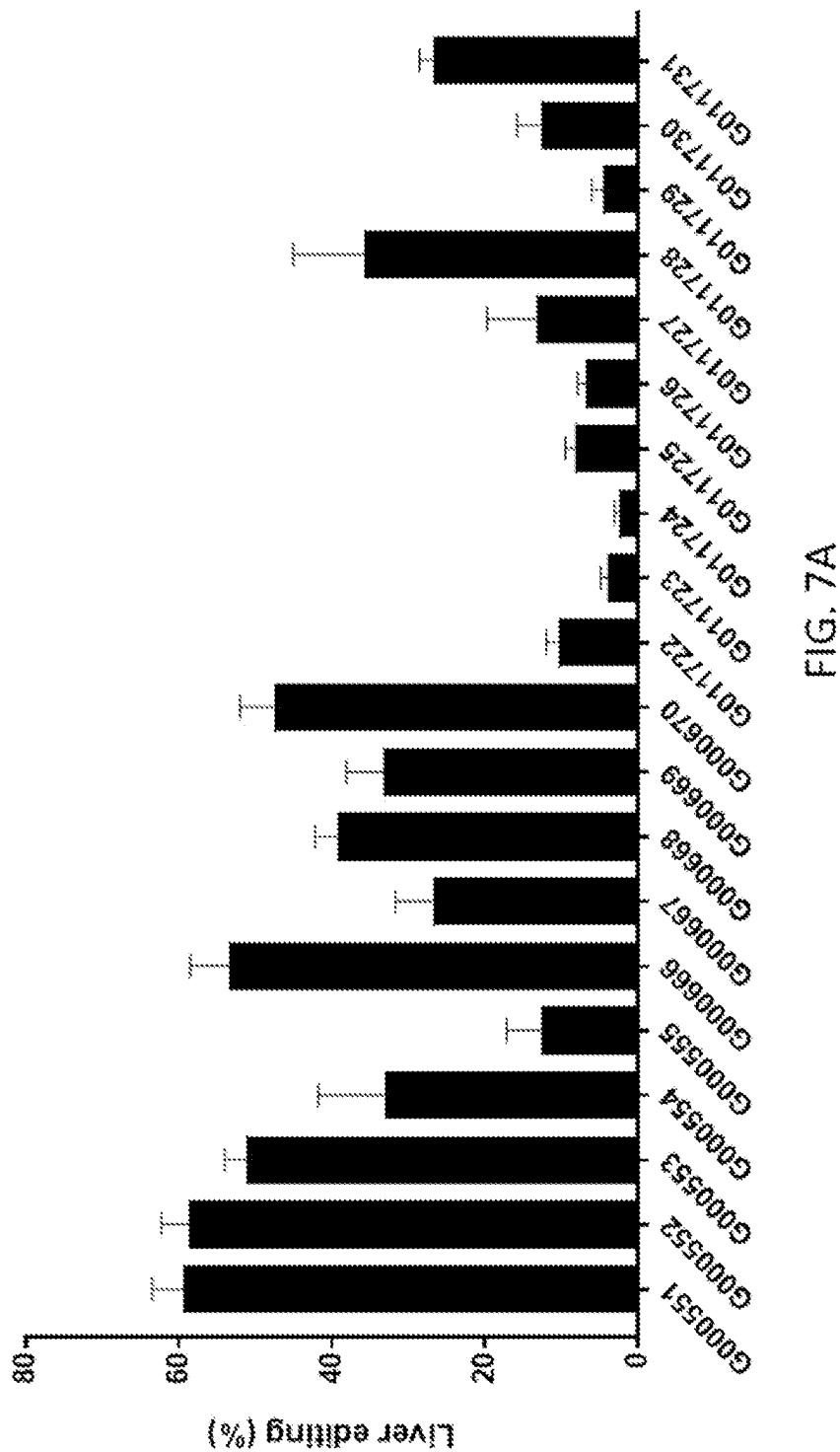
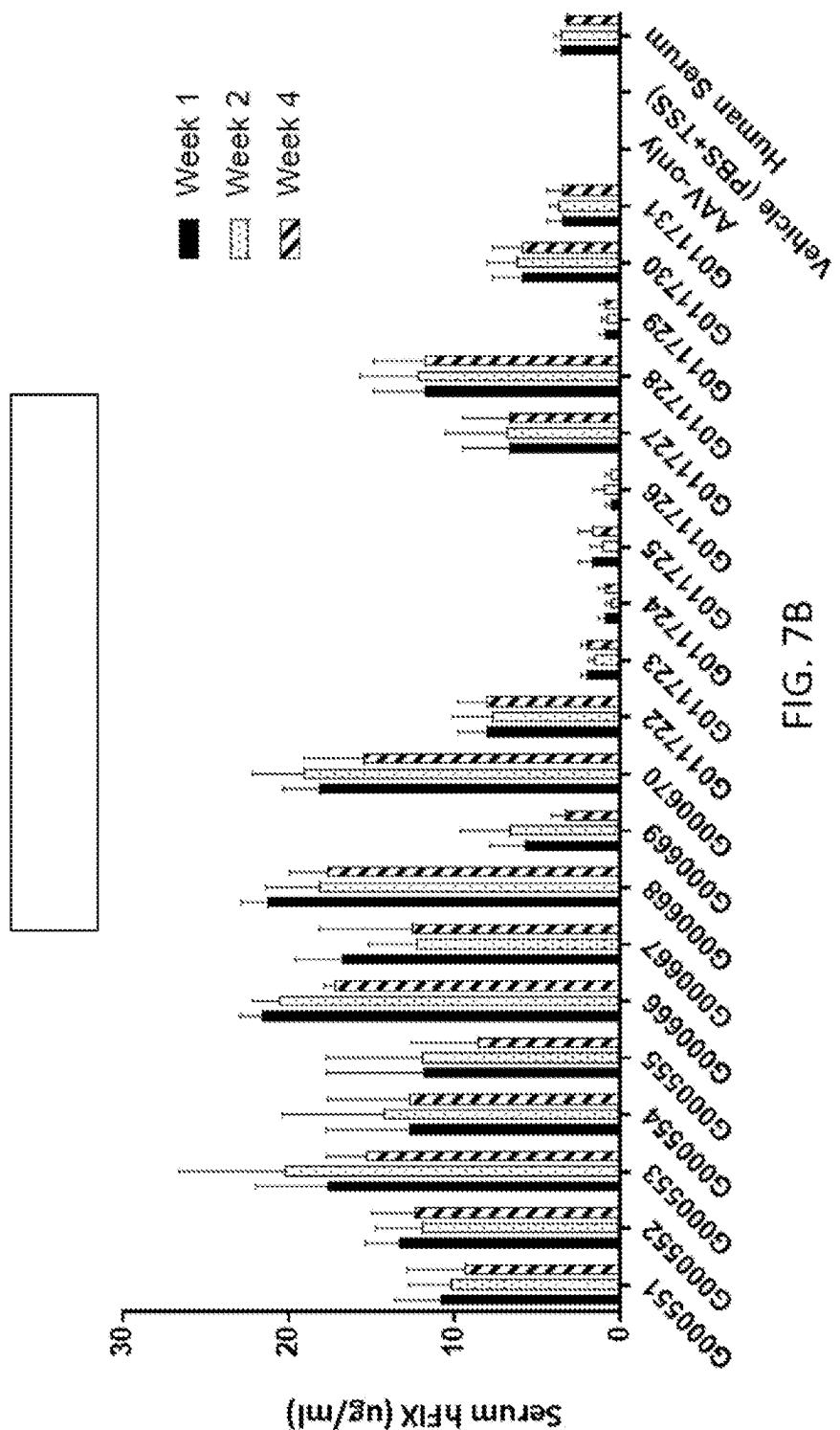


FIG. 5D







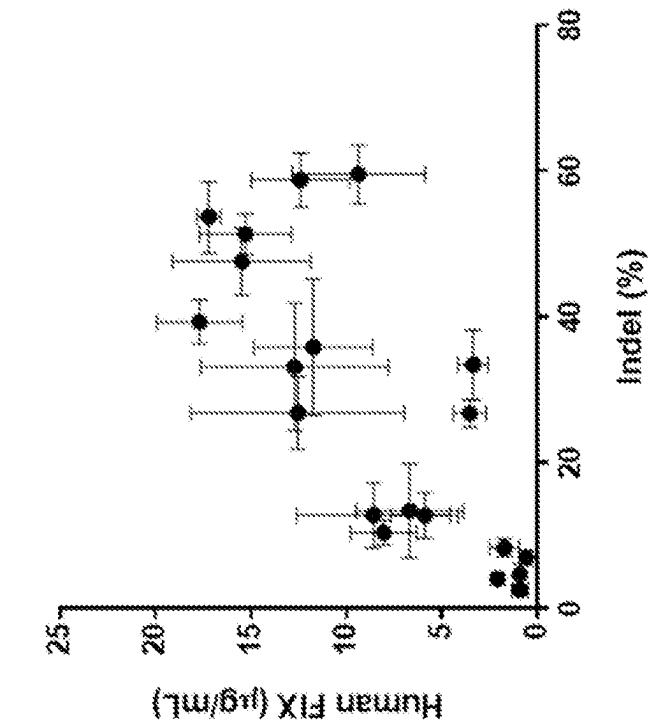


FIG. 7D

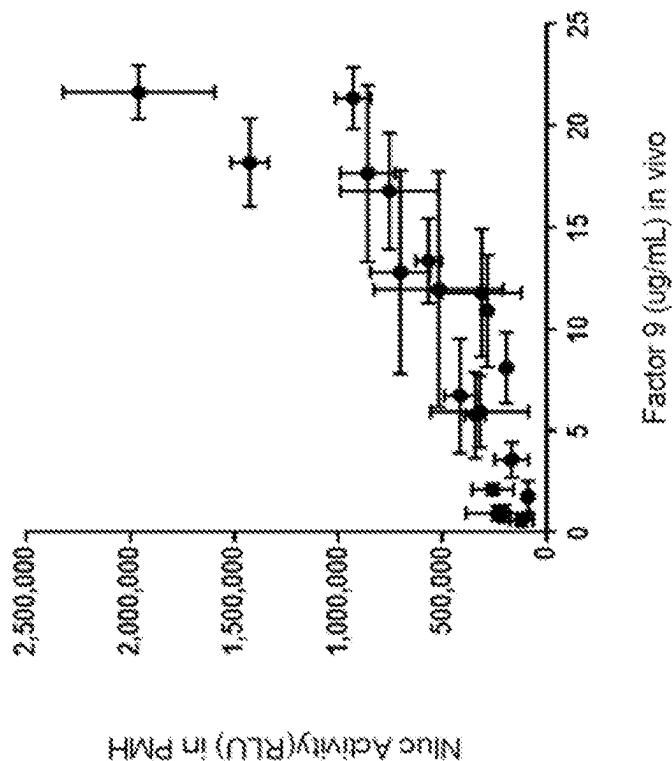


FIG. 7C

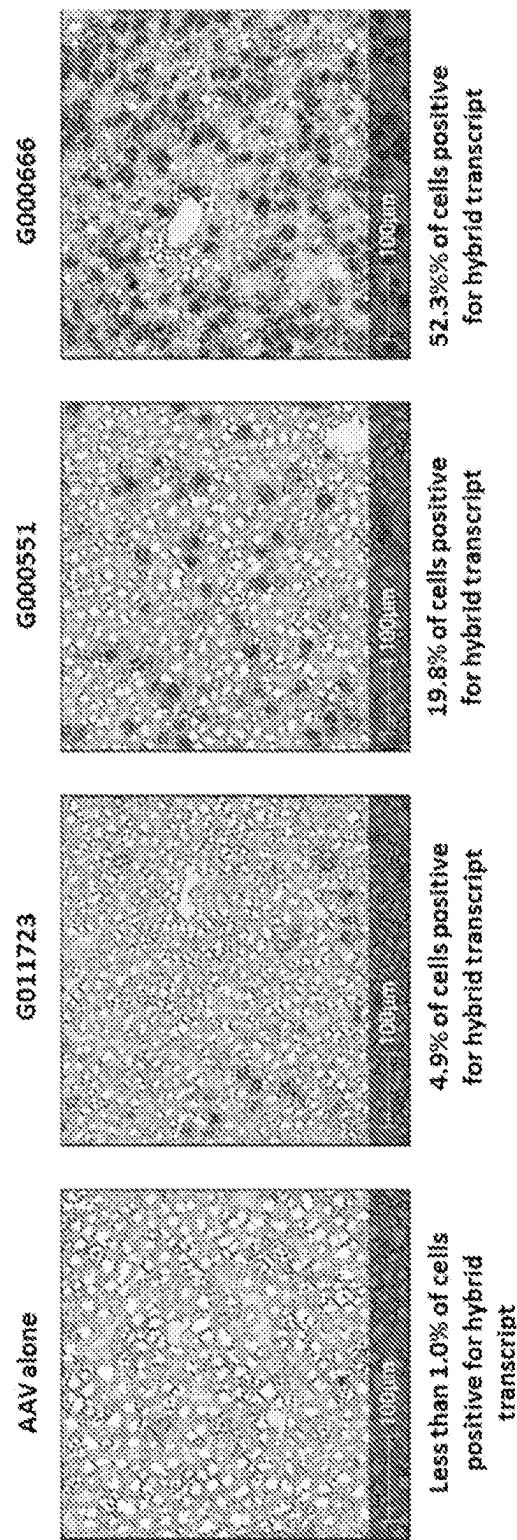
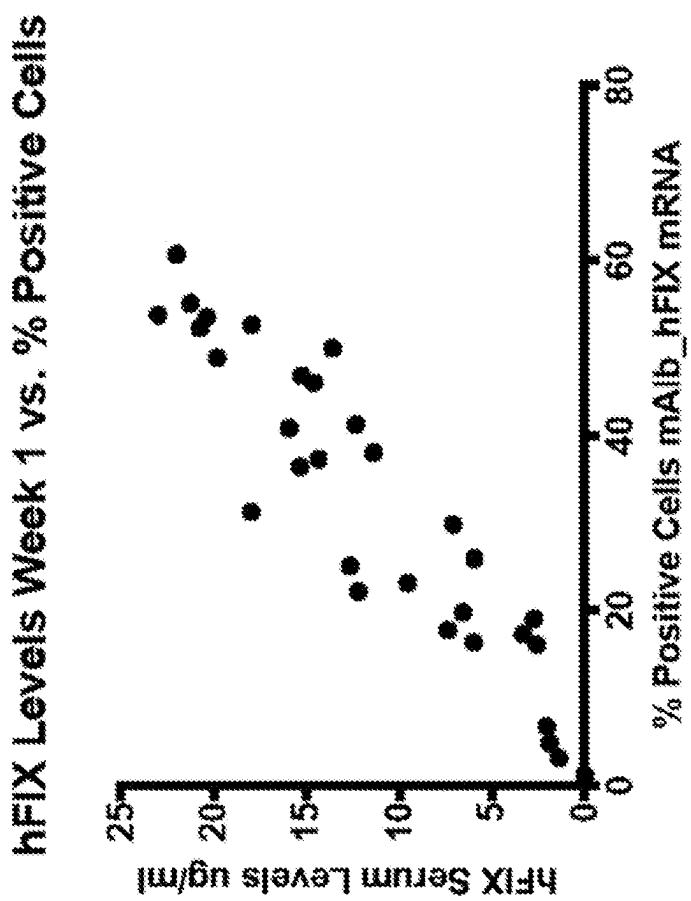


FIG. 8A

FIG. 8B



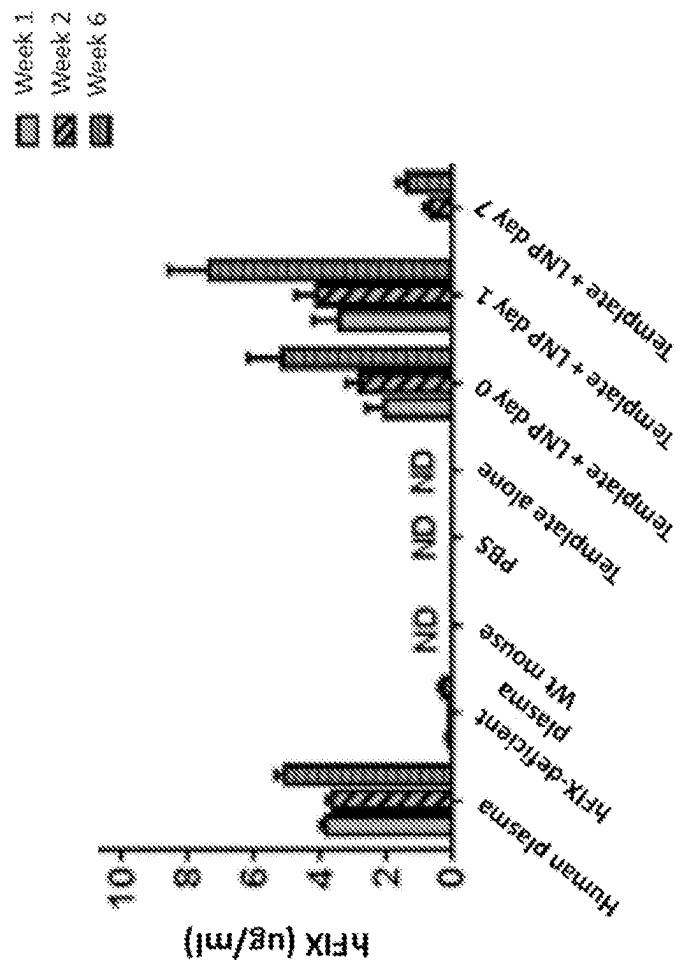
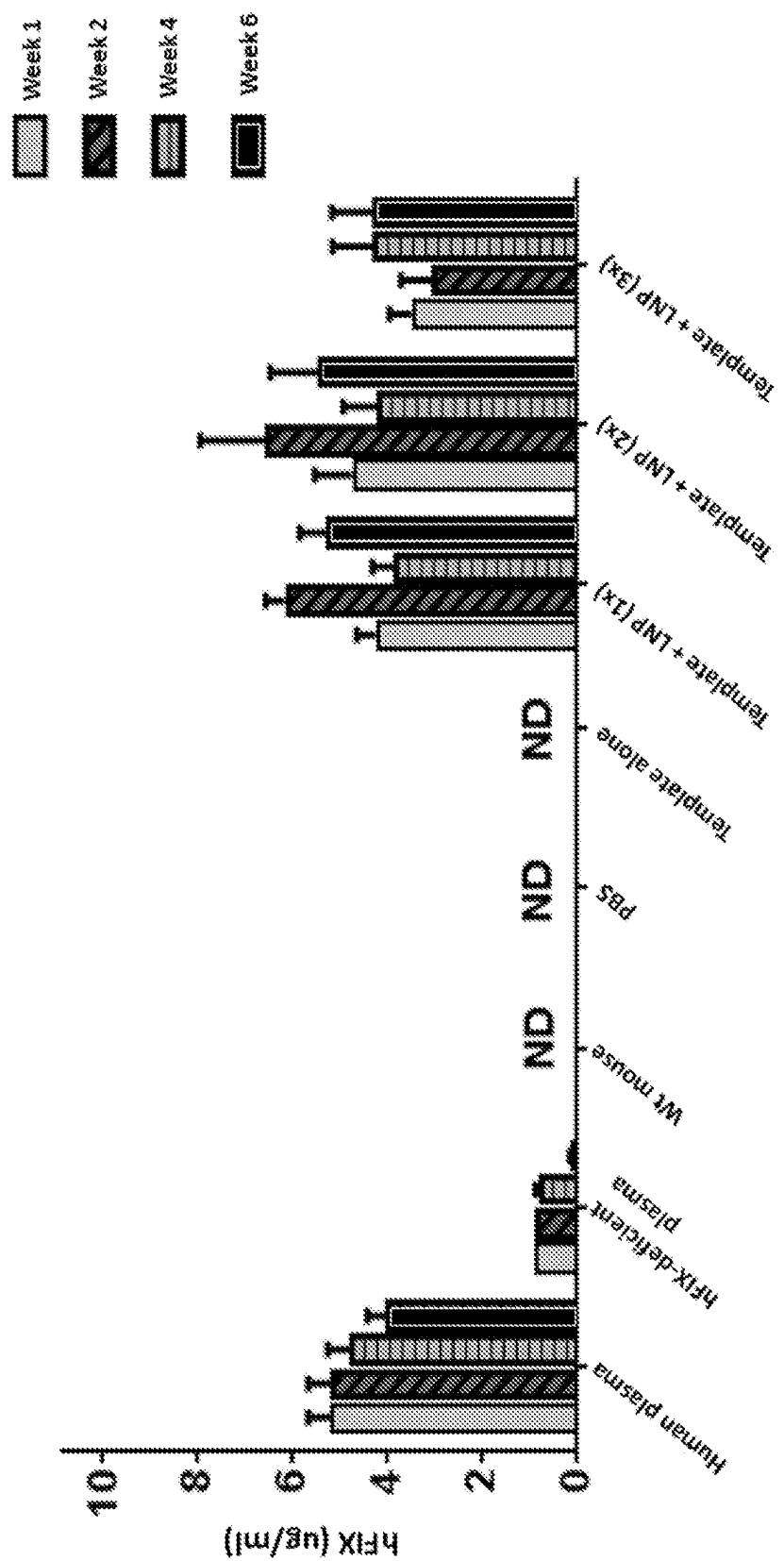
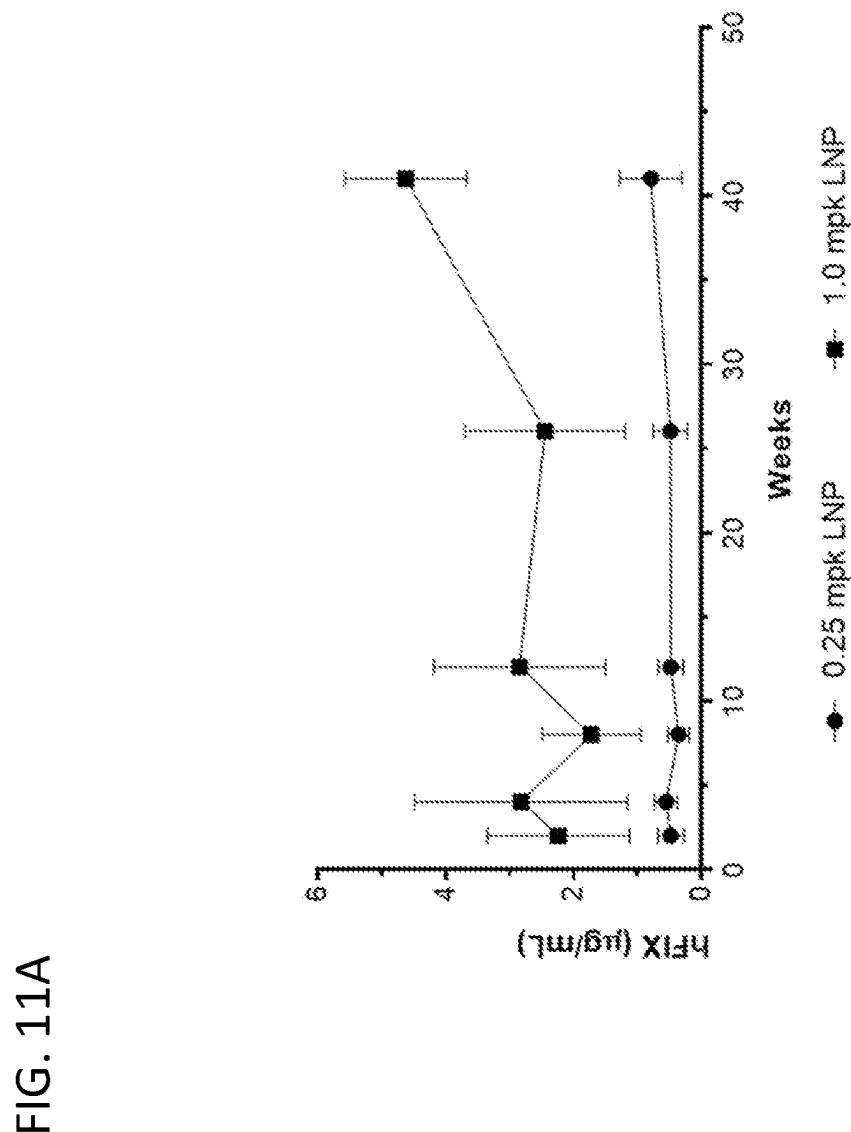


FIG. 9

FIG. 10





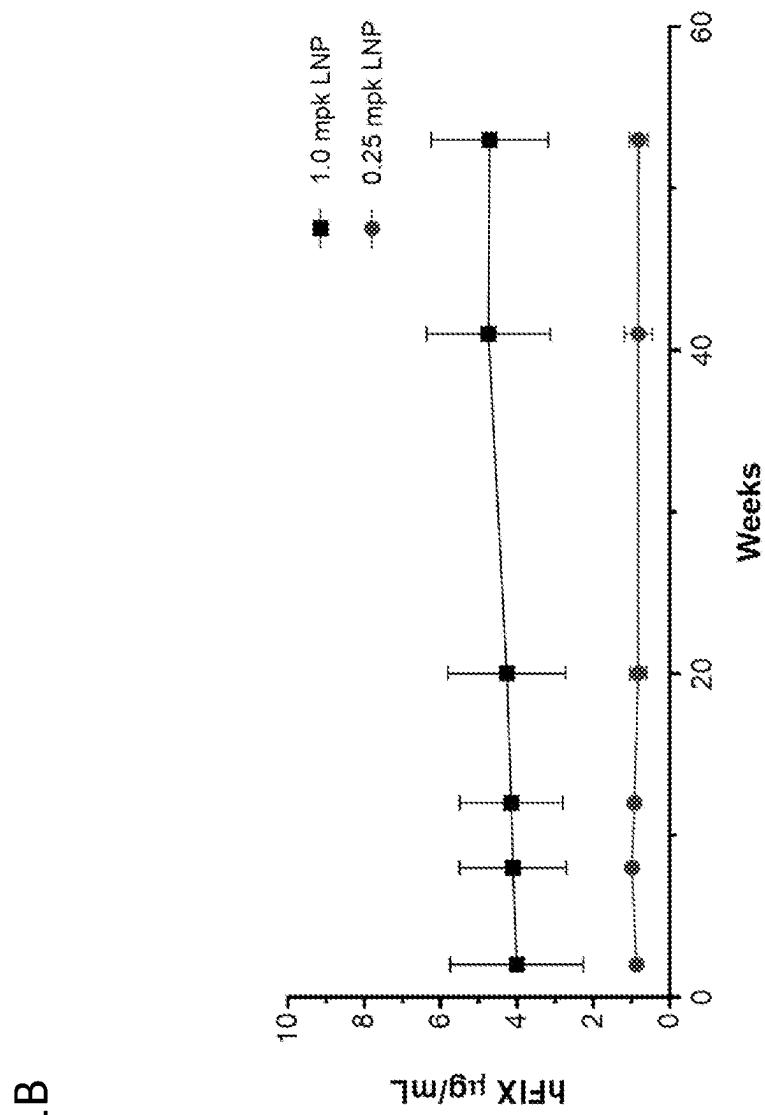
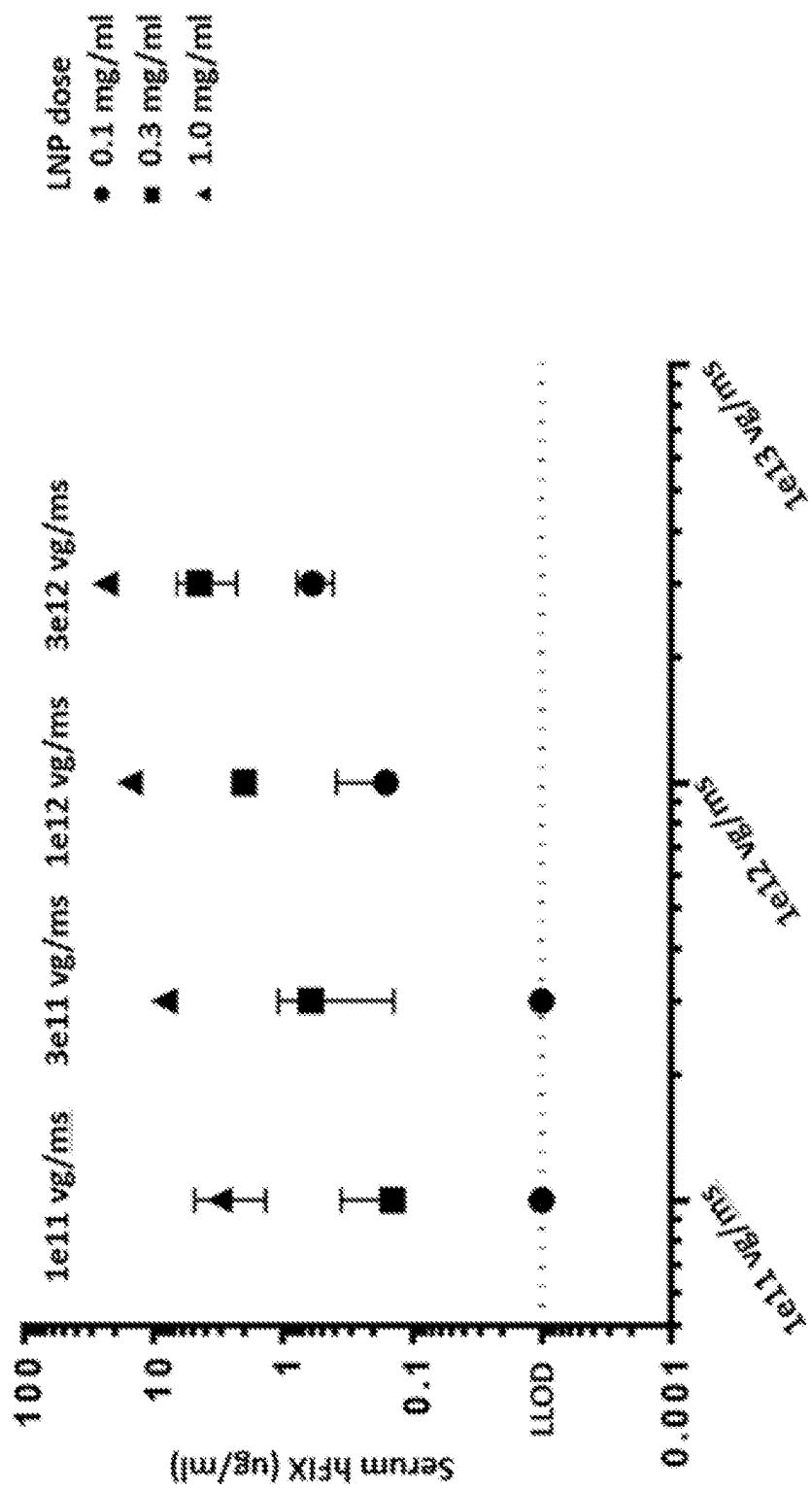
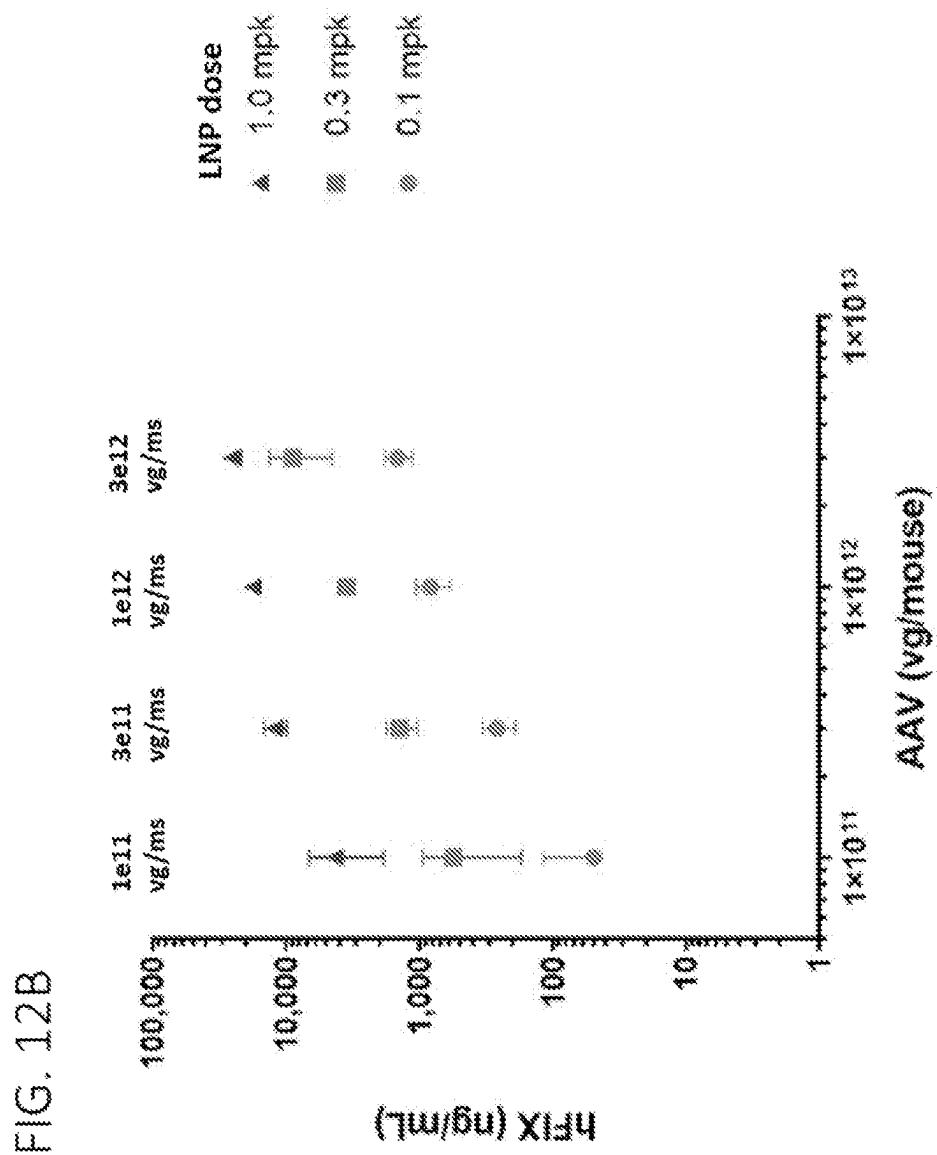
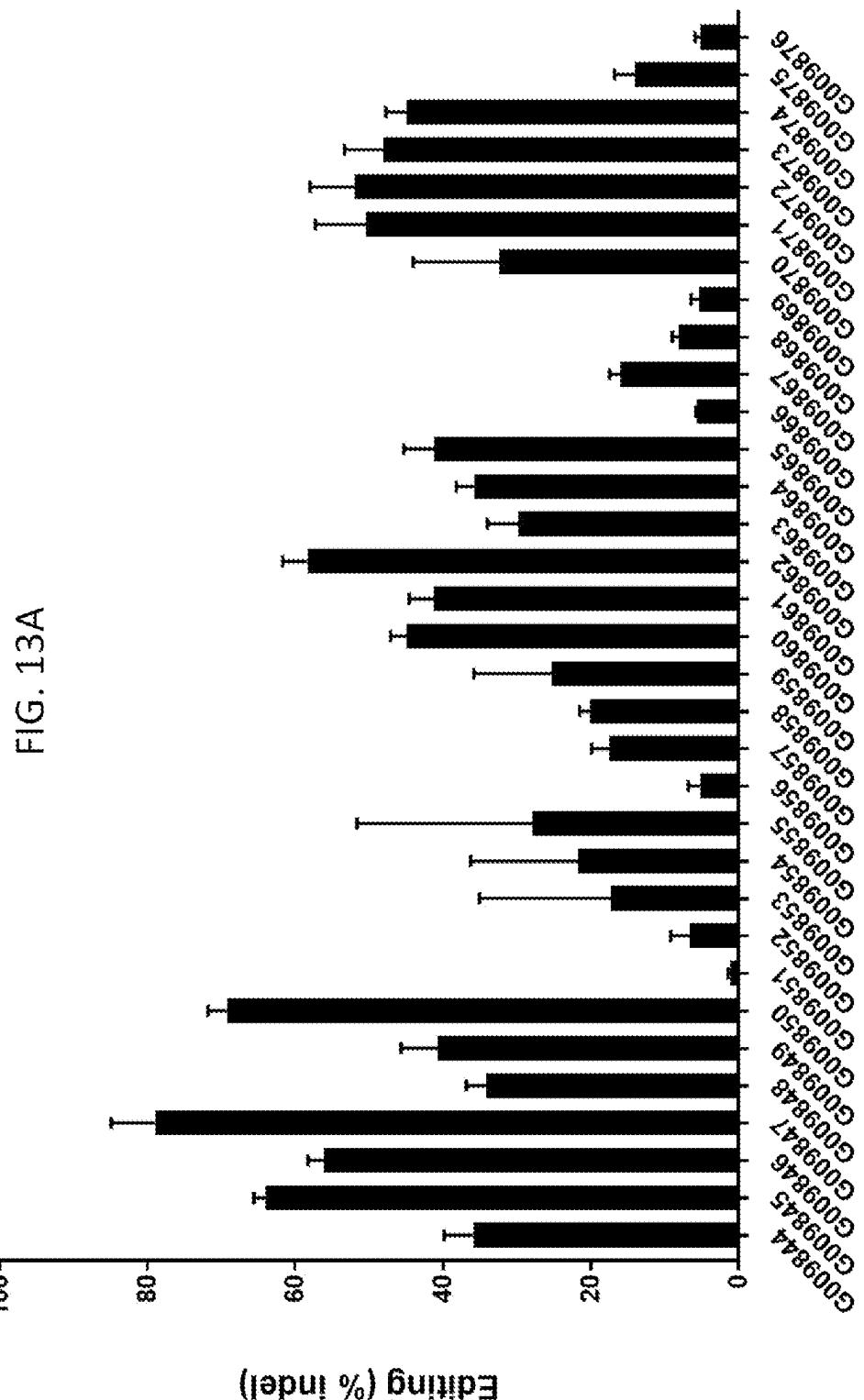


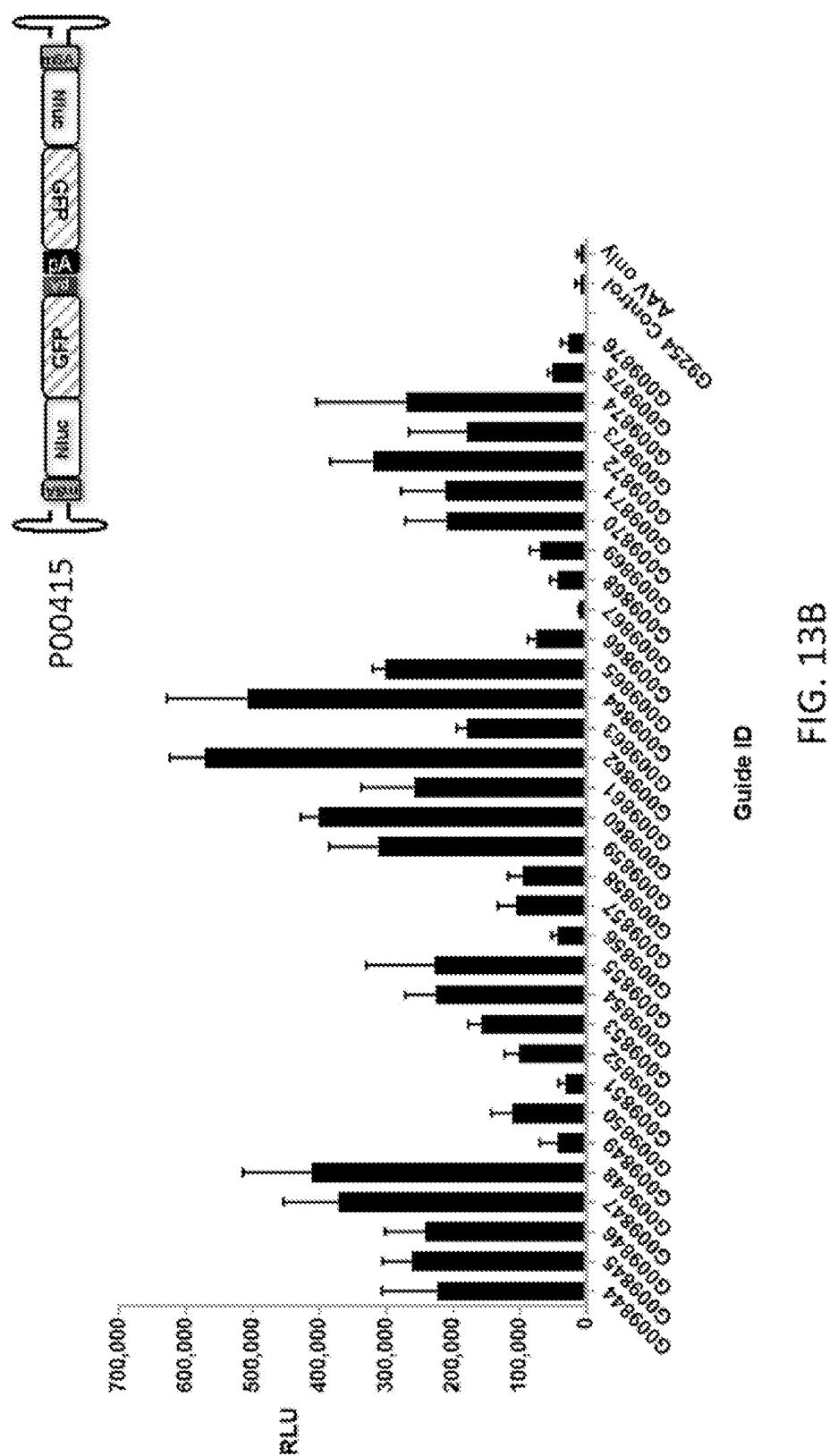
FIG. 11B

FIG. 12A









138

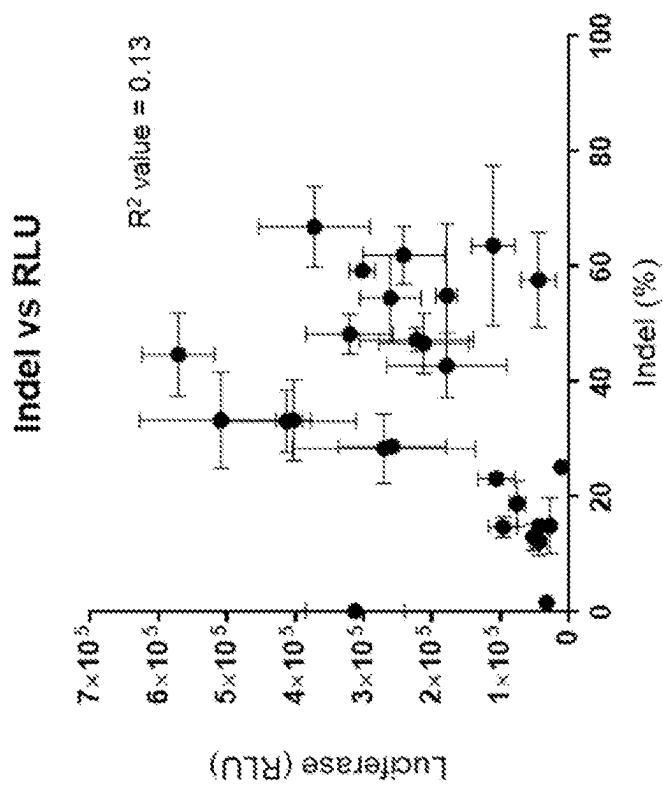
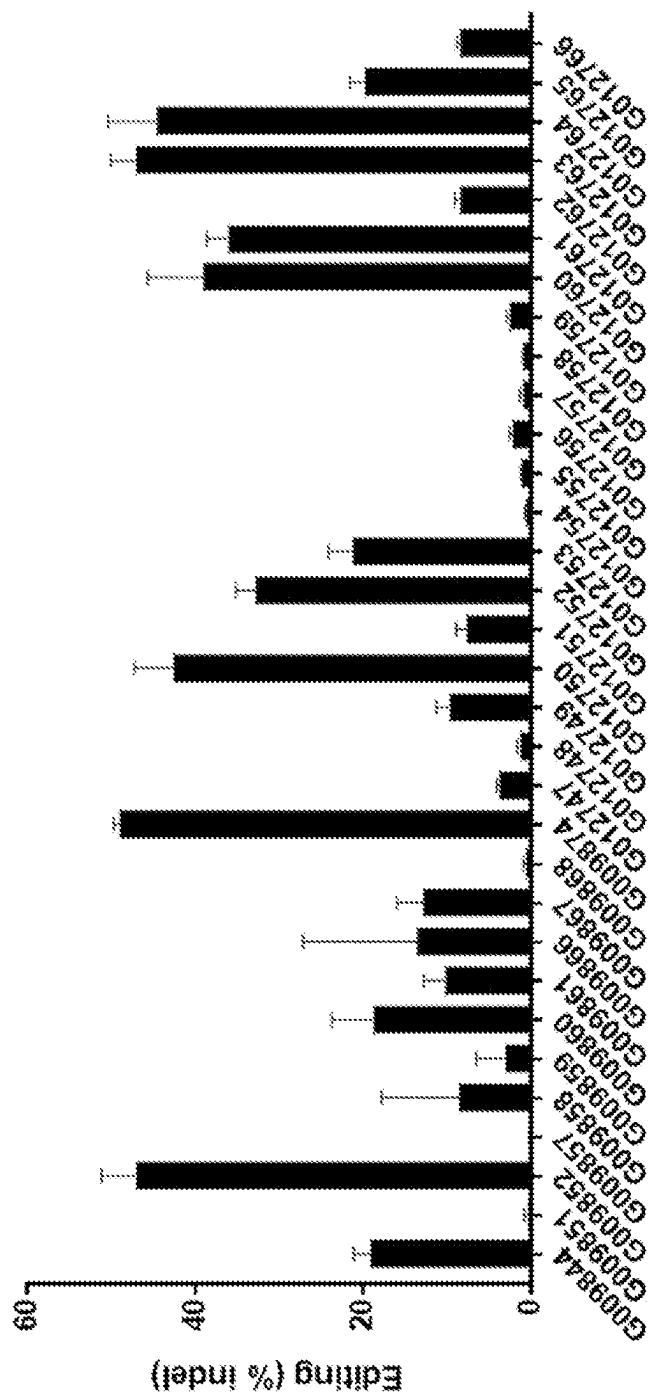
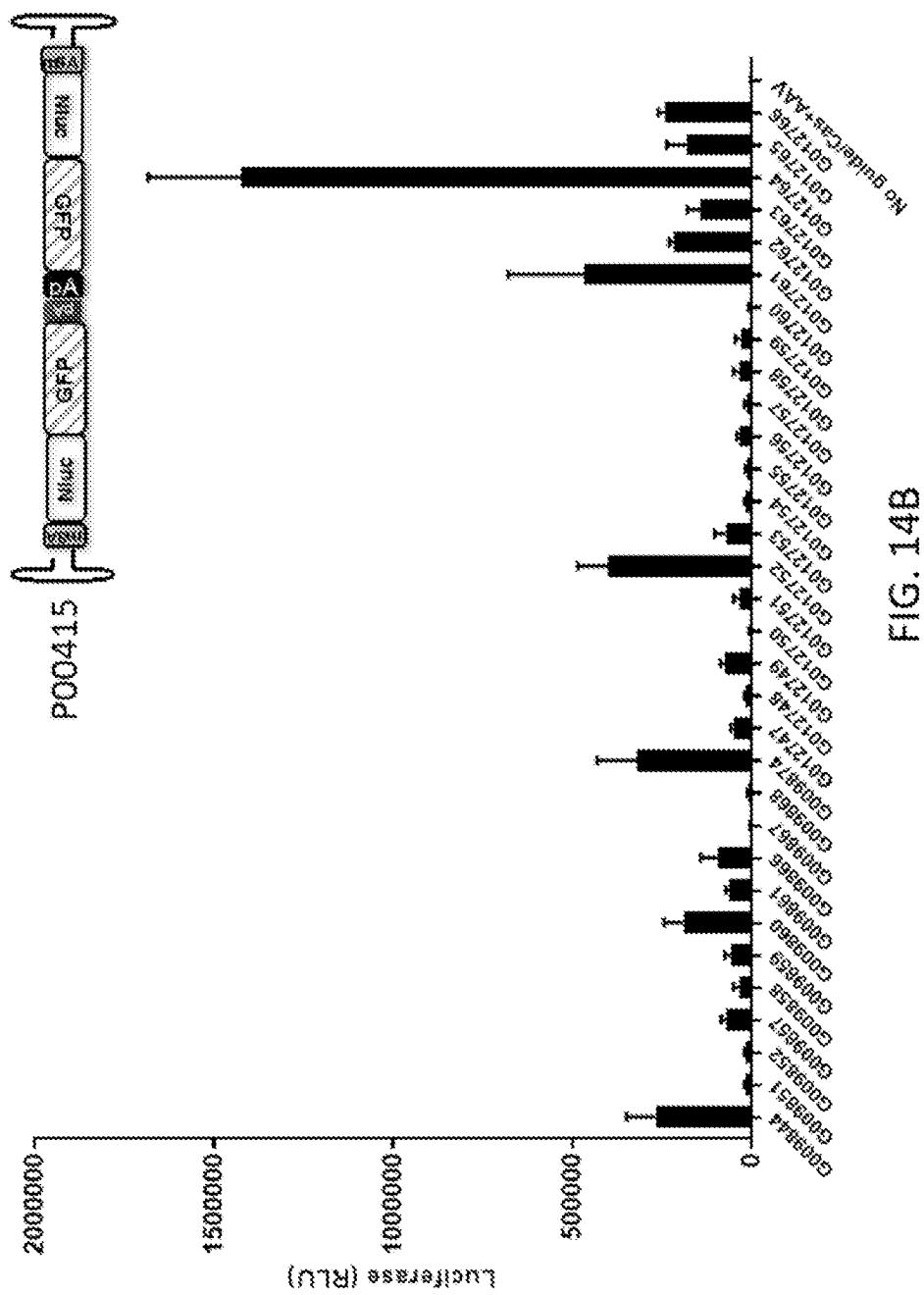


FIG. 13C





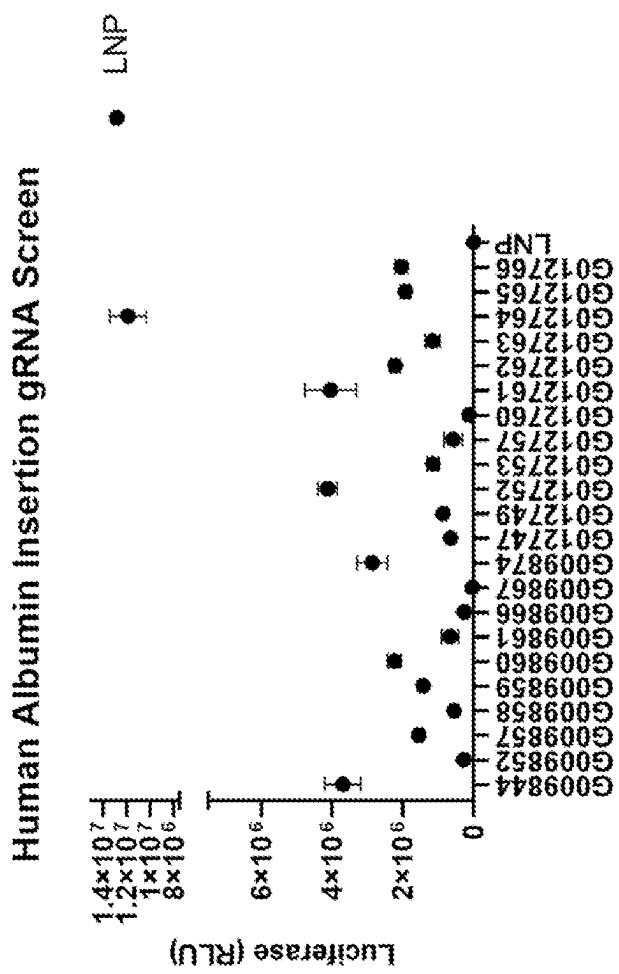


FIG. 14C

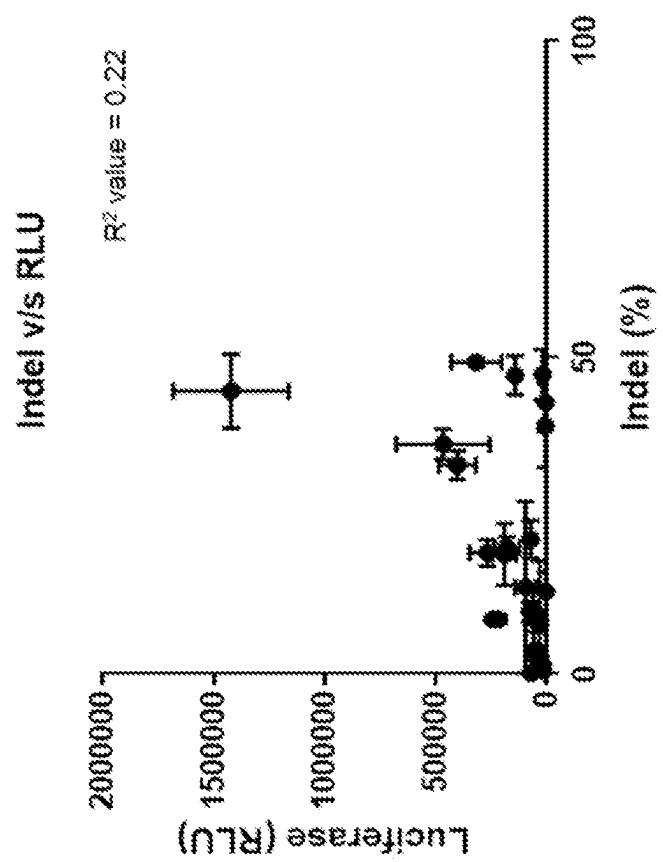
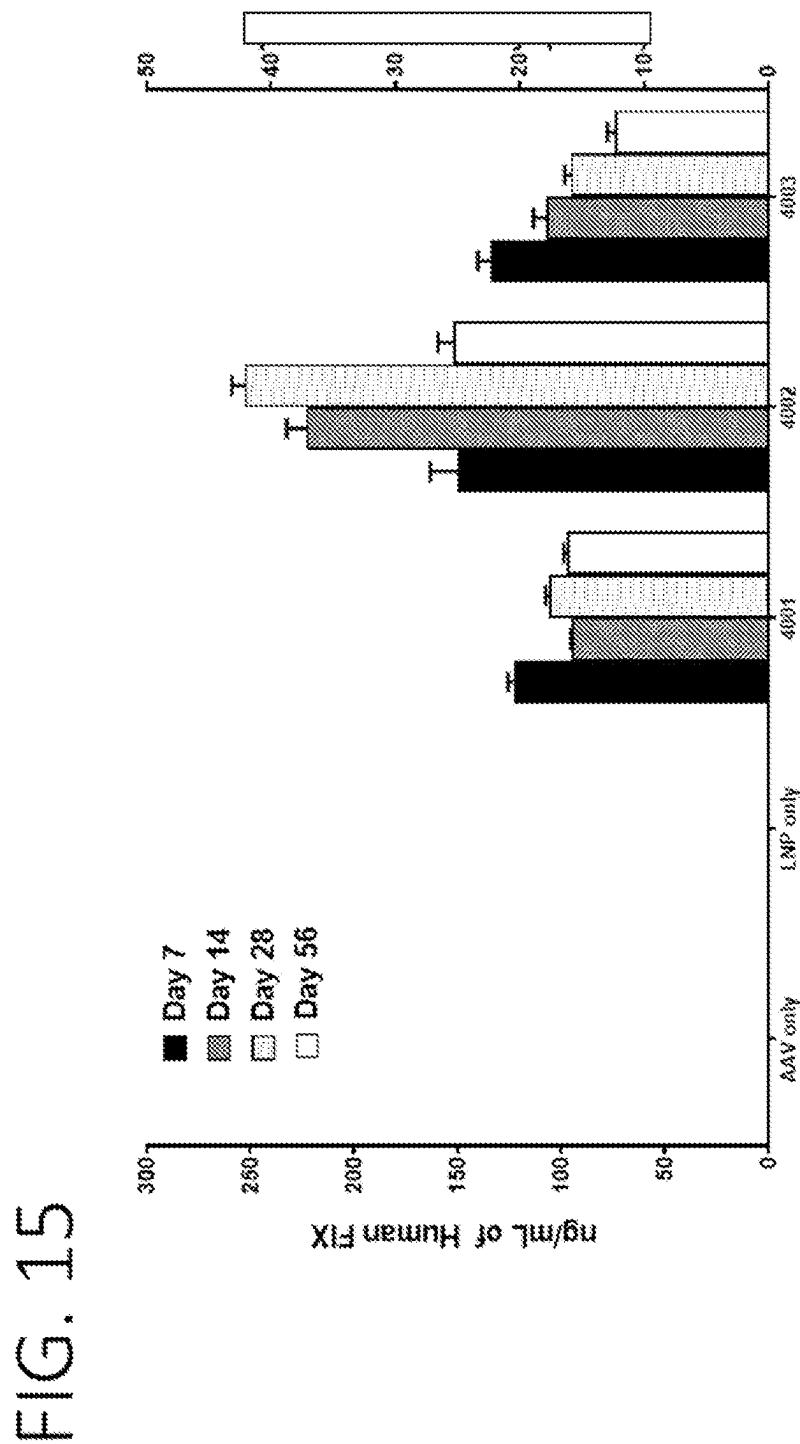


FIG. 14D



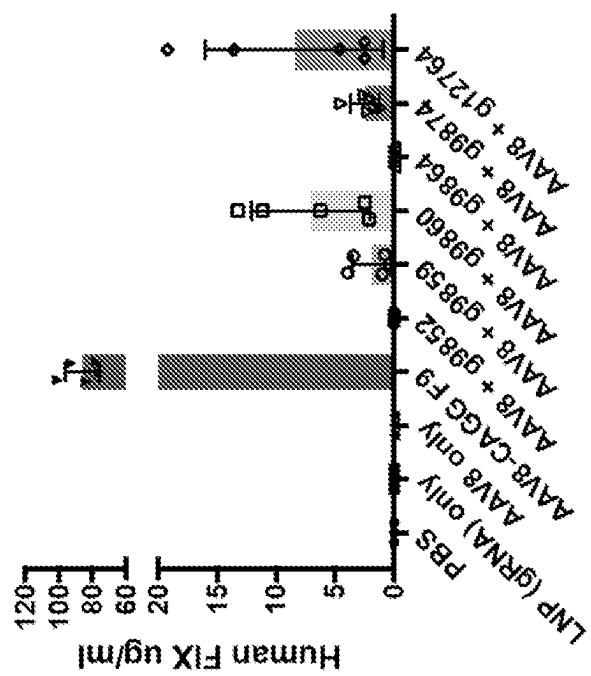


FIG. 16A

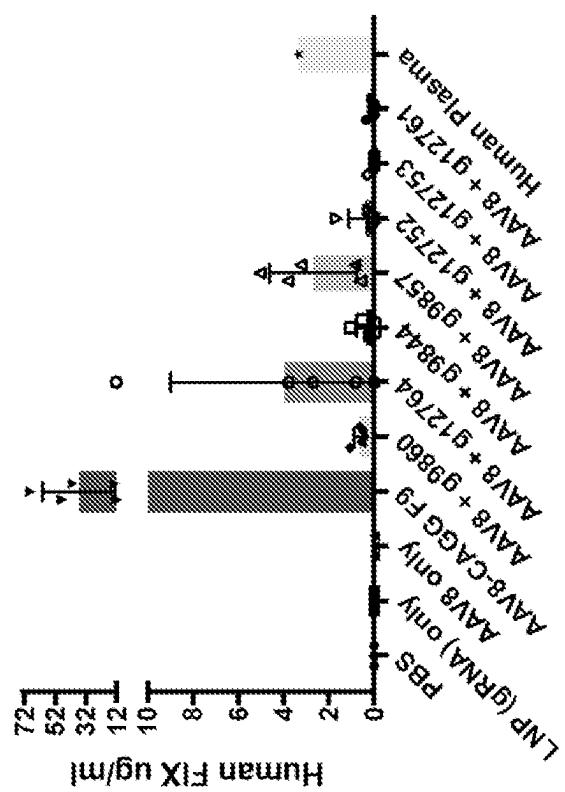


FIG. 16B

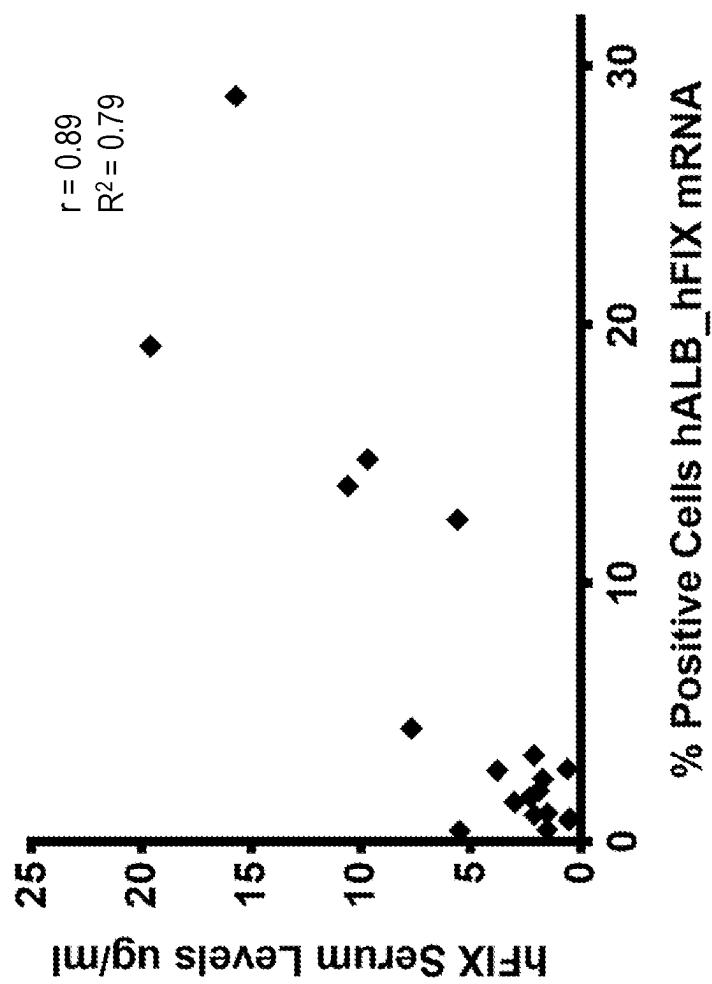


FIG. 17

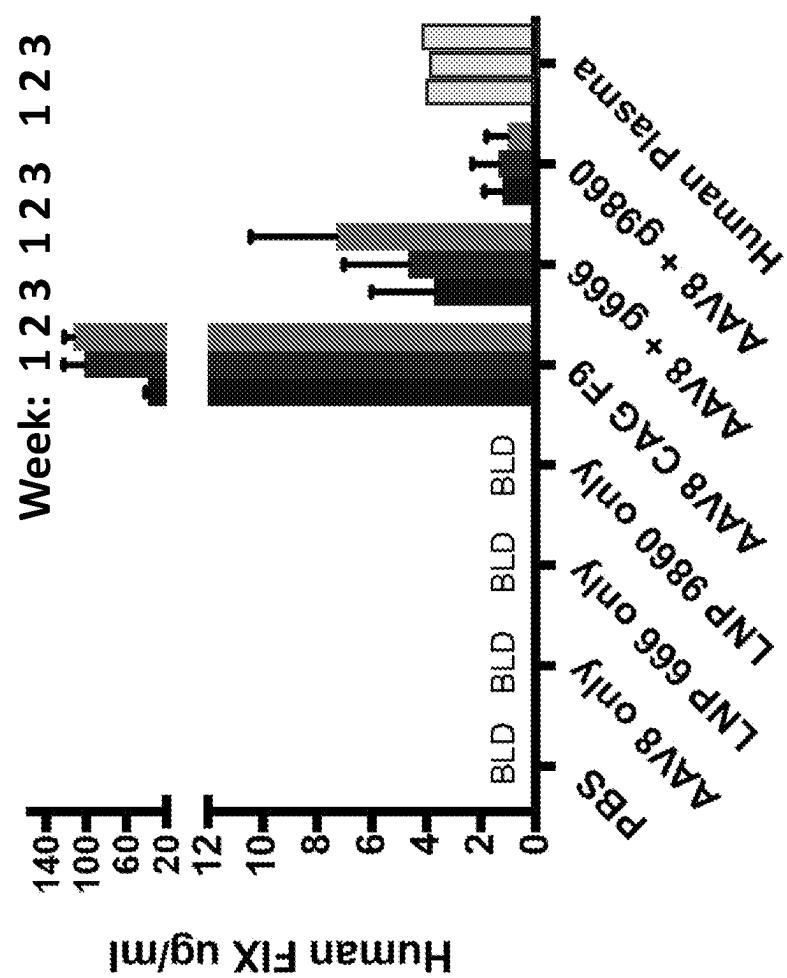


FIG. 18

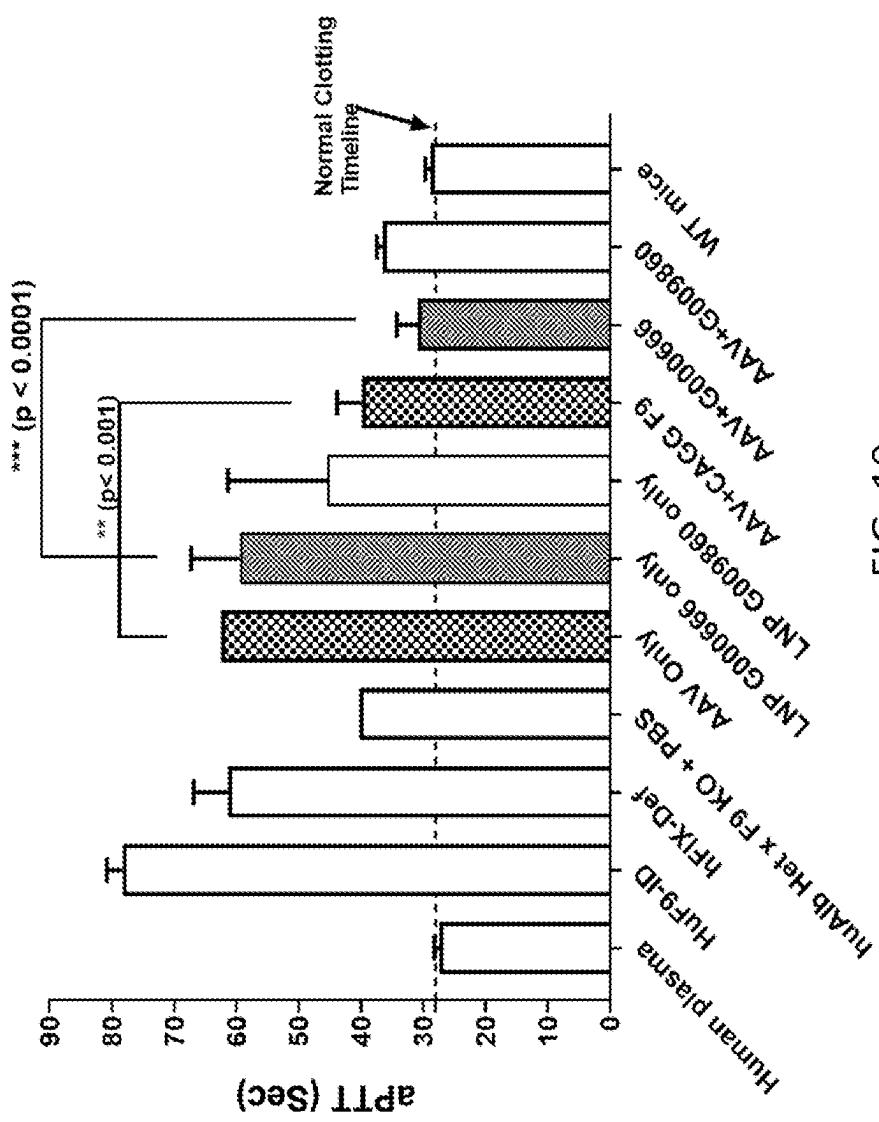


FIG. 19

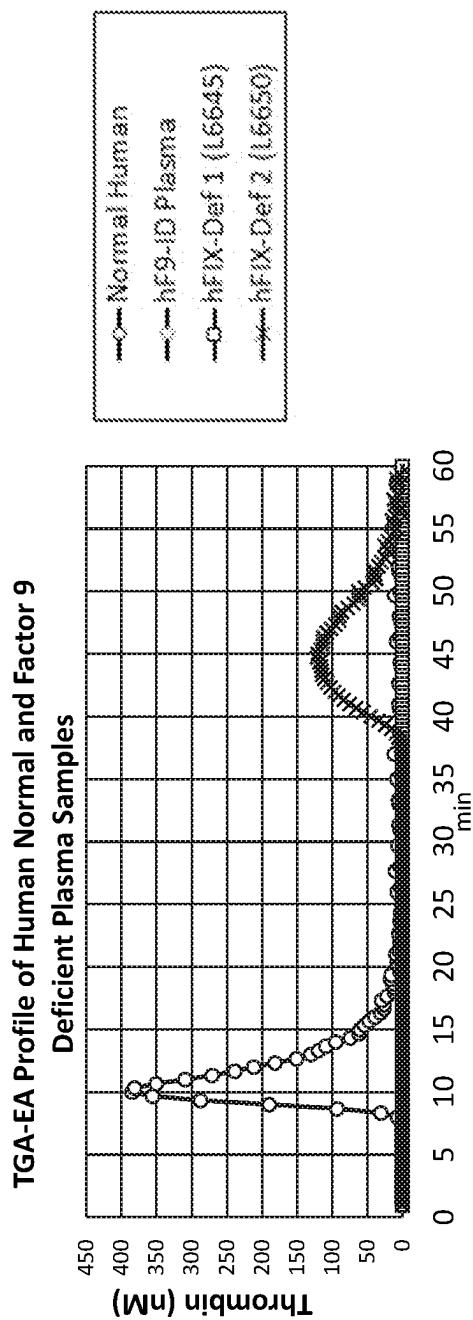


FIG. 20A
TGA-EA Profile of Mice Plasma from hF9 Insertion in huAlb Het X F9 KO Mice Study

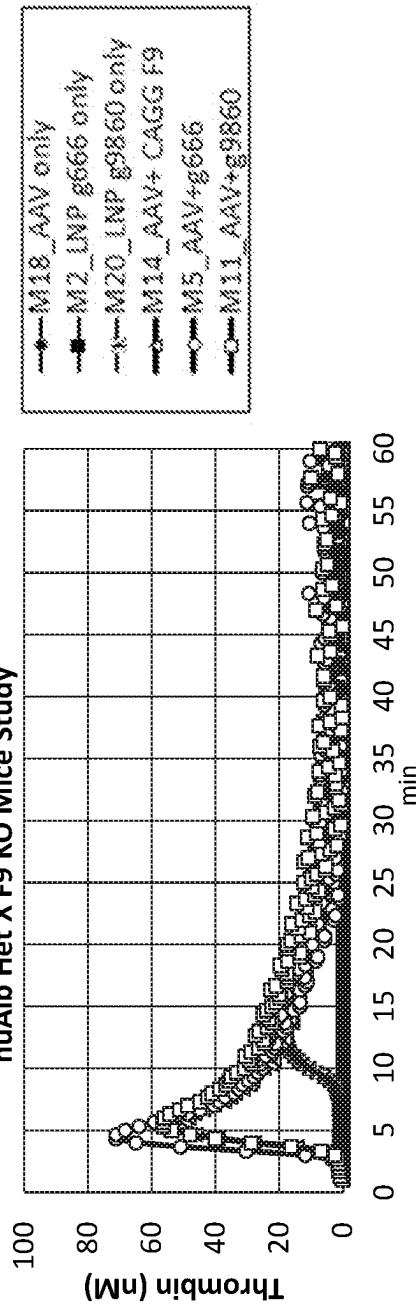


FIG. 20B

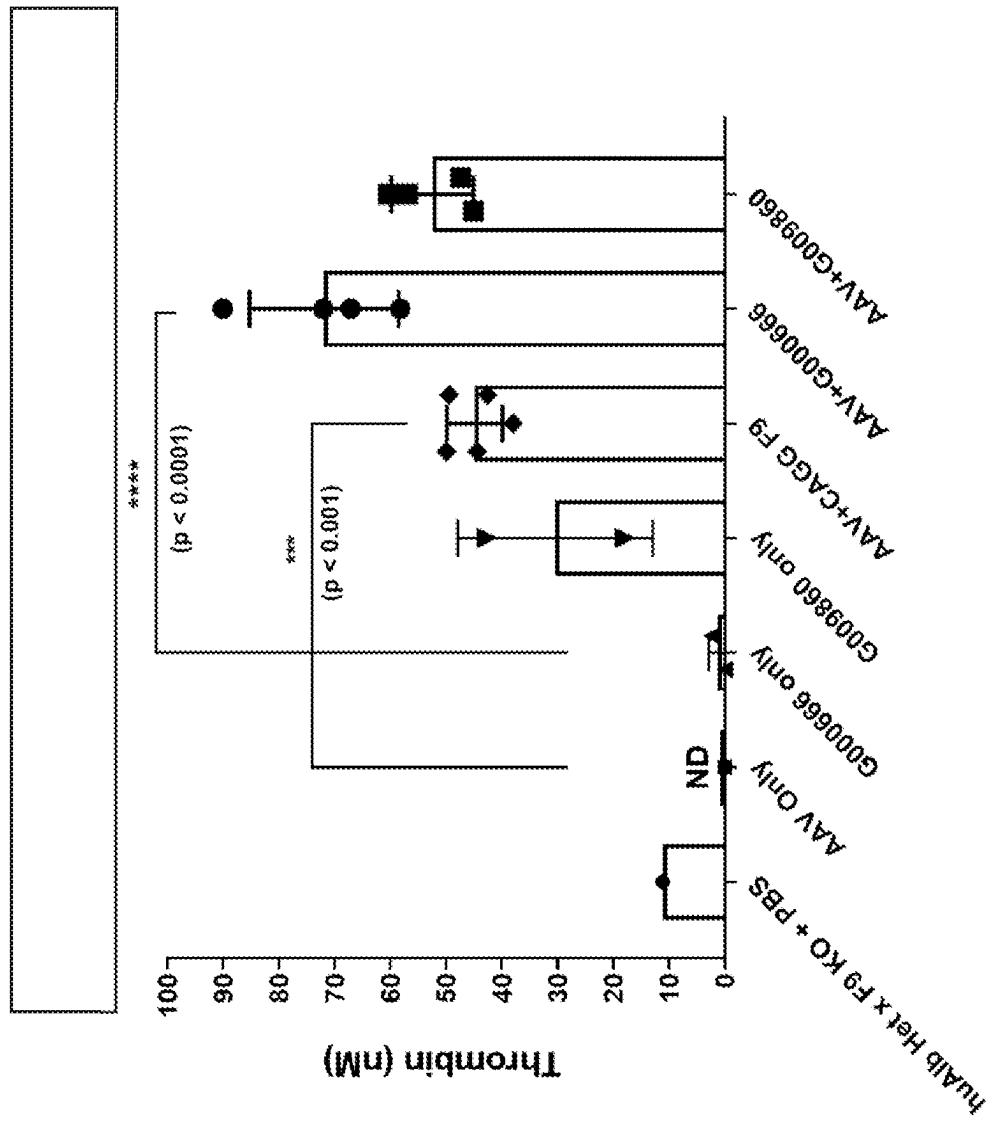


FIG. 21

COMPOSITIONS AND METHODS FOR EXPRESSING FACTOR IX

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a Continuation of U.S. Non-Provisional application Ser. No. 16/657,961, filed Oct. 18, 2019, which claims the benefit of priority from U.S. Provisional Application No. 62/747,509, filed on Oct. 18, 2018, U.S. Provisional Application No. 62/829,009, filed on Apr. 3, 2019, U.S. Provisional Application No. 62/829,621, filed on Apr. 4, 2019 and U.S. Provisional Application No. 62/840,352, filed on Apr. 29, 2019, each of which is hereby incorporated by reference in its entirety.

REFERENCE TO A SEQUENCE LISTING SUBMITTED AS AN XML FILE

[0002] The Sequence Listing written in file 625637SEQLIST.xml is 484,218 bytes, was created on May 1, 2025, and is hereby incorporated by reference in its entirety.

BACKGROUND AND SUMMARY

[0003] Bleeding disorders are caused by inadequate blood clotting. This deficiency may be caused by congenital coagulation disorders, acquired coagulation disorders, or hemorrhagic conditions induced by trauma. Bleeding is one of the most serious and significant manifestations of disease, and may occur from a local site or be generalized. Localized bleeding may be associated with lesions and may be further complicated by a defective haemostatic mechanism. Congenital or acquired deficiencies of any of the coagulation factors may be associated with a hemorrhagic tendency. Classic examples of bleeding disorders include hemophilia, such as hemophilia A, which results from a deficiency in factor VIII, or hemophilia B (Christmas Disease), which results from a deficiency in factor IX. Hemophilia occurs in all racial and ethnic groups, and affects many people in the United States and worldwide.

[0004] Traditional therapy for bleeding disorders includes parenteral replacement of deficient clotting factors, such as factor VII, factor VIII or factor IX. For example, current treatments for Hemophilia B rely on chronic, repeated intravenous infusions of purified recombinant Factor IX. However, those treatments suffer from a number of drawbacks including the need for repeated intravenous infusions, being associated with inhibitor formation, and generally being more prophylactic rather than curative. See, e.g., Petrini 2001, *Hemophilia* 7:99; Fischer et al. 2002, *Blood* 99 (7):2337.

[0005] Gene therapy, which involves introducing a copy of a missing or defective gene into a patient, provide one possible method of introducing Factor IX to patients for a longer duration. However, there exists a need for additional compositions and methods that offer improved, long term expression of Factor IX.

[0006] The present disclosure provides compositions and methods useful for expressing Factor IX in a host cell or a population of host cells (in vitro or in vivo), and for treating hemophilia (e.g., hemophilia B). Provided herein are guide RNAs for use in targeted insertion of a sequence encoding Factor IX into a human genomic locus, e.g., a safe harbor site, such as an albumin safe harbor site. Also provided are

donor constructs (e.g., bidirectional constructs), comprising a sequence encoding Factor IX, for use in targeted insertion into a safe harbor site, such as intron 1 of the albumin safe harbor site. In some embodiments, the guide RNA disclosed herein can be used in combination with an RNA-guided DNA binding agent (e.g., Cas nuclease) and a donor construct (e.g., bidirectional construct) comprising a Factor IX transgene. In some embodiments, the donor construct (e.g., bidirectional construct) can be used with a gene editing system (e.g., CRISPR/Cas system; zinc finger nuclelease (ZFN) system; transcription activator-like effector nuclelease (TALEN) system). In some embodiments, the guide RNA disclosed herein can be used in combination with an RNA-guided DNA binding agent (e.g., Cas nuclease) and a donor construct (e.g., bidirectional construct) that comprises a Factor IX transgene. The following embodiments are provided.

[0007] In some aspects, provided herein is a method of introducing a Factor IX nucleic acid to a cell or a population of cells, comprising administering: i) a nucleic acid construct comprising a Factor IX protein coding sequence; ii) an RNA-guided DNA binding agent; and iii) a guide RNA (gRNA) comprising a sequence. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNA comprises a sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNA comprises a sequence that is a sequence selected from the group consisting of SEQ ID NOS: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNA comprises a sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNA comprises a sequence that is selected from the group consisting of SEQ ID NOS: 34-97. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 98-119. In some embodiments, the guide RNA comprises a sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 98-119. In some embodiments, the guide RNA comprises a sequence that is a sequence selected from the group consisting of SEQ ID NOS: 120-163.

[0008] In some aspects, provided herein is a method of expressing Factor IX in a cell or population of cells, comprising administering: i) a nucleic acid construct comprising a Factor IX protein coding sequence; ii) an RNA-guided DNA binding agent; and iii) a guide RNA (gRNA) comprising a sequence. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNA comprises a sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID

NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNA comprises a sequence that is selected from the group consisting of SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNA comprises a sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNA comprises a sequence selected from the group consisting of SEQ ID NOs: 34-97. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 98-119. In some embodiments, the guide RNA comprises a sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 98-119. In some embodiments, the guide RNA comprises a sequence that is selected from the group consisting of SEQ ID NOs: 120-163.

[0009] In some aspects, provided herein is a method of introducing or expressing Factor IX in a cell or population of cells, comprising administering: i) a nucleic acid construct comprising a Factor IX protein coding sequence; ii) an RNA-guided DNA binding agent; and iii) a guide RNA (gRNA) comprising a sequence wherein the administration is *in vitro*.

[0010] In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNA comprises at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNA comprises a sequence selected from the group consisting of SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNA comprises at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNA comprises a sequence selected from the group consisting of SEQ ID NOs: 34-97.

[0011] In some embodiments, the nucleic acid construct is administered in a nucleic acid vector and/or a lipid nanoparticle. In some embodiments, the RNA-guided DNA binding agent and/or gRNA is administered in a nucleic acid vector and/or lipid nanoparticle. In some embodiments, the nucleic acid vector is a viral vector. In some embodiments, the viral vector is selected from the group consisting of an adeno associate viral (AAV) vector, adenovirus vector, retrovirus vector, and lentivirus vector. In some embodiments, the AAV vector is selected from the group consisting of AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV-DJ, and AAV2/8.

[0012] In some embodiments, the nucleic acid construct, RNA-guided DNA binding agent, and gRNA are administered sequentially, in any order and/or in any combination. In some embodiments, wherein the nucleic acid construct, RNA-guided DNA binding agent, and gRNA, individually

or in any combination, are administered simultaneously. In some embodiments, the RNA-guided DNA binding agent, or RNA-guided DNA binding agent and gRNA in combination, is administered prior to administering the nucleic acid construct. In some embodiments, the nucleic acid construct is administered prior to administering the gRNA and/or RNA-guided DNA binding agent.

[0013] In some embodiments, the RNA-guided DNA binding agent is a Cas nuclease. In some embodiments, the Cas nuclease is a class 2 Cas nuclease. In some embodiments the Cas nuclease is Cas9. In some embodiments, the Cas nuclease is an *S. pyogenes* Cas9 nuclease. In some embodiments, the Cas nuclease is a nickase.

[0014] In some embodiments, the nucleic acid construct is a bidirectional nucleic acid construct. In some embodiments, the nucleic acid construct is single-stranded or double-stranded. In some embodiments, the nucleic acid construct is a single-stranded DNA or a double-stranded DNA. In some embodiments, the bidirectional construct does not comprise a promoter that drives the expression of the Factor IX protein. In some embodiments, the cell or population of cells expresses Factor IX with a heterologous peptide, such as an albumin signal peptide.

[0015] In some embodiments, the cell or population of cells includes a liver cell. In some embodiments, the liver cell is a hepatocyte.

[0016] In some embodiments, the nucleic acid encodes a wild-type Factor IX protein. In some embodiments, the nucleic acid encodes a mutant Factor IX protein. In some embodiments, the nucleic acid encodes a Factor IX protein having a mutation R338L.

[0017] In some aspects, provided herein is a method of introducing a Factor IX nucleic acid to a cell or population of cells, comprising administering to the cell or population of cells a bidirectional nucleic acid construct comprising a Factor IX protein coding sequence, thereby expressing Factor IX in the cell or population of cells. Provided herein is a method of expressing Factor IX in a cell or population of cells, comprising administering to the cell or population of cells a bidirectional nucleic acid construct comprising a Factor IX protein coding sequence, thereby expressing Factor IX expression in the cell or population of cells.

[0018] In some embodiments, the bidirectional nucleic acid construct comprises: a) a first segment comprising a coding sequence for Factor IX; and b) a second segment comprising a reverse complement of a coding sequence of Factor IX, wherein the construct does not comprise a promoter that drives the expression of Factor IX. In some embodiments, the bidirectional nucleic acid construct comprises: a) a first segment comprising a coding sequence for Factor IX; and b) a second segment comprising a reverse complement of a coding sequence of a second polypeptide, wherein the construct does not comprise a promoter that drives the expression of the polypeptide.

[0019] In some embodiments, the method of introducing a Factor IX nucleic acid to a cell or population of cells further comprises administering an RNA-guided DNA binding agent. In some embodiments, the method further comprises administering a gRNA. In some embodiments, the bidirectional nucleic acid construct is administered in a nucleic acid vector and/or a lipid nanoparticle. In some embodiments, the RNA-guided DNA binding agent is administered in a nucleic acid vector and/or lipid nanoparticle. In some embodiments, the gRNA is administered in a nucleic acid

vector and/or lipid nanoparticle. In some embodiments, the nucleic acid vector is a viral vector. In some embodiments, the viral vector is selected from the group consisting of an adeno associate viral (AAV) vector, adenovirus vector, retrovirus vector, and lentivirus vector. In some embodiments, the AAV vector is selected from the group consisting of AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV-DJ, and AAV2/8.

[0020] In some embodiments, the bidirectional nucleic acid construct, RNA-guided DNA binding agent, and gRNA are administered sequentially, in any order and/or in any combination. In some embodiments, the bidirectional nucleic acid construct, RNA-guided DNA binding agent, and gRNA, in any combination, are administered simultaneously. In some embodiments, the RNA-guided DNA binding agent, or RNA-guided DNA binding agent and gRNA in combination, is administered prior to administering the bidirectional nucleic acid construct. In some embodiments, the bidirectional nucleic acid construct is administered prior to administering the gRNA and/or RNA-guided DNA binding agent.

[0021] In some embodiments, the RNA-guided DNA binding agent is a Cas nuclease. In some embodiments, the Cas nuclease is a class 2 Cas nuclease. In some embodiments, the Cas nuclease is selected from the group consisting of *S. pyogenes* nuclease, *S. aureus* nuclease, *C. jejuni* nuclease, *S. thermophilus* nuclease, *N. meningitidis* nuclease, and variants thereof. In some embodiments, the Cas nuclease is Cas9. In some embodiments, the Cas nuclease is a nickase.

[0022] In some embodiments, the bidirectional construct does not comprise a promoter that drives the expression of the Factor IX protein. In some embodiments, the bidirectional construct is single-stranded or double-stranded. In some embodiments, the nucleic acid construct is a single-stranded DNA or a double-stranded DNA. In some embodiments, the gRNA comprises at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2-33 or a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-33.

[0023] In some aspects, provided herein is a composition for use in expressing Factor IX in a cell, wherein the composition comprises: i) a nucleic acid construct comprising a Factor IX protein coding sequence; ii) an RNA-guided DNA binding agent; and iii) a guide RNA (gRNA) comprising a guide sequence selected from the group consisting of SEQ ID NOs: 2-33 or a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-33. Provided herein is a composition for use in expressing Factor IX in a cell or population of cells, wherein the composition comprises a bidirectional nucleic acid construct comprising a Factor IX protein coding sequence. In some embodiments, a host cell is made by the method of any preceding embodiment.

[0024] In some embodiments, the host cell is a liver cell. In some embodiments, the host cell is a non-dividing cell type. In some embodiments, the host cell expresses the Factor IX polypeptide encoded by the bidirectional construct. In some embodiments, the host cell is a hepatocyte.

[0025] In some embodiments of the method, construct, or host cell of any above method, the gRNA comprises SEQ ID NO: 401.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] FIG. 1 shows construct formats as represented in AAV genomes. SA=slice acceptor; pA=polyA signal sequence; HA=homology arm; LHA=left homology arm; RHA=right homology arm

[0027] FIG. 2 shows vectors without homology arms are not effective in an immortalized liver cell line (Hepa1-6). An scAAV derived from plasmid P00204 comprising 200 bp homology arms resulted in expression of hFIX in the dividing cells. Use of the AAV vectors derived from P00123 (scAAV lacking homology arms) and P00147 (ssAAV bidirectional construct lacking homology arms) did not result in detectable expression of hFIX.

[0028] FIGS. 3A and 3B show results from in vivo testing of insertion templates with and without homology arms using vectors derived from P00123, P00147, or P00204. FIG. 3A shows liver editing levels as measured by indel formation of ~60% were detected in each group of animals treated with LNPs comprising CRISPR/Cas9 components. FIG. 3B shows animals receiving the ssAAV vectors without homology arms (derived from P00147) in combination with LNP treatment resulted in the highest level of hFIX expression in serum.

[0029] FIGS. 4A and 4B show results from in vivo testing of ssAAV insertion templates with and without homology arms. FIG. 4A compares targeted insertion with vectors derived from plasmids P00350, P00356, P00362 (having asymmetrical homology arms as shown), and P00147 (bidirectional construct as shown in FIG. 4B). FIG. 4B compares insertion into a second site targeted with vectors derived from plasmids P00353, P00354 (having symmetrical homology arms as shown), and P00147.

[0030] FIGS. 5A-5D show results of targeted insertion of bidirectional constructs across 20 target sites in primary mouse hepatocytes. FIG. 5A shows the schematics of each of the vectors tested. FIG. 5B shows editing as measured by indel formation for each of the treatment groups across each combination tested. FIG. 5C and FIG. 5D show that significant levels of editing (as indel formation at a specific target site) did not necessarily result in more efficient insertion or expression of the transgenes. hSA=human F9 splice acceptor; mSA=mouse albumin splice acceptor; HiBit=tag for luciferase based detection; pA=polyA signal sequence; Nluc=nanoluciferase reporter; GFP=green fluorescent reporter.

[0031] FIG. 6 shows results from in vivo screening of targeted insertion with bidirectional constructs across 10 target sites using with ssAAV derived from P00147. As shown, significant levels of indel formation do not necessarily result in high levels of transgene expression.

[0032] FIGS. 7A-7D show results from in vivo screening of bidirectional constructs across 20 target sites using ssAAV derived from P00147. FIG. 7A shows varied levels of editing as measured by indel formation were detected for each of the treatment groups across each LNP/vector combination tested. FIG. 7B provides corresponding targeted insertion data. The results show poor correlation between indel formation and insertion or expression of the bidirectional constructs (FIG. 7B and FIG. 7D), and a positive correlation between in vitro and in vivo results (FIG. 7C).

[0033] FIGS. 8A and 8B show insertion of the bidirectional construct at the cellular level using in situ hybridization method using probes that can detect the junctions between the hFIX transgene and the mouse albumin exon 1

sequence (FIG. 8A). Circulating hFIX levels correlated with the number of cells that were positive for the hybrid transcript (FIG. 8B).

[0034] FIG. 9 shows the effect on targeted insertion of varying the timing between delivery of the ssAAV comprising the bidirectional hFIX construct and LNP.

[0035] FIG. 10 shows the effect on targeted insertion of varying the number of LNP doses (e.g., 1, 2, or 3) following delivery of the bidirectional hFIX construct.

[0036] FIG. 11A shows the durability of hFIX expression in vivo. FIG. 11B demonstrates expression from intron 1 of albumin was sustained.

[0037] FIG. 12A and FIG. 12B show that varying AAV or LNP dose can modulate the amount of expression of hFIX from intron 1 of the albumin gene in vivo.

[0038] FIGS. 13A-13C show results from screening bidirectional constructs across target sites in primary cynomolgus hepatocytes. FIG. 13A shows varied levels of editing as measured by indel formation detected for each of the samples. FIG. 13B and FIG. 13C show that significant levels of indel formation was not predictive for insertion or expression of the bidirectional constructs into intron 1 of albumin.

[0039] FIGS. 14A-14C show results from screening bidirectional constructs across target sites in primary human hepatocytes. FIG. 14A shows editing as measured by indel formation detected for each of the samples. FIG. 14B, FIGS. 14C, and 14D show that significant levels of indel formation was not predictive for insertion or expression of the bidirectional constructs into intron 1 of the albumin gene.

[0040] FIG. 15 shows the results of in vivo studies where non-human primates were dosed with LNPs along with a bi-directional hFIX insertion template (derived from P00147). Systemic hFIX levels were achieved only in animals treated with both LNPs and AAV, with no hFIX detectable using AAV or LNPs alone.

[0041] FIG. 16A and FIG. 16B show human Factor IX expression levels in the plasma samples at week 6 post-injection.

[0042] FIG. 17 shows week 7 serum levels and % positive cells across the multiple lobes for each animal.

[0043] FIG. 18 shows human Factor IX expression levels in the plasma samples in each group at weeks 1, 2, and 4 post-injection.

[0044] FIG. 19 shows insertion of the hF9 transgene and clotting function in the APTT assay.

[0045] FIG. 20A and FIG. 20B show insertion of the hF9 transgene and thrombin generation in TGA-EA analysis.

[0046] FIG. 21 shows insertion of the hF9 transgene and thrombin generation.

DETAILED DESCRIPTION

[0047] Reference will now be made in detail to certain embodiments of the invention, examples of which are illustrated in the accompanying drawings. While the invention is described in conjunction with the illustrated embodiments, it will be understood that they are not intended to limit the invention to those embodiments. On the contrary, the invention is intended to cover all alternatives, modifications, and equivalents, which may be included within the invention as defined by the appended embodiments.

[0048] Before describing the present teachings in detail, it is to be understood that the disclosure is not limited to specific compositions or process steps, as such may vary. It should be noted that, as used in this specification and the

appended embodiments, the singular form "a" "an" and "the" include plural references unless the context clearly dictates otherwise. Thus, for example, reference to "a conjugate" includes a plurality of conjugates and reference to "a cell" includes a plurality or population of cells and the like. As used herein, the term "include" and its grammatical variants are intended to be non-limiting, such that recitation of items in a list is not to the exclusion of other like items that can be substituted or added to the listed items.

[0049] Numeric ranges are inclusive of the numbers defining the range. Measured and measureable values are understood to be approximate, taking into account significant digits and the error associated with the measurement. Also, the use of "comprise", "comprises", "comprising", "contain", "contains", "containing", "include", "includes", and "including" are not intended to be limiting. It is to be understood that both the foregoing general description and detailed description are exemplary and explanatory only and are not restrictive of the teachings.

[0050] Unless specifically noted in the specification, embodiments in the specification that recite "comprising" various components are also contemplated as "consisting of" or "consisting essentially of" the recited components; embodiments in the specification that recite "consisting of" various components are also contemplated as "comprising" or "consisting essentially of" the recited components; and embodiments in the specification that recite "consisting essentially of" various components are also contemplated as "consisting of" or "comprising" the recited components (this interchangeability does not apply to the use of these terms in the claims). The term "or" is used in an inclusive sense, i.e., equivalent to "and/or," unless the context clearly indicates otherwise. The term "about", when used before a list, modifies each member of the list. The term "about" or "approximately" means an acceptable error for a particular value as determined by one of ordinary skill in the art, which depends in part on how the value is measured or determined.

[0051] The term "about", when used before a list, modifies each member of the list. The term "about" or "approximately" means an acceptable error for a particular value as determined by one of ordinary skill in the art, which depends in part on how the value is measured or determined.

[0052] The section headings used herein are for organizational purposes only and are not to be construed as limiting the desired subject matter in any way. In the event that any material incorporated by reference contradicts any term defined in this specification or any other express content of this specification, this specification controls.

I. Definitions

[0053] Unless stated otherwise, the following terms and phrases as used herein are intended to have the following meanings:

[0054] "Polynucleotide" and "nucleic acid" are used herein to refer to a multimeric compound comprising nucleosides or nucleoside analogs which have nitrogenous heterocyclic bases or base analogs linked together along a backbone, including conventional RNA, DNA, mixed RNA-DNA, and polymers that are analogs thereof. A nucleic acid "backbone" can be made up of a variety of linkages, including one or more of sugar-phosphodiester linkages, peptide-nucleic acid bonds ("peptide nucleic acids" or PNA; PCT No. WO 95/32305), phosphorothioate linkages, methylphosphonate linkages, or combinations thereof. Sugar

moieties of a nucleic acid can be ribose, deoxyribose, or similar compounds with substitutions, e.g., 2' methoxy or 2' halide substitutions. Nitrogenous bases can be conventional bases (A, G, C, T, U), analogs thereof (e.g., modified uridines such as 5-methoxyuridine, pseudouridine, or N1-methylpseudouridine, or others); inosine; derivatives of purines or pyrimidines (e.g., N⁴-methyl deoxyguanosine, deaza- or aza-purines, deaza- or aza-pyrimidines, pyrimidine bases with substituent groups at the 5 or 6 position (e.g., 5-methylcytosine), purine bases with a substituent at the 2, 6, or 8 positions, 2-amino-6-methylaminopurine, O⁶-methylguanine, 4-thio-pyrimidines, 4-amino-pyrimidines, 4-dimethylhydrazine-pyrimidines, and O⁴-alkyl-pyrimidines; U.S. Pat. No. 5,378,825 and PCT No. WO 93/13121). For general discussion see *The Biochemistry of the Nucleic Acids* 5-36, Adams et al., ed., 11th ed., 1992). Nucleic acids can include one or more “abasic” residues where the backbone includes no nitrogenous base for position(s) of the polymer (U.S. Pat. No. 5,585,481). A nucleic acid can comprise only conventional RNA or DNA sugars, bases and linkages, or can include both conventional components and substitutions (e.g., conventional bases with 2' methoxy linkages, or polymers containing both conventional bases and one or more base analogs). Nucleic acid includes “locked nucleic acid” (LNA), an analogue containing one or more LNA nucleotide monomers with a bicyclic furanose unit locked in an RNA mimicking sugar conformation, which enhance hybridization affinity toward complementary RNA and DNA sequences (Vester and Wengel, 2004, *Biochemistry* 43(42): 13233-41). RNA and DNA have different sugar moieties and can differ by the presence of uracil or analogs thereof in RNA and thymine or analogs thereof in DNA.

[0055] “Guide RNA”, “gRNA”, and simply “guide” are used herein interchangeably to refer to either a guide that comprises a guide sequence, e.g., crRNA (also known as CRISPR RNA), or the combination of a crRNA and a trRNA (also known as tracrRNA). The crRNA and trRNA may be associated as a single RNA molecule (single guide RNA, sgRNA) or, for example, in two separate RNA molecules (dual guide RNA, dgRNA). “Guide RNA” or “gRNA” refers to each type. The trRNA may be a naturally-occurring sequence, or a trRNA sequence with modifications or variations compared to naturally-occurring sequences. Guide RNAs, such as sgRNAs or dgRNAs, can include modified RNAs as described herein.

[0056] As used herein, a “guide sequence” refers to a sequence within a guide RNA that is complementary to a target sequence and functions to direct a guide RNA to a target sequence for binding or modification (e.g., cleavage) by an RNA-guided DNA binding agent. A “guide sequence” may also be referred to as a “targeting sequence,” or a “spacer sequence.” A guide sequence can be 20 base pairs in length, e.g., in the case of *Streptococcus pyogenes* (i.e., Spy Cas9) and related Cas9 homologs/orthologs. Shorter or longer sequences can also be used as guides, e.g., 15-, 16-, 17-, 18-, 19-, 21-, 22-, 23-, 24-, or 25-nucleotides in length. For example, in some embodiments, the guide sequence comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs:2-33. In some embodiments, the target sequence is in a gene or on a chromosome, for example, and is complementary to the guide sequence. In some embodiments, the degree of complementarity or identity between a guide sequence and its corresponding target sequence may be about 75%, 80%,

85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%. For example, in some embodiments, the guide sequence comprises a sequence with about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 2-33. In some embodiments, the guide sequence and the target region may be 100% complementary or identical. In other embodiments, the guide sequence and the target region may contain at least one mismatch. For example, the guide sequence and the target sequence may contain 1, 2, 3, or 4 mismatches, where the total length of the target sequence is at least 15, 16, 17, 18, 19, 20 or more base pairs. In some embodiments, the guide sequence and the target region may contain 1-4 mismatches where the guide sequence comprises at least 15, 16, 17, 18, 19, 20 or more nucleotides. In some embodiments, the guide sequence and the target region may contain 1, 2, 3, or 4 mismatches where the guide sequence comprises 20 nucleotides.

[0057] Target sequences for RNA-guided DNA binding agents include both the positive and negative strands of genomic DNA (i.e., the sequence given and the sequence's reverse complement), as a nucleic acid substrate for an RNA-guided DNA binding agent is a double stranded nucleic acid. Accordingly, where a guide sequence is said to be “complementary to a target sequence”, it is to be understood that the guide sequence may direct a guide RNA to bind to the reverse complement of a target sequence. Thus, in some embodiments, where the guide sequence binds the reverse complement of a target sequence, the guide sequence is identical to certain nucleotides of the target sequence (e.g., the target sequence not including the PAM) except for the substitution of U for T in the guide sequence.

[0058] As used herein, an “RNA-guided DNA-binding agent” means a polypeptide or complex of polypeptides having RNA and DNA binding activity, or a DNA-binding subunit of such a complex, wherein the DNA binding activity is sequence-specific and depends on the sequence of the RNA. The term RNA-guided DNA binding-agent also includes nucleic acids encoding such polypeptides. Exemplary RNA-guided DNA-binding agents include Cas cleavases/nickases.

[0059] Exemplary RNA-guided DNA-binding agents may include inactivated forms thereof (“dCas DNA-binding agents”), e.g. if those agents are modified to permit DNA cleavage, e.g. via fusion with a FokI cleavase domain. “Cas nuclease”, as used herein, encompasses Cas cleavases and Cas nickases. Cas cleavases and Cas nickases include a Csm or Cmr complex of a type III CRISPR system, the Cas10, Csm1, or Cmr2 subunit thereof, a Cascade complex of a type I CRISPR system, the Cas3 subunit thereof, and Class 2 Cas nucleases. As used herein, a “Class 2 Cas nuclease” is a single-chain polypeptide with RNA-guided DNA binding activity. Class 2 Cas nucleases include Class 2 Cas cleavases/nickases (e.g., H840A, D10A, or N863A variants), which further have RNA-guided DNA cleavases or nickase activity, and Class 2 dCas DNA-binding agents, in which cleavase/nickase activity is inactivated”), if those agents are modified to permit DNA cleavage. Class 2 Cas nucleases include, for example, Cas9, Cpf1, C2c1, C2c2, C2c3, HF Cas9 (e.g., N497A, R661A, Q695A, Q926A variants), HypaCas9 (e.g., N692A, M694A, Q695A, H698A variants), eSPCas9(1.0) (e.g., K810A, K1003A, R1060A variants), and eSPCas9(1.1) (e.g., K848A, K1003A, R1060A variants)

proteins and modifications thereof. Cpf1 protein, Zetsche et al., *Cell*, 163: 1-13 (2015), also contains a RuvC-like nuclease domain. Cpf1 sequences of Zetsche are incorporated by reference in their entirety. See, e.g., Zetsche, Tables Si and S3. See, e.g., Makarova et al., *Nat Rev Microbiol*, 13(11): 722-36 (2015); Shmakov et al., *Molecular Cell*, 60:385-397 (2015). As used herein, delivery of an RNA-guided DNA-binding agent (e.g. a Cas nuclease, a Cas9 nuclease, or an *S. pyogenes* Cas9 nuclease) includes delivery of the polypeptide or mRNA.

[0060] As used herein, “ribonucleoprotein” (RNP) or “RNP complex” refers to a guide RNA together with an RNA-guided DNA binding agent, such as a Cas nuclease, e.g., a Cas cleavage, Cas nickase, or dCas DNA binding agent (e.g., Cas9). In some embodiments, the guide RNA guides the RNA-guided DNA binding agent such as Cas9 to a target sequence, and the guide RNA hybridizes with and the agent binds to the target sequence; in cases where the agent is a cleavage or nickase, binding can be followed by cleaving or nicking.

[0061] As used herein, a first sequence is considered to “comprise a sequence with at least X % identity to” a second sequence if an alignment of the first sequence to the second sequence shows that X % or more of the positions of the second sequence in its entirety are matched by the first sequence. For example, the sequence AAGA comprises a sequence with 100% identity to the sequence AAG because an alignment would give 100% identity in that there are matches to all three positions of the second sequence. The differences between RNA and DNA (generally the exchange of uridine for thymidine or vice versa) and the presence of nucleoside analogs such as modified uridines do not contribute to differences in identity or complementarity among polynucleotides as long as the relevant nucleotides (such as thymidine, uridine, or modified uridine) have the same complement (e.g., adenine for all of thymidine, uridine, or modified uridine; another example is cytosine and 5-methylcytosine, both of which have guanosine or modified guanosine as a complement). Thus, for example, the sequence 5'-AXG where X is any modified uridine, such as pseudouridine, N1-methyl pseudouridine, or 5-methoxyuridine, is considered 100% identical to AUG in that both are perfectly complementary to the same sequence (5'-CAU). Exemplary alignment algorithms are the Smith-Waterman and Needleman-Wunsch algorithms, which are well-known in the art. One skilled in the art will understand what choice of algorithm and parameter settings are appropriate for a given pair of sequences to be aligned; for sequences of generally similar length and expected identity >50% for amino acids or >75% for nucleotides, the Needleman-Wunsch algorithm with default settings of the Needleman-Wunsch algorithm interface provided by the EBI at the www.ebi.ac.uk web server is generally appropriate.

[0062] As used herein, a first sequence is considered to be “X % complementary to” a second sequence if X % of the bases of the first sequence base pairs with the second sequence. For example, a first sequence 5'AAGA3' is 100% complementary to a second sequence 3'TTCT5', and the second sequence is 100% complementary to the first sequence. In some embodiments, a first sequence 5'AAGA3' is 100% complementary to a second sequence 3'TTCTGTGA5', whereas the second sequence is 50% complementary to the first sequence.

[0063] As used herein, “mRNA” is used herein to refer to a polynucleotide that is entirely or predominantly RNA or modified RNA and comprises an open reading frame that can be translated into a polypeptide (i.e., can serve as a substrate for translation by a ribosome and amino-acylated tRNAs). mRNA can comprise a phosphate-sugar backbone including ribose residues or analogs thereof, e.g., 2'-methoxy ribose residues. In some embodiments, the sugars of an mRNA phosphate-sugar backbone consist essentially of ribose residues, 2'-methoxy ribose residues, or a combination thereof.

[0064] Guide sequences useful in the guide RNA compositions and methods described herein are shown in Table 1 throughout the application.

[0065] As used herein, “indels” refer to insertion/deletion mutations consisting of a number of nucleotides that are either inserted or deleted at the site of double-stranded breaks (DSBs) in a target nucleic acid.

[0066] As used herein, “Factor IX” is used interchangeably with “FIX” or “F9”, and is also known as Christmas Factor. The human wild-type Factor IX protein sequence is available at NCBI NP_000124; gene sequence is available at NCBI NM_000133. Examples of the Factor IX protein sequence are described herein (e.g. SEQ ID NO: 700, SEQ ID NO: 701, and/or SEQ ID NO: 702). As used herein, Factor IX also encompasses a variant of Factor IX, e.g., a variant that possesses increased coagulation activity as compared to wild type Factor IX. A hyperactive variant of Factor IX may comprise a substitution of R338. An example of such a variant Factor IX comprises the mutation R338L relative to SEQ ID NO: 701. The terms hyperactive and hyperfunctional are being used interchangeably herein. Further examples of variant Factor IX comprise an amino acid at residue 338 chosen from alanine, leucine, valine, isoleucine, phenylalanine, tryptophan, methionine, serine, and threonine. Further Factor IX variants comprise an amino acid at residue 338 chosen from leucine, cysteine, aspartic acid, glutamic acid, histidine, lysine, asparagine, glutamine, or tyrosine. As used herein, Factor IX also encompasses a variant that is 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 60%, 70%, 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild-type Factor IX. As used herein, Factor IX also encompasses a variant that is 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 60%, 70%, 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to SEQ ID NO: 701 or SEQ ID NO: 702. As used herein, Factor IX also encompasses a fragment that possesses at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX. In some embodiments, a Factor IX variant may be a hyperactive Factor IX variant. In certain instances, the Factor IX variant possesses between about 80% and about 100%, 120%, 140%, 160%, 180%, or 200% of the activity as compared to the wild-type Factor IX. The specific activity of the Factor IX variant can be used to calculate its functionally normalized activity, for example as described in Example 13. The specific activities of Factor IX variants, e.g. R338L, are known in the literature and can be calculated using known methods. A hyperfunctional Factor IX variant may have about 1.2, 1.5, 2, 3, 4, 5, 6, 8, 10, 12, or 15 times the specific activity of a corresponding wild type Factor IX protein. In one embodiment, the hyperfunctional Factor IX may have

about 8-12 times the specific activity of a corresponding wild type Factor IX protein. In another embodiment, the hyperfunctional Factor IX may have 1.2-5 times the specific activity of a corresponding wild type Factor IX protein. Exemplary sequences are known in the art, and include sequences in U.S. Pat. Nos. 4,770,999, 4,994,371, 5,521,070, 6,046,380, 6,531,298, and 8,383,388, for example.

[0067] As used herein, a “target sequence” refers to a sequence of nucleic acid in a target gene that has complementarity to the guide sequence of the gRNA. The interaction of the target sequence and the guide sequence directs an RNA-guided DNA binding agent to bind, and potentially nick or cleave (depending on the activity of the agent), within the target sequence.

[0068] As used herein, “hemophilia” refers to a disorder caused by a missing or defective Factor IX gene or polypeptide. The disorder includes conditions that are inherited and/or acquired (e.g., caused by a spontaneous mutation in the gene), and includes hemophilia B. In some embodiments, the defective Factor IX gene or polypeptide results in reduced Factor IX level in the plasma and/or a reduced coagulation activity of Factor IX. As used herein, hemophilia includes mild, moderate, and severe hemophilia. For example, individuals with less than about 1% active factor are classified as having severe haemophilia, those with about 1-5% active factor have moderate haemophilia, and those with mild haemophilia have between about 5-40% of normal levels of active clotting factor.

[0069] As used herein, “normal” or “healthy” individuals include those having between 50 and 160% of normal pooled plasma level of Factor IX activity and antigen levels. Based on its purification from human plasma, the concentration of Factor IX in the normal adult (normal pooled plasma level of Factor IX) is about 300-400 µg/ml of plasma. In some embodiments, the level of Factor IX, e.g., circulating Factor IX, can be measured by a coagulation and/or an immunologic assay, e.g., an sandwich immunoassay, ELISA (see, e.g., Example 13), MSD (see, e.g., Example 14). Factor IX procoagulant activity is determined by the ability of the patient’s plasma to correct the clotting time of Factor IX-deficient plasma.

[0070] As used herein, “treatment” refers to any administration or application of a therapeutic for disease or disorder in a subject, and includes inhibiting the disease, arresting its development, relieving one or more symptoms of the disease, curing the disease, or preventing reoccurrence of one or more symptoms of the disease. For example, treatment of hemophilia may comprise alleviating symptoms of hemophilia.

[0071] As used herein, a “bidirectional nucleic acid construct” (interchangeably referred to herein as a “bidirectional construct”) comprises at least two nucleic acid segments, wherein one segment (the first segment) comprises a coding sequence that encodes a polypeptide of interest (the coding sequence may be referred to herein as “transgene” or a first transgene), while the other segment (the second segment) comprises a sequence wherein the complement of the sequence encodes a polypeptide of interest, or a second transgene. That is, the at least two segments can encode identical or different polypeptides. When the two segments encode the identical polypeptide, the coding sequence of the first segment need not be identical to the complement of the sequence of the second segment. In some embodiments, the sequence of the second segment is a reverse complement of

the coding sequence of the first segment. A bidirectional construct can be single-stranded or double-stranded. The bidirectional construct disclosed herein encompasses a construct that is capable of expressing any polypeptide of interest.

[0072] In some embodiments, a bidirectional nucleic acid construct comprises a first segment that comprises a coding sequence that encodes a first polypeptide (a first transgene), and a second segment that comprises a sequence wherein the complement of the sequence encodes a second polypeptide (a second transgene). In some embodiments, the first and the second polypeptides are at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical. In some embodiments, the first and the second polypeptides comprise an amino acid sequence that is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical, e.g. across 50, 100, 200, 500, 1000 or more amino acid residues.

[0073] As used herein, a “reverse complement” refers to a sequence that is a complement sequence of a reference sequence, wherein the complement sequence is written in the reverse orientation. For example, for a hypothetical sequence 5'CTGGACCGA3' (SEQ ID NO: 500), the “perfect” complement sequence is 3'GACCTGGCT5' (SEQ ID NO: 501), and the “perfect” reverse complement is written 5'TCGGTCCAG3' (SEQ ID NO: 502). A reverse complement sequence need not be “perfect” and may still encode the same polypeptide or a similar polypeptide as the reference sequence. Due to codon usage redundancy, a reverse complement can diverge from a reference sequence that encodes the same polypeptide. As used herein, “reverse complement” also includes sequences that are, e.g., 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the reverse complement sequence of a reference sequence.

[0074] As used herein, “polypeptide” refers to a wild-type or variant protein (e.g., mutant, fragment, fusion, or combinations thereof). A variant polypeptide may possess at least or about 5%, 10%, 15%, 20%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% functional activity of the wild-type polypeptide. In some embodiments, the variant is at least 70%, 75%, 80%, 85%, 90%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the sequence of the wild-type polypeptide. In some embodiments, a variant polypeptide may be a hyperactive variant. In certain instances, the variant possesses between about 80% and about 120%, 140%, 160%, 180%, 200%, 300%, 400%, 500%, or more of a functional activity of the wild-type polypeptide.

[0075] As used herein, a “heterologous gene” refers to a gene that has been introduced as an exogenous source to a site within a host cell genome (e.g., at a genomic locus such as a safe harbor locus including an albumin intron 1 site). That is, the introduced gene is heterologous with respect to its insertion site. A polypeptide expressed from such heterologous gene is referred to as a “heterologous polypeptide.” The heterologous gene can be naturally-occurring or engineered, and can be wild type or a variant. The heterologous gene may include nucleotide sequences other than the sequence that encodes the heterologous polypeptide (e.g., an internal ribosomal entry site). The heterologous gene can be a gene that occurs naturally in the host genome, as a wild type or a variant (e.g., mutant). For example, although the

host cell contains the gene of interest (as a wild type or as a variant), the same gene or variant thereof can be introduced as an exogenous source for, e.g., expression at a locus that is highly expressed. The heterologous gene can also be a gene that is not naturally occurring in the host genome, or that expresses a heterologous polypeptide that does not naturally occur in the host genome. “Heterologous gene”, “exogenous gene”, and “transgene” are used interchangeably. In some embodiments, the heterologous gene or transgene includes an exogenous nucleic acid sequence, e.g. a nucleic acid sequence is not endogenous to the recipient cell. In some embodiments, the heterologous gene or transgene includes an exogenous nucleic acid sequence, e.g. a nucleic acid sequence that does not naturally occur in the recipient cell. For example, a heterologous gene may be heterologous with respect to its insertion site and with respect to its recipient cell.

[0076] A “safe harbor” locus is a locus within the genome wherein a gene may be inserted without significant deleterious effects on the host cell, e.g. hepatocyte, e.g., without causing apoptosis, necrosis, and/or senescence, or without causing more than 5%, 10%, 15%, 20%, 25%, 30%, or 40% apoptosis, necrosis, and/or senescence as compared to a control cell. See, e.g., Hsin et al., “Hepatocyte death in liver inflammation, fibrosis, and tumorigenesis,” 2017. In some embodiments, a safe harbor locus allows overexpression of an exogenous gene without significant deleterious effects on the host cell, e.g. hepatocyte, e.g., without causing apoptosis, necrosis, and/or senescence, or without causing more than 5%, 10%, 15%, 20%, 25%, 30%, or 40% apoptosis, necrosis, and/or senescence as compared to a control cell. In some embodiments, a desirable safe harbor locus may be one in which expression of the inserted gene sequence is not perturbed by read-through expression from neighboring genes. The safe harbor may be within an albumin gene, such as a human albumin gene. The safe harbor may be within an albumin intron 1 region, e.g., human albumin intron 1. The safe harbor may be a human safe harbor, e.g., for a liver tissue or hepatocyte host cell. In some embodiments, a safe harbor allows overexpression of an exogenous gene without significant deleterious effects on the host cell or cell population, such as hepatocytes or liver cells, e.g. without causing apoptosis, necrosis, and/or senescence, or without causing more than 5%, 10%, 15%, 20%, 25%, 30%, or 40% apoptosis, necrosis, and/or senescence as compared to a control cell or cell population.

II. Compositions

A. Compositions Comprising Guide RNA (gRNAs)

[0077] Provided herein are guide RNA compositions and methods useful for inserting and expressing a Factor IX gene within a genomic locus, e.g., a safe harbor site of a host cell or a population of host cells. In particular, as exemplified herein, targeting and inserting an exogenous gene at the albumin locus (e.g., at intron 1) allows the use of albumin’s endogenous promoter to drive robust expression of the exogenous gene. The present disclosure is based, in part, on the identification of guide RNAs that specifically target sites within intron 1 of the albumin gene, and which provide efficient insertion and expression of the Factor IX gene. As shown in the Examples and further described herein, the ability of identified gRNAs to mediate high levels of editing as measured through indel forming activity, unexpectedly does not necessarily correlate with use of the same gRNAs

to mediate efficient insertion of transgenes as measured through, e.g., expression of the transgene. That is, certain gRNAs that are able to achieve a high level of indel formation are not necessarily able to mediate efficient insertion, and conversely, some gRNAs shown to achieve low levels of indel formation may mediate efficient insertion and expression of a transgene.

[0078] In some embodiments, provided herein are compositions and methods useful for inserting and expressing a Factor IX gene within a region of an albumin locus (e.g., intron 1) of a host cell. In some embodiments, disclosed herein are compositions useful for introducing or inserting a heterologous Factor IX nucleic acid within an albumin locus of a host cell, e.g., using a guide RNA disclosed herein with an RNA-guided DNA binding agent, and a construct (e.g., donor construct or template) comprising a heterologous Factor IX nucleic acid (“Factor IX transgene”). In some embodiments, disclosed herein are compositions useful for expressing a heterologous Factor IX from an albumin locus of a host cell, e.g., using a guide RNA disclosed herein with an RNA-guided DNA binding agent and a construct (e.g., donor) comprising a heterologous Factor IX nucleic acid. In some embodiments, disclosed herein are compositions useful for expressing a heterologous Factor IX from an albumin locus of a host cell, e.g., using a guide RNA disclosed herein with an RNA-guided DNA binding agent and a bidirectional construct comprising a heterologous Factor IX nucleic acid. In some embodiments, disclosed herein are compositions useful for inducing a break (e.g., double-stranded break (DSB) or single-stranded break (nick)) within the serum albumin gene of a host cell, e.g., using a guide RNA disclosed herein with an RNA-guided DNA binding agent (e.g., a CRISPR/Cas system). The compositions may be used in vitro or in vivo for, e.g., treating hemophilia.

[0079] In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that binds, or is capable of binding, within an intron of an albumin locus. In some embodiments, the guide RNAs disclosed herein bind within a region of intron 1 of the human albumin gene (SEQ ID NO: 1). It will be appreciated that not every base of the guide sequence must bind within the recited regions. For example, in some embodiments, 15, 16, 17, 18, 19, 20, or more bases of the guide RNA sequence bind with the recited regions. For example, in some embodiments, 15, 16, 17, 18, 19, 20, or more contiguous bases of the guide RNA sequence bind with the recited regions.

[0080] In some embodiments, the guide RNAs disclosed herein mediate a target-specific cutting by an RNA-guided DNA binding agent (e.g., Cas nuclease) at a site within human albumin intron 1 (SEQ ID NO: 1). It will be appreciated that, in some embodiments, the guide RNAs comprise guide sequences that bind to, or are capable of binding to, said regions.

[0081] In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33.

[0082] In some embodiments, the guide RNAs disclosed herein comprise a guide sequence having at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected

from the group consisting of a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID Nos: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOS: 34-97.

[0083] In some embodiments, the guide RNAs disclosed herein comprise a guide sequence having at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOS: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97.

[0084] In some embodiments, the guide RNAs disclosed herein mediate a target-specific cutting resulting in a double-stranded break (DSB). In some embodiments, the guide RNAs disclosed herein mediate a target-specific cutting resulting in a single-stranded break (nick).

[0085] In some embodiments, the guide RNAs disclosed herein bind to a region upstream of a propospacer adjacent motif (PAM). As would be understood by those of skill in the art, the PAM sequence occurs on the strand opposite to the strand that contains the target sequence. That is, the PAM sequence is on the complement strand of the target strand (the strand that contains the target sequence to which the guide RNA binds). In some embodiments, the PAM is

selected from the group consisting of NGG, NNGRRT, NNGRR(N), NNAGAAW, NNNNG(A/C)TT, and NNNN-RYAC.

[0086] In some embodiments, the guide RNA sequences provided herein are complementary to a sequence adjacent to a PAM sequence.

[0087] In some embodiments, the guide RNA sequence comprises a sequence that is complementary to a sequence within a genomic region selected from the tables herein according to coordinates in human reference genome hg38. In some embodiments, the guide RNA sequence comprises a sequence that is complementary to a sequence that comprises 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 consecutive nucleotides from within a genomic region selected from the tables herein. In some embodiments, the guide RNA sequence comprises a sequence that is complementary to a sequence that comprises 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 consecutive nucleotides spanning a genomic region selected from the tables herein.

[0088] The guide RNAs disclosed herein mediate a target-specific cutting resulting in a double-stranded break (DSB). The guide RNAs disclosed herein mediate a target-specific cutting resulting in a single-stranded break (SSB or nick).

[0089] In some embodiments, the guide RNAs disclosed herein mediate target-specific cutting by an RNA-guided DNA binding agent (e.g., a Cas nuclease, as disclosed herein), resulting in insertion of a heterologous Factor IX nucleic acid within intron 1 of an albumin gene. In some embodiments, the guide RNA and/or cutting results in a rate of between 30 and 35%, 35 and 40%, 40 and 45%, 45 and 50%, 50 and 55%, 55 and 60%, 60 and 65%, 65 and 70%, 70 and 75%, 75 and 80%, 80 and 85%, 85 and 90%, 90 and 95%, or 95 and 99% insertion of a heterologous Factor IX gene. In some embodiments, the guide RNA and/or cutting results in a rate of at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% insertion of a heterologous Factor IX nucleic acid. Insertion rates can be measured in vitro or in vivo. For example, in some embodiments, rate of insertion can be determined by detecting and measuring the inserted Factor IX nucleic acid within a population of cells, and calculating a percentage of the population that contains the inserted Factor IX nucleic acid. Methods of measuring insertion rates are known and available in the art. In some embodiments, the guide RNA allows between 5 and 10%, 10 and 15%, 15 and 20%, 20 and 25%, 25 and 30%, 30 and 35%, 35 and 40%, 40 and 45%, 45 and 50%, 50 and 55%, 55 and 60%, 60 and 65%, 65 and 70%, 70 and 75%, 75 and 80%, 80 and 85%, 85 and 90%, 90 and 95%, 95 and 99% or more increased expression of a heterologous Factor IX gene. Increased expression of a heterologous Factor IX gene can be measured in vitro or in vivo. For example, in some embodiments, increased expression can be determined by detecting and measuring the Factor IX polypeptide level and comparing the level against the Factor IX polypeptide level before, e.g., treating the cells or administration to a subject. In some embodiments, the guide RNA allows between 5 and 10%, 10 and 15%, 15 and 20%, 20 and 25%, 25 and 30%, 30 and 35%, 35 and 40%, 40 and 45%, 45 and 50%, 50 and 55%, 55 and 60%, 60 and 65%, 65 and 70%, 70 and 75%, 75 and 80%, 80 and 85%, 85 and 90%, 90 and 95%, 95 and 99% or more increased activity that results from expression of a heterologous Factor IX gene. For example, increased

activity can be determined by detecting and measuring the coagulation activity and comparing the activity against the the coagulation activity before, e.g., treating the cells or administration to a subject. In some embodiments, increased activity can be determined using by assessing clotting function in an aPTT assay and/or thrombin generation in an TGA-EA assay. Such methods are available and known in the art (e.g. Simioni et al, NEJM 2009).

[0090] Each of the guide sequences shown in Table 1 at SEQ ID NOS:2-33 may further comprise additional nucleotides to form a crRNA and/or guide RNA, e.g., with the following exemplary nucleotide sequence following the guide sequence at its 3' end: GUUUUAGAGCUAGUUAAAUAAGGCUAGUCCGUUAUC (SEQ ID NO: 400) in 5' to 3' orientation. Genomic coordinates are according to human reference genome hg38. In the case of a sgRNA, the above guide sequences may further comprise additional nucleotides to form a sgRNA, e.g., with the following exemplary nucleotide sequence following the 3' end of the guide sequence:

(SEQ ID NO: 401)
GUUUUAGAGCUAGAAAUAAGCAAGUUAAAUAAGGCUAGUCCGUUAUC
ACUUGAAAAAGUGGCACCGAGUCGGUCUUUU
or

(SEQ ID NO: 402)
GUUUUAGAGCUAGAAAUAAGCAAGUUAAAUAAGGCUAGUCCGUUAUC
ACUUGAAAAAGUGGCACCGAGUCGGUGC in
5' to 3' orientation.

[0091] Each of the guide sequences in Table 1 at SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33 may further comprise additional nucleotides to form a crRNA, e.g., with the following exemplary nucleotide sequence following the guide sequence at its 3' end: GUUUUAGAGCUAGCUAUGCU-GUUUG (SEQ ID NO: 400) in 5' to 3' orientation. In the case of a sgRNA, the above guide sequences may further comprise additional nucleotides to form a sgRNA, e.g., with the following exemplary nucleotide sequence following the 3' end of the guide sequence:

(SEQ ID NO: 401)
GUUUUAGAGCUAGAAAUAAGCAAGUUAAAUAAGGCUAGUCCGUUAUC
ACUUGAAAAAGUGGCACCGAGUCGGUCUUUU
or

(SEQ ID NO: 402)
GUUUUAGAGCUAGAAAUAAGCAAGUUAAAUAAGGCUAGUCCGUUAUC
ACUUGAAAAAGUGGCACCGAGUCGGUGC in
5' to 3' orientation.

TABLE 1

Human guide RNA sequences and chromosomal coordinates			
Guide ID	Guide Sequence	Human Genomic Coordinates (hg38)	SEQ ID NO:
G009844	GAGCAACCUC ACUCUUGUCU	chr4: 73405113- 73405133	2
G009851	AUGCAUUUGU UUCAAAAUAU	chr4: 73405000- 73405020	3
G009852	UGCAUUGUU UCAAAAUAUU	chr4: 73404999- 73405019	4

TABLE 1-continued

Human guide RNA sequences and chromosomal coordinates			
Guide ID	Guide Sequence	Human Genomic Coordinates (hg38)	SEQ ID NO:
G009857	AUUUAUGAGA UCAACAGCAC	chr4: 73404761- 73404781	5
G009858	GAUCAACAGC ACAGGUUUG	chr4: 73404753- 73404773	6
G009859	UUAAAUAAG CAUAGUGCAA	chr4: 73404727- 73404747	7
G009860	UAAAGCAUAG UGCAAUGGAU	chr4: 73404722- 73404742	8
G009861	UAGUGCAAUG GAUAGGUCUU	chr4: 73404715- 73404735	9
G009866	UACAAAACU UUAUUUUACU	chr4: 73404452- 73404472	10
G009867	AAAGUUGAAC AAUAGAAAAA	chr4: 73404418- 73404438	11
G009868	AAUGCAUAAU CUAAGUCAAA	chr4: 73405013- 73405033	12
G009874	UAAUAAAUU CAAACAUCCU	chr4: 73404561- 73404581	13
G012747	GCAUCUUAAA AGAAUUAUUU	chr4: 73404478- 73404498	14
G012748	UUUGGCAUUU AUUUCAAAAA	chr4: 73404496- 73404516	15
G012749	UGUAUUUUGUG AAGUCUUACA	chr4: 73404529- 73404549	16
G012750	UCCUAGGUAA AAAAAAAAAA	chr4: 73404577- 73404597	17
G012751	UAAAAUUUCUU UUGCGCACUA	chr4: 73404620- 73404640	18
G012752	UGACUGAAC UUCACAGAAU	chr4: 73404664- 73404684	19
G012753	GACUGAAACU UCACAGAAUA	chr4: 73404665- 73404685	20
G012754	UUCAUUUAG UCUGUCUUUC	chr4: 73404803- 73404823	21
G012755	AUUACUAAAG UUUGAAUUA	chr4: 73404859- 73404879	22
G012756	AAUUUUUAAA AUAGUAUUCU	chr4: 73404897- 73404917	23
G012757	UGAAUUAUUC UUCUGUUAAA	chr4: 73404924- 73404944	24
G012758	AUCAUCCUGA GUUUUUCUGU	chr4: 73404965- 73404985	25
G012759	UUACUAAAAC UUUAAUUUAC	chr4: 73404453- 73404473	26
G012760	ACCUUUUUUU UUUUUUACCU	chr4: 73404581- 73404601	27

TABLE 1-continued

Human guide RNA sequences and chromosomal coordinates		Human Genomic Coordinates (hg38)	SEQ ID NO.
Guide ID	Guide Sequence		
G012761	AGUGCAAUAGG AUAGGUUU	chr4: 73404714- 73404734	28
G012762	UGAUUCCUAC AGAAAAACUC	chr4: 73404973- 73404993	29
G012763	UGGGCAAGGG AAGAAAAAAA	chr4: 73405094- 73405114	30
G012764	CCUCACUCUU GUCUGGGCAA	chr4: 73405107- 73405127	31
G012765	ACCUCACUCU UGUCUGGGCA	chr4: 73405108- 73405128	32
G012766	UGAGCAACCU CACCUUGUC	chr4: 73405114- 73405134	33

[0092] The guide RNA may further comprise a trRNA. In each composition and method embodiment described herein, the crRNA and trRNA may be associated as a single RNA (sgRNA) or may be on separate RNAs (dgRNA). In the context of sgRNAs, the crRNA and trRNA components may be covalently linked, e.g., via a phosphodiester bond or other covalent bond. In some embodiments, the sgRNA comprises one or more linkages between nucleotides that is not a phosphodiester linkage.

[0093] In each of the composition, use, and method embodiments described herein, the guide RNA may comprise two RNA molecules as a “dual guide RNA” or “dgRNA”. The dgRNA comprises a first RNA molecule comprising a crRNA comprising, e.g., a guide sequence shown in Table 1, and a second RNA molecule comprising a trRNA. The first and second RNA molecules may not be covalently linked, but may form a RNA duplex via the base pairing between portions of the crRNA and the trRNA.

[0094] In each of the composition, use, and method embodiments described herein, the guide RNA may comprise a single RNA molecule as a “single guide RNA” or “sgRNA”. The sgRNA may comprise a crRNA (or a portion thereof) comprising a guide sequence shown in Table 1 covalently linked to a trRNA. The sgRNA may comprise 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a guide sequence shown in Table 1. In some embodiments, the crRNA and the trRNA are covalently linked via a linker. In some embodiments, the sgRNA forms a stem-loop structure via the base pairing between portions of the crRNA and the trRNA. In some embodiments, the crRNA and the trRNA are covalently linked via one or more bonds that are not a phosphodiester bond.

[0095] In some embodiments, the trRNA may comprise all or a portion of a trRNA sequence derived from a naturally-occurring CRISPR/Cas system. In some embodiments, the trRNA comprises a truncated or modified wild type trRNA. The length of the trRNA depends on the CRISPR/Cas system used. In some embodiments, the trRNA comprises or consists of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or more than 100 nucleotides. In some embodiments, the trRNA may com-

prise certain secondary structures, such as, for example, one or more hairpin or stem-loop structures, or one or more bulge structures.

[0096] In some embodiments, the target sequence or region within intron 1 of a human albumin locus (SEQ ID NO: 1) may be complementary to the guide sequence of the guide RNA. In some embodiments, the degree of complementarity or identity between a guide sequence of a guide RNA and its corresponding target sequence may be at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%. In some embodiments, the target sequence and the guide sequence of the gRNA may be 100% complementary or identical. In other embodiments, the target sequence and the guide sequence of the gRNA may contain at least one mismatch. For example, the target sequence and the guide sequence of the gRNA may contain 1, 2, 3, 4, or 5 mismatches, where the total length of the guide sequence is about 20, or 20. In some embodiments, the target sequence and the guide sequence of the gRNA may contain 1-4 mismatches where the guide sequence is about 20, or 20 nucleotides.

[0097] In some embodiments, a composition or formulation disclosed herein comprises an mRNA comprising an open reading frame (ORF) encoding an RNA-guided DNA binding agent, such as a Cas nuclease as described herein. In some embodiments, an mRNA comprising an ORF encoding an RNA-guided DNA binding agent, such as a Cas nuclease, is provided, used, or administered.

B. Modified gRNAs and mRNAs

[0098] In some embodiments, the gRNA is chemically modified. A gRNA comprising one or more modified nucleosides or nucleotides is called a “modified” gRNA or “chemically modified” gRNA, to describe the presence of one or more non-naturally and/or naturally occurring components or configurations that are used instead of or in addition to the canonical A, G, C, and U residues. In some embodiments, a modified gRNA is synthesized with a non-canonical nucleoside or nucleotide, is here called “modified.” 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 (an exemplary backbone modification); (ii) alteration, e.g., replacement, of a constituent of the ribose sugar, e.g., of the 2' hydroxyl on the ribose sugar (an exemplary sugar modification); (iii) wholesale replacement of the phosphate moiety with “dephospho” linkers (an exemplary backbone modification); (iv) modification or replacement of a naturally occurring nucleobase, including with a non-canonical nucleobase (an exemplary base modification); (v) replacement or modification of the ribose-phosphate backbone (an exemplary backbone modification); (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, cap or linker (such 3' or 5' cap modifications may comprise a sugar and/or backbone modification); and (vii) modification or replacement of the sugar (an exemplary sugar modification).

[0099] Chemical modifications such as those listed above can be combined to provide modified gRNAs and/or mRNAs comprising nucleosides and nucleotides (collectively “residues”) that can have two, three, four, or more modifications. For example, a modified residue can have a modified sugar and a modified nucleobase. In some embodi-

ments, every base of a gRNA is modified, e.g., all bases have a modified phosphate group, such as a phosphorothioate group. In certain embodiments, all, or substantially all, of the phosphate groups of an gRNA molecule are replaced with phosphorothioate groups. In some embodiments, modified gRNAs comprise at least one modified residue at or near the 5' end of the RNA. In some embodiments, modified gRNAs comprise at least one modified residue at or near the 3' end of the RNA. Certain gRNAs comprise at least one modified residue at or near the 5' end and 3' end of the RNA.

[0100] In some embodiments, the gRNA comprises one, two, three or more modified residues. In some embodiments, at least 5% (e.g., at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or 100%) of the positions in a modified gRNA are modified nucleosides or nucleotides.

[0101] Unmodified nucleic acids can be prone to degradation by, e.g., intracellular nucleases or those found in serum. For example, nucleases can hydrolyze nucleic acid phosphodiester bonds. Accordingly, in one aspect the gRNAs described herein can contain one or more modified nucleosides or nucleotides, e.g., to introduce stability toward intracellular or serum-based nucleases. In some embodiments, the modified gRNA molecules 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, which involves the induction of cytokine expression and release, particularly the interferons, and cell death.

[0102] In some embodiments of a backbone modification, the phosphate group of a modified residue can be modified by replacing one or more of the oxygens with a different substituent. Further, the modified residue, e.g., modified residue present in a modified nucleic acid, can include the wholesale replacement of an unmodified phosphate moiety with a modified phosphate group as described herein. In some embodiments, the backbone modification of the phosphate backbone can include alterations that result in either an uncharged linker or a charged linker with unsymmetrical charge distribution.

[0103] Examples of modified phosphate groups include, phosphorothioate, phosphoroselenates, borano phosphates, borano phosphate esters, hydrogen phosphonates, phosphoroamidates, alkyl or aryl phosphonates and phosphotriesters. 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. The stereogenic phosphorous atom can possess either the “R” configuration (herein Rp) or the “S” configuration (herein Sp). The backbone 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.

[0104] The phosphate group can be replaced by non-phosphorus containing connectors in certain backbone modifications. In some embodiments, the charged 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.

[0105] 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. Such modifications may comprise backbone and sugar modifications. 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.

[0106] The modified nucleosides and modified nucleotides can include one or more modifications to the sugar group, i.e. at sugar modification. For example, the 2' hydroxyl group (OH) can be modified, e.g. 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.

[0107] Examples of 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(CH₂CH₂O)_nCH₂CH₂OR 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 2' hydroxyl group modification can be 2'-O-Me. In some embodiments, the 2' hydroxyl group modification can be a 2'-fluoro modification, which replaces the 2' hydroxyl group with a fluoride. In some embodiments, the 2' hydroxyl group modification can be a 2'-H, which replaces the 2' hydroxyl group with a hydrogen. In some embodiments, the 2' hydroxyl group modification can include “locked” nucleic acids (LNA) in which the 2' hydroxyl can be connected, e.g., by a C₁₋₆ alkylene or C₁₋₆ 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., NH₂; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroaryl amino, or diheteroaryl amino, ethylenediamine, or polyamino) and aminoalkoxy, O(CH₂)_n-amino, (wherein amino can be, e.g., NH₂; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroaryl amino, or diheteroaryl amino, ethylenediamine, or polyamino). In some embodiments, the 2' hydroxyl group modification can include “unlocked” nucleic acids (UNA) in which the ribose ring lacks the C2'-C3' bond. In some embodiments, the 2' hydroxyl group modification can include the methoxyethyl group (MOE), (OCH₂CH₂OCH₃, e.g., a PEG derivative).

[0108] “Deoxy” 2' modifications can include hydrogen (i.e. deoxyribose sugars, e.g., at the overhang portions of partially dsRNA); halo (e.g., bromo, chloro, fluoro, or iodo); amino (wherein amino can be, e.g., NH₂; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroaryl amino, diheteroaryl amino, or amino acid); NH(CH₂CH₂NH)_nCH₂CH₂— 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.

[0109] The sugar modification can comprise a sugar group which may 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 modified nucleic acids can also include abasic sugars. 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.

[0110] The modified nucleosides and modified nucleotides described herein, which can be incorporated into a modified nucleic acid, can include a modified base, also called a 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 residues that can be incorporated into modified nucleic acids. The nucleobase of the nucleotide can be independently selected from a purine, a pyrimidine, a purine analog, or pyrimidine analog. In some embodiments, the nucleobase can include, for example, naturally-occurring and synthetic derivatives of a base.

[0111] In embodiments employing a dual guide RNA, each of the crRNA and the tracr RNA can contain modifi-

cations. Such modifications may be at one or both ends of the crRNA and/or tracr RNA. In embodiments comprising an sgRNA, one or more residues at one or both ends of the sgRNA may be chemically modified, and/or internal nucleosides may be modified, and/or the entire sgRNA may be chemically modified. Certain embodiments comprise a 5' end modification. Certain embodiments comprise a 3' end modification.

[0112] In some embodiments, the guide RNAs disclosed herein comprise one of the modification patterns disclosed in WO2018/107028 A1, filed Dec. 8, 2017, titled “Chemically Modified Guide RNAs,” the contents of which are hereby incorporated by reference in their entirety. In some embodiments, the guide RNAs disclosed herein comprise one of the structures/modification patterns disclosed in US20170114334, the contents of which are hereby incorporated by reference in their entirety. In some embodiments, the guide RNAs disclosed herein comprise one of the structures/modification patterns disclosed in WO2017/136794, WO2017004279, US2018187186, US2019048338, the contents of which are hereby incorporated by reference in their entirety.

[0113] In some embodiments, the sgRNA of the present disclosure comprises the modification patterns shown below in Table 2. “Full Sequence” in Table 2 refers to an sgRNA sequence for each of the guides listed in Table 1. “Full Sequence Modified” shows a modification pattern for each sgRNA.

TABLE 2

sgRNA and modification patterns to sgRNA of human albumin guide sequences					
Guide ID	Full Sequence	SEQ ID NO:	Full Sequence Modified	SEQ ID NO:	
G009844	GAGCAACCUCACUCUUGUCUGUUUU AGAGCUAGAAAAGCAAGUUAAA AAGGCUGGUCCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	34	mG*mA*mG*CAACCUCACUCUUGUCUGU UUUAGAmGmCmUmAmGmAmAmUm AmGmCAAGUAAAUAAGGCUAGUCC GUUAUCAmAmCmUmUmGmAmAmAm AmGmUmGmGmCmAmCmCmGmAmGmUm CmGmGmUmGmCmU*mU*mU*mU	66	
G009851	AUGCAUUUGUUUCAAAAUAUGUUUU AGAGCUAGAAAAGCAAGUUAAA AAGGCUGGUCCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	35	mA*mU*mG*CAUUUGUUUCAAAAUAUG UUUAGAmGmCmUmAmGmAmAmUm AmGmCAAGUAAAUAAGGCUAGUCC UUUAUCAmAmCmUmUmGmAmAmAmAm GmUmGmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	67	
G009852	UGCAUUUGUUUCAAAAUAUGUUUU AGAGCUAGAAAAGCAAGUUAAA AAGGCUGGUCCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	36	mU*mG*mC*AUUUGUUUCAAAAUAUGU UUUAGAmGmCmUmAmGmAmAmAmUmAm GmCAAGUAAAUAAGGCUAGUCCGUUA UCAmAmCmUmUmGmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCmGmUm GmCmU*mU*mU*mU	68	
G009857	AUUUAUGAGAUCAACAGCACGUUUU AGAGCUAGAAAAGCAAGUUAAA AAGGCUGGUCCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	37	mA*mU*mU*UAUGAGAUCAACAGCACGU UUUAGAmGmCmUmAmGmAmAmAmUmAm GmCAAGUAAAUAAGGCUAGUCCGUUA UCAmAmCmUmUmGmAmAmAmAmAmGm UmGmGmCmAmCmCmGmAmGmUmCmGmUm GmGmCmU*mU*mU*mU	69	

TABLE 2 -continued

		sgRNA and modification patterns to sqRNA of human albumin guide sequences			
Guide ID	Full Sequence	SEQ ID NO:	Full Sequence	Modified	SEQ ID NO:
G009858	GAUCAACAGCACAGGUUUUGGUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	38	mG*mA*mU*CAACAGCACAGGUUUUGGU UUUAGAmGmCmUmAmGmAmAmUmAm GmCAAGGUAAAAGGUAGUCGUUA UCAmAmCmUmUmGmAmAmAmAmGm UmGmGmCmAmCmCmGmAmGmUmCmGm GmUmGmCmU*mU*mU*mU		70
G009859	UUAAAAGCAUAGUGCAAGGUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	39	mU*mU*mA*AAUAAAGCAUAGUGCAAGGU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		71
G009860	UAAAGCAUAGUGCAAUGGAUGUUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	40	mU*mA*mA*AGCAUAGUGCAAUGGAUGUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		72
G009861	UAGUGCAAUGGUAGGUUAGGUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	41	mU*mA*mG*UGCAAUGGUAGGUUAGGU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		73
G009866	UACUAAAACUUUACUUUACUGUUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	42	mU*mA*mC*UAAAACUUUACUUUACUGUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		74
G009867	AAAGGUAGAACAAUAGAAAAAGUUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	43	mA*mA*mA*GUUGAACAAUAGAAAAAGUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		75
G009868	AAUGCAUAAUCUAAGUCAAAGUUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	44	mA*mA*mU*GCAUAAUCUAAGUCAAAGUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		76
G009874	UAAAUAUUAACAAACAUCCUGUUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	45	mU*mA*mA*UAAAUAUCAACAUCCUGUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		77
G012747	GCAUCUUAAAAGAAUUAUUGUUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	46	mG*mC*mA*UCUUAAAAGAAUUAUUGUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		78
G012748	UUUGGCAUUUACUUUACUUAAAAGUUUU AGAGCUAGAAAAGCAAGGUAAAA AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	47	mU*mU*mU*GGCAUUUACUUAAAAGUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGGUAAAAGGUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmGmUmGmCm U*mU*mU*mU		79

TABLE 2 -continued

		sgRNA and modification patterns to sqRNA of human albumin guide sequences			
Guide ID	Full Sequence	SEQ ID NO:	Full Sequence	Modified	SEQ ID NO:
G012749	UGUAUUUGUGAAGUCUUACAGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	48	mU*mG*mU*AUUUGUGAAGUCUUACAGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		80
G012750	UCCUAGGUAAAAAAAAAGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	49	mU*mC*mC*UAGGUAAAAAAAAAGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		81
G012751	UAUUUUUUUUUGCGCACUAGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	50	mU*mA*mA*UUUUUUUUUGCGCACUAGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		82
G012752	UGACUGAAACUUCACAGAAUGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	51	mU*mG*mA*CUGAACUUUCACAGAAUGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		83
G012753	GACUGAAACUUCACAGAAUAGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	52	mG*mA*mC*UGAACUUUCACAGAAUAGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		84
G012754	UUCAUUUUAGCUGUCUUUCUGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	53	mU*mU*mC*AUUUUAGUCUGCUUCUGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		85
G012755	AUUAUCUAAGUUUGAAUUAUGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	54	mA*mU*mU*AUCUAAGUUUGAAUUAUGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		86
G012756	AAUUUUUUAAAUAGUAAUCUGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	55	mA*mA*mU*UUUUUUAAAAGUAUUUCUGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		87
G012757	UGAAUUUUUUUCUUCUGUUUAAGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	56	mU*mG*mA*AUUAUUUUUCUGUUUAAGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		88
G012758	AUCAUCCUGAGUUUUUCUGUGUUUU AGAGCUAGAAAUGCAGUAAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	57	mA*mU*mC*AUCCUGAGUUUUUCUGUGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		89

TABLE 2 -continued

Guide ID	Full Sequence	sgRNA and modification patterns to sqRNA of human albumin guide sequences			SEQ ID NO:
		SEQ ID NO:	Full Sequence	Modified	
G012759	UUACUAAAACUUUAAAAACGUUUU AGAGCUAGAAAUAAGCAAGGUAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	58	mU*mU*mA*CUAAAACUUUAAAAACGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		90
G012760	ACCUUUUUUUUUUUUUUACCUGUUUU AGAGCUAGAAAUAAGCAAGGUAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	59	mA*mC*mC*UUUUUUUUUUUUACCUGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		91
G012761	AGUGCAAUGGAUAGGUUUUUGUUUU AGAGCUAGAAAUAAGCAAGGUAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	60	mA*mG*mU*GCAAUGGAUAGGUUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		92
G012762	UGAUUCCUACAGAAAAACUCGUUUU AGAGCUAGAAAUAAGCAAGGUAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	61	mU*mG*mA*UUCCUACAGAAAAACUCGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		93
G012763	UGGGCAAGGGAGAGAAAAAGUUUU AGAGCUAGAAAUAAGCAAGGUAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	62	mU*mG*mG*GCAAGGGAGAGAAAAAGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		94
G012764	CCUCACUCUUGUCUGGGCAAGUUUU AGAGCUAGAAAUAAGCAAGGUAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	63	mC*mC*mU*CACUCUUGUCUGGGCAAGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		95
G012765	ACCUCACUCUUGUCUGGGCAGUUUU AGAGCUAGAAAUAAGCAAGGUAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	64	mA*mC*mC*UACUCUUGUCUGGGCAGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		96
G012766	UGAGCAACCUCACUCUUGUCGUUUU AGAGCUAGAAAUAAGCAAGGUAAAAU AAGGCUAGUCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCUUUU	65	mU*mG*mA*GCAACCUCACUCUUGUCGUUU UAGAmGmCmUmAmGmAmAmUmAmGmC AAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAmAmGmUmGmGm CmAmCmCmGmAmGmUmCmGmUmGmCm U*mU*mU*mU		97

[0114] In some embodiments, the modified sgRNA comprises the following sequence: mN*mN*mN*NNNNNNNNNNNNNNNNNNGUUUU AGAmGmCmUmAmGmAmAmAmUmA mGmCAAGUUAAAUAAGGCUAGUCGUUAUCAm AmCmUmUmGmAmAmAmAm GmUmGmGmC mAmCmCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU (SEQ ID NO: 300), where "N" may be any natural or non-natural nucleotide, and wherein the totality of N's comprise an albumin intron 1 guide

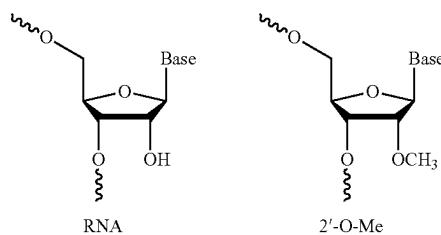
sequence as described in Table 1. For example, encompassed herein is SEQ ID NO: 300, where the N's are replaced with any of the guide sequences disclosed herein in Table 1 (SEQ ID Nos: 2-33).

[0115] For example, encompassed herein is SEQ ID NO: 300, where the N's are replaced with any of the guide sequences disclosed herein in Table 1 (SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33).

[0116] Any of the modifications described below may be present in the gRNAs and mRNAs described herein.

[0117] The terms “mA,” “mC,” “mU,” or “mG” may be used to denote a nucleotide that has been modified with 2'-O-Me.

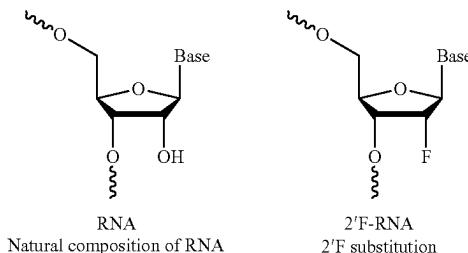
[0118] Modification of 2'-O-methyl can be depicted as follows:



[0119] Another chemical modification that has been shown to influence nucleotide sugar rings is halogen substitution. For example, 2'-fluoro (2'-F) substitution on nucleotide sugar rings can increase oligonucleotide binding affinity and nuclease stability.

[0120] In this application, the terms “fA,” “fC,” “fU,” or “fG” may be used to denote a nucleotide that has been substituted with 2'-F.

[0121] Substitution of 2'-F can be depicted as follows:

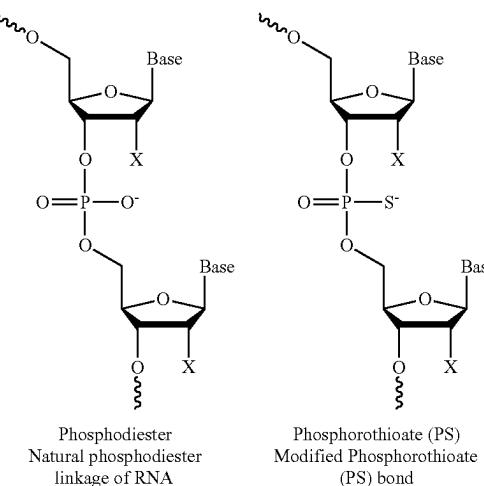


[0122] Phosphorothioate (PS) linkage or bond refers to a bond where a sulfur is substituted for one nonbridging phosphate oxygen in a phosphodiester linkage, for example in the bonds between nucleotides bases. When phosphorothioates are used to generate oligonucleotides, the modified oligonucleotides may also be referred to as S-oligos.

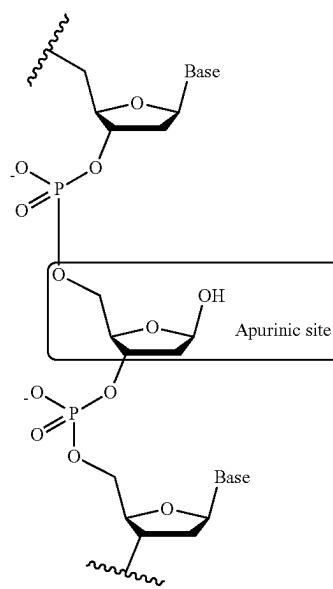
[0123] A “*” may be used to depict a PS modification. In this application, the terms A*, C*, U*, or G* may be used to denote a nucleotide that is linked to the next (e.g., 3') nucleotide with a PS bond.

[0124] In this application, the terms “mA*,” “mC*,” “mU*,” or “mG*” may be used to denote a nucleotide that has been substituted with 2'-O-Me and that is linked to the next (e.g., 3') nucleotide with a PS bond.

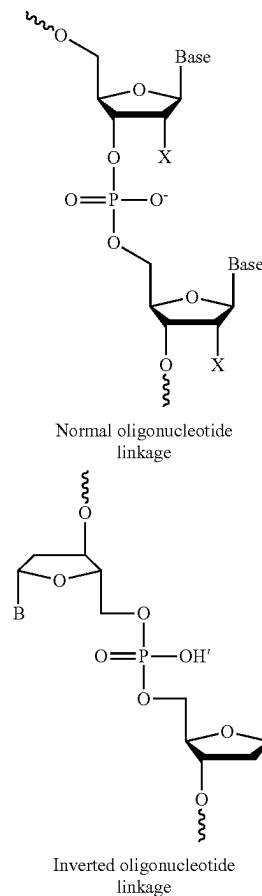
[0125] The diagram below shows the substitution of S— into a nonbridging phosphate oxygen, generating a PS bond in lieu of a phosphodiester bond:



[0126] A basic nucleotides refer to those which lack nitrogenous bases. The figure below depicts an oligonucleotide with an abasic (also known as apurinic) site that lacks a base:



[0127] Inverted bases refer to those with linkages that are inverted from the normal 5' to 3' linkage (i.e., either a 5' to 5' linkage or a 3' to 3' linkage). For example:



[0128] An abasic nucleotide can be attached with an inverted linkage. For example, an abasic nucleotide may be attached to the terminal 5' nucleotide via a 5' to 5' linkage, or an abasic nucleotide may be attached to the terminal 3' nucleotide via a 3' to 3' linkage. An inverted abasic nucleotide at either the terminal 5' or 3' nucleotide may also be called an inverted abasic end cap.

[0129] In some embodiments, one or more of the first three, four, or five nucleotides at the 5' terminus, and one or more of the last three, four, or five nucleotides at the 3' terminus are modified. In some embodiments, the modification is a 2'-O-Me, 2'-F, inverted abasic nucleotide, PS bond, or other nucleotide modification well known in the art to increase stability and/or performance.

[0130] In some embodiments, the first four nucleotides at the 5' terminus, and the last four nucleotides at the 3' terminus are linked with phosphorothioate (PS) bonds.

[0131] In some embodiments, the first three nucleotides at the 5' terminus, and the last three nucleotides at the 3' terminus comprise a 2'-O-methyl (2'-O-Me) modified nucleotide. In some embodiments, the first three nucleotides at the 5' terminus, and the last three nucleotides at the 3' terminus comprise a 2'-fluoro (2'-F) modified nucleotide. In some embodiments, the first three nucleotides at the 5'

terminus, and the last three nucleotides at the 3' terminus comprise an inverted abasic nucleotide.

[0132] In some embodiments, the guide RNA comprises a modified sgRNA. In some embodiments, the sgRNA comprises the modification pattern shown in SEQ ID No: 300, where N is any natural or non-natural nucleotide, and where the totality of the N's comprise a guide sequence that directs a nuclease to a target sequence in human albumin intron 1, e.g., as shown in Table 1.

[0133] In some embodiments, the guide RNA comprises a sgRNA shown in any one of SEQ ID No: 34-97. In some embodiments, the guide RNA comprises a sgRNA comprising any one of the guide sequences of SEQ ID No: 2-33 and the nucleotides of SEQ ID No: 300 wherein the nucleotides of SEQ ID No: 300 are on the 3' end of the guide sequence, and wherein the sgRNA may be modified, e.g., as shown in SEQ ID NO: 300.

[0134] In some embodiments, the guide RNA comprises a sgRNA shown in any one of SEQ ID NOs: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the guide RNA comprises a sgRNA comprising any one of the guide sequences of SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33 and the nucleotides of SEQ ID No: 300 wherein the nucleotides of SEQ ID NO: 300 are on the 3' end of the guide sequence, and wherein the sgRNA may be modified, e.g., as shown in SEQ ID NO: 300.

[0135] As noted above, in some embodiments, a composition or formulation disclosed herein comprises an mRNA comprising an open reading frame (ORF) encoding an RNA-guided DNA binding agent, such as a Cas nuclease as described herein. In some embodiments, an mRNA comprising an ORF encoding an RNA-guided DNA binding agent, such as a Cas nuclease, is provided, used, or administered. As described below, the mRNA comprising a Cas nuclease may comprise a Cas9 nuclease, such as an *S. pyogenes* Cas9 nuclease having cleavage, nickase, and/or site-specific DNA binding activity. In some embodiments, the ORF encoding an RNA-guided DNA nuclease is a “modified RNA-guided DNA binding agent ORF” or simply a “modified ORF,” which is used as shorthand to indicate that the ORF is modified.

[0136] Cas9 ORFs, including modified Cas9 ORFs, are provided herein and are known in the art. As one example, the Cas9 ORF can be codon optimized, such that coding sequence includes one or more alternative codons for one or more amino acids. An “alternative codon” as used herein refers to variations in codon usage for a given amino acid, and may or may not be a preferred or optimized codon (codon optimized) for a given expression system. Preferred codon usage, or codons that are well-tolerated in a given system of expression, is known in the art. The Cas9 coding sequences, Cas9 mRNAs, and Cas9 protein sequences of WO2013/176772, WO2014/065596, WO2016/106121, and WO2019/067910 are hereby incorporated by reference. In particular, the ORFs and Cas9 amino acid sequences of the table at paragraph [0449] WO2019/067910, and the Cas9 mRNAs and ORFs of paragraphs [0214]-[0234] of WO2019/067910 are hereby incorporated by reference.

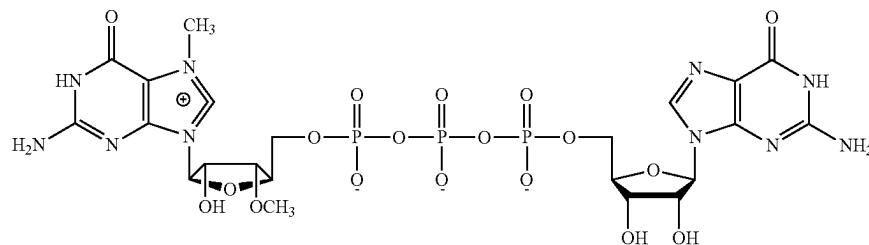
[0137] In some embodiments, the modified ORF may comprise a modified uridine at least at one, a plurality of, or all uridine positions. In some embodiments, the modified uridine is a uridine modified at the 5 position, e.g., with a halogen, methyl, or ethyl. In some embodiments, the modified uridine is a pseudouridine modified at the 1 position, e.g., with a halogen, methyl, or ethyl. The modified uridine can be, for example, pseudouridine, N1-methyl-pseudouridine, 5-methoxyuridine, 5-iodouridine, or a combination thereof. In some embodiments, the modified uridine is

5-methoxyuridine. In some embodiments, the modified uridine is 5-iodouridine. In some embodiments, the modified uridine is pseudouridine. In some embodiments, the modified uridine is N1-methyl-pseudouridine. In some embodiments, the modified uridine is a combination of pseudouridine and N1-methyl-pseudouridine. In some embodiments, the modified uridine is a combination of pseudouridine and 5-methoxyuridine. In some embodiments, the modified uridine is a combination of N1-methyl pseudouridine and 5-methoxyuridine. In some embodiments, the modified uridine is a combination of 5-iodouridine and N1-methyl-pseudouridine. In some embodiments, the modified uridine is a combination of pseudouridine and 5-iodouridine. In some embodiments, the modified uridine is a combination of 5-iodouridine and 5-methoxyuridine.

[0138] In some embodiments, an mRNA disclosed herein comprises a 5' cap, such as a Cap0, Cap1, or Cap2. A 5' cap is generally a 7-methylguanine ribonucleotide (which may be further modified, as discussed below e.g. with respect to ARCA) linked through a 5'-triphosphate to the 5' position of the first nucleotide of the 5'-to-3' chain of the mRNA, i.e., the first cap-proximal nucleotide. In Cap0, the riboses of the first and second cap-proximal nucleotides of the mRNA both

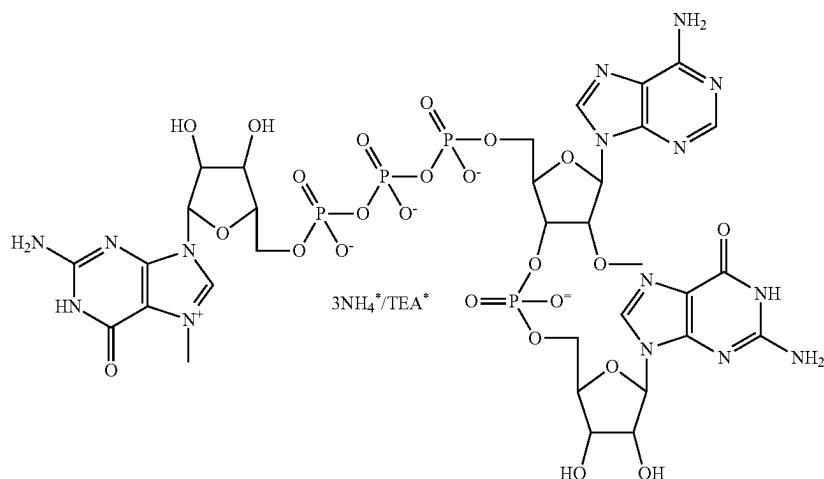
mammalian mRNAs such as human mRNAs, comprise Cap1 or Cap2. Cap0 and other cap structures differing from Cap1 and Cap2 may be immunogenic in mammals, such as humans, due to recognition as "non-self" by components of the innate immune system such as IFIT-1 and IFIT-5, which can result in elevated cytokine levels including type I interferon. Components of the innate immune system such as IFIT-1 and IFIT-5 may also compete with eIF4E for binding of an mRNA with a cap other than Cap1 or Cap2, potentially inhibiting translation of the mRNA.

[0139] A cap can be included co-transcriptionally. For example, ARCA (anti-reverse cap analog; Thermo Fisher Scientific Cat. No. AM8045) is a cap analog comprising a 7-methylguanine 3'-methoxy-5'-triphosphate linked to the 5' position of a guanine ribonucleotide which can be incorporated in vitro into a transcript at initiation. ARCA results in a Cap0 cap in which the 2' position of the first cap-proximal nucleotide is hydroxyl. See, e.g., Stepinski et al., (2001) "Synthesis and properties of mRNAs containing the novel 'anti-reverse' cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl(3'deoxy)GpppG," *RNA* 7: 1486-1495. The ARCA structure is shown below.



comprise a 2'-hydroxyl. In Cap1, the riboses of the first and second transcribed nucleotides of the mRNA comprise a 2'-methoxy and a 2'-hydroxyl, respectively. In Cap2, the riboses of the first and second cap-proximal nucleotides of the mRNA both comprise a 2'-methoxy. See, e.g., Katibah et al. (2014) *Proc Natl Acad Sci USA* 111(33):12025-30; Abbas et al. (2017) *Proc Natl Acad Sci USA* 114(11):E2106-E2115. Most endogenous higher eukaryotic mRNAs, including

[0140] CleanCap™ AG (m7G(5')ppp(5')(2'OMeA)pG; TriLink Biotechnologies Cat. No. N-7113) or CleanCap™ GG (m7G(5')ppp(5')(2'OMeG)pG; TriLink Biotechnologies Cat. No. N-7133) can be used to provide a Cap1 structure co-transcriptionally. 3'-O-methylated versions of CleanCap™ AG and CleanCap™ GG are also available from TriLink Biotechnologies as Cat. Nos. N-7413 and N-7433, respectively. The CleanCap™ AG structure is shown below.



[0141] Alternatively, a cap can be added to an RNA post-transcriptionally. For example, Vaccinia capping enzyme is commercially available (New England Biolabs Cat. No. M2080S) and has RNA triphosphatase and guanyltransferase activities, provided by its D1 subunit, and guanine methyltransferase, provided by its D12 subunit. As such, it can add a 7-methylguanine to an RNA, so as to give Cap0, in the presence of S-adenosyl methionine and GTP. See, e.g., Guo, P. and Moss, B. (1990) *Proc. Natl. Acad. Sci. USA* 87, 4023-4027; Mao, X. and Shuman, S. (1994) *J. Biol. Chem.* 269, 24472-24479.

[0142] In some embodiments, the mRNA further comprises a poly-adenylated (poly-A) tail. In some embodiments, the poly-A tail comprises at least 20, 30, 40, 50, 60, 70, 80, 90, or 100 adenines, optionally up to 300 adenines. In some embodiments, the poly-A tail comprises 95, 96, 97, 98, 99, or 100 adenine nucleotides.

C. Donor Constructs

[0143] The compositions and methods described herein include the use of a nucleic acid construct that comprises a sequence encoding a heterologous Factor IX gene to be inserted into a cut site created by a guide RNA of the present disclosure and an RNA-guided DNA binding agent. As used herein, such a construct is sometimes referred to as a “donor construct/template”. In some embodiments, the construct is a DNA construct. Methods of designing and making various functional/structural modifications to donor constructs are known in the art. In some embodiments, the construct may comprise any one or more of a polyadenylation tail sequence, a polyadenylation signal sequence, splice acceptor site, or selectable marker. In some embodiments, the polyadenylation tail sequence is encoded, e.g., as a “poly-A” stretch, at the 3' end of the coding sequence. Methods of designing a suitable polyadenylation tail sequence and/or polyadenylation signal sequence are well known in the art. For example, the polyadenylation signal sequence AAUAAA (SEQ ID NO: 800) is commonly used in mammalian systems, although variants such as UAUAAA (SEQ ID NO: 801) or AU/GUAAA (SEQ ID NO: 802) have been identified. See, e.g., NJ Proudfoot, *Genes & Dev.* 25(17): 1770-82, 2011.

[0144] In some embodiments, the donor construct comprises a sequence encoding Factor IX, wherein the Factor IX sequence is wild type Factor IX, e.g., SEQ ID NO: 700. In some embodiments, the donor construct comprises a sequence encoding Factor IX, wherein the Factor IX sequence is wild type Factor IX, e.g., SEQ ID NO: 701. In some embodiments, the sequence encodes a variant of Factor IX. For example, the variant can possess increased coagulation activity than wild type Factor IX. For example, the variant Factor IX can comprise one or mutations, such as an amino acid substitution in position R338 (e.g., R338L), relative to SEQ ID NO: 701. In some embodiments, the sequence encodes a Factor IX variant that is 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, SEQ ID NO: 701, or SEQ ID NO: 702, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX. In some embodiments, the sequence encodes a fragment of Factor IX, wherein the fragment possesses at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX.

[0145] In some embodiments, the donor construct comprises a sequence encoding a Factor IX variant, wherein the Factor IX variant activates coagulation in the absence of its cofactor, Factor VIII. Such Factor IX variants can further maintain the activity of wild type Factor IX. Such Factor IX variants can be used to treat hemophilia, such as hemophilia B. For example, such a Factor IX variant can comprise an amino acid substitution at position L6, V181, K265, 1383, E185, or a combination thereof relative to wild type Factor IX (e.g., relative to SEQ ID NO: 701). For example, such a Factor IX variant can comprise an L6F mutation, a V181I mutation, a K265A mutation, an 1383V mutation, an E185D mutation, or a combination thereof relative to wild type Factor IX (e.g., relative to SEQ ID NO: 701).

[0146] In one example, the Factor IX protein can comprise amino acid substitutions at positions L6 and V181. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6 and K265. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6 and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6 and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181 and K265. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181 and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181 and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions K265 and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions K265 and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions 1383 and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, and K265. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, K265, and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, K265, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, 1383, and E186. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, 1383, and E186. In another example, the Factor IX protein can comprise amino acid substitutions at positions K265, 1383, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, K265, and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, K265, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, 1383, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, K265, 1383, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, 1383, and E185.

[0147] In a specific example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, and 1383. In another specific example, the Factor IX protein

can comprise amino acid substitutions at positions V181, K265, 1383, and E185. In another specific example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, K265, and 1383.

[0148] In one example, the Factor IX protein can comprise an L6F mutation and a V181I mutation. In another example, the Factor IX protein can comprise an L6F mutation and a K265A mutation. In another example, the Factor IX protein can comprise an L6F mutation and an I383V mutation. In another example, the Factor IX protein can comprise an L6F mutation and an E185D mutation. In another example, the Factor IX protein can comprise a V181I mutation and a K265A mutation. In another example, the Factor IX protein can comprise a V181I mutation and an I383V mutation. In another example, the Factor IX protein can comprise a V181I mutation and an E185D mutation. In another example, the Factor IX protein can comprise a K265A mutation and an E185D mutation. In another example, the Factor IX protein can comprise an I383V mutation and an E185D mutation. In another example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, and a K265A mutation. In another example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, and an I383V mutation. In another example, the Factor IX protein can comprise an L6F mutation, a K265A mutation, and an E185D mutation. In another example, the Factor IX protein can comprise an L6F mutation, an I383V mutation, and an E186D mutation. In another example, the Factor IX protein can comprise a V181I mutation, a K265A mutation, and an I383V mutation. In another example, the Factor IX protein can comprise a V181I mutation, a K265A mutation, and an E185D mutation. In another example, the Factor IX protein can comprise a V181I mutation, an I383V mutation, and an E186D mutation. In another example, the Factor IX protein can comprise a K265A mutation, an I383V mutation, and an E185D mutation. In another example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, a K265A mutation, and an I383V mutation. In another example, the Factor IX protein can comprise an L6F mutation, a K265A mutation, an I383V mutation, and an E185D mutation. In some embodiments, the Factor IX variant is at least 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX. In certain embodiments, the Factor IX variant is at least 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type

Factor IX and comprises a V181I mutation, a K265A mutation, an I383V mutation, and/or an E185D mutation. In another specific example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, a K265A mutation, and an I383V mutation. In some embodiments, the Factor IX variant is at least 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX and comprises an L6F mutation, a V181I mutation, a K265A mutation, and/or an I383V mutation.

[0150] The length of the construct can vary, depending on the size of the gene to be inserted, and can be, for example, from 200 base pairs (bp) to about 5000 bp, such as about 200 bp to about 2000 bp, such as about 500 bp to about 1500 bp. In some embodiments, the length of the DNA donor template is about 200 bp, or is about 500 bp, or is about 800 bp, or is about 1000 base pairs, or is about 1500 base pairs. In other embodiments, the length of the donor template is at least 200 bp, or is at least 500 bp, or is at least 800 bp, or is at least 1000 bp, or is at least 1500 bp. In other embodiments, the length of the donor template is at least 200 bp, or is at least 500 bp, or is at least 800 bp, or is at least 1000 bp, or is at least 1500 bp, or at least 2000, or at least 2500, or at least 3000, or at least 3500, or at least 4000, or at least 4500, or at least 5000.

[0151] The construct can be DNA or RNA, single-stranded, double-stranded or partially single- and partially double-stranded and can be introduced into a host cell in linear or circular (e.g., minicircle) form. See, e.g., U.S. Patent Publication Nos. 2010/0047805, 2011/0281361, 2011/0207221. If introduced in linear form, the ends of the donor sequence can be protected (e.g., from exonucleolytic degradation) by methods known to those of skill in the art. For example, one or more dideoxynucleotide residues are added to the 3' terminus of a linear molecule and/or self-complementary oligonucleotides are ligated to one or both ends. See, for example, Chang et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4959-4963; Nehls et al. (1996) *Science* 272: 886-889. Additional methods for protecting exogenous polynucleotides from degradation include, but are not limited to, addition of terminal amino group(s) and the use of modified internucleotide linkages such as, for example, phosphorothioates, phosphoramidates, and O-methyl ribose or deoxyribose residues. A construct can be introduced into a cell as part of a vector molecule having additional sequences such as, for example, replication origins, promoters and genes encoding antibiotic resistance. A construct may omit viral elements. Moreover, donor constructs can be introduced as naked nucleic acid, as nucleic acid complexed with an agent such as a liposome or poloxamer, or can be delivered by viruses (e.g., adenovirus, AAV, herpesvirus, retrovirus, lentivirus).

[0152] In some embodiments, the construct may be inserted so that its expression is driven by the endogenous promoter at the insertion site (e.g., the endogenous albumin promoter when the donor is integrated into the host cell's albumin locus). In such cases, the transgene may lack control elements (e.g., promoter and/or enhancer) that drive its expression (e.g., a promoterless construct). Nonetheless, it will be apparent that in other cases the construct may comprise a promoter and/or enhancer, for example a constitutive promoter or an inducible or tissue specific (e.g., liver- or platelet-specific) promoter that drives expression of

[0149] In a specific example, the Factor IX protein can comprise a V181I mutation, an K265A mutation, and an I383V mutation. In another specific example, the Factor IX protein can comprise a V181I mutation, a K265A mutation, an I383V mutation, and an E185D mutation. In some embodiments, the Factor IX variant is at least 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX. In certain embodiments, the Factor IX variant is at least 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type

the functional protein upon integration. The construct may comprise a sequence encoding a heterologous Factor IX protein downstream of and operably linked to a signal sequence encoding a signal peptide. In some embodiments, the nucleic acid construct works in homology-independent insertion of a nucleic acid that encodes a Factor IX protein. In some embodiments, the nucleic acid construct works in non-dividing cells, e.g., cells in which NHEJ, not HR, is the primary mechanism by which double-stranded DNA breaks are repaired. The nucleic acid may be a homology-independent donor construct.

[0153] Some donor constructs comprising a heterologous Factor IX nucleic acid (Factor IX transgene) are capable of insertion into a cut site in a target DNA sequence for a gene editing system (e.g., capable of insertion into a safe harbor gene, such as an albumin locus) by non-homologous end joining. In some cases, such constructs do not comprise homology arms. For example, such constructs can be inserted into a blunt end double-strand break following cleavage with a gene editing system (e.g., CRISPR/Cas system) as disclosed herein. In a specific example, the construct can be delivered via AAV and can be capable of insertion by non-homologous end joining (e.g., the construct can be one that does not comprise homology arms).

[0154] In a specific example, the construct can be inserted via homology-independent targeted integration. For example, the heterologous Factor IX nucleic acid in the construct can be flanked on each side by a target site for a gene editing system (e.g., the same target site as in the target DNA sequence for targeted insertion (e.g., in a safe harbor gene), and the same gene editing system being used to cleave the target DNA sequence for targeted insertion). The gene editing system can then cleave the target sites flanking the heterologous Factor IX nucleic acid. In a specific example, the construct is delivered AAV-mediated delivery, and cleavage of the target sites flanking the heterologous Factor IX nucleic acid can remove the inverted terminal repeats (ITRs) of the AAV. In some methods, the target DNA sequence for targeted insertion (e.g., target DNA sequence in a safe harbor locus, e.g., a gRNA target sequence including the flanking protospacer adjacent motif) is no longer present if the heterologous Factor IX nucleic acid is inserted into the cut site or target DNA sequence in the correct orientation but it is reformed if the heterologous Factor IX nucleic acid is inserted into the cut site or target DNA sequence in the opposite orientation. This can help ensure that the heterologous Factor IX nucleic acid is inserted in the correct orientation for expression.

[0155] Also described herein are bidirectional nucleic acid constructs that allow enhanced insertion and expression of a Factor IX gene. Briefly, various bidirectional constructs disclosed herein comprise at least two nucleic acid segments, wherein one segment (the first segment) comprises a coding sequence that encodes Factor IX (sometimes interchangeably referred to herein as “transgene”), while the other segment (the second segment) comprises a sequence wherein the complement of the sequence encodes Factor IX.

[0156] In one embodiment, a bidirectional construct comprise at least two nucleic acid segments in *cis*, wherein one segment (the first segment) comprises a coding sequence (sometimes interchangeably referred to herein as “transgene”), while the other segment (the second segment) comprises a sequence wherein the complement of the sequence encodes a transgene. The first transgene and the second

transgene may be the same or different. The bidirectional constructs may comprise at least two nucleic acid segments in *cis*, wherein one segment (the first segment) comprises a coding sequence that encodes a heterologous gene in one orientation, while the other segment (the second segment) comprises a sequence wherein its complement encodes the heterologous gene in the other orientation. That is, the first segment is a complement of the second segment (not necessarily a perfect complement); the complement of the second segment is the reverse complement of the first segment (not necessarily a perfect reverse complement though both encode the same heterologous protein). A bidirectional construct may comprise a first coding sequence that encodes a heterologous gene linked to a splice acceptor and a second coding sequence wherein the complement encodes a heterologous gene in the other orientation, also linked to a splice acceptor.

[0157] When used in combination with a gene editing system (e.g., CRISPR/Cas system; zinc finger nuclease (ZFN) system; transcription activator-like effector nuclease (TALEN) system) as described herein, the bidirectionality of the nucleic acid constructs allows the construct to be inserted in either direction (is not limited to insertion in one direction) within a target insertion site, allowing the expression of Factor IX from either a) a coding sequence of one segment (e.g., the left segment encoding “Human F9” of FIG. 1 upper left ssAAV construct), or b) a complement of the other segment (e.g., the complement of the right segment encoding “Human F9” indicated upside down in the upper left ssAAV construct FIG. 1), thereby enhancing insertion and expression efficiency, as exemplified herein. Various known gene editing systems can be used in the practice of the present disclosure, including, e.g., CRISPR/Cas system; zinc finger nuclease (ZFN) system; transcription activator-like effector nuclease (TALEN) system.

[0158] The bidirectional constructs disclosed herein can be modified to include any suitable structural feature as needed for any particular use and/or that confers one or more desired function. In some embodiments, the bidirectional nucleic acid construct disclosed herein does not comprise a homology arm. In some embodiments, the bidirectional nucleic acid construct disclosed herein is a homology-independent donor construct. In some embodiments, owing in part to the bidirectional function of the nucleic acid construct, the bidirectional construct can be inserted into a genomic locus in either direction as described herein to allow for efficient insertion and/or expression of a polypeptide of interest (e.g., Factor IX).

[0159] In some embodiments, the bidirectional nucleic acid construct does not comprise a promoter that drives the expression of Factor IX. For example, the expression of Factor IX is driven by a promoter of the host cell (e.g., the endogenous albumin promoter when the transgene is integrated into a host cell's albumin locus).

[0160] In some embodiments, the bidirectional nucleic acid construct comprises a first segment comprising a coding sequence for Factor IX and a second segment comprising a reverse complement of a coding sequence of Factor IX. Thus, the coding sequence in the first segment is capable of expressing Factor IX, while the complement of the reverse complement in the second segment is also capable of expressing Factor IX. As used herein, “coding sequence” when referring to the second segment comprising a reverse complement sequence refers to the complementary (coding)

strand of the second segment (i.e., the complement coding sequence of the reverse complement sequence in the second segment).

[0161] In some embodiments, the coding sequence that encodes Factor IX in the first segment is less than 100% complementary to the reverse complement of a coding sequence that also encodes Factor IX. That is, in some embodiments, the first segment comprises a coding sequence (1) for Factor IX, and the second segment is a reverse complement of a coding sequence (2) for Factor IX, wherein the coding sequence (1) is not identical to the coding sequence (2). For example, coding sequence (1) and/or coding sequence (2) that encodes for Factor IX can be codon optimized, such that coding sequence (1) and the reverse complement of coding sequence (2) possess less than 100% complementarity. In some embodiments, the coding sequence of the second segment encodes Factor IX using one or more alternative codons for one or more amino acids of the same (i.e., same amino acid sequence) Factor IX encoded by the coding sequence in the first segment. An “alternative codon” as used herein refers to variations in codon usage for a given amino acid, and may or may not be a preferred or optimized codon (codon optimized) for a given expression system. Preferred codon usage, or codons that are well-tolerated in a given system of expression is known in the art.

[0162] In some embodiments, the second segment comprises a reverse complement sequence that adopts different codon usage from that of the coding sequence of the first segment in order to reduce hairpin formation. Such a reverse complement forms base pairs with fewer than all nucleotides of the coding sequence in the first segment, yet it optionally encodes the same polypeptide. In such cases, the coding sequence, e.g. for Polypeptide A, of the first segment many be homologous to, but not identical to, the coding sequence, e.g. for Polypeptide A of the second half of the bidirectional construct. In some embodiments, the second segment comprises a reverse complement sequence that is not substantially complementary (e.g., not more than 70% complementary) to the coding sequence in the first segment. In some embodiments, the second segment comprises a reverse complement sequence that is highly complementary (e.g., at least 90% complementary) to the coding sequence in the first segment. In some embodiments, the second segment comprises a reverse complement sequence having at least about 30%, about 35%, about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, about 97%, or about 99% complementarity to the coding sequence in the first segment.

[0163] In some embodiments, the second segment comprises a reverse complement sequence having 100% complementarity to the coding sequence in the first segment. That is, the sequence in the second segment is a perfect reverse complement of the coding sequence in the first segment. By way of example, the first segment comprises a hypothetical sequence 5' CTGGACCGA 3' (SEQ ID NO: 500) and the second segment comprises the reverse complement of SEQ ID NO: 1—i.e., 5' TCGGTCCAG 3' (SEQ ID NO: 502).

[0164] In some embodiments, the bidirectional nucleic acid construct comprises a first segment comprising a coding sequence for Factor IX (a first polypeptide) and a second segment comprising a reverse complement of a coding sequence of a (second) polypeptide. In some embodiments,

the first and second segments each comprise a coding sequence that encodes the same polypeptide (e.g., Factor IX), as described above. In some embodiments, the first and second segments each comprise a coding sequence that encodes different polypeptides. For example, the first polypeptide is Factor IX and the second polypeptide is Polypeptide B. As a further example, the first polypeptide is Factor IX and the second polypeptide is a variant (e.g., a fragment, mutant, fusion) of Factor IX (e.g., having R338L mutation described herein). A coding sequence that encodes a polypeptide may optionally comprise one or more additional sequences, such as sequences encoding amino- or carboxy-terminal amino acid sequences such as a signal sequence, label sequence (e.g. HiBit), or heterologous functional sequence (e.g. nuclear localization sequence (NLS) or self-cleaving peptide) linked to the polypeptide. A coding sequence that encodes a polypeptide may optionally comprise sequences encoding one or more amino-terminal signal peptide sequences. Each of these additional sequences can be the same or different in the first segment and second segment of the construct.

[0165] In some embodiments, the bidirectional nucleic acid construct is linear. For example, the first and second segments are joined in a linear manner through a linker sequence. In some embodiments, the 5' end of the second segment that comprises a reverse complement sequence is linked to the 3' end of the first segment. In some embodiments, the 5' end of the first segment is linked to the 3' end of the second segment that comprises a reverse complement sequence. In some embodiments, the linker sequence is about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 500, 1000, 1500, 2000 or more nucleotides in length. As would be appreciate by those of skill in the art, other structural elements in addition to, or instead of a linker sequence, can be inserted between the first and second segments.

[0166] The bidirectional constructs disclosed herein can be modified to include any suitable structural feature as needed for any particular use and/or that confers one or more desired function. In some embodiments, the bidirectional nucleic acid construct disclosed herein does not comprise a homology arm. In some embodiments, owing in part to the bidirectional function of the nucleic acid construct, the bidirectional construct can be inserted into a genomic locus in either direction (orientation) as described herein to allow for efficient insertion and/or expression of a polypeptide of interest (e.g., a heterologous Factor IX).

[0167] In some embodiments, one or both of the first and second segment comprises a polyadenylation tail sequence. Methods of designing a suitable polyadenylation tail sequence are well known in the art.

[0168] In some embodiments, one or both of the first and second segment comprises a polyadenylation tail sequence and/or a polyadenylation signal sequence downstream of an open reading frame. In some embodiments, the polyadenylation tail sequence is encoded, e.g., as a “poly-A” stretch, at the 3' end of the first and/or second segment. In some embodiments, a polyadenylation tail sequence is provided co-transcriptionally as a result of a polyadenylation signal sequence that is encoded at or near the 3' end of the first and/or second segment. In some embodiments, a poly-A tail comprises at least 20, 30, 40, 50, 60, 70, 80, 90, or 100 adenines, optionally up to 300 adenines. In some embodi-

ments, the poly-A tail comprises 95, 96, 97, 98, 99, or 100 adenine nucleotides. Methods of designing a suitable polyadenylation tail sequence and/or polyadenylation signal sequence are well known in the art. Suitable splice acceptor sequences are disclosed and exemplified herein, including mouse albumin and human FIX splice acceptor sites.

[0169] In some embodiments, the polyadenylation signal sequence AAUAAA (SEQ ID NO: 800) is commonly used in mammalian systems, although variants such as UAUAAA (SEQ ID NO: 801) or AU/GUAAA (SEQ ID NO: 802) have been identified. See, e.g., NJ Proudfoot, *Genes & Dev.* 25(17):1770-82, 2011. In some embodiments, a polyA tail sequence is included.

[0170] In some embodiments, the constructs disclosed herein can be DNA or RNA, single-stranded, double-stranded, or partially single- and partially double-stranded. For example, the constructs can be single- or double-stranded DNA. In some embodiments, the nucleic acid can be modified (e.g., using nucleoside analogs), as described herein.

[0171] In some embodiments, the constructs disclosed herein comprise a splice acceptor site on either or both ends of the construct, e.g., 5' of an open reading frame in the first and/or second segments, or 5' of one or both transgene sequences. In some embodiments, the splice acceptor site comprises NAG. In further embodiments, the splice acceptor site consists of NAG. In some embodiments, the splice acceptor is an albumin splice acceptor, e.g., an albumin splice acceptor used in the splicing together of exons 1 and 2 of albumin. In some embodiments, the splice acceptor is derived from the human albumin gene. In some embodiments, the splice acceptor is derived from the mouse albumin gene. In some embodiments, the splice acceptor is a F9 (or "FIX") splice acceptor, e.g., the F9 splice acceptor used in the splicing together of exons 1 and 2 of F9. In some embodiments, the splice acceptor is derived from the human F9 gene. In some embodiments, the splice acceptor is derived from the mouse F9 gene. Additional suitable splice acceptor sites useful in eukaryotes, including artificial splice acceptors are known and can be derived from the art. See, e.g., Shapiro, et al., 1987, *Nucleic Acids Res.*, 15, 7155-7174, Burset, et al., 2001, *Nucleic Acids Res.*, 29, 255-259.

[0172] In some embodiments, the bidirectional constructs disclosed herein can be modified on either or both ends to include one or more suitable structural features as needed, and/or to confer one or more functional benefit. For example, structural modifications can vary depending on the method(s) used to deliver the constructs disclosed herein to a host cell—e.g., use of viral vector delivery or packaging into lipid nanoparticles for delivery. Such modifications include, without limitation, e.g., terminal structures such as inverted terminal repeats (ITR), hairpin, loops, and other structures such as toroid. In some embodiments, the constructs disclosed herein comprise one, two, or three ITRs. In some embodiments, the constructs disclosed herein comprise no more than two ITRs. Various methods of structural modifications are known in the art.

[0173] In some embodiments, one or both ends of the construct can be protected (e.g., from exonucleolytic degradation) by methods known in the art. For example, one or more dideoxynucleotide residues are added to the 3' terminus of a linear molecule and/or self-complementary oligonucleotides are ligated to one or both ends. See, for example, Chang et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4959-

4963; Nehls et al. (1996) *Science* 272:886-889. Additional methods for protecting the constructs from degradation include, but are not limited to, addition of terminal amino group(s) and the use of modified internucleotide linkages such as, for example, phosphorothioates, phosphoramidates, and O-methyl ribose or deoxyribose residues.

[0174] In some embodiments, the constructs disclosed herein can be introduced into a cell as part of a vector having additional sequences such as, for example, replication origins, promoters and genes encoding antibiotic resistance. In some embodiments, the constructs can be introduced as naked nucleic acid, as nucleic acid complexed with an agent such as a liposome, polymer, or poloxamer, or can be delivered by viral vectors (e.g., adenovirus, AAV, herpesvirus, retrovirus, lentivirus).

[0175] In some embodiments, although not required for expression, the constructs disclosed herein may also include transcriptional or translational regulatory sequences, for example, promoters, enhancers, insulators, internal ribosome entry sites, sequences encoding peptides, and/or polyadenylation signals.

[0176] In some embodiments, the constructs comprising a coding sequence for Factor IX may include one or more of the following modifications: codon optimization (e.g., to human codons) and/or addition of one or more glycosylation sites. See, e.g., McIntosh et al. (2013) *Blood* (17):3335-44.

D. Gene Editing System

[0177] Various known gene editing systems can be used for targeted insertion of the Factor IX gene in the practice of the present disclosure, including, e.g., CRISPR/Cas system; zinc finger nuclease (ZFN) system; transcription activator-like effector nuclease (TALEN) system. Generally, the gene editing systems involve the use of engineered cleavage systems to induce a double strand break (DSB) or a nick (e.g., a single strand break, or SSB) in a target DNA sequence. Cleavage or nicking can occur through the use of specific nucleases such as engineered ZFN, TALENs, or using the CRISPR/Cas system with an engineered guide RNA to guide specific cleavage or nicking of a target DNA sequence. Further, targeted nucleases are being developed based on the Argonaute system (e.g., from *T. thermophilus*, known as 'TtAgo', see Swarts et al (2014) *Nature* 507 (7491): 258-261), which also may have the potential for uses in genome editing and gene therapy.

[0178] It will be appreciated that for methods that use the guide RNAs disclosed herein, the methods include the use of the CRISPR/Cas system (and any of the donor construct disclosed herein that comprises a sequence encoding Factor IX). It will also be appreciated that the present disclosure contemplates methods of targeted insertion and expression of Factor IX using the bidirectional constructs disclosed herein, which can be performed with or without the guide RNAs disclosed herein (e.g., using a ZFN system to cause a break in a target DNA sequence, creating a site for insertion of the bidirectional construct).

[0179] In some embodiments, a CRISPR/Cas system (e.g., a guide RNA and RNA-guided DNA binding agent) can be used to create a site of insertion at a desired locus within a host genome, at which site a donor construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX disclosed herein can be inserted to express Factor IX. The Factor IX may be heterologous with respect to its insertion site or locus, for example a safe harbor locus from which

Factor IX is not normally expressed, as described herein. Alternatively, in some embodiments, Factor IX may be non-heterologous with respect to its insertion site, for example, insertion of a wild type Factor IX into the endogenous locus to correct a defective Factor IX gene. The safe harbor may be within an albumin gene, such as a human albumin gene. The safe harbor may be within an albumin intron 1 region, e.g., human albumin intron 1. The safe harbor may be a human safe harbor, e.g., for a liver tissue or hepatocyte host cell. In some embodiments, a guide RNA described herein can be used according to the present methods with an RNA-guided DNA binding agent (e.g., Cas nuclease) to create a site of insertion, at which site a donor construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX can be inserted to express Factor IX. The guide RNAs useful for targeted insertion of Factor IX into intron 1 of the human albumin locus are exemplified and described herein (see, e.g., Table 1).

[0180] Methods of using various RNA-guided DNA-binding agents, e.g., a nuclease, such as a Cas nuclease, e.g., Cas9, are also well known in the art. While the use of a bidirectional nucleic acid with a CRISPR/Cas system is exemplified herein, it will be appreciated that suitable variations to the system can also be used. It will be appreciated that, depending on the context, the RNA-guided DNA-binding agent can be provided as a nucleic acid (e.g., DNA or mRNA) or as a protein. In some embodiments, the present method can be practiced in a host cell that already comprises and/or expresses an RNA-guided DNA-binding agent.

[0181] In some embodiments, the RNA-guided DNA-binding agent, such as a Cas9 nuclease, has cleavage activity, which can also be referred to as double-strand endonuclease activity. In some embodiments, the RNA-guided DNA-binding agent, such as a Cas9 nuclease, has nuclease activity, which can also be referred to as single-strand endonuclease activity. In some embodiments, the RNA-guided DNA-binding agent comprises a Cas nuclease. Examples of Cas nucleases include those of the type II CRISPR systems of *S. pyogenes*, *S. aureus*, and other prokaryotes (see, e.g., the list in the next paragraph), and variant or mutant (e.g., engineered, non-naturally occurring, naturally occurring, or other variant) versions thereof. See, e.g., US2016/0312198 A1; US 2016/0312199 A1.

[0182] Non-limiting exemplary species that the Cas nuclease can be derived from include *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus* sp., *Staphylococcus aureus*, *Listeria innocua*, *Lactobacillus gasseri*, *Francisella novicida*, *Wolinella succinogenes*, *Sutterella wadsworthensis*, *Gammaproteobacterium*, *Neisseria meningitidis*, *Campylobacter jejuni*, *Pasteurella multocida*, *Fibrobacter succinogenes*, *Rhodospirillum rubrum*, *Nocardiopsis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Streptosporangium roseum*, *Ali-cyclobacillus acidocaldarius*, *Bacillus pseudomycoides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Lactobacillus buchneri*, *Treponema denticola*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas* sp., *Crocosphaera watsonii*, *Cyanobacteria* sp., *Microcystis aeruginosa*, *Synechococcus* sp., *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor beccii*, *Candidatus Desulforudis*, *Clostridium botulinum*, *Clostridium difficile*, *Finegoldia*

magna, *Natranaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochromatium vinosum*, *Marinobacter* sp., *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc* sp., *Arthrobacteria maxima*, *Arthrobacteria platensis*, *Arthrobacteria* sp., *Lynghya* sp., *Microcoleus chthonoplastes*, *Oscillatoria* sp., *Petrotoga mobilis*, *Thermosiphon africanus*, *Streptococcus pasteurianus*, *Neisseria cinerea*, *Campylobacter lari*, *Parvibaculum lavamentivorans*, *Corynebacterium diphtheriae*, *Acidaminococcus* sp., *Lachnospiraceae* bacterium ND2006, and *Acaryochloris marina*.

[0183] In some embodiments, the Cas nuclease is the Cas9 nuclease from *Streptococcus pyogenes*. In some embodiments, the Cas nuclease is the Cas9 nuclease from *Streptococcus thermophilus*. In some embodiments, the Cas nuclease is the Cas9 nuclease from *Neisseria meningitidis*. In some embodiments, the Cas nuclease is the Cas9 nuclease is from *Staphylococcus aureus*. In some embodiments, the Cas nuclease is the Cpf1 nuclease from *Francisella novicida*. In some embodiments, the Cas nuclease is the Cpf1 nuclease from *Acidaminococcus* sp. In some embodiments, the Cas nuclease is the Cpf1 nuclease from *Lachnospiraceae bacterium* ND2006. In further embodiments, the Cas nuclease is the Cpf1 nuclease from *Francisella tularensis*, *Lachnospiraceae bacterium*, *Butyrivibrio proteo-clasticus*, *Perigrinibacteria bacterium*, *Parcubacteria bacterium*, *Smithella*, *Acidaminococcus*, *Candidatus Methanoplasma termitum*, *Eubacterium eligens*, *Moraxella bovoculi*, *Lepotspira inadai*, *Porphyromonas crevioricanis*, *Prevotella disiens*, or *Porphyromonas macacae*. In certain embodiments, the Cas nuclease is a Cpf1 nuclease from an *Acidaminococcus* or *Lachnospiraceae*.

[0184] In some embodiments, the gRNA together with an RNA-guided DNA-binding agent is called a ribonucleoprotein complex (RNP). In some embodiments, the RNA-guided DNA-binding agent is a Cas nuclease. In some embodiments, the gRNA together with a Cas nuclease is called a Cas RNP. In some embodiments, the RNP comprises Type-I, Type-II, or Type-III components. In some embodiments, the Cas nuclease is the Cas9 protein from the Type-II CRISPR/Cas system. In some embodiment, the gRNA together with Cas9 is called a Cas9 RNP.

[0185] Wild type Cas9 has two nuclease domains: RuvC and HNH. The RuvC domain cleaves the non-target DNA strand, and the HNH domain cleaves the target strand of DNA. In some embodiments, the Cas9 protein comprises more than one RuvC domain and/or more than one HNH domain. In some embodiments, the Cas9 protein is a wild type Cas9. In each of the composition, use, and method embodiments, the Cas induces a double strand break in target DNA.

[0186] In some embodiments, chimeric Cas nucleases are used, where one domain or region of the protein is replaced by a portion of a different protein. In some embodiments, a Cas nuclease domain may be replaced with a domain from a different nuclease such as Fok1. In some embodiments, a Cas nuclease may be a modified nuclease.

[0187] In other embodiments, the Cas nuclease may be from a Type-I CRISPR/Cas system. In some embodiments, the Cas nuclease may be a component of the Cascade complex of a Type-I CRISPR/Cas system. In some embodiments, the Cas nuclease may be a Cas3 protein. In some

embodiments, the Cas nuclease may be from a Type-III CRISPR/Cas system. In some embodiments, the Cas nuclease may have an RNA cleavage activity.

[0188] In some embodiments, the RNA-guided DNA-binding agent has single-strand nickase activity, i.e., can cut one DNA strand to produce a single-strand break, also known as a “nick.” In some embodiments, the RNA-guided DNA-binding agent comprises a Cas nickase. A nickase is an enzyme that creates a nick in dsDNA, i.e., cuts one strand but not the other of the DNA double helix. In some embodiments, a Cas nickase is a version of a Cas nuclease (e.g., a Cas nuclease discussed above) in which an endonucleolytic active site is inactivated, e.g., by one or more alterations (e.g., point mutations) in a catalytic domain. See, e.g., U.S. Pat. No. 8,889,356 for discussion of Cas nickases and exemplary catalytic domain alterations. In some embodiments, a Cas nickase such as a Cas9 nickase has an inactivated RuvC or HNH domain.

[0189] In some embodiments, the RNA-guided DNA-binding agent is modified to contain only one functional nuclelease domain. For example, the agent protein may be modified such that one of the nuclelease domains is mutated or fully or partially deleted to reduce its nucleic acid cleavage activity. In some embodiments, a nickase is used having a RuvC domain with reduced activity. In some embodiments, a nickase is used having an inactive RuvC domain. In some embodiments, a nickase is used having an HNH domain with reduced activity. In some embodiments, a nickase is used having an inactive HNH domain.

[0190] In some embodiments, a conserved amino acid within a Cas protein nuclease domain is substituted to reduce or alter nuclease activity. In some embodiments, a Cas nuclease may comprise an amino acid substitution in the RuvC or RuvC-like nuclease domain. Exemplary amino acid substitutions in the RuvC or RuvC-like nuclease domain include D10A (based on the *S. pyogenes* Cas9 protein). See, e.g., Zetsche et al. (2015) *Cell* Oct 22:163(3): 759-771. In some embodiments, the Cas nuclease may comprise an amino acid substitution in the HNH or HNH-like nuclease domain. Exemplary amino acid substitutions in the HNH or HNH-like nuclease domain include E762A, H840A, N863A, H983A, and D986A (based on the *S. pyogenes* Cas9 protein). See, e.g., Zetsche et al. (2015). Further exemplary amino acid substitutions include D917A, E1006A, and D1255A (based on the *Francisella novicida* U112 Cpf1 (FnCpf1) sequence (UniProtKB-AOQ7Q2 (CPF1_FRATN)).

[0191] In some embodiments, a nickase is provided in combination with a pair of guide RNAs that are complementary to the sense and antisense strands of the target sequence, respectively. In this embodiment, the guide RNAs direct the nickase to a target sequence and introduce a DSB by generating a nick on opposite strands of the target sequence (i.e., double nicking). In some embodiments, a nickase is used together with two separate guide RNAs targeting opposite strands of DNA to produce a double nick in the target DNA. In some embodiments, a nickase is used together with two separate guide RNAs that are selected to be in close proximity to produce a double nick in the target DNA.

[0192] In some embodiments, the RNA-guided DNA-binding agent comprises one or more heterologous functional domains (e.g., is or comprises a fusion polypeptide).

[0193] In some embodiments, the heterologous functional domain may facilitate transport of the RNA-guided DNA-binding agent into the nucleus of a cell. For example, the heterologous functional domain may be a nuclear localization signal (NLS). In some embodiments, the RNA-guided DNA-binding agent may be fused with 1-10 NLS(s). In some embodiments, the RNA-guided DNA-binding agent may be fused with 1-5 NLS(s). In some embodiments, the RNA-guided DNA-binding agent may be fused with one NLS. Where one NLS is used, the NLS may be linked at the N-terminus or the C-terminus of the RNA-guided DNA-binding agent sequence. It may also be inserted within the RNA-guided DNA-binding agent sequence. In other embodiments, the RNA-guided DNA-binding agent may be fused with more than one NLS. In some embodiments, the RNA-guided DNA-binding agent may be fused with 2, 3, 4, or 5 NLSs. In some embodiments, the RNA-guided DNA-binding agent may be fused with two NLSs. In certain circumstances, the two NLSs may be the same (e.g., two SV40 NLSs) or different. In some embodiments, the RNA-guided DNA-binding agent is fused to two SV40 NLS sequences linked at the carboxy terminus. In some embodiments, the RNA-guided DNA-binding agent may be fused with two NLSs, one linked at the N-terminus and one at the C-terminus. In some embodiments, the RNA-guided DNA-binding agent may be fused with 3 NLSs. In some embodiments, the RNA-guided DNA-binding agent may be fused with no NLS. In some embodiments, the NLS may be a monopartite sequence, such as, e.g., the SV40 NLS, PKKKRKV (SEQ ID NO: 600) or PKKKRRV (SEQ ID NO: 601). In some embodiments, the NLS may be a bipartite sequence, such as the NLS of nucleoplasmin, KRPAATK-KAGQAKKKK (SEQ ID NO: 602). In a specific embodiment, a single PKKKRKV (SEQ ID NO: 600) NLS may be linked at the C-terminus of the RNA-guided DNA-binding agent. One or more linkers are optionally included at the fusion site.

III. Delivery Methods

[0194] The guide RNA, RNA-guided DNA binding agents (e.g., Cas nuclease), and nucleic acid constructs (e.g., bidirectional construct) disclosed herein can be delivered to a host cell or population of host cells or a subject, in vivo or ex vivo, using various known and suitable methods available in the art. The guide RNA, RNA-guided DNA binding agents, and nucleic acid constructs can be delivered individually or together in any combination, using the same or different delivery methods as appropriate.

[0195] Conventional viral and non-viral based gene delivery methods can be used to introduce the guide RNA disclosed herein as well as the RNA-guided DNA binding agent and donor construct in cells (e.g., mammalian cells) and target tissues. As further provided herein, non-viral vector delivery systems nucleic acids such as non-viral vectors, plasmid vectors, and, e.g. naked nucleic acid, and nucleic acid complexed with a delivery vehicle such as a liposome, lipid nanoparticle (LNP), or polyplex. Viral vector delivery systems include DNA and RNA viruses.

[0196] Methods and compositions for non-viral delivery of nucleic acids include electroporation, lipofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, LNPs, polycation or lipid:nucleic acid conjugates, naked nucleic acid (e.g., naked DNA/RNA), artificial viruses, and agent-enhanced uptake of DNA. Sonoporation

using, e.g., the Sonitron 2000 system (Rich-Mar) can also be used for delivery of nucleic acids.

[0197] Additional exemplary nucleic acid delivery systems include those provided by AmaxaBiosystems (Cologne, Germany), Maxcyte, Inc. (Rockville, Md.), BTX Molecular Delivery Systems (Holliston, Ma.) and Copernicus Therapeutics Inc., (see for example U.S. Pat. No. 6,008,336). Lipofection is described in e.g., U.S. Pat. Nos. 5,049,386; 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (e.g., Transfectam™ and Lipofectin™). The preparation of lipid:nucleic acid complexes, including targeted liposomes such as immunolipid complexes, is well known in the art, and as described herein.

[0198] Various delivery systems (e.g., vectors, liposomes, LNPs) containing the guide RNAs, RNA-guided DNA binding agent, and donor construct, singly or in combination, can also be administered to an organism for delivery to cells *in vivo* or administered to a cell or cell culture *ex vivo*. Administration is by any of the routes normally used for introducing a molecule into ultimate contact with blood, fluid, or cells including, but not limited to, injection, infusion, topical application and electroporation. Suitable methods of administering such nucleic acids are available and well known to those of skill in the art.

[0199] In certain embodiments, the present disclosure provides DNA or RNA vectors encoding any of the compositions disclosed herein—e.g., a guide RNA comprising any one or more of the guide sequences described herein; or a construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX. In some embodiments, the vector also comprises a sequence encoding an RNA-guided DNA binding agent. In certain embodiments, the invention comprises DNA or RNA vectors encoding any one or more of the compositions described herein, or in any combination. In some embodiments, the vectors further comprise, e.g., promoters, enhancers, and regulatory sequences. In some embodiments, the vector that comprises a bidirectional construct comprising a sequence that encodes Factor IX does not comprise a promoter that drives Factor IX expression. For example, the expression of the Factor IX polypeptide is driven by a promoter of the host cell (e.g., the endogenous albumin promoter when the transgene is integrated into a host cell's albumin locus). In some embodiments, the bidirectional nucleic acid construct includes a first segment and a second segment, each having a splice acceptor upstream of a transgene. In certain embodiments, the splice acceptor is compatible with the splice donor sequence of the host cell's safe harbor site, e.g. the splice donor of intron 1 of a human albumin gene. In some embodiments, the vector that comprises a guide RNA comprising any one or more of the guide sequences described herein also comprises one or more nucleotide sequence(s) encoding a crRNA, a trRNA, or a crRNA and trRNA, as disclosed herein.

[0200] In some embodiments, the vector comprises a nucleotide sequence encoding a guide RNA described herein. In some embodiments, the vector comprises one copy of the guide RNA. In other embodiments, the vector comprises more than one copy of the guide RNA. In embodiments with more than one guide RNA, the guide RNAs may be non-identical such that they target different target sequences, or may be identical in that they target the same target sequence. In some embodiments where the vectors comprise more than one guide RNA, each guide

RNA may have other different properties, such as activity or stability within a complex with an RNA-guided DNA nuclease, such as a Cas RNP complex. In some embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to at least one transcriptional or translational control sequence, such as a promoter, a 3' UTR, or a 5' UTR. In one embodiment, the promoter may be a tRNA promoter, e.g., tRNA^{Lys3}, or a tRNA chimera. See Mefford et al., *RNA*. 2015 21:1683-9; Scherer et al., *Nucleic Acids Res.* 2007 35: 2620-2628. In some embodiments, the promoter may be recognized by RNA polymerase III (Pol III). Non-limiting examples of Pol III promoters include U6 and H1 promoters. In some embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to a mouse or human U6 promoter. In other embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to a mouse or human H1 promoter. In embodiments with more than one guide RNA, the promoters used to drive expression may be the same or different. In some embodiments, the nucleotide encoding the crRNA of the guide RNA and the nucleotide encoding the trRNA of the guide RNA may be provided on the same vector. In some embodiments, the nucleotide encoding the crRNA and the nucleotide encoding the trRNA may be driven by the same promoter. In some embodiments, the crRNA and trRNA may be transcribed into a single transcript. For example, the crRNA and trRNA may be processed from the single transcript to form a double-molecule guide RNA. Alternatively, the crRNA and trRNA may be transcribed into a single-molecule guide RNA (sgRNA). In other embodiments, the crRNA and the trRNA may be driven by their corresponding promoters on the same vector. In yet other embodiments, the crRNA and the trRNA may be encoded by different vectors.

[0201] In some embodiments, the nucleotide sequence encoding the guide RNA may be located on the same vector comprising the nucleotide sequence encoding an RNA-guided DNA binding agent such as a Cas protein. In some embodiments, expression of the guide RNA and of the RNA-guided DNA binding agent such as a Cas protein may be driven by their own corresponding promoters. In some embodiments, expression of the guide RNA may be driven by the same promoter that drives expression of the RNA-guided DNA binding agent such as a Cas protein. In some embodiments, the guide RNA and the RNA-guided DNA binding agent such as a Cas protein transcript may be contained within a single transcript. For example, the guide RNA may be within an untranslated region (UTR) of the RNA-guided DNA binding agent such as a Cas protein transcript. In some embodiments, the guide RNA may be within the 5' UTR of the transcript. In other embodiments, the guide RNA may be within the 3' UTR of the transcript. In some embodiments, the intracellular half-life of the transcript may be reduced by containing the guide RNA within its 3' UTR and thereby shortening the length of its 3' UTR. In additional embodiments, the guide RNA may be within an intron of the transcript. In some embodiments, suitable splice sites may be added at the intron within which the guide RNA is located such that the guide RNA is properly spliced out of the transcript. In some embodiments, expression of the RNA-guided DNA binding agent such as a Cas protein and the guide RNA from the same vector in close temporal proximity may facilitate more efficient formation of the CRISPR RNP complex.

[0202] In some embodiments, the nucleotide sequence encoding the guide RNA and/or RNA-guided DNA binding agent may be located on the same vector comprising the construct that comprises a Factor IX gene. In some embodiments, proximity of the construct comprising the Factor IX gene and the guide RNA (and/or the RNA-guided DNA binding agent) on the same vector may facilitate more efficient insertion of the construct into a site of insertion created by the guide RNA/RNA-guided DNA binding agent.

[0203] In some embodiments, the vector comprises one or more nucleotide sequence(s) encoding a sgRNA and an mRNA encoding an RNA-guided DNA binding agent, which can be a Cas protein, such as Cas9 or Cpf1. In some embodiments, the vector comprises one or more nucleotide sequence(s) encoding a crRNA, a trRNA, and an mRNA encoding an RNA-guided DNA binding agent, which can be a Cas protein, such as, Cas9 or Cpf1. In one embodiment, the Cas9 is from *Streptococcus pyogenes* (i.e., Spy Cas9). In some embodiments, the nucleotide sequence encoding the crRNA, trRNA, or crRNA and trRNA (which may be a sgRNA) comprises or consists of a guide sequence flanked by all or a portion of a repeat sequence from a naturally-occurring CRISPR/Cas system. The nucleic acid comprising or consisting of the crRNA, trRNA, or crRNA and trRNA may further comprise a vector sequence wherein the vector sequence comprises or consists of nucleic acids that are not naturally found together with the crRNA, trRNA, or crRNA and trRNA.

[0204] In some embodiments, the crRNA and the trRNA are encoded by non-contiguous nucleic acids within one vector. In other embodiments, the crRNA and the trRNA may be encoded by a contiguous nucleic acid. In some embodiments, the crRNA and the trRNA are encoded by opposite strands of a single nucleic acid. In other embodiments, the crRNA and the trRNA are encoded by the same strand of a single nucleic acid.

[0205] In some embodiments, the vector comprises a donor construct (e.g., the bidirectional nucleic acid construct) comprising a sequence that encodes Factor IX, as disclosed herein. In some embodiments, in addition to the donor construct (e.g., bidirectional nucleic acid construct) disclosed herein, the vector may further comprise nucleic acids that encode the guide RNAs described herein and/or nucleic acid encoding an RNA-guided DNA-binding agent (e.g., a Cas nuclease such as Cas9). In some embodiments, a nucleic acid encoding an RNA-guided DNA-binding agent are each or both on a separate vector from a vector that comprises the donor construct (e.g., bidirectional construct) disclosed herein. In any of the embodiments, the vector may include other sequences that include, but are not limited to, promoters, enhancers, regulatory sequences, as described herein. In some embodiments, the promoter does not drive the expression of Factor IX of the donor construct (e.g., bidirectional construct). In some embodiments, the vector comprises one or more nucleotide sequence(s) encoding a crRNA, a trRNA, or a crRNA and trRNA. In some embodiments, the vector comprises one or more nucleotide sequence(s) encoding a sgRNA and an mRNA encoding an RNA-guided DNA nuclease, which can be a Cas nuclease (e.g., Cas9). In some embodiments, the vector comprises one or more nucleotide sequence(s) encoding a crRNA, a trRNA, and an mRNA encoding an RNA-guided DNA nuclease, which can be a Cas nuclease, such as, Cas9. In some embodiments, the Cas9 is from *Streptococcus pyo-*

genes (i.e., Spy Cas9). In some embodiments, the nucleotide sequence encoding the crRNA, trRNA, or crRNA and trRNA (which may be a sgRNA) comprises or consists of a guide sequence flanked by all or a portion of a repeat sequence from a naturally-occurring CRISPR/Cas system. The nucleic acid comprising or consisting of the crRNA, trRNA, or crRNA and trRNA may further comprise a vector sequence wherein the vector sequence comprises or consists of nucleic acids that are not naturally found together with the crRNA, trRNA, or crRNA and trRNA.

[0206] In some embodiments, the vector may be circular. In other embodiments, the vector may be linear. In some embodiments, the vector may be enclosed in a lipid nanoparticle, liposome, non-lipid nanoparticle, or viral capsid. Non-limiting exemplary vectors include plasmids, phage-mids, cosmids, artificial chromosomes, minichromosomes, transposons, viral vectors, and expression vectors.

[0207] In some embodiments, the vector may be a viral vector. In some embodiments, the viral vector may be genetically modified from its wild type counterpart. For example, the viral vector may comprise an insertion, deletion, or substitution of one or more nucleotides to facilitate cloning or such that one or more properties of the vector is changed. Such properties may include packaging capacity, transduction efficiency, immunogenicity, genome integration, replication, transcription, and translation. In some embodiments, a portion of the viral genome may be deleted such that the virus is capable of packaging exogenous sequences having a larger size. In some embodiments, the viral vector may have an enhanced transduction efficiency. In some embodiments, the immune response induced by the virus in a host may be reduced. In some embodiments, viral genes (such as, e.g., integrase) that promote integration of the viral sequence into a host genome may be mutated such that the virus becomes non-integrating. In some embodiments, the viral vector may be replication defective. In some embodiments, the viral vector may comprise exogenous transcriptional or translational control sequences to drive expression of coding sequences on the vector. In some embodiments, the virus may be helper-dependent. For example, the virus may need one or more helper virus to supply viral components (such as, e.g., viral proteins) required to amplify and package the vectors into viral particles. In such a case, one or more helper components, including one or more vectors encoding the viral components, may be introduced into a host cell or population of host cells along with the vector system described herein. In other embodiments, the virus may be helper-free. For example, the virus may be capable of amplifying and packaging the vectors without a helper virus. In some embodiments, the vector system described herein may also encode the viral components required for virus amplification and packaging.

[0208] Non-limiting exemplary viral vectors include adeno-associated virus (AAV) vector, lentivirus vectors, adenovirus vectors, helper dependent adenoviral vectors (HDAd), herpes simplex virus (HSV-1) vectors, bacteriophage T4, baculovirus vectors, and retrovirus vectors. In some embodiments, the viral vector may be an AAV vector. In other embodiments, the viral vector may a lentivirus vector.

[0209] In some embodiments, "AAV" refers all serotypes, subtypes, and naturally-occurring AAV as well as recombinant AAV. "AAV" may be used to refer to the virus itself or

a derivative thereof. The term "AAV" includes AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AA VRh.64R1, AA Vhu.37, AA VRh.8, AA VRh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AA VRh10, AA VLK03, AV10, AAV11, AAV12, rh10, and hybrids thereof, avian AAV, bovine AAV, canine AAV, equine AAV, primate AAV, nonprimate AAV, and ovine AAV. The genomic sequences of various serotypes of AAV, as well as the sequences of the native terminal repeats (TRs), Rep proteins, and capsid subunits are known in the art. Such sequences may be found in the literature or in public databases such as GenBank. A "AAV vector" as used herein refers to an AAV vector comprising a heterologous sequence not of AAV origin (i.e., a nucleic acid sequence heterologous to AAV), typically comprising a sequence encoding a heterologous polypeptide of interest. The construct may comprise an AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AA VRh.64R1, AA Vhu.37, AA VRh.8, AA VRh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AA VRh10, AA VLK03, AV10, AAV11, AAV12, rh10, and hybrids thereof, avian AAV, bovine AAV, canine AAV, equine AAV, primate AAV, nonprimate AAV, and ovine AAV capsid sequence. In general, the heterologous nucleic acid sequence (the transgene) is flanked by at least one, and generally by two, AAV inverted terminal repeat sequences (ITRs). An AAV vector may either be single-stranded (ssAAV) or self-complementary (scAAV).

[0210] In some embodiments, the lentivirus may be non-integrating. In some embodiments, the viral vector may be an adenovirus vector. In some embodiments, the adenovirus may be a high-cloning capacity or "gutless" adenovirus, where all coding viral regions apart from the 5' and 3' inverted terminal repeats (ITRs) and the packaging signal ('I') are deleted from the virus to increase its packaging capacity. In yet other embodiments, the viral vector may be an HSV-1 vector. In some embodiments, the HSV-1-based vector is helper dependent, and in other embodiments it is helper independent. For example, an amplicon vector that retains only the packaging sequence requires a helper virus with structural components for packaging, while a 30kb-deleted HSV-1 vector that removes non-essential viral functions does not require helper virus. In additional embodiments, the viral vector may be bacteriophage T4. In some embodiments, the bacteriophage T4 may be able to package any linear or circular DNA or RNA molecules when the head of the virus is emptied. In further embodiments, the viral vector may be a baculovirus vector. In yet further embodiments, the viral vector may be a retrovirus vector. In embodiments using AAV or lentiviral vectors, which have smaller cloning capacity, it may be necessary to use more than one vector to deliver all the components of a vector system as disclosed herein. For example, one AAV vector may contain sequences encoding an RNA-guided DNA binding agent such as a Cas protein (e.g., Cas9), while a second AAV vector may contain one or more guide sequences.

[0211] In some embodiments, the vector system may be capable of driving expression of one or more nuclease components in a cell. In some embodiments, the bidirectional construct, optionally as part of a vector system, may comprise a promoter capable of driving expression of a coding sequence in a cell. In some embodiments, the vector does not comprise a promoter that drives expression of one or more coding sequences once it is integrated in a cell (e.g.,

uses the host cell's endogenous promoter such as when inserted at intron 1 of an albumin locus, as exemplified herein). In some embodiments, the cell may be a eukaryotic cell, such as, e.g., a yeast, plant, insect, or mammalian cell. In some embodiments, the eukaryotic cell may be a mammalian cell. In some embodiments, the eukaryotic cell may be a rodent cell. In some embodiments, the eukaryotic cell may be a human cell. Suitable promoters to drive expression in different types of cells are known in the art. In some embodiments, the promoter may be wild type. In other embodiments, the promoter may be modified for more efficient or efficacious expression. In yet other embodiments, the promoter may be truncated yet retain its function. For example, the promoter may have a normal size or a reduced size that is suitable for proper packaging of the vector into a virus.

[0212] In some embodiments, the vector may comprise a nucleotide sequence encoding an RNA-guided DNA binding agent such as a Cas protein (e.g., Cas9) described herein. In some embodiments, the nuclease encoded by the vector may be a Cas protein. In some embodiments, the vector system may comprise one copy of the nucleotide sequence encoding the nuclease. In other embodiments, the vector system may comprise more than one copy of the nucleotide sequence encoding the nuclease. In some embodiments, the nucleotide sequence encoding the nuclease may be operably linked to at least one transcriptional or translational control sequence. In some embodiments, the nucleotide sequence encoding the nuclease may be operably linked to at least one promoter.

[0213] In some embodiments, the vector may comprise any one or more of the constructs comprising a heterologous Factor IX gene described herein. In some embodiments, the Factor IX gene may be operably linked to at least one transcriptional or translational control sequence. In some embodiments, the Factor IX gene may be operably linked to at least one promoter. In some embodiments, the Factor IX gene is not linked to a promoter that drives the expression of the heterologous gene.

[0214] In some embodiments, the promoter may be constitutive, inducible, or tissue-specific. In some embodiments, the promoter may be a constitutive promoter. Non-limiting exemplary constitutive promoters include cytomegalovirus immediate early promoter (CMV), simian virus (SV40) promoter, adenovirus major late (MLP) promoter, Rous sarcoma virus (RSV) promoter, mouse mammary tumor virus (MMTV) promoter, phosphoglycerate kinase (PGK) promoter, elongation factor-alpha (EF1a) promoter, ubiquitin promoters, actin promoters, tubulin promoters, immunoglobulin promoters, a functional fragment thereof, or a combination of any of the foregoing. In some embodiments, the promoter may be a CMV promoter. In some embodiments, the promoter may be a truncated CMV promoter. In other embodiments, the promoter may be an EF1a promoter. In some embodiments, the promoter may be an inducible promoter. Non-limiting exemplary inducible promoters include those inducible by heat shock, light, chemicals, peptides, metals, steroids, antibiotics, or alcohol. In some embodiments, the inducible promoter may be one that has a low basal (non-induced) expression level, such as, e.g., the Tet-On® promoter (Clontech).

[0215] In some embodiments, the promoter may be a tissue-specific promoter, e.g., a promoter specific for expression in the liver.

[0216] In some embodiments, the compositions comprise a vector system. In some embodiments, the vector system may comprise one single vector. In other embodiments, the vector system may comprise two vectors. In additional embodiments, the vector system may comprise three vectors. When different guide RNAs are used for multiplexing, or when multiple copies of the guide RNA are used, the vector system may comprise more than three vectors.

[0217] In some embodiments, the vector system may comprise inducible promoters to start expression only after it is delivered to a target cell. Non-limiting exemplary inducible promoters include those inducible by heat shock, light, chemicals, peptides, metals, steroids, antibiotics, or alcohol. In some embodiments, the inducible promoter may be one that has a low basal (non-induced) expression level, such as, e.g., the Tet-On® promoter (Clontech).

[0218] In additional embodiments, the vector system may comprise tissue-specific promoters to start expression only after it is delivered into a specific tissue.

[0219] The vector comprising: a guide RNA, RNA-binding DNA binding agent, or donor construct comprising a sequence encoding Factor IX, individually or in any combination, may be delivered by liposome, a nanoparticle, an exosome, or a microvesicle. The vector may also be delivered by a lipid nanoparticle (LNP). One or more guide RNA, RNA-binding DNA binding agent (e.g. mRNA), or donor construct comprising a sequence encoding a heterologous protein, individually or in any combination, may be delivered by liposome, a nanoparticle, an exosome, or a microvesicle. One or more guide RNA, RNA-binding DNA binding agent (e.g. mRNA), or donor construct comprising a sequence encoding a heterologous protein, individually or in any combination, may be delivered by LNP.

[0220] Lipid nanoparticles (LNPs) are a well-known means for delivery of nucleotide and protein cargo, and may be used for delivery of any of the guide RNAs, RNA-guided DNA binding agent, and/or donor construct (e.g., bidirectional construct) disclosed herein. In some embodiments, the LNPs deliver the compositions in the form of nucleic acid (e.g., DNA or mRNA), or protein (e.g., Cas nuclease), or nucleic acid together with protein, as appropriate.

[0221] In some embodiments, provided herein is a method for delivering any of the guide RNAs described herein and/or donor construct (e.g., bidirectional construct) disclosed herein, alone or in combination, to a host cell or a population of host cells or a subject, wherein any one or more of the components is associated with an LNP. In some embodiments, the method further comprises an RNA-guided DNA binding agent (e.g., Cas9 or a sequence encoding Cas9).

[0222] In some embodiments, provided herein is a composition comprising any of the guide RNAs described herein and/or donor construct (e.g., bidirectional construct) disclosed herein, alone or in combination, with an LNP. In some embodiments, the composition further comprises an RNA-guided DNA binding agent (e.g., Cas9 or a sequence encoding Cas9).

[0223] In some embodiments, the LNPs comprise cationic lipids. In some embodiments, the LNPs comprise (9Z,12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl octadeca-9,12-dienoate, also called 3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl (9Z,12Z)-octadeca-9,12-dienoate) or another ionizable lipid.

See, e.g., lipids of PCT/US2018/053559 (filed Sep. 28, 2018), WO/2017/173054, WO2015/095340, and WO2014/136086, as well as references provided therein. In some embodiments, the LNPs comprise molar ratios of a cationic lipid amine to RNA phosphate (N:P) of about 4.5, 5.0, 5.5, 6.0, or 6.5. In some embodiments, the term cationic and ionizable in the context of LNP lipids is interchangeable, e.g., wherein ionizable lipids are cationic depending on the pH.

[0224] In some embodiments, LNPs associated with the bidirectional construct disclosed herein are for use in preparing a medicament for treating a disease or disorder. The disease or disorder may be a Factor IX deficiency such as hemophilia B.

[0225] In some embodiments, any of the guide RNAs described herein, RNA-guided DNA binding agents, and/or donor construct (e.g., bidirectional construct) disclosed herein, alone or in combination, whether naked or as part of a vector, is formulated in or administered via a lipid nanoparticle; see e.g., WO/2017/173054 the contents of which are hereby incorporated by reference in their entirety.

[0226] In some embodiments, an LNP composition is encompassed comprising: an RNA component and a lipid component, wherein the lipid component comprises an amine lipid such as a biodegradable, ionizable lipid. In some instances, the lipid component comprises biodegradable, ionizable lipid, cholesterol, DSPC, and PEG-DMG.

[0227] It will be apparent that a guide RNA disclosed herein, an RNA-guided DNA binding agent (e.g., Cas nuclease or a nucleic acid encoding a Cas nuclease), and a donor construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX can be delivered using the same or different systems. For example, the guide RNA, Cas nuclease, and construct can be carried by the same vector (e.g., AAV). Alternatively, the Cas nuclease (as a protein or mRNA) and/or gRNA can be carried by a plasmid or LNP, while the donor construct can be carried by a vector such as AAV. Furthermore, the different delivery systems can be administered by the same or different routes (e.g. by infusion; by injection, such as intramuscular injection, tail vein injection, or other intravenous injection; by intraperitoneal administration and/or intramuscular injection).

[0228] The different delivery systems can be delivered in vitro or in vivo simultaneously or in any sequential order. In some embodiments, the donor construct, guide RNA, and Cas nuclease can be delivered in vitro or in vivo simultaneously, e.g., in one vector, two vectors, individual vectors, one LNP, two LNPs, individual LNPs, or a combination thereof. In some embodiments, the donor construct can be delivered in vivo or in vitro, as a vector and/or associated with a LNP, prior to (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or more days) delivering the guide RNA and/or Cas nuclease, as a vector and/or associated with a LNP singly or together as a ribonucleoprotein (RNP). As a further example, the guide RNA and Cas nuclease, as a vector and/or associated with a LNP singly or together as a ribonucleoprotein (RNP), can be delivered in vivo or in vitro, prior to (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or more days) delivering the construct, as a vector and/or associated with a LNP.

[0229] In some embodiments, the present disclosure also provides pharmaceutical formulations for administering any of the guide RNAs disclosed herein. In some embodiments, the pharmaceutical formulation includes an RNA-guided

DNA binding agent (e.g., Cas nuclease) and a donor construct comprising a coding sequence of a therapeutic heterologous gene, as disclosed herein. Pharmaceutical formulations suitable for delivery into a subject (e.g., human subject) are well known in the art.

IV. Methods of Use

[0230] The gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents described herein are useful for introducing a Factor IX nucleic acid to a host cell or population of host cells, *in vivo* or *in vitro*. In some embodiments, the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents described herein are useful for expressing Factor IX in a host cell or population of host cells, or in a subject in need thereof. In some embodiments, the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents described herein are useful for treating hemophilia (e.g., hemophilia B) in a subject in need thereof. Administration of any one or more of the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents described herein will increase Factor IX protein levels and/or Factor IX activity levels, e.g. circulating, serum, or plasma levels. In some embodiments, the effectiveness of the treatment can be assessed by measuring serum or plasma Factor IX activity, wherein an increase in the subject's plasma level and/or activity of Factor IX indicates effectiveness of the treatment. In some embodiments, the effectiveness of the treatment can be assessed by measuring serum or plasma Factor IX protein and/or activity levels, wherein an increase in the subject's plasma level and/or activity of Factor IX indicates effectiveness of the treatment. In some embodiments, effectiveness of the treatment can be determined by assessing clotting function in an aPTT assay and/or thrombin generation in an TGA-EA assay. In some embodiments, effectiveness of the treatment can be determined by assessing the level of Factor IX, e.g., circulating Factor IX, can be measured by a coagulation and/or an immunologic assay, e.g., an sandwich immunoassay, ELISA (see, e.g., Example 13), MSD (see, e.g., Example 14).

[0231] In normal or healthy individuals, Factor IX activity and antigen levels vary between about 50 and 160% of normal pooled plasma which is about 3-5 µg/ml, based on its purification from adult human plasma Amiral et al., Clin. Chem. 30(9), 1512-16, 1984 at Table 2; see also Osterud et al., 1978. Individuals having less than 50% of normal plasma level of Factor IX activity and/or antigen levels are classified as having hemophilia. In particular, individuals with less than about 1% active factor are classified as having severe haemophilia, while those with about 1-5% active factor have moderate haemophilia. Individuals with mild haemophilia have between about 6-49% of normal levels of active clotting factor. In some embodiments, the level of circulating factor IX can be measured by a coagulation and/or an immunologic assay, which methods are well known in the art (e.g. Simioni et al, NEJM 2009, Adcock et al., Coagulation Handbook, Esoterix Laboratory Services, 2006). An immunologic method for detecting hFIX protein, and a method of functionally normalizing Factor IX activity of a hyperfunctional hFIX variant is found in Example 13.

In some embodiments, Factor IX, e.g., circulating Factor IX, can be measured by a coagulation and/or an immunologic assay, e.g., an sandwich immunoassay, ELISA (see, e.g., Example 13), MSD (see, e.g., Example 14).

[0232] Accordingly, in some embodiments, the compositions and methods disclosed herein are useful for increasing plasma levels of Factor IX or Factor IX activity levels in a subject having hemophilia to about 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, or more, of normal level.

[0233] In some embodiments, the compositions and methods disclosed herein are useful for increasing Factor IX activity and/or levels, for example increasing circulating FIX protein levels to about 0.05, 0.1, 0.2, 0.5, 1, 2, 3, or 4 µg/ml. FIX protein levels may reach about 150 µg/ml, or more. In some embodiments, the compositions and methods disclosed herein are useful for increasing Factor IX protein levels to about 4 µg/ml. In some embodiments, the compositions and methods disclosed herein are useful for increasing Factor IX protein levels to about 4 µg/ml to about 5 µg/ml, about 4 µg/ml to 6 µg/ml, about 4 µg/ml to 8 µg/ml, about 4 µg/ml to about 10 µg/ml, or more. In some embodiments, the compositions and methods disclosed herein are useful for increasing Factor IX protein levels to about 0.1 µg/ml to about 10 µg/ml, about 1 µg/ml to about 10 µg/ml, about 0.1 µg/ml to about 6 µg/ml, about 1 µg/ml to about 6 µg/ml, about 2 µg/ml to about 5 µg/ml, or about 3 µg/ml to about 5 µg/ml. For example, the compositions and methods disclosed herein are useful for increasing plasma levels of Factor IX in a subject having hemophilia to about 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150 µg/ml, or more.

[0234] In some embodiments, the compositions and methods disclosed herein are useful for increasing plasma levels of Factor IX activity and/or levels in a subject having hemophilia by about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 1%, 6%, 17%, %18%, %19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, or more, as compared to the subject's plasma level and/or activity of Factor IX before administration.

[0235] In some embodiments, the compositions and methods disclosed herein are useful for increasing Factor IX protein and/or Factor IX activity in a host cell or population of host cells by about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, or more as compared to a Factor IX level and/or activity before administration to the host cell or population of host cells, e.g. a normal level. In some embodiments, the cell is a liver cell or a population of liver

cells. In some embodiments, the liver cell is hepatocyte or the population of liver cells are hepatocytes.

[0236] In some embodiments, the method comprises administering a guide RNA and an RNA-guided DNA binding agent (such as an mRNA encoding a Cas9 nuclease) in an LNP. In further embodiments, the method comprises administering an AAV nucleic acid construct encoding a Factor IX protein, such as an bidirectional FIX construct. CRISPR/Cas9 LNP, comprising guide RNA and an mRNA encoding a Cas9, can be administered intravenously. AAV FIX donor construct can be administered intravenously. Exemplary dosing of CRISPR/Cas9 LNP includes about 0.1, 0.25, 0.3, 0.5, 1, 2, 3, 4, 5, 6, 8, or 10 mpk (RNA). The units mg/kg and mpk are being used interchangeably herein. Exemplary dosing of AAV comprising a nucleic acid encoding a FIX protein includes an MOI of about 10^{11} , 10^{12} , 10^{13} , and 10^{14} vg/kg, optionally the MOI may be about 1×10^{13} to 1×10^{14} vg/kg.

[0237] In some embodiments, the method comprises expressing a therapeutically effective amount of the Factor IX protein. In some embodiments, the method comprises achieving a therapeutically effective level of circulating Factor IX coagulation activity in an individual. In particular embodiments, the method comprises achieving Factor IX activity of at least about 5% to about 50% of normal. The method may comprise achieving Factor IX activity of at least about 50% to about 150% of normal. In certain embodiments, the method comprises achieving an increase in Factor IX activity over the patient's baseline Factor IX activity of at least about 1% to about 50% of normal Factor IX activity, or at least about 5% to about 50% of normal Factor IX activity, or at least about 50% to about 150% of normal Factor IX activity.

[0238] In some embodiments, the method further comprises achieving a durable effect, e.g. at least 1 month, 2 months, 6 months, 1 year, or 2 year effect. In some embodiments, the method further comprises achieving the therapeutic effect in a durable and sustained manner, e.g. at least 1 month, 2 months, 6 months, 1 year, or 2 year effect. In some embodiments, the level of circulating Factor IX activity and/or level is stable for at least 1 month, 2 months, 6 months, 1 year, or more. In some embodiments a steady-state activity and/or level of FIX protein is achieved by at least 7 days, at least 14 days, or at least 28 days. In additional embodiments, the method comprises maintaining Factor IX activity and/or levels after a single dose for at least 1, 2, 4, or 6 months, or at least 1, 2, 3, 4, or 5 years.

[0239] In additional embodiments involving insertion into the albumin locus, the individual's circulating albumin levels are normal. The method may comprise maintaining the individual's circulating albumin levels within +5%, +10%, +15%, +20%, or +50% of normal circulating albumin levels. In certain embodiments, the individual's albumin levels are unchanged as compared to the albumin levels of untreated individuals by at least week 4, week 8, week 12, or week 20. In certain embodiments, the individual's albumin levels transiently drop then return to normal levels. In particular, the methods may comprise detecting no significant alterations in levels of plasma albumin.

[0240] In some embodiments, the invention comprises a method or use of modifying (e.g., creating a double strand break in) an albumin gene, such as a human albumin gene, comprising, administering or delivering to a host cell or population of host cells any one or more of the gRNAs,

donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents (e.g., Cas nuclease) described herein. In some embodiments, the invention comprises a method or use of modifying (e.g., creating a double strand break in) an albumin intron 1 region, such as a human albumin intron 1, comprising, administering or delivering to a host cell or population of host cells any one or more of the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents (e.g., Cas nuclease) described herein. In some embodiments, the invention comprises a method or use of modifying (e.g., creating a double strand break in) a human safe harbor, such as liver tissue or hepatocyte host cell, comprising, administering or delivering to a host cell or population of host cells any one or more of the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents (e.g., Cas nuclease) described herein. Insertion within a safe harbor locus, such as an albumin locus, allows overexpression of the Factor IX gene without significant deleterious effects on the host cell or cell population, such as hepatocytes or liver cells. In some embodiments, the invention comprises a method or use of modifying (e.g., creating a double strand break in) intron 1 of a human albumin locus comprising, administering or delivering to a host cell or population of host cells any one or more of the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents (e.g., Cas nuclease) described herein. In some embodiments, the guide RNA comprises a guide sequence that contains at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides that bind within intron 1 of a human albumin locus (SEQ ID NO: 1). In some embodiments, the guide RNA comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOS: 34-97. In some embodiments, the guide RNA comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNA

comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOS: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the method is performed *in vitro*. In some embodiments, the method is performed *in vivo*. In some embodiments, the donor construct is a bidirectional construct that comprises a sequence encoding Factor IX. In some embodiments, the host cell is a liver cell, such as. In additional embodiments, the liver cell is a hepatocyte.

[0241] In some embodiments, the invention comprises a method or use of introducing a Factor IX nucleic acid to a host cell or population of host cells comprising, administering or delivering any one or more of the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents (e.g., Cas nuclease) described herein. In some embodiments, the guide RNA comprises a guide sequence that contains at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides that are capable of binding to a region within intron 1 of human albumin locus (SEQ ID NO: 1). In some embodiments, the guide RNA comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17,

18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOs: 34-97. In some embodiments, the guide RNA comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOs: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOs: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOs: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the method is *in vitro*. In some embodiments, the method is *in vivo*. In some embodiments, the donor construct is a bidirectional construct that comprises a sequence encoding Factor IX. In some embodiments, the host cell is a liver cell, or the population of host cells are liver cells, such as hepatocyte.

[0242] In some embodiments, the invention comprises a method or use of expressing Factor IX in a host cell or a population of host cells comprising, administering or delivering any one or more of the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents (e.g., Cas nuclease) described herein. In some embodiments, the guide RNA comprises a guide sequence that contains at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides that are capable of binding to a region within intron 1 of human albumin locus (SEQ ID NO: 1). In some embodiments, the guide RNA comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS:2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected

from the group consisting of SEQ ID NOS: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOS: 34-97. In some embodiments, the guide RNA comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOS: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the method is *in vitro*. In some embodiments, the method is *in vivo*. In some embodiments, the donor construct is a bidirectional construct that comprises a sequence encoding Factor IX. In some embodiments, the host cell is a liver cell, or the population of host cells are liver cells, such as hepatocyte.

[0243] In some embodiments, the invention comprises a method or use of treating hemophilia (e.g., hemophilia B) comprising, administering or delivering any one or more of the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents (e.g., Cas nuclease) described herein to a subject in need thereof. In some embodiments, the guide RNA comprises a guide sequence that contains at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides that are capable of binding to a region within intron 1 of human albumin locus (SEQ ID NO: 1). In some embodiments, the guide RNA comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-33. In

some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2, 8, 13, 19, 28, 29, 31, 32, 33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOS: 34-97. In some embodiments, the guide RNA comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNA comprises a sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence selected from the group consisting of SEQ ID NOS: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 95%, 90%, 85%, 80%, or 75% identical to a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 2-5, 10-17, 21-27, and 29-33. In some embodiments, the guide RNAs disclosed herein comprise a guide sequence that is selected from the group consisting of SEQ ID NOS: 34-37, 42-49, 53-59, 61-69, 74-81, 85-91, and 93-97. In some embodiments, the donor construct is a bidirectional construct that comprises a sequence encoding Factor IX. In some embodiments, the host cell is a liver cell, or the population of host cells are liver cells, such as hepatocytes.

[0244] As described herein, the donor construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX, guide RNA, and RNA-guided DNA binding agent can be delivered using any suitable delivery system and method known in the art. The compositions can be delivered *in vitro* or *in vivo* simultaneously or in any sequential order. In some embodiments, the donor construct, guide RNA, and Cas nuclease can be delivered *in vitro* or *in vivo* simultaneously, e.g., in one vector, two vectors, individual vectors, one LNP, two LNPs, individual LNPs, or a combination

thereof. In some embodiments, the donor construct can be delivered in vivo or in vitro, as a vector and/or associated with a LNP, prior to (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or more days) delivering the guide RNA and/or Cas nuclease, as a vector and/or associated with a LNP singly or together as a ribonucleoprotein (RNP). In some embodiments, the donor construct can be delivered in multiple administrations, e.g., every day, every two days, every three days, every four days, every week, every two weeks, every three weeks, or every four weeks. In some embodiments, the donor construct can be delivered at one-week intervals, e.g., at week 1, week 2, and week 3, etc. As a further example, the guide RNA and Cas nuclease, as a vector and/or associated with a LNP singly or together as a ribonucleoprotein (RNP), can be delivered in vivo or in vitro, prior to (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or more days) delivering the construct, as a vector and/or associated with a LNP. In some embodiments, the albumin guide RNA can be delivered in multiple administrations, e.g., every day, every two days, every three days, every four days, every week, every two weeks, every three weeks, or every four weeks. In some embodiments, the the albumin guide RNA can be delivered at one-week intervals, e.g., at week 1, week 2, and week 3, etc. In some embodiments, the Cas nuclease can be delivered in multiple administrations, e.g., can be delivered every day, every two days, every three days, every four days, every week, every two weeks, every three weeks, or every four weeks. In some embodiments, the Cas nuclease can be delivered at one-week intervals, e.g., at week 1, week 2, and week 3, etc. In some embodiments, the guide RNA and Cas nuclease are associated with an LNP and delivered to the host cell or the population of host cells prior to delivering the Factor IX donor construct.

[0245] In some embodiments, the donor construct comprises a sequence encoding Factor IX, wherein the Factor IX sequence is wild type Factor IX, e.g., SEQ ID NO: 700. In some embodiments, the donor construct comprises a sequence encoding Factor IX, wherein the Factor IX sequence is wild type Factor IX, e.g., SEQ ID NO: 701. In some embodiments, the sequence encodes a variant of Factor IX. For example, the variant possesses increased coagulation activity than wild type Factor IX. For example, the variant Factor IX comprises one or more mutations, such as an amino acid substitution in position R338 (e.g., R338L), relative to SEQ ID NO: 701. In some embodiments, the sequence encodes a Factor IX variant that is 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, SEQ ID NO: 701, or SEQ ID NO: 702, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX. In some embodiments, the sequence encodes a fragment of Factor IX, wherein the fragment possesses at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX.

[0246] In one example, the Factor IX protein can comprise amino acid substitutions at positions L6 and V181. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6 and K265. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6 and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6 and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions

V181 and K265. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181 and an 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181 and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions K265 and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions K265 and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions 1383 and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, and K265. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, K265, and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, K265, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, 1383, and E186. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, 1383, and E186. In another example, the Factor IX protein can comprise amino acid substitutions at positions K265, 1383, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, K265, and 1383. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, 1383, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions L6, K265, 1383, and E185. In another example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, 1383, and E185.

[0247] In some embodiments, the donor construct comprises a sequence encoding a Factor IX variant, wherein the Factor IX variant activates coagulation in the absence of its cofactor, Factor VIII (expression results in therapeutically relevant FVIII mimetic activity). Such Factor IX variants can further maintain the activity of wild type Factor IX. For example, such a Factor IX variant can comprise an amino acid substitution at position L6, V181, K265, 1383, E185, or a combination thereof relative to wild type Factor IX (e.g., relative to SEQ ID NO: 701). For example, such a Factor IX variant can comprise an L6F mutation, a V181I mutation, a K265A mutation, an I383V mutation, an E185D mutation, or a combination thereof relative to wild type Factor IX (e.g., relative to SEQ ID NO: 701).

[0248] In a specific example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, and 1383. In another specific example, the Factor IX protein can comprise amino acid substitutions at positions V181, K265, 1383, and E185. In another specific example, the Factor IX protein can comprise amino acid substitutions at positions L6, V181, K265, and 1383.

[0249] In one example, the Factor IX protein can comprise an L6F mutation and a V181I mutation. In another example, the Factor IX protein can comprise an L6F mutation and a K265A mutation. In another example, the Factor IX protein can comprise an L6F mutation and an I383V mutation. In

another example, the Factor IX protein can comprise an L6F mutation and an E185D mutation. In another example, the Factor IX protein can comprise a V181I mutation and a K265A mutation. In another example, the Factor IX protein can comprise a V181I mutation and an I383V mutation. In another example, the Factor IX protein can comprise a V181I mutation and an E185D mutation. In another example, the Factor IX protein can comprise a K265A mutation and an I383V mutation. In another example, the Factor IX protein can comprise a K265A mutation and an E185D mutation. In another example, the Factor IX protein can comprise an I383V mutation and an E185D mutation. In another example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, and a K265A mutation. In another example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, and an I383V mutation. In another example, the Factor IX protein can comprise an L6F mutation, a V181I mutation and an E185D mutation. In another example, the Factor IX protein can comprise an L6F mutation, a K265A mutation, and an I383V mutation. In another example, the Factor IX protein can comprise an L6F mutation, a K265A mutation, and an E185D mutation. In another example, the Factor IX protein can comprise an L6F mutation, an I383V mutation, and an E186D mutation. In another example, the Factor IX protein can comprise a V181I mutation, a K265A mutation, and an I383V mutation. In another example, the Factor IX protein can comprise a V181I mutation, an I383V mutation, and an E186D mutation. In another example, the Factor IX protein can comprise a K265A mutation, an I383V mutation, and an E185D mutation. In another example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, a K265A mutation, and an I383V mutation. In another example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, an I383V mutation, and an E185D mutation. In another example, the Factor IX protein can comprise a V181I mutation, a K265A mutation, an I383V mutation, and an E185D mutation.

[0250] In a specific example, the Factor IX protein can comprise a V181I mutation, a K265A mutation, and an I383V mutation. In another specific example, the Factor IX protein can comprise a V181I mutation, a K265A mutation, an I383V mutation, and an E185D mutation. In some embodiments, the Factor IX protein is at least 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX. In certain embodiments, the Factor IX variant is at least 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity as compared to wild type Factor IX and comprises a V181I mutation, a K265A mutation, an I383V mutation, and/or an E185D mutation. In another specific example, the Factor IX protein can comprise an L6F mutation, a V181I mutation, a K265A mutation, and an I383V mutation. In certain embodiments, the Factor IX variant is at least 80%, 85%, 90%, 93%, 95%, 97%, 99% identical to SEQ ID NO: 700, having at least 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%, or more, activity

as compared to wild type Factor IX and comprises a V181I mutation, a K265A mutation, and an I383V mutation.

[0251] In some embodiments, the host cell is a liver cell, or the population of host cells are liver cells. In some embodiments, the host cell is, or the population of host cells are, any suitable non-dividing cell. As used herein, a “non-dividing cell” refers to cells that are terminally differentiated and do not divide, as well as quiescent cells that do not divide but retains the ability to re-enter cell division and proliferation. Liver cells, for example, retain the ability to divide (e.g., when injured or resected), but do not typically divide. During mitotic cell division, homologous recombination is a mechanism by which the genome is protected and double-stranded breaks are repaired. In some embodiments, a “non-dividing” cell refers to a cell in which homologous recombination (HR) is not the primary mechanism by which double-stranded DNA breaks are repaired in the cell, e.g., as compared to a control dividing cell. In some embodiments, a “non-dividing” cell refers to a cell in which non-homologous end joining (NHEJ) is the primary mechanism by which double-stranded DNA breaks are repaired in the cell, e.g., as compared to a control dividing cell. Non-dividing cell types have been described in the literature, e.g. by active NHEJ double-stranded DNA break repair mechanisms. See, e.g. Iyama, DNA Repair (Amst.) 2013, 12(8): 620-636. In some embodiments, the host cell includes, but is not limited to, a liver cell, a muscle cell, or a neuronal cell. In some embodiments, the host cell, or the population of host cells are, is a hepatocyte, such as a mouse, cyno, or human hepatocyte. In some embodiments, the host cell is a myocyte, such as a mouse, cyno, or human myocyte. In some embodiments, provided herein is a host cell composition comprising any one or more guide RNA described herein, alone or in combination with an RNA-guided DNA binding protein. In some embodiments, provided herein is a host cell composition comprising any one or more of the vectors described herein.

[0252] In some embodiments, the donor construct (e.g., bidirectional construct) is administered in a nucleic acid vector, such as an AAV vector, e.g., AAV8. In some embodiments, the donor construct does not comprise a homology arm.

[0253] In some embodiments, the subject is a mammal. In some embodiments, the subject is human. In some embodiments, the subject is cow, pig, monkey, sheep, dog, cat, fish, or poultry.

[0254] In some embodiments, the donor construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX, guide RNA, and RNA-guided DNA binding agent are administered intravenously. In some embodiments, the donor construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX, guide RNA, and RNA-guided DNA binding agent are administered into the hepatic circulation.

[0255] In some embodiments, a single administration of a donor construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX, guide RNA, and RNA-guided DNA binding agent is sufficient to increase expression of Factor IX to a desirable level. In other embodiments, more than one administration of a composition comprising a donor construct (e.g., bidirectional construct) comprising a sequence encoding Factor IX, guide RNA, and RNA-guided DNA binding agent may be beneficial to maximize therapeutic effects.

[0256] In some embodiments, the present disclosure includes combination therapies comprising any one or more of the gRNAs, donor construct (e.g., bidirectional construct comprising a sequence encoding Factor IX), and RNA-guided DNA binding agents (e.g., Cas nuclease) described herein together with an additional therapy suitable for treating hemophilia, as described above. For example, the methods of the present disclosure can be combined with the use of other hemostatic agents, blood factors, and medications. For example, the subject may be administered a therapeutically effective amount of one or more factors selected from the group consisting of factor XI, factor XII, prekallikrein, high molecular weight kininogen (HMWK), factor V, factor VII, factor VIII, factor X, factor XIII, factor II, factor VIIa, and von Willebrands factor.

[0257] In some embodiments, treatment may further comprise administering a procoagulant, such as an activator of the intrinsic coagulation pathway, including factor Xa, factor IXa, factor Xla, factor XIIa, and VIIIa, prekallekrein, and

high-molecular weight kininogen; or an activator of the extrinsic coagulation pathway, including tissue factor, factor VIIa, factor Va, and factor Xa.

[0258] This description and exemplary embodiments should not be taken as limiting. For the purposes of this specification and appended embodiments, unless otherwise indicated, all numbers expressing quantities, percentages, or proportions, and other numerical values used in the specification and embodiments, are to be understood as being modified in all instances by the term "about," to the extent they are not already so modified. Accordingly, unless indicated to the contrary, the numerical parameters set forth in the following specification and attached embodiments are approximations that may vary depending upon the desired properties sought to be obtained. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the embodiments, each numerical parameter should at least be construed in light of the number of reported significant digits and by applying ordinary rounding techniques.

Human Factor IX Protein Sequence (SEQ ID NO: 700) NCBI Ref: NP_000124:
 MQRVNIMIAESPLGITICLLGYLSSAETVFLDHENAKILNRPKRYNSGKLEEFVQGNL
 ERECMEEKCSFEEAREVFENTERTTTEFWKQVYDGDQCESNPCLNGGCKDDINSYECWCP
 FGFGEGKNCELDVTCNIKNGRCEQFCKNSADNKVVCSTEGYRLAENQKSCEPAVPPCGR
 VSVSQTSKLTRAETVFPDVYVNSTEAEILDNTQSTQSFNDFTRVGGEDAKPGQFPW
 QVVLNGKVDAGCGGSIVNEKWIVTAAHCVETGVKITVVAEHNIEETEHTEQKRNVIRII
 PHHNYNAAINKYNHDIALLELDEPLVLSNVYVTPICIADKEYTNIFLKFGSGYVSGWGRVF
 HKGRSLVLYQLRPLVDRATCLRSTKPTIYNNMFCAGFHEGGRDSQGDSGPHVTEVE
 GTSFLTGIIISWGEECAMKGKYIYTKVSRYNWIKEKTKL

Human Factor IX Nucleotide Sequence (SEQ ID NO: 706) NCBI Ref: NM_000133:
 1 accactttca caatctgcta gcaaaggta tgcagcgcgt gaacatgatc atggcagaat
 61 caccaggcct catcaccatc tgcccttttag gatatctact cagtgctgaa tgtacagtt
 121 ttcttgatca tggaaaacgc aacaaaattc tgaatcgcc aaagaggat aattcaggta
 181 aatttggaaa gttttgttcaa gggAACCTTG agagagaatg tatggaaagaa aagtgttagt
 241 ttgaaagaagc acggaaatgtt tttggaaaatc ctggaaaagac aactgaattt tggaaagcagt
 301 atgggtatgg agatcagtgt gagtccaaatc catgtttaa tggccggcagt tgcaggatg
 361 acattaattc ctatgaatgt tgggtccctt ttggatttga agggaaagaaac tggtaattag
 421 atgttaacatg taatctttag aatggcaatg ggcggcgtt ttgtaaaaat agtgtctgata
 481 acaagggtt ttgtctctgt ttgcggggat atcgacttgc agaaaaccag aagtccctgt
 541 aaccgcgcgt gcattttcca tggtaaagag ttctgtttc acaaaacttc aagtcaccc
 601 gtgtctgagac tttttttctt gatgtggact atgttaattt tactgaagct gaaaccattt
 661 tggataacatc actcaaaacatcat ttaatgactt cactccgggtt ttgggtggag
 721 aagatgcacca accaggcttta tttcccttggc aggttgggtttaa gatgtttaaa gttgtatc
 781 tctgttgagg ctctatcggtt aatggaaaat ggttggtaac tggtgttttccatc ttttttt
 841 ctgtgtttaa aatttacagtt gtgcgggtt aacataatat tgaggagaca aacatacac
 901 agcaaaaagc aaatgtgtt cgttatttttcc tccacaccaat ctacaatgc gctttaata
 961 agtacaacca tggatgttcc ctcttggacat tggacccaaatc cttagtgcata aacagctac
 1021 ttacacctat ttgatgttgc gacaaggaaat acacgaaatcat ttcccttcaaa ttttgcatt
 1081 gctatgttaa tggctgggaa aggttccatc acaaaaggag atcagttta gtttttgcatt
 1141 accttagatgtt tcacttgcgtt gacccggccca catgttttcg atcttacaaag ttccatcat
 1201 atacaacatgtt gtttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1261 gtttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1321 ggggttggaaat gtttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1381 tcaactggat taaggaaaaaa acaaaatgttca cttatggaaat gatggatttc caaggtaat
 1441 tcatttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1501 agatttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1561 atttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1621 aatttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1681 ctgtccatcat gatactatgg ttctccacta tggcaactaa ctcaacttcaat ttcccttcat
 1741 tagcaggattt ccattttccc gatctttttt gtttttttttcaatc ctttttttttcaatc
 1801 agtttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1861 tggatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1921 ctttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 1981 ctcttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 2041 catcatttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 2101 ctttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt
 2161 ggg
 2221 taatataacaa tataaatata tagtgtgtgtt gatgttttttgcgtt gatgttttttgcgtt
 2281 acacatataaa tggaaagcaat aagccatttctt aagagcttgcgtt gatgttttttgcgtt
 2341 aggcatgattt tcacaaagggc aagatttggca ttttttttttgcgtt gatgttttttgcgtt
 2401 cccacacataa ttgttttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt gatgttttttgcgtt

-continued

```

2461 ccgttcgttt gcaatctaca gctagtagag actttgagga agaattcaac agtgtgtctt
2521 caggcgtt cagagccaaq caagaqgtt aagtgccta gaccagggcataaagtatc
2581 atgtctccct taactagcat acccccgaat ggagaagggt gcagcggct caaaggcata
2641 agtcattcca atcagccaaq taagttgtcc tttctgggt tcgtgttac catggacat
2701 ttgattata gtaatcctt ctatcttcaa tcttcgtttag agttgtgtac caactgacgt
2761 atgtttccct ttgtgaatta ataaactggt gtctgggtt at

```

Human Factor IX polypeptide (SEQ ID No: 701)
 YNSGKLEEFVQGNLERECMEEKSFEAREVFENTERTEFWKQYVDGDQCESNPCLNGGSCKDDINSYE
 CWCPFGFEGKNCEDVTCNIKGRCCEQFCCKNSADNKVVCSCTEGYRLAENQKSCPEAVPPPCGRVSVSQT
 SKLTRAETVFPDVYVNTEAETILDNITQSTQSFDTRVGGEDAOKPGQFPWQVVLNGKVDACCGGS
 VNEKWIVTAAHCVETGVKITVVAEGHNIETEHTEOKRNRVIRIIPHJHNYNAAINKYNHIALLEDEPLV
 LNSYVTPICIADKEYTNIFLKFGSGYVSGWGRVFHKGRSLALVQYLRVPLVDRATCLRSTKFTIYNNMFC
 AGFHEGGRDSCQDGGPHVTEVEGTSFLTGIISWGEECAMKGKYGIYTAKVSRVNWIEKTKLT

EXAMPLES

[0259] The following examples are provided to illustrate certain disclosed embodiments and are not to be construed as limiting the scope of this disclosure in any way.

Example 1—Materials and Methods

Cloning and Plasmid Preparation

[0260] A bidirectional insertion construct flanked by ITRs was synthesized and cloned into pUC57-Kan by a commercial vendor. The resulting construct (P00147) was used as the parental cloning vector for other vectors. The other insertion constructs (without ITRs) were also commercially synthesized and cloned into pUC57. Purified plasmid was digested with BglII restriction enzyme (New England Bio-Labs, cat #R0144S), and the insertion constructs were cloned into the parental vector. Plasmid was propagated in Stbl3™ Chemically Competent *E. coli* (Thermo Fisher, Cat #C737303).

AAV Production

[0261] Triple transfection in HEK293 cells was used to package genomes with constructs of interest for AAV8 and AAV-DJ production and resulting vectors were purified from both lysed cells and culture media through iodixanol gradient ultracentrifugation method (See, e.g., Lock et al., Hum Gene Ther. 2010 October; 21(10):1259-71). The plasmids used in the triple transfection that contained the genome with constructs of interest are referenced in the Examples by a “PXXXX” number, see also e.g., Table 9. Isolated AAV was dialyzed in storage buffer (PBS with 0.001% Pluronic F68). AAV titer was determined by qPCR using primers/probe located within the ITR region.

In Vitro Transcription (“IVT”) of Nuclease mRNA

[0262] Capped and polyadenylated *Streptococcus pyogenes* (“Spy”) Cas9 mRNA containing N1-methyl pseudo-U was generated by in vitro transcription using a linearized plasmid DNA template and T7 RNA polymerase. Generally, plasmid DNA containing a T7 promoter and a 100 nt poly (A/T) region was linearized by incubating at 37° C. with XbaI to complete digestion followed by heat inactivation of XbaI at 65° C. The linearized plasmid was purified from enzyme and buffer salts. The IVT reaction to generate Cas9 modified mRNA was incubated at 37° C. for 4 hours in the following conditions: 50 ng/µL linearized plasmid; 2 mM each of GTP, ATP, CTP, and N1-methyl pseudo-UTP (Trilink); 10 mM ARCA (Trilink); 5 U/µL T7 RNA polymerase (NEB); 1 U/µL Murine Rnase inhibitor (NEB); 0.004 U/µL Inorganic *E. coli* pyrophosphatase (NEB); and 1× reaction

buffer. TURBO Dnase (ThermoFisher) was added to a final concentration of 0.01 U/µL, and the reaction was incubated for an additional 30 minutes to remove the DNA template. The Cas9 mRNA was purified using a MegaClear Transcription Clean-up kit according to the manufacturer’s protocol (ThermoFisher). Alternatively, the Cas9 mRNA was purified using LiCl precipitation, ammonium acetate precipitation, and sodium acetate precipitation or using a LiCl precipitation method followed by further purification by tangential flow filtration. The transcript concentration was determined by measuring the light absorbance at 260 nm (Nanodrop), and the transcript was analyzed by capillary electrophoresis by Bioanalyzer (Agilent).

[0263] Cas9 mRNAs below comprise Cas9 ORF SEQ ID NO: 703 or SEQ ID NO: 704 or a sequence of Table 24 of PCT/US2019/053423 (which is hereby incorporated by reference).

Lipid Formulations for Delivery of Cas9 mRNA and gRNA

[0264] Cas9 mRNA and gRNA were delivered to cells and animals utilizing lipid formulations comprising ionizable lipid ((9Z,12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl octadeca-9,12-dienoate, also called 3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl (9Z,12Z)-octadeca-9,12-dienoate), cholesterol, DSPC, and PEG2k-DMG.

[0265] For experiments utilizing pre-mixed lipid formulations (referred to herein as “lipid packets”), the components were reconstituted in 100% ethanol at a molar ratio of ionizable lipid:cholesterol:DSPC:PEG2k-DMG of 50:38:9:3, prior to being mixed with RNA cargos (e.g., Cas9 mRNA and gRNA) at a lipid amine to RNA phosphate (N:P) molar ratio of about 6.0, as further described herein.

[0266] For experiments utilizing the components formulated as lipid nanoparticles (LNPs), the components were dissolved in 100% ethanol at various molar ratios. The RNA cargos (e.g., Cas9 mRNA and gRNA) were dissolved in 25 mM citrate, 100 mM NaCl, pH 5.0, resulting in a concentration of RNA cargo of approximately 0.45 mg/mL.

[0267] For the experiments described in Example 2, the LNPs were formed by microfluidic mixing of the lipid and RNA solutions using a Precision Nanosystems NanoAssembler™ Benchtop Instrument, according to the manufacturer’s protocol. A 2:1 ratio of aqueous to organic solvent was maintained during mixing using differential flow rates. After mixing, the LNPs were collected, diluted in water (approximately 1:1 v/v), held for 1 hour at room temperature, and further diluted with water (approximately 1:1 v/v) before final buffer exchange. The final buffer exchange into 50 mM Tris, 45 mM NaCl, 5% (w/v) sucrose, pH 7.5 (TSS) was

completed with PD-10 desalting columns (GE). If required, formulations were concentrated by centrifugation with Amicon 100 kDa centrifugal filters (Millipore). The resulting mixture was then filtered using a 0.2 m sterile filter. The final LNP was stored at -80° C. until further use. The LNPs were formulated at a molar ratio of ionizable lipid:cholesterol: DSPC:PEG2k-DMG of 45:44:9:2, with a lipid amine to RNA phosphate (N:P) molar ratio of about 4.5, and a ratio of gRNA to mRNA of 1:1 by weight.

[0268] For the experiments described in other examples, the LNPs were prepared using a cross-flow technique utilizing impinging jet mixing of the lipid in ethanol with two volumes of RNA solutions and one volume of water. The lipid in ethanol was mixed through a mixing cross with the two volumes of RNA solution. A fourth stream of water was mixed with the outlet stream of the cross through an inline tee (See WO2016010840 FIG. 2.). The LNPs were held for 1 hour at room temperature, and further diluted with water (approximately 1:1 v/v). Diluted LNPs were concentrated using tangential flow filtration on a flat sheet cartridge (Sartorius, 100kD MWCO) and then buffer exchanged by diafiltration into 50 mM Tris, 45 mM NaCl, 5% (w/v) sucrose, pH 7.5 (TSS). Alternatively, the final buffer exchange into TSS was completed with PD-10 desalting columns (GE). If required, formulations were concentrated by centrifugation with Amicon 100 kDa centrifugal filters (Millipore). The resulting mixture was then filtered using a 0.2 m sterile filter. The final LNP was stored at 4° C. or -80° C. until further use. The LNPs were formulated at a molar ratio of ionizable lipid:cholesterol:DSPC:PEG2k-DMG of 50:38:9:3, with a lipid amine to RNA phosphate (N:P) molar ratio of about 6.0, and a ratio of gRNA to mRNA of 1:1 by weight.

Cell Culture and In Vitro Delivery of Cas9 mRNA, gRNA, and Insertion Constructs

Hepa1-6 cells

[0269] Hepa 1-6 cells were plated at density of 10,000 cells/well in 96-well plates. 24 hours later, cells were treated with LNP and AAV. Before treatment the media was aspirated off from the wells. LNP was diluted to 4 ng/ul in DMEM+10% FBS media and further diluted to 2 ng/ul in 10% FBS (in DMEM) and incubated at 37° C. for 10 min (at a final concentration of 5% FBS). Target MOI of AAV was 1e6, diluted in DMEM+10% FBS media. 50 µl of the above diluted LNP at 2 ng/ul was added to the cells (delivering a total of 100 ng of RNA cargo) followed by 50 µl of AAV. The treatment of LNP and AAV were minutes apart. Total volume of media in cells was 100 µl. After 72 hours post-treatment and 30 days post-treatment, supernatant from these treated cells were collected for human FIX ELISA analysis as described below.

Primary Hepatocytes

[0270] Primary mouse hepatocytes (PMH), primary cyno hepatocytes (PCH) and primary human hepatocytes (PHH) were thawed and resuspended in hepatocyte thawing medium with supplements (ThermoFisher) followed by centrifugation. The supernatant was discarded, and the pelleted cells resuspended in hepatocyte plating medium plus supplement pack (ThermoFisher). Cells were counted and plated on Bio-coat collagen I coated 96-well plates at a density of 33,000 cells/well for PHH and 50,000 cells/well for PCH and 15,000 cells/well for PMH. Plated cells were allowed to settle and adhere for 5 hours in a tissue culture incubator at

37° C. and 5% CO₂ atmosphere. After incubation cells were checked for monolayer formation and were washed thrice with hepatocyte maintenance prior and incubated at 37° C.

[0271] For experiments utilizing lipid packet delivery, Cas9 mRNA and gRNA were each separately diluted to 2 mg/ml in maintenance media and 2.9 µl of each were added to wells (in a 96-well Eppendorf plate) containing 12.5 µl of 50 mM sodium citrate, 200 mM sodium chloride at pH 5 and 6.9 µl of water. 12.5 µl of lipid packet formulation was then added, followed by 12.5 µl of water and 150 µl of TSS. Each well was diluted to 20 ng/µl (with respect to total RNA content) using hepatocyte maintenance media, and then diluted to 10 ng/µl (with respect to total RNA content) with 6% fresh mouse serum. Media was aspirated from the cells prior to transfection and 40 µl of the lipid packet/RNA mixtures were added to the cells, followed by addition of AAV (diluted in maintenance media) at an MOI of 1e5. Media was collected 72 hours post-treatment for analysis and cells were harvested for further analysis, as described herein.

Luciferase Assays

[0272] For experiments involving NanoLuc detection in cell media, one volume of Nano-Glo® Luciferase Assay Substrate was combined with 50 volumes of Nano-Glo® Luciferase Assay Buffer. The assay was run on a Promega Glomax runner at an integration time of 0.5 sec using 1:10 dilution of samples (50 µl of reagent+40 µl water+10 µl cell media).

[0273] For experiments involving detection of the HiBit tag in cell media, LgBiT Protein and Nano-GloR HiBiT Extracellular Substrate were diluted 1:100 and 1:50, respectively, in room temperature Nano-GloR HiBiT Extracellular Buffer. The assay was run on a Promega Glomax runner at an integration time of 1.0 sec using 1:10 dilution of samples (50 µl of reagent+40 µl water+10 µl cell media).

In Vivo Delivery of LNP and/or AAV

[0274] Mice were dosed with AAV, LNP, both AAV and LNP, or vehicle (PBS+0.001% Pluronic for AAV vehicle, TSS for LNP vehicle) via the lateral tail vein. AAV were administered in a volume of 0.1 mL per animal with amounts (vector genomes/mouse, "vg/ms") as described herein. LNPs were diluted in TSS and administered at amounts as indicated herein, at about 5 µl/gram body weight. Typically, mice were injected first with AAV and then with LNP, if applicable. At various times points post-treatment, serum and/or liver tissue was collected for certain analyses as described further below.

Human Factor IX (hFIX) ELISA Analysis

[0275] For in vitro studies, total human Factor IX levels secreted in cell media were determined using a Human Factor IX ELISA Kit (Abcam, Cat #ab188393) according to manufacturer's protocol. Secreted hFIX levels were quantitated off a standard curve using 4 parameter logistic fit and expressed as ng/ml of media.

[0276] For in vivo studies, blood was collected and the serum or plasma was isolated as indicated. The total human Factor IX levels were determined using a Human Factor IX ELISA Kit (Abcam, Cat #ab188393) according to manufacturer's protocol. Serum or plasma hFIX levels were quantitated off a standard curve using 4 parameter logistic fit and expressed as µg/mL of serum.

Next-Generation Sequencing (“NGS”) and Analysis for On-Target Cleavage Efficiency

[0277] Deep sequencing was utilized to identify the presence of insertions and deletions introduced by gene editing, e.g., within intron 1 of albumin. PCR primers were designed around the target site and the genomic area of interest was amplified. Primer sequence design was done as is standard in the field.

[0278] Additional PCR was performed according to the manufacturer’s protocols (Illumina) to add chemistry for sequencing. The amplicons were sequenced on an Illumina MiSeq instrument. The reads were aligned to the reference genome after eliminating those having low quality scores. The resulting files containing the reads were mapped to the reference genome (BAM files), where reads that overlapped the target region of interest were selected and the number of wild type reads versus the number of reads which contain an insertion or deletion (“indel”) was calculated.

[0279] The editing percentage (e.g., the “editing efficiency” or “percent editing”) is defined as the total number of sequence reads with insertions or deletions (“indels”) over the total number of sequence reads, including wild type.

In Situ Hybridization Analysis

[0280] BaseScope (ACDbio, Newark, CA) is a specialized RNA in situ hybridization technology that can provide specific detection of exon junctions, e.g., in a hybrid mRNA transcript that contains an insertion transgene (hFIX) and coding sequence from the site of insertion (e.g. exon 1 of albumin). BaseScope was used to measure the percentage of liver cells expressing the hybrid mRNA.

[0281] To detect the hybrid mRNA, two probes against the hybrid mRNAs that may arise following insertion of a bidirectional construct were designed by ACDbio (Newark, CA). One of the probes was designed to detect a hybrid mRNA resulting from insertion of the construct in one orientation, while the other probe was designed to detect a hybrid mRNA resulting from insertion of the construct in the other orientation. Livers from different groups of mice were collected and fresh-frozen sectioned. The BaseScope assay, using a single probe or pooled probes was performed according to the manufacturer’s protocol. Slides were scanned and analyzed by the HALO software. The background (saline treated group) of this assay was 0.58%.

Example 2—In Vitro Testing of Insertion Templates with and without Homology Arms

[0282] In this Example, Hepa1-6 cells were cultured and treated with AAV harboring insertion templates of various forms (e.g., having either a single-stranded genome (“ssAAV”) or a self-complementary genome (“scAAV”)), in the presence or absence of LNP delivering Cas9 mRNA and G000551 e.g., as described in Example 1 (n=3). The AAV and LNP were prepared as described in Example 1. Following treatment, the media was collected for human Factor IX levels as described in Example 1.

[0283] Hepa1-6 cells are an immortalized mouse liver cell line that continues to divide in culture. As shown in FIG. 2 (72 hour post-treatment time point), only the vector (scAAV derived from plasmid P00204) comprising 200 bp homology arms resulted in detectable expression of hFIX. Use of the AAV vectors derived from P00123 (scAAV lacking homol-

ogy arms) and P00147 (ssAAV bidirectional construct lacking homology arms) did not result in any detectable expression of hFIX in this experiment. The cells were kept in culture and these results were confirmed when re-assayed at 30 days post-treatment (data not shown).

Example 3—In Vivo Testing of Insertion Templates with and without Homology Arms

[0284] In this Example, mice were treated with AAV derived from the same plasmids (P00123, P00204, and P00147) as tested in vitro in Example 2. The dosing materials were prepared and dosed as described in Example 1. C57Bl/6 mice were dosed (n=5 for each group) with 3e11 vector genomes each (vg/ms) followed by LNP comprising G000551 (“G551”) at a dose of 4 mg/kg (with respect to total RNA cargo content). Four weeks post dose, the animals were euthanized and liver tissue and sera were collected for editing and hFIX expression, respectively.

[0285] As shown in FIG. 3A and Table 12, liver editing levels of ~60% were detected in each group of animals treated with LNP comprising gRNA targeting intron 1 of murine albumin. However, despite robust and consistent levels of editing in each treatment group, animals receiving the bi-directional vector without homology arms (ssAAV vector derived from P00147) in combination with LNP treatment resulted in the highest level of hFIX expression in serum (FIG. 3B and Table 13).

TABLE 12

Template	% Indel	
	Average Indel (%)	St. Dev Indel (%)
scAAV Blunt (P00123)	66.72	4.09
ssAAV Blunt (P00147)	68.10	2.27
ssAAV HR (P00204)	70.16	3.68
LNP only	68.24	6.47
Vehicle	0.28	0.08

TABLE 13

Template	Factor IX Levels	
	Average Factor IX (ug/mL)	St.Dev Factor IX (ug/mL)
scAAV Blunt (P00123)	0.75	0.28
ssAAV Blunt (P00147)	2.92	1.04
ssAAV HR (P00204)	0.96	0.35
LNP only	0	0
Vehicle	0	0

Example 4—In Vivo Testing of ssAAV Insertion Templates with and without Homology Arms

[0286] The experiment described in this example examined the effect of incorporating homology arms into ssAAV vectors *in vivo*.

[0287] The dosing materials used in this experiment were prepared and dosed as described in Example 1. C57Bl/6 mice were dosed (n=5 for each group) with 3e11 vg/ms followed by LNP comprising G000666 (“G666”) or G000551 (“G551”) at a dose of 0.5 mg/kg (with respect to total RNA cargo content). Four weeks post dose, the animals sera was collected for hFIX expression.

[0288] As shown in FIG. 4A and Table 14, use of the ssAAV vectors with asymmetrical homology arms (300/600 bp arms, 300/2000 bp arms, and 300/1500 bp arms for vectors derived from plasmids P00350, P00356, and P00362, respectively) for insertion into the site targeted by G551 resulted in levels of circulating hFIX that were below the lower limit of detection for the assay. However, use of the ssAAV vector (derived from P00147) without homology arms and having two hFIX open reading frames (ORF) in a bidirectional orientation resulted in detectable levels of circulating hFIX in each animal.

[0289] Similarly, use of the ssAAV vectors with asymmetrical homology arms (500 bp arms and 800 bp arms for vectors derived from plasmids P00353 and P00354, respectively) for insertion into the site targeted by G666 resulted in lower but detectable levels, as compared to use of the bidirectional vector without homology arms (derived from P00147) (see FIG. 4B and Table 15).

TABLE 14

Serum hFIX Levels		
AAV	Average Serum FIX (ug/mL)	St.Dev Serum FIX (ug/mL)
P00147	5.13	1.31
P00350	-0.22	0.08
P00356	-0.23	0.04
P00362	-0.09	0.16

TABLE 15

Serum hFIX Levels		
AAV	Average Serum FIX (ug/mL)	St.Dev Serum FIX (ug/mL)
P00147	7.72	4.67
P00353	0.20	0.23
P00354	0.46	0.26

Example 5—In Vitro Screening of Bidirectional Constructs Across Target Sites in Primary Mouse Hepatocytes

[0290] Having demonstrated that bidirectional constructs lacking homology arms outperformed vectors with other configurations, the experiment described in this Example examined the effects of altering the splice acceptors used to form the hybrid transcript between hFIX and exon 1 of albumin and altering the gRNAs for targeting CRISPR/Cas9-mediated insertion. These varied bidirectional constructs were tested across a panel of target sites utilizing 20 different gRNAs targeting intron 1 of murine albumin in primary mouse hepatocytes (PMH).

[0291] The ssAAV and lipid packet delivery materials tested in this Example were prepared and delivered to PMH as described in Example 1, with the AAV at an MOI of 1e5. Following treatment, isolated genomic DNA and cell media was collected for editing and transgene expression analysis, respectively. Each of the vectors comprised a reporter that can be measured through luciferase-based fluorescence detection as described in Example 1, plotted in FIG. 5C as relative luciferase units (“RLU”). The vectors comprised a HiBit peptide fused at the 3' ends of the hFIX ORF, which allows for sensitive detection of relative expression. Schematics of each vector tested are provided in FIG. 5A. The gRNAs tested are shown in FIGS. 5B and 5C, using a shortened number for those listed in Table 5 (e.g., where the leading zeros are omitted, for example where “G551” corresponds to “G000551” in Table 5).

[0292] As shown in FIG. 5B and Table 16, consistent but varied levels of editing were detected for each of the treatment groups across each combination tested. Transgene expression using various combinations of template and guide RNA is shown in FIG. 5C and Table 17. As shown in FIG. 5D, a significant level of indel formation did not necessarily result in more efficient expression of the transgenes. Using P00411- and P00418-derived templates, the R² values were 0.54 and 0.37, respectively, when guides with less than 10% editing are not included. The mouse albumin splice acceptor and human FIX splice acceptor each resulted in effective transgene expression.

TABLE 16

Guide ID	% Indel					
	P00411		P00418		P00415	
	Average Indel (%)	St. Dev Indel (%)	Average Indel (%)	St. Dev Indel (%)	Average Indel (%)	St. Dev Indel (%)
G000551	67.4	1.42	70.67	2.29	66.73	4.90
G000552	90.93	0.15	91.10	2.43	90.37	1.01
G000553	77.80	3.83	77.47	1.87	80.50	0.85
G000554	72.37	6.49	70.53	3.16	70.60	2.91
G000555	35.37	2.63	35.77	9.34	40.47	4.75
G000666	62.47	3.87	50.90	19.41	65.90	3.99
G000667	30.57	2.73	25.30	3.67	31.67	2.29
G000668	63.60	2.02	66.65	4.60	68.30	4.90
G000669	19.10	2.51	19.33	1.53	18.70	1.25
G000670	47.80	3.27	49.10	4.42	51.97	2.06
G011722	4.20	0.72	4.27	1.20	4.20	0.26
G011723	5.63	1.27	6.07	0.15	5.93	0.15
G011724	6.10	1.28	8.50	2.69	7.13	1.27
G011725	1.93	0.29	2.60	0.79	2.53	0.65
G011726	10.73	1.46	11.70	0.50	12.43	1.33
G011727	14.20	1.56	14.80	2.36	16.20	2.69
G011728	10.55	1.20	13.65	0.92	15.50	1.56
G011729	5.00	0.10	5.63	0.25	6.00	1.01

TABLE 16-continued

Guide ID	% Indel					
	P00411		P00418		P00415	
	Average Indel (%)	St. Dev Indel (%)	Average Indel (%)	St. Dev Indel (%)	Average Indel (%)	St. Dev Indel (%)
G011730	7.83	0.97	9.13	0.59	7.33	0.59
G011731	23.70	0.66	25.27	1.21	24.87	1.01
AAV Only	0.15	0.07	0.05	0.07	0.10	0.00

TABLE 17

Guide ID	Luciferase Levels					
	P00411		P00418		P00415	
	Average Luciferase (RLU)	St. Dev Luciferase (RLU)	Average Luciferase (RLU)	St. Dev Luciferase (RLU)	Average Luciferase (RLU)	St. Dev Luciferase (RLU)
G000551	58000.00	4331.28	41800.00	2165.64	78633.33	20274.70
G000552	95700.00	10573.08	80866.67	27911.35	205333.33	30664.86
G000553	205333.33	52993.71	177333.33	32929.22	471666.67	134001.00
G000554	125333.33	55949.38	91933.33	19194.10	232666.67	67002.49
G000555	59933.33	11566.04	77733.33	11061.80	155666.67	15947.83
G000666	88500.00	28735.87	93266.67	30861.19	313000.00	15394.80
G000667	75333.33	22653.11	68966.67	27222.11	153000.00	30805.84
G000668	164000.00	56320.51	133400.00	65111.29	429000.00	120751.80
G000669	28933.33	11636.29	22033.33	2413.16	46466.67	6543.19
G000670	162666.67	32959.57	200000.00	33867.39	424666.67	36473.73
G011722	16766.67	3384.28	8583.33	4103.10	24000.00	8915.16
G011723	22733.33	7252.82	17133.33	4905.44	26100.00	8109.87
G011724	17300.00	2400.00	28033.33	9091.94	30933.33	3365.02
G011725	8253.33	1163.20	8890.00	1429.27	20366.67	13955.05
G011726	12223.33	3742.54	11610.00	2490.44	14950.00	8176.03
G011727	35600.00	8128.35	36300.00	12301.22	86700.00	5023.94
G011728	14900.00	5011.99	22466.67	7130.45	38166.67	13829.08
G011729	10460.00	2543.95	11223.33	2220.28	26966.67	16085.50
G011730	14833.33	2307.24	21700.00	8681.59	41233.33	25687.03
G011731	16433.33	3274.65	22566.67	2205.30	20756.67	13096.20
AAV Only	217.00	15.56	215.00	15.56	207.00	1.41

Example 6—In Vivo Screening of Bidirectional Constructs Across Target Sites

[0293] The ssAAV and LNPs tested in this Example were prepared and delivered to C57Bl/6 mice as described in Example 1 to assess the performance of the bidirectional constructs across target sites in vivo. Four weeks post dose, the animals were euthanized and liver tissue and sera were collected for editing and hFIX expression, respectively.

[0294] In an initial experiment, 10 different LNP formulations containing 10 different gRNA targeting intron 1 of albumin were delivered to mice along with ssAAV derived from P00147. The AAV and LNP were delivered at 3e11 vg/ms and 4 mg/kg (with respect to total RNA cargo content), respectively (n=5 for each group). The gRNAs tested in this experiment are shown in FIG. 6 and tabulated in Table 18. As shown in FIG. 6 and as observed in vitro, a significant level of indel formation was not predictive for insertion or expression of the transgenes.

[0295] In a separate experiment, a panel of 20 gRNAs targeting the 20 different target sites tested in vitro in Example 5 were tested in vivo. To this end, LNP formulations containing the 20 gRNAs targeting intron 1 of albumin were delivered to mice along with ssAAV derived from P00147. The AAV and LNP were delivered at 3e11 vg/ms

and 1 mg/kg (with respect to total RNA cargo content), respectively. The gRNAs tested in this experiment are shown in FIGS. 7A and 7B.

[0296] As shown, in FIG. 7A and tabulated in Table 19, varied levels of editing were detected for each of the treatment groups across each LNP/vector combination tested. However, as shown in FIG. 7B and Table 20 and consistent with the in vitro data described in Example 5, higher levels of editing did not necessarily result in higher levels of expression of the transgenes in vivo, indicating a lack of correlation between editing and insertion/expression of the bidirectional hFIX constructs. Indeed, very little correlation exists between the amount of editing achieved and the amount of hFIX expression as viewed in the plot provided in FIG. 7D. In particular, an R² value of only 0.34 is calculated between the editing and expression data sets for this experiment, when those gRNAs achieving less than 10% editing are removed from the analysis. Interestingly, as shown in FIG. 7C, a correlation plot is provided comparing the levels of expression as measured in RLU from the in vitro experiment of Example 5 to the transgene expression levels in vivo detected in this experiment, with an R² value of 0.70, demonstrating a positive correlation between the primary cell screening and the in vivo treatments.

[0297] To assess insertion of the bidirectional construct at the cellular level, liver tissues from treated animals were assayed using an *in situ* hybridization method (BaseScope), e.g., as described in Example 1. This assay utilized probes that can detect the junctions between the hFIX transgene and the mouse albumin exon 1 sequence, as a hybrid transcript. As shown in FIG. 8A, cells positive for the hybrid transcript were detected in animals that received both AAV and LNP. Specifically, when AAV alone is administered, less than 1.0% of cells were positive for the hybrid transcript. With administration of LNPs comprising G011723, G000551, or G000666, 4.9%, 19.8%, or 52.3% of cells were positive for the hybrid transcript. Additionally, as shown in FIG. 8B, circulating hFIX levels correlated with the number of cells that were positive for the hybrid transcript. Lastly, the assay utilized pooled probes that can detect insertion of the bidirectional hFIX construct in either orientation. However, when a single probe was used that only detects a single orientation, the amount of cells that were positive for the hybrid transcript was about half that detected using the pooled probes (in one example, 4.46% vs 9.68%), suggesting that the bidirectional construct indeed is capable of inserting in either orientation giving rise to expressed hybrid transcripts that correlate with the amount of transgene expression at the protein level. These data show that the circulating hFIX levels achieved are dependent on the guide used for insertion.

TABLE 18

hFIX Serum Levels and % Indel				
Guide	Average Indel (%)	St. Dev Indel (%)	Average hFIX Serum Levels	St. Dev hFIX Serum Levels
G000551	75.02	1.27	3.82	3.38
G000555	51.18	1.19	32.56	9.05
G000553	62.78	2.64	25.07	4.04
G000667	52.96	4.96	32.03	6.74
G000554	55.24	2.28	29.48	7.34
G000552	67.56	1.73	14.79	5.34
G000668	43.14	5.78	26.72	7.97
G000669	50.68	2.97	10.70	4.43
G000666	64.62	1.34	26.19	5.56
G000670	55.90	1.30	30.96	8.44

TABLE 19

% Liver Editing		
Guide	Average Liver Editing (%)	St. Dev Liver Editing (%)
G000551	59.48	4.02
G000555	58.72	3.65
G000553	51.26	2.81
G000554	33.04	8.76
G000555	12.72	4.46
G000666	53.60	4.92
G000667	26.74	4.98
G000668	39.22	3.04
G000669	33.34	4.77
G000670	47.50	5.58
G011722	10.34	1.68
G011723	4.02	0.84
G011724	2.46	0.64
G011725	8.26	1.24
G011726	6.90	1.01
G011727	13.33	6.43
G011728	35.78	9.34

TABLE 19-continued

Guide	% Liver Editing	
	Average Liver Editing (%)	St. Dev Liver Editing (%)
G011729	4.62	1.46
G011730	12.68	3.14
G011731	26.70	1.86

TABLE 20

Guide	Serum hFIX Levels					
	Week 1		Week 2		Week 4	
	Average FIX (ug/mL)	St. Dev FIX (ug/mL)	Average FIX (ug/mL)	St. Dev FIX (ug/mL)	Average FIX (ug/mL)	St. Dev FIX (ug/mL)
G000551	10.88	2.74	10.25	2.51	9.39	3.48
G000555	13.34	2.09	12.00	2.75	12.43	2.57
G000553	17.64	4.34	20.27	6.35	15.31	2.43
G000554	12.79	4.99	14.29	6.09	12.74	4.93
G000555	11.94	5.79	11.99	5.76	8.61	4.02
G000666	21.63	1.32	20.65	1.55	17.23	0.62
G000667	16.77	2.86	12.35	2.85	12.57	5.60
G000668	21.35	1.51	18.20	3.18	17.72	2.25
G000669	5.76	2.10	6.72	2.93	3.39	0.78
G000670	18.18	2.17	19.16	3.05	15.49	3.61
G011722	8.07	1.74	7.74	2.41	8.07	1.74
G011723	2.11	0.28	1.65	0.28	2.11	0.28
G011724	0.92	0.43	0.60	0.30	0.92	0.43
G011725	1.75	0.77	1.14	0.67	1.75	0.77
G011726	0.59	0.30	1.01	0.64	0.59	0.30
G011727	6.71	2.80	6.90	3.68	6.71	2.80
G011728	11.77	3.12	12.29	3.43	11.77	3.12
G011729	0.94	0.35	0.89	0.29	0.94	0.35
G011730	5.93	1.77	6.33	1.73	5.93	1.77
G011731	3.56	0.87	3.78	0.50	3.56	0.87
AAV Only	0.00	0.00	0.00	0.00	0.00	0.00
Vehicle	0.00	0.00	0.00	0.00	0.00	0.00
Human Serum	3.63	0.32	3.61	0.35	3.28	0.03

Example 7—Timing of AAV and LNP Delivery In Vivo

[0298] In this Example, the timing between delivery of ssAAV comprising the bidirectional hFIX construct and LNP was examined in C57Bl/6 mice.

[0299] The ssAAV and LNPs tested in this Example were prepared and delivered to mice as described in Example 1. The LNP formulation contained G000551 and the bidirectional template was delivered as ssAAV derived from P00147. The AAV and LNP were delivered at 3e11 vg/ms and 4 mg/kg (with respect to total RNA cargo content), respectively (n=5 for each group). A “Template only” cohort received AAV only, and a “PBS” cohort received no AAV or LNP. One cohort received AAV and LNP sequentially (minutes apart) at day 0 (“Template+LNP day 0”); another cohort received AAV at day 0 and LNP at day 1 (“Template+LNP day 1”); and a final cohort received AAV at day 0 and LNP at day 7 (“Template+LNP day 7”). At 1 week, 2 weeks and 6 weeks, plasma was collected for hFIX expression analysis.

[0300] As shown in FIG. 9, hFIX was detected in each cohort at each time assayed, except for the 1 week timepoint for the cohort that received the LNP at day 7 post AAV delivery.

Example 8—Multiple Dosing of LNP Following Delivery of AAV

[0301] In this Example, the effects of repeat dosing of LNP following administration of ssAAV was examined.

[0302] The ssAAV and LNPs tested in this Example were prepared and delivered to C57Bl/6 mice as described in Example 1. The LNP formulation contained G000551 and the ssAAV was derived from P00147. The AAV and LNP were delivered at 3e11 vg/ms and 0.5 mg/kg (with respect to total RNA cargo content), respectively (n=5 for each group). A “Template only” cohort received AAV only, and a “PBS” cohort received no AAV or LNP. One cohort received AAV and LNP sequentially (minutes apart) at day 0 with no further treatments (“Template+LNP(1x)” in FIG. 10); another cohort received AAV and LNP sequentially (minutes apart) at day 0 and a second dose at day 7 (“Template+LNP(2x)” in FIG. 10); and a final cohort received AAV and LNP sequentially (minutes apart) at day 0, a second dose of LNP at day 7 and a third dose of LNP at day 14 (“Template+LNP(3x)” in FIG. 10). At 1, 2, 4 and 6 weeks post-administration of AAV, plasma was collected for hFIX expression analysis.

[0303] As shown in FIG. 10, hFIX was detected in each cohort at each time assayed, and multiple subsequent doses of LNP did not significantly increase the amount of hFIX expression.

Example 9—Durability of hFIX Expression In Vivo

[0304] The durability of hFIX expression over time in treated animals was assessed in this Example. To this end, hFIX was measured in the serum of treated animals post-dose, as part of a one-year durability study.

[0305] The ssAAV and LNPs tested in this Example were prepared and delivered to C57Bl/6 mice as described in Example 1. The LNP formulation contained G000551 and the ssAAV was derived from P00147. The AAV was delivered at 3e11 vg/ms and the LNP was delivered at either 0.25 or 1.0 mg/kg (with respect to total RNA cargo content) (n=5 for each group).

[0306] As shown in FIG. 11A and Table 21, hFIX expression was sustained at each time point assessed for both groups out to 41 weeks. A drop in the levels observed at 8 weeks is believed to be due to the variability of the ELISA assay. Serum albumin levels were measured by ELISA at week 2 and week 41, showing that circulating albumin levels are maintained across the study.

[0307] As shown in FIG. 11B and Table 22, hFIX expression was sustained at each time point assessed for both groups out to 52 weeks.

TABLE 21

Week	FIX Levels			
	Dose			
	0.25 mpk LNP		1 mpk LNP	
Week	Average hFIX (ug/mL)	StDev hFIX (ug/mL)	Average hFIX (ug/mL)	StDev hFIX (ug/mL)
2	0.48	0.21	2.24	1.12
4	0.55	0.18	2.82	1.67
8	0.40	0.17	1.72	0.77
12	0.48	0.20	2.85	1.34

TABLE 21-continued

Week	FIX Levels			
	Dose			
	0.25 mpk LNP		1 mpk LNP	
Week	Average hFIX (ug/mL)	StDev hFIX (ug/mL)	Average hFIX (ug/mL)	StDev hFIX (ug/mL)
20	0.48	0.27	2.45	1.26
41	0.79	0.49	4.63	0.95

TABLE 22

Week	FIX Levels			
	Dose			
	0.25 mpk LNP		1 mpk LNP	
Week	Average hFIX (ug/mL)	StDev hFIX (ug/mL)	Average hFIX (ug/mL)	StDev hFIX (ug/mL)
2	0.87	0.15	4.02	1.75
8	0.99	0.15	4.11	1.41
12	0.93	0.14	4.15	1.35
20	0.83	0.22	4.27	1.54
41	0.83	0.37	4.76	1.62
52	0.82	0.25	4.72	1.54

Example 10—Effects of Varied Doses of AAV and LNP to Modulate hFIX Expression In Vivo

[0308] In this Example, the effects of varying the dose of both AAV and LNP to modulate expression of hFIX was assessed in C57Bl/6 mice.

[0309] The ssAAV and LNPs tested in this Example were prepared and delivered to mice as described in Example 1. The LNP formulation contained G000553 and the ssAAV was derived from P00147. The AAV was delivered at 1e11, 3e11, 1e12 or 3e12 vg/ms and the LNP was delivered at 0.1, 0.3, or 1.0 mg/kg (with respect to total RNA cargo content) (n=5 for each group). Two weeks post-dose, the animals were euthanized. Sera were collected at two timepoints for hFIX expression analysis.

[0310] As shown in FIG. 12A (1 week), FIG. 12B (2 weeks) and Table 23, varying the dose of either AAV or LNP can modulate the amount of expression of hFIX in vivo.

TABLE 23

Timepoint	Serum hFIX				
	RNP Dose (mg/kg)	AAV Dose (MOI)	Mean FIX (ng/ml)	SD	N
Week 1	0.1	1E+11	0.08	0.02	2
		3E+11	0.11	0.04	5
		1E+12	0.41	0.15	5
		3E+12	0.61	0.17	5
	0.3	1E+11	0.36	0.14	5
		3E+11	0.67	0.26	5
	1E+12	1E+12	1.76	0.14	5
		3E+12	4.70	2.40	5

TABLE 23-continued

Serum hFIX					
Timepoint	RNP Dose (mg/kg)	AAV Dose (MOI)	Mean FIX (ng/ml)	SD	N
Week 2	1.0	1E+11	3.71	0.31	4
		3E+11	8.00	0.51	5
		1E+12	14.17	1.38	5
		3E+12	20.70	2.79	5
	0.1	Human serum 1:1000	6.62	—	1
		1E+11	0.12	0.01	2
		3E+11	0.26	0.07	5
		1E+12	0.83	0.24	5
	0.3	3E+12	1.48	0.35	5
		1E+11	0.70	0.26	4
		3E+11	1.42	0.37	5
		1E+12	3.53	0.49	5
	1.0	3E+12	8.94	4.39	5
		1E+11	5.40	0.47	4
		3E+11	12.31	2.45	5
		1E+12	17.89	1.95	5
	Human serum 1:1000	3E+12	25.52	3.62	5
			4.47	—	1

Example 11—In Vitro Screening of Bidirectional Constructs Across Target Sites in Primary Cynomolgus and Primary Human Hepatocytes

[0311] In this Example, ssAAV vectors comprising a bidirectional construct were tested across a panel of target sites utilizing gRNAs targeting intron 1 of cynomolgus (“cyno”) and human albumin in primary cyno (PCH) and primary human hepatocytes (PHH), respectively.

[0312] The ssAAV and lipid packet delivery materials tested in this Example were prepared and delivered to PCH and PHH as described in Example 1. Following treatment, isolated genomic DNA and cell media was collected for editing and transgene expression analysis, respectively. Each of the vectors comprised a reporter that can be measured through luciferase-based fluorescence detection as described in Example 1 (derived from plasmid P00415), plotted in FIGS. 13B and 14B as relative luciferase units (“RLU”). For example, the AAV vectors contained the NanoLuc ORF (in addition to GFP). Schematics of the vectors tested are provided in FIGS. 13B and 14B. The gRNAs tested are shown in each of the Figures using a shortened number for those listed in Table 1 and Table 7.

[0313] As shown in FIG. 13A for PCH and FIG. 14A for PHH, varied levels of editing were detected for each of the combinations tested (editing data for some combinations tested in the PCH experiment are not reported in FIG. 13A and Table 3 due to failure of certain primer pairs used for the amplicon based sequencing). The editing data shown in FIGS. 13A and 14A graphically, are reproduced numerically in Table 3 and Table 4 below. However, as shown in FIGS. 13B, 13C and FIGS. 14B and 14C, a significant level of indel formation was not predictive for insertion or expression of the transgenes, indicating little correlation between editing and insertion/expression of the bidirectional constructs in PCH and PHH, respectively. As one measure, the R² value calculated in FIG. 13C is 0.13, and the R² value of FIG. 14D is 0.22.

[0314] Additionally, ssAAV vectors comprising a bidirectional construct were tested across a panel of target sites utilizing single guide RNAs targeting intron 1 of human albumin in primary human hepatocytes (PHH).

[0315] The ssAAV and LNP materials were prepared and delivered to PHH as described in Example 1. Following treatment, isolated genomic DNA and cell media was collected for editing and transgene expression analysis, respectively. Each of the vectors comprised a reporter that can be measured through luciferase-based fluorescence detection as described in Example 1 (derived from plasmid P00415), plotted in FIG. 14D as relative luciferase units (“RLU”) and tabulated in Table 24 below. For example, the AAV vectors contained the NanoLuc ORF (in addition to GFP). Schematics of the vectors tested are provided in FIGS. 13B and 14B. The gRNAs tested are shown in FIG. 14D using a shortened number for those listed in Table 1 and Table 7.

TABLE 3

Albumin intron 1 editing data for sgRNAs delivered to primary cynomolgus hepatocytes		
GUIDE ID	Avg % Edit	Std Dev % Edit
G009867	25.05	0.21
G009866	18.7	3.96
G009876	14.85	4.88
G009875	12.85	2.33
G009874	28.25	6.01
G009873	42.65	5.59
G009865	59.15	0.21
G009872	48.15	3.46
G009871	46.5	5.23
G009864	33.2	8.34
G009863	54.8	12.45
G009862	44.6	7.21
G009861	28.65	0.21
G009860	33.2	7.07
G009859	0.05	0.07
G009858	14.65	1.77
G009857	23	0.99
G009856	14.8	0.99
G009851	1.5	0.42
G009868	12.15	2.47
G009850	63.45	13.93
G009849	57.55	8.27
G009848	33	5.37
G009847	66.75	7
G009846	61.85	5.02
G009845	54.4	7.5
G009844	47.15	2.05

TABLE 4

Albumin intron 1 editing data for sgRNAs delivered to primary human hepatocytes		
GUIDE ID	Avg % Edit	Std Dev % Edit
G009844	19.07	2.07
G009851	0.43	0.35
G009852	47.20	3.96
G009857	0.10	0.14
G009858	8.63	9.16
G009859	3.07	3.50
G009860	18.80	4.90
G009861	10.27	2.51
G009866	13.60	13.55
G009867	12.97	3.04
G009868	0.63	0.32
G009874	49.13	0.60
G012747	3.83	0.23
G012748	1.30	0.35
G012749	9.77	1.50
G012750	42.73	4.58
G012751	7.77	1.16

TABLE 4-continued

Albumin intron 1 editing data for sgRNAs delivered to primary human hepatocytes		
GUIDE ID	Avg % Edit	Std Dev % Edit
G012752	32.93	2.27
G012753	21.20	2.95
G012754	0.60	0.10
G012755	1.10	0.10
G012756	2.17	0.40
G012757	1.07	0.25
G012758	0.90	0.10
G012759	2.60	0.35
G012760	39.10	6.58
G012761	36.17	2.43
G012762	8.50	0.57
G012763	47.07	3.07
G012764	44.57	5.83
G012765	19.90	1.68
G012766	8.50	0.28

TABLE 24

hAlb Guide Screen Luciferase		
Guide	Average Luciferase (RLU)	St. Dev Luciferase (RLU)
G009844	3700000	509116.9
G009852	281000	69296.46
G009857	1550000	127279.2
G009858	551000	108894.4
G009859	1425000	77781.75
G009860	2240000	183847.8
G009861	663500	238295
G009866	274000	11313.71
G009867	44700	565.6854
G009874	2865000	431335.1
G012747	651000	59396.97
G012749	867000	93338.1
G012752	4130000	268700.6
G012753	1145000	162634.6
G012757	579000	257386.9
G012760	129000	36769.55
G012761	4045000	728320
G012762	2220000	127279.2
G012763	1155000	205061
G012764	11900000	1555635
G012765	1935000	134350.3
G012766	2050000	169705.6
LNP	8430	212.132

Example 12—In Vivo Testing of Factor IX Expression from an Alternative Safe Harbor Locus

[0316] In this Example, insertion of ssAAV comprising a bidirectional hFIX construct at an alternative safe harbor locus was evaluated. To test the insertion into an alternative safe harbor locus, AAV was prepared as described above. Mice were administered with AAVs at a dose of 3e11 vg/mouse immediately followed by administration of LNPs formulated with Cas9 mRNAs and guide RNAs at a dose of 0.3 mg/kg. Animals were sacrificed 4 weeks post-dose, and liver and blood samples were collected. Editing in the liver samples was determined by NGS. Human hFIX levels in the serum was determined by ELISA. The NGS and ELISA data showed effective insertion and expression of hFIX within the alternative safe harbor locus.

Example 13—In Vivo Testing of the Human Factor IX Gene Insertion in Non-Human Primates

[0317] In this example, an 8 week study was performed to evaluate the human Factor IX gene insertion and hFIX protein expression in cynomolgus monkeys through administration of adeno-associated virus (AAV) and/or lipid nanoparticles (LNP) with various guides. This study was conducted with LNP formulations and AAV formulations prepared as described above. Each LNP formulation contained Cas9 mRNA and guide RNA (gRNA) with an mRNA: gRNA ratio of 2:1 by weight. The ssAAV was derived from P00147.

[0318] Male cynomolgus monkeys were treated in cohorts of n=3. Animals were dosed with AAV by slow bolus injection or infusion in the doses described in Table 10. Following AAV treatment, animals received buffer or LNP as described in Table 10 by slow bolus or infusion.

[0319] Two weeks post-dose, liver specimens were collected through single ultrasound-guided percutaneous biopsy. Each biopsy specimen was flash frozen in liquid nitrogen and stored at -86 to -60°C. Editing analysis of the liver specimens was performed by NGS Sequencing as previously described.

[0320] For Factor IX ELISA analysis, blood samples were collected from the animals on days 7, 14, 28, and 56 post-dose. Blood samples were collected and processed to plasma following blood draw and stored at -86 to -60°C until analysis.

[0321] The total human Factor IX levels were determined from plasma samples by ELISA. Briefly, Reacti-Bind 96-well microplate (VWR Cat #PI15041) were coated with capture antibody (mouse mAB to human Factor IX antibody (HTI, Cat #AHIX-5041)) at a concentration of 1 µg/ml then blocked using 1xPBS with 5% Bovine Serum Albumin. Test samples or standards of purified human Factor IX protein (ERL, Cat #HFIX 1009, Lot #HFIX4840) diluted in Cynomolgus monkey plasma were next incubated in individual wells. The detection antibody (Sheep anti-human Factor 9 polyclonal antibody, Abcam, Cat #ab128048) was adsorbed at a concentration of 100 ng/ml. The secondary antibody (Donkey anti-Sheep IgG pAbs with HRP, Abcam, Cat #ab97125) was used at 100 ng/mL. TMB Substrate Reagent set (BD OptEIA Cat #555214) was used to develop the plate. Optical density was assessed spectrophotometrically at 450 nm on a microplate reader (Molecular Devices i3 system) and analyzed using SoftMax pro 6.4.

[0322] Indel formation was detected, confirming that editing occurred. The NGS data showed effective indel formation. Expression of hFIX from the albumin locus in NHPs was measured by ELISA and is depicted in Table 11 and FIG. 15. Plasma levels of hFIX reached levels previously described as therapeutically effective (George, et al., NEJM 377(23), 2215-27, 2017).

[0323] As measured, circulating hFIX protein levels were sustained through the eight week study (see FIG. 15, showing day 7, 14, 28, and 56 average levels of ~135, ~140, ~150, and ~110 ng/mL, respectively), achieving protein levels ranging from ~75 ng/mL to ~250 ng/mL. Plasma hFIX levels were calculated using a specific activity of ~8 fold higher for the R338L hyperfunctional hFIX variant (Simioni et al., NEJM 361(17), 1671-75, 2009) (which reports a protein-specific activity of hFIX-R338L of 390±28 U per milligram, and a protein-specific activity for wild-type factor IX of 45±2.4 U per milligram). Calculating the func-

tionally normalized Factor IX activity for the hyperfunctional Factor IX variant tested in this example, the experiment achieved stable levels of human Factor IX protein in the NHPs over the 8 week study that correspond to about 20-40% of wild type Factor IX activity (range spans 12-67% of wild type Factor IX activity).

TABLE 10

Editing in liver					
Animal ID	Guide ID	F9-AAV (vg/kg)	F9-AAV Volume (mL/kg)	LNP (mg/kg)	LNP Volume (mL/kg)
4001	G009860	3E+13	1	3	2
4002	G009860	3E+13	1	3	2
4003	G009860	3E+13	1	3	2
5001	TSS	3E+13	1	0	0
5002	TSS	3E+13	1	0	0
5003	TSS	3E+13	1	0	0
6001	G009862	0	0	3	2
6002	G009862	0	0	3	2
6003	G009862	0	0	3	2

#ab128048), and donkey anti-Sheep IgG pAbs with HRP (Abcam, Cat #ab97125), as described in Example 13. Human FIX protein levels >3 fold higher than those achieved in the experiment of Example 13 were obtained from the bidirectional template using alternative CRISPR/Cas9 LNP. In the study, ELISA assay results indicate that circulating hFIX protein levels at or above the normal range of human FIX levels (3-5 ug/mL; Amiral et al., Clin. Chem., 30(9), 1512-16, 1984) were achieved using G009860 in the NHPs by at least the day 14 and 28 timepoints. Initial data indicated circulating human FIX protein levels of ~3-4 g/mL at day 14 after a single dose, with levels sustained through the first 28 days (~3-5 g/mL) of the study. The human FIX levels were measured at the conclusion of the study by the same method and data are presented in the Table 25. Additional guides G009847, G009862, and G009864 were also tested and shown to facilitate insertion of a FIX-expressing template in the NHP study.

TABLE 25

Serum human Factor IX protein levels—ELISA Method of Example 13									
	Day 7	Day 14	Day 28	Day 42	Day 56		Day 7	Day 14	Day 28
	FIX ng/mL	STD DEV	FIX ng/mL						
3001	2532.8	145.6	2562.6	99.0	3011.7	62.7	2936.7	72.4	2748.5
3002	2211.4	95.8	2958.5	119.2	3350.2	98.4	3049.7	112.7	3036.7
3003	3195.1	475.6	4433.9	238.7	3367.2	157.7	3746.1	95.6	3925.0
									86.0
									90.6
									157.4

TABLE 11

hFIX expression				
Animal ID	Day 7 Factor IX (ng/mL)	Day 14 Factor IX (ng/mL)	Day 28 Factor IX (ng/mL)	Day 56 Factor IX (ng/mL)
4001	122.84/+2.85	94.93/+0.56	105.65/+1.94	97.31/+1.49
4002	149.77/+13.5	222.92/+9.61	252.49/+6.46	152.05/+7.46
4003	134.06/+6.17	107.04/+6.46	95.30/+3.18	74.23/+3.53
5001	ND	ND	ND	ND
5002	ND	ND	ND	ND
5003	ND	ND	ND	ND
6001	ND	ND	ND	ND
6002	ND	ND	ND	ND
6003	ND	ND	ND	ND

Example 14 In Vivo Testing of Factor IX Insertion in Non-Human Primates

[0324] In this example, a study was performed to evaluate the Factor IX gene insertion and hFIX protein expression in cynomolgus monkeys following administration of ssAAV derived from P00147 and/or CRISPR/Cas9 lipid nanoparticles (LNP) with various guides including G009860 and various LNP components.

[0325] Indel formation was measured by NGS, confirming that editing occurred. Total human Factor IX levels were determined from plasma samples by ELISA using a mouse mAB to human Factor IX antibody (HTI, Cat #AHIX-5041), sheep anti-human Factor 9 polyclonal antibody (Abcam, Cat

[0326] Circulating albumin levels were measured by ELISA, indicating that baseline albumin levels are maintained at 28 days. Tested albumin levels in untreated animals varied ±~15% in the study. In treated animals, circulating albumin levels changed minimally and did not drop out of the normal range, and the levels recovered to baseline within one month.

[0327] Circulating human FIX protein levels were also determined by a sandwich immunoassay with a greater dynamic range. Briefly, an MSD GOLD 96-well Streptavidin SECTOR Plate (Meso Scale Diagnostics, Cat. L15SA-1) was blocked with 1% ECL Blocking Agent (Sigma, GERPN2125). After tapping out the blocking solution, biotinylated capture antibody (Sino Biological, 11503-R044) was immobilized on the plate. Recombinant human FIX protein (Enzyme Research Laboratories, HFIX 1009) was used to prepare a calibration standard in 0.5% ECL Blocking Agent. Following a wash, calibration standards and plasma samples were added to the plate and incubated. Following a wash, a detection antibody (Haematologic Technologies, AHIX-5041) conjugated with a sulfo-tag label was added to the wells and incubated. After washing away any unbound detection antibody, Read Buffer T was applied to the wells. Without any additional incubation, the plate was imaged with an MSD Quick Plex SQ120 instrument and data was analyzed with Discovery Workbench 4.0 software package (Meso Scale Discovery). Concentrations are expressed as mean calculated concentrations in ug/m. For the samples, N=3 unless indicated with an asterisk, in which case N=2.

Expression of hFIX from the albumin locus in the treated study group as measured by the MSD ELISA is depicted in Table 26.

TABLE 26

Serum human Factor IX protein levels—MSD ELISA					
Mean Calc. Conc. (μg/mL)					
	3001	3002	3003		
Time Point	Conc.	Inter-Assay CV	Conc.	Inter-Assay CV	Inter-Assay CV
Day 7	7.85	20%	5.63	14%	11.20
Day 14	8.65	15%	11.06	18%	14.70
Day 28	9.14	7%	14.12	7%	10.85
Day 42	9.03	10%	33.12*	0%	13.22
Day 56	10.24	13%	16.72	12%	33.84*

Example 15—Off-Target Analysis of Albumin Human Guides

[0328] A biochemical method (See, e.g., Cameron et al., *Nature Methods*. 6, 600-606; 2017) was used to determine potential off-target genomic sites cleaved by Cas9 targeting Albumin. In this experiment, 13 sgRNA targeting human Albumin and two control guides with known off-target profiles were screened using isolated HEK293 genomic DNA. The number of potential off-target sites detected using a guide concentration of 16 nM in the biochemical assay were shown in Table 27. The assay identified potential off-target sites for the sgRNAs tested.

TABLE 27

Off-Target Analysis			
sgRNA ID	Target	Guide Sequence (SEQ ID NO.:)	Off-Target Site Count
G012753	Albumin	GACUGAAACUUCACAGAAUA (SEQ ID NO: 20)	62
G012761	Albumin	AGUGCAAUGGAUAGGUUUU (SEQ ID NO: 28)	75
G012752	Albumin	UGACUGAAACUUCACAGAAU (SEQ ID NO: 19)	223
G012764	Albumin	CCUCACUCUUGUCUGGGCAA (SEQ ID NO: 31)	3985
G012763	Albumin	UGGGCAAGGGAAGAAAAAAA (SEQ ID NO: 30)	5443
G009857	Albumin	AUUUAUGAGAUCAACAGCAC (SEQ ID NO: 5)	131
G009859	Albumin	UUAAAUAAGCAUAGUGCAA (SEQ ID NO: 7)	91
G009860	Albumin	UAAAGCAUAGUGCAAUGGAU (SEQ ID NO: 8)	133
G012762	Albumin	UGAUUCCUACAGAAAAACUC (SEQ ID NO: 29)	68

TABLE 27-continued

Off-Target Analysis			
sgRNA ID	Target	Guide Sequence (SEQ ID NO.:)	Off-Target Site Count
G009844	Albumin	GAGCAACCUCACUCUUGUCU (SEQ ID NO: 2)	107
G012765	Albumin	ACCUCACUCUUGUCUGGGCA (SEQ ID NO: 32)	41
G012766	Albumin	UGAGCAACCUCACUCUUGUC (SEQ ID NO: 33)	78
G009874	Albumin	UAAUAAAUAUCAAACAUCCU (SEQ ID NO: 13)	53
G000644	EMX1	GAGUCCGAGCAGAAGAAGAA (SEQ ID NO: 1129)	304
G000645	VEGFA	GACCCCCUCCACCCGCCUC (SEQ ID NO: 1130)	1641

[0329] In known off-target detection assays such as the biochemical method used above, a large number of potential off-target sites are typically recovered, by design, so as to “cast a wide net” for potential sites that can be validated in other contexts, e.g., in a primary cell of interest. For example, the biochemical method typically overrepresents the number of potential off-target sites as the assay utilizes purified high molecular weight genomic DNA free of the cell environment and is dependent on the dose of Cas9 RNP used. Accordingly, potential off-target sites identified by these methods may be validated using targeted sequencing of the identified potential off-target sites.

Example 16. Use of Humanized Albumin Mice to Screen Guide RNAs for Human F9 Insertion In Vivo

[0330] We aimed to identify effective guide RNAs for hF9 insertion into the human albumin locus. To this end, we utilized mice in which the mouse albumin locus was replaced with the corresponding human albumin genomic sequence, including the first intron (ALB^{hu/hu} mice). This allowed us to test the insertion efficiency of guide RNAs targeting the first intron of human albumin in the context of an adult liver in vivo. Two separate mouse experiments were set up using the ALB^{hu/hu} mice to screen a total of 11 guide RNAs, each targeting the first intron of the human albumin locus. All mice were weighed and injected via tail vein at day 0 of the experiment. Blood was collected at weeks 1, 3, 4, and 6 via tail bleed, and plasma was separated. Mice were terminated at week 7. Blood was collected via the vena cava, and plasma was separated. Livers and spleens were dissected as well.

[0331] In the first experiment, 6 LNPs comprising Cas9 mRNA and the following guides were prepared as in Example 1 and tested: G009852, G009859, G009860, G009864, G009874, and G012764. LNPs were diluted to 0.3 mg/kg (using an average weight of 30 grams) and co-injected with AAV8 packaged with the bi-directional hF9 insertion template at a dose of 3E11 viral genomes per mouse. Five ALB^{hu/hu} male mice between 12 and 14 weeks old were injected per group. Five mice from same cohort were injected with AAV8 packaged with a CAGG promoter

operably linked to hF9, which leads to episomal expression of hF9 (at 3E11 viral genomes per mouse). There were three negative control groups with three mice per group that were injected with buffer alone, AAV8 packaged with the bi-directional hF9 insertion template alone, or LNP-G009874 alone.

[0332] In the experiment, the following LNPs comprising Cas9 mRNA and the following guides were prepared as in Example 1 and tested: G009860, G012764, G009844, G009857, G012752, G012753, and G012761. All were diluted to 0.3 mg/kg (using an average weight of 40 grams) and co-injected with AAV8 packaged with the bi-directional hF9 insertion template at a dose of 3E11 viral genomes per mouse. Five ALB^{hu/hu} male mice 30 weeks old were injected per group. Five mice from same cohort were injected with AAV8 packaged with a CAGG promoter operably linked to hF9, which leads to episomal expression of hF9 (at 3E11 viral genomes per mouse). There were three negative control groups with three mice per group that were injected with buffer alone, AAV8 packaged with the bi-directional hF9 insertion template alone, or LNP-G009874 alone.

[0333] For analysis, an ELISA was performed to measure levels of hFIX circulating in the mice at each timepoint. Human Factor IX ELISA Kits (ab188393) were used for this purpose, and all plates were run with human pooled normal plasma from George King Bio-Medical as a positive assay control. Human Factor IX expression levels in the plasma samples in each group at week 6 post-injection are shown in FIG. 16A and FIG. 16B. Consistent with the in vitro insertion data, low to no Factor IX serum levels were detected when guide RNA G009852 was used. Consistent with the lack of an adjacent PAM sequence in human albumin, Factor IX serum levels were not detectable when guide RNA G009864 was used. Factor IX expression in the serum was observed for the groups using guide RNAs G009859, G009860, G009874, and G0012764.

[0334] Spleens and a portion of the left lateral lobe of all livers were submitted for next-generation sequencing (NGS) analysis. NGS was used to assess the percentage of liver cells with insertions/deletions (indels) at the humanized albumin locus at week 7 post-injection with AAV-hF9 donor and LNP-CRISPR/Cas9. Consistent with the lack of an adjacent PAM sequence in human albumin, no editing was detectable in the liver when guide RNA G009864 was used. Editing in the liver was observed for the groups using guide RNAs G009859, G009860, G009874, and G012764 (data not shown).

[0335] The remaining liver was fixed for 24 hours in 10% neutral buffered formalin and then transferred to 70% ethanol. Four to five samples from separate lobes were cut and shipped to HistoWisz and were processed and embedded in paraffin blocks. Five-micron sections were then cut from each paraffin block, and BASESCOPE™ was performed on the Ventana Ultra Discovery (Roche) using the universal BASESCOPE™ procedure and reagents by Advanced Cell Diagnostics and a custom designed probe that targets the unique mRNA junction formed between the human albumin signal sequence from the first intron of the ALB^{hu/hu} albumin locus and the hF9 transgene when successful integration and transcription is achieved. HALO imaging software (Indica Labs) was then used to quantify the percentage of positive cells in each sample. The average of percentage positive cells across the multiple lobes for each animal was then correlated to the hFIX levels in the serum at week 7. The

results are shown in FIG. 17 and Table 28. The week 7 serum levels and the 00 positive cells for the hALB-hF9 mRNA strongly correlated ($r=0.89$; $R^2=0.79$).

TABLE 28

Week 7 hFIX and BASESCOPE™ Data.					
Mouse	Guide	hFIX ug/mL (Week 7)	% mRNA Probe (4-5 Sections)	STD % mRNA Probe	Total Cells Counted
1	Buffer	ND	0.09	0.03	152833
4	AAV	ND	0.53	0.67	351084
7	Only	ND	0.48	0.33	75160
7	LNP	ND	0.48	0.33	75160
10	CAG F9	211.8	0.20	0.22	190277
15	G009852	ND	0.30	0.09	144518
20	G009859	0.5	0.82	0.45	143817
21	G009859	0.5	0.88	0.43	160172
22	G009859	2.3	1.71	1.54	26015
23	G009859	3.8	2.74	0.59	183085
24	G009859	0.6	2.78	1.96	152424
25	G009860	5.6	12.46	5.80	78935
26	G009860	10.6	13.76	5.32	112252
27	G009860	9.7	14.80	5.45	201592
28	G009860	2.1	3.32	0.76	84710
29	G009860	3.0	1.52	0.35	203277
30	G009864	ND	1.94	1.78	145807
35	G009874	1.7	2.42	1.14	126665
36	G009874	1.5	1.08	0.53	195861
37	G009874	2.1	1.02	1.29	181679
38	G009874	5.5	0.40	0.43	175359
39	G009874	1.5	0.44	0.18	205417
40	G012764	15.7	28.85	7.11	167824
41	G012764	19.6	19.17	8.23	70081
42	G012764	1.9	1.95	1.79	154742
43	G012764	7.7	4.38	0.68	114060
44	G012764	3.0	1.64	1.04	238623
43	DapB(-)	—	0.12	0.07	144730

Example 17—Use of Humanized Albumin Mice
Crossed with F9 Knockout Mice to Assess
Functionality of Inserted Human F9 In Vivo

[0336] For a next study, functionality of inserted hF9 was tested in male ALB^{ms/hu}xF9^{-/-} mice. LNPs comprising Cas9 mRNA and the following guides were prepared as in Example 1 and tested: G009860 (targeting the first intron of the human albumin locus) and G000666 (targeting the first intron of the mouse albumin locus). G009860 was diluted to 0.3 mg/kg, and G000666 was diluted to 1.0 mg/kg (using an average weight of 31.2 grams), and both were co-injected with AAV8 packaged with the bi-directional hF9 insertion template at a dose of 3E11 viral genomes per mouse. Five ALB^{ms/hu}xF9^{-/-} male mice (16 weeks old) were injected per group. Five mice from same cohort were injected with AAV8 packaged with a CAGG promoter operably linked to hF9, which leads to episomal expression of hF9 (at 3E11 viral genomes per mouse). There were six negative control animals with one mouse per group that was injected with buffer alone or AAV8 packaged with the bi-directional hF9 insertion template alone, and two mice per group that were injected with LNP-G009860 or LNP-G000666 alone at 0.3 mg/kg and 1.0 mg/kg, respectively.

[0337] For analysis, an ELISA was performed to measure levels of hFIX circulating in the mice at each timepoint. Human Factor IX ELISA Kits (ab188393) were used for this purpose, and all plates were run with human pooled normal plasma from George King Bio-Medical as a positive assay

control. Spleens and a portion of the left lateral lobe of all livers were submitted for NGS analysis.

[0338] Human Factor IX expression levels in the plasma samples in each group at weeks 1, 2, and 4 post-injection are shown in FIG. 18 and in Table 29. In addition, NGS results showing insertion and deletion (indel) levels at the albumin locus in the liver and spleen are shown in Table 29. As shown in FIG. 18 and Table 29, hFIX was detected in the plasma of treated Alb^{+hu/F9^{-/-}}

TABLE 29

Sample	Week 1 (μ g/mL)	Week 3 (μ g/mL)	Week 4 (μ g/mL)	INDEL Liver	INDEL Spleen
S1 PBS	BLD	BLD	BLD	6.12	0.12
S18 AAV8 only	BLD	BLD	BLD	0.73	0.10
S2 G000666 only	BLD	BLD	BLD	37.48	0.92
S4 G000666 only	BLD	BLD	BLD	30.67	1.17
S19 G009860 only	BLD	BLD	BLD	12.25	0.31
S20 G009860 only	BLD	BLD	BLD	10.73	0.45
S10 CAG	42.60	129.83	117.74	1.45	0.12
S14 CAG	35.55	82.25	100.95	0.08	0.11
S15 CAG	37.30	115.51	107.26	0.10	0.05
S16 CAG	36.39	81.27	116.24	0.05	0.10
S17 CAG	40.50	101.38	124.15	0.16	0.06
S5 AAV8 + G000666	2.90	5.00	8.79	41.46	1.43
S6 AAV8 + G000666	4.67	6.11	10.29	33.81	1.59
S7 AAV8 + G000666	2.88	3.15	3.01	33.47	1.04
S8 AAV8 + G000666	0.94	1.61	No sample	36.54	1.34
S9 AAV8 + G000666	7.14	7.53	7.23	30.63	1.38
S11 AAV8 + G009860	0.73	0.62	0.86	11.15	0.52
S12 AAV8 + G009860	0.52	0.43	0.47	7.05	0.39
S13 AAV8 + G009860	1.71	1.89	0.93	18.38	0.57
S21 AAV8 + G009860	1.21	2.79	0.59	13.44	0.22
S22 AAV8 + G009860	2.06	1.03	2.37	18.06	0.19
Human	4.00	3.91	4.12	N/A	N/A

[0339] The remaining liver was fixed for 24 hours in 1000 neutral buffered formalin and then transferred to 700% ethanol. Four to five samples from separate lobes were cut and shipped to HistoWiz and were processed and embedded in paraffin blocks. Five-micron sections were then cut from each paraffin block for analysis via BASESCOPE™ on the Ventana Ultra Discovery (Roche) using the universal BASESCOPE™ procedure and reagents by Advanced Cell Diagnostics and a custom designed probe that targets the unique mRNA junction formed between either the human or the mouse albumin signal sequence from the first intron of each respective albumin locus in the ALB^{ms/hu} mouse and the hF9 transgene when successful integration and transcription is achieved. HALO imaging software (Indica Labs) is used to quantify the percentage of positive cells in each sample.

[0340] Next, terminal blood was used for assessment of functional coagulation activity by activated partial thromboplastin time (aPTT) and Thrombin Generation Assay (TGA). Activated partial thromboplastin time (aPTT) is a clinical measurement of intrinsic pathway clotting activity in plasma. Plasma is induced to clot by the addition of ellagic acid or kaolin, both of which activate coagulation factor XII in the intrinsic pathway (as known as the contact pathway)

of coagulation, that subsequently results in the generation of fibrin from fibrinogen once thrombin is activated. The aPTT assay provides an estimation of an individual's ability to generate a clot, and this information can be used to determine risk of bleeding or thrombosis. To test aPTT, a semi-automated benchtop system (Diagnostica Stago STart 4) with an electro-mechanical clot detection method (viscosity-based detection system) was used to assess clotting in plasma. To each cuvette with a steel ball, 50 μ L of citrated plasma was added and incubated at 37° C. for 5 min, and then clotting was triggered with the addition of 50 μ L of ellagic acid (final concentration of 30 μ M) at 37° C. for 300 seconds. Following final activation of clotting by adding 50 μ L of 0.025 M calcium chloride (final concentration of 8 mM) to each cuvette, the steel ball began to oscillate back and forth between the two drive coils. The movement of the ball was detected by the receiver coil. The generation of fibrin increased plasma viscosity until the ball ceased to move, which was recorded as the clotting time. The only parameter measured was clotting time. Runs were conducted in duplicate.

[0341] Thrombin generation assay (TGA) is a non-clinical assessment of the kinetics of thrombin generation in activated plasma. Thrombin generation is an essential process of coagulation because thrombin is responsible for activation of other coagulation factors and propagation of additional thrombin (via FXI activation) for the conversion of fibrinogen to fibrin. Thrombin generation assay provides an estimation of an individual's ability to generate thrombin, and this information can be used to determine risk of bleeding or thrombosis. To perform the TGA, a calibrated automated thrombogram was used to assess thrombin generation levels in a spectrophotometer (Thrombinograph™, Thermo Scientific). For high throughput experimentation, 96-well plates (Immulon II HB) were used. To each well, 55 μ L of citrated plasma (4x diluted with saline for mouse plasma) was added and incubated at 37° C. for 30 min. Thrombin generation is triggered with the addition of 15 μ L of 2 M ellagic acid (final concentration of 0.33 μ M) at 37° C. for 45 min. Thrombin generation was determined following the automated injection of 15 μ L of the fluorogenic substrate with 16 mM CaCl₂ (FluCa; Thrombinoscope BV) into each well. The fluorogenic substrate reacted with the generated thrombin, which was measured continuously in the plasma every 33 sec for 90 min at 460 nm. The fluorescence intensity was proportional to the proteolytic activity of thrombin. The main parameters measured in the tracing were lag time, peak thrombin generation, time to peak thrombin generation, and endogenous thrombin potential (ETP). The lag time provides an estimation of time required for initial detection of thrombin in plasma. The peak is the maximum amount of thrombin generated at a given time after activation. Time to peak thrombin generation is time from initiation of the coagulation cascade to the peak generation of thrombin. ETP is the total amount of thrombin generated during the 60 minutes measured. Runs were conducted in duplicate.

[0342] As shown in FIG. 19 and Table 30, insertion of the hF9 transgene using for example G000666 showed recovered clotting function in the aPTT assay. AAV only and LNP only negative control samples showed prolonged aPTT times of 45-60 seconds in saline. The positive control CAGG and test samples AAV8+LNP were closer to the normal human aPTT of 28-34 seconds.

TABLE 30

aPTT and TGA-EA.

Sample #	I.V. Injection	Week 4 F9 μg/mL	Average aPTT (sec)	TGA-EA Peak (nM)
1	PBS	BLD	40.2	11.13
18	AAV Only	BLD	62.5	-1
2	LNP g666 only	BLD	53.9	-1
4	LNP g666 only	BLD	65.0	2.45
19	LNP G009860 only	BLD	34.1	42.83
20	LNP G009860 only	BLD	56.7	18.07
10	AAV + CAGG F9	117.74	41.1	42.65
14	AAV + CAGG F9	100.95	34.1	49.96
15	AAV + CAGG F9	107.26	42.2	49.49
16	AAV + CAGG F9	116.24	37.9	44.46
17	AAV + CAGG F9	124.15	44.1	38.02
5	AAV + g666	8.79	31.3	72.11
6	AAV + g666	10.29	32.6	90.14
7	AAV + g666	3.01	33.5	58.33
8	AAV + g666	no sample	NA	NA
9	AAV + g666	7.23	25.9	67.23
11	AAV + G009860	0.86	36.8	56.92
12	AAV + G009860	0.47	37.7	45.16
13	AAV + G009860	0.93	35.3	60.45
21	AAV + G009860	0.59	36.1	47.44
22	AAV + G009860	2.37	>300	Clots in tube

[0343] As shown in FIG. 20A, FIG. 20B, and FIG. 21 and in Table 30, insertion of the hF9 transgene using for example G000666 showed increased thrombin generation in TGA-EA analysis. Thrombin concentrations were higher in the positive control CAGG and AAV8+LNP as compared to the negative control samples.

[0344] In conclusion, hFIX was detected in the plasma of Alb^{+/-}/F9^{-/-} mice at 1, 3, and 4 weeks, and the expressed

hFIX-R338L was found to be functional since thrombin was generated in a TGA assay, and aPTT clotting time was improved.

Human albumin intron 1:

(SEQ ID NO: 1)

GTAAGAAATCCATTTCATTGTTCAACTTTTATTCTATTCCCAG
 TAAAATAAGTTTAGTAAACTCTGCATCTTAAAGAATTATTTGGC
 ATTTATTTCTAAATGCCATAGTATTGTATTGAAAGTCTTACAA
 GGTTATCTTATAATAAAATTCAAACATCCTAGGTAAAAAAAAAAAA
 GGTCAAGATTGTTAGTGACTGTAATTCTTTGCGCACTAAGGAAA
 GTGCAAAGTAACCTAGAGTGACTGAAACTCACAGAATAGGGTGAAG
 ATTGAAATTCAACTATCCAAAGACCTATCCATTGCACTATGCTTTA
 TTTAAAAACACAAAACCTGTGCTGTGATCTCATAAATAGAACTTGT
 ATTTATTTATTTCTATTAGTCTGCTTCTGGTGCTGTTGATA
 GACACTAAAAGAGTATTAGATATTCTAAGTTGAATATAAGGCTAT
 AAATATTTAATAATTAAAGTATTCTGGTAATTGAATTATTC
 TTCTGTTAAAGGCAAGAGAAATAATTGAACATCATCCTGAGTTTC
 TGAGGAATCAGAGCCAATATTGAAACAAATGCATAATCTAAGTC
 AAATGGAAAGAAATATAAAAGTAACATTATTACTTCTGTTCTTC
 AGTATTTAACAACTCTTTCTTCTGCCAG

TABLE 5

Mouse albumin guide RNA			
Guide ID	Guide Sequence	Mouse Genomic Coordinates (mm10)	SEQ ID NO:
G000551	AUUUGCAUCUGAGAACCUU	chr5:90461148-90461168	98
G000552	AUCGGGAACUGGCAUCUUA	chr5:90461590-90461610	99
G000553	GUUACAGGAAAUCUGAAGG	chr5:90461569-90461589	100
G000554	GAUCGGGAACUGGCAUCUUC	chr5:90461589-90461609	101
G000555	UGCAUCUGAGAACCCUUAGG	chr5:90461151-90461171	102
G000666	CACUCUUGUCUGUGGAAACA	chr5:90461709-90461729	103
G000667	AUCGUUACAGGAAAUCUGA	chr5:90461572-90461592	104
G000668	GCAUCUUCAGGGAGUAGCUU	chr5:90461601-90461621	105
G000669	CAAUCUUAAAUAUGUUGUG	chr5:90461674-90461694	106
G000670	UCACUCUUGUCUGUGGAAAC	chr5:90461710-90461730	107
G011722	UGCUUGUAUUUUUCAUGUAA	chr5:90461039-90461059	108
G011723	GUAAAUAUCUACUAAGACAA	chr5:90461425-90461445	109
G011724	UUUUUCUAGUAAUGGAAGCC	chr5:90461047-90461067	110
G011725	UUUAUUAUUGAUUAUUUU	chr5:90461174-90461194	111
G011726	GCACAGAUAAAACACUAAA	chr5:90461480-90461500	112

TABLE 5-continued

Mouse albumin guide RNA			
Guide ID	Guide Sequence	Mouse Genomic Coordinates (mm10)	SEQ ID NO:
G011727	CACAGAUUAACACUUUAC	chr5:90461481-90461501	113
G011728	GGUUUUAAAAAUAAUAAAUGU	chr5:90461502-90461522	114
G011729	UCAGAUUUCCUGUAACGAU	chr5:90461572-90461592	115
G011730	CAGAUUUCCUGUAACGAUC	chr5:90461573-90461593	116
G011731	CAAUGGUAAAUAAGAAAUA	chr5:90461408-90461428	117
G013018	GGAAAAUCUGAAGGUGGCAA	chr5:90461563-90461583	118
G013019	GGCGAUCUCACUCUUGUCUG	chr5:90461717-90461737	119

TABLE 6

Mouse albumin sqRNAs and modification pattern			
Guide ID	Full Sequence	SEQ ID NO: Full Sequence Modified	SEQ ID NO:
G000551	AUUUGCAUCUGAGAACCUUUG UUAGAGCUAGAAUAGCAAGU UAAAUAAGCUAGGCCUUAU CAACUUGAAAAGUGCACC GUCGGGCGUUU	120 mA*mU*mU*UGCAUCUGAGAACCCUUGUUUAGAm GmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU AAGGCUAGUCGUUAUCAmCmUmUmGmAmAm AmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	142
G000552	AUCGGGAACUGGCACUUCA GUUUUAGAGCUAGAAAAGC AAGUAAAUAAGCUAGUC CGUUUACUACUUGAAAAGU GGCACCGAGUCGGGCGUUU	121 mA*mU*mC*GGGACUGGCACUUCAGUUUUAGAm GmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU AAGGCUAGUCGUUAUCAmCmUmUmGmAmAm AmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	143
G000553	GUUACAGGAAAUCUGAAGG GUUUUAGAGCUAGAAAAGC AAGUAAAUAAGCUAGUC CGUUUACUACUUGAAAAGU GGCACCGAGUCGGGCGUUU	122 mG*mU*mU*ACAGGAAAUCUGAAGGGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU UAAGGCUAGUCGUUAUCAmCmUmUmGmAmAm mAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmG GmGmUmGmCmU*mU*mU*mU	144
G000554	GAUCGGGAACUGGCACUUUC GUUUUAGAGCUAGAAAAGC AAGUAAAUAAGCUAGUC CGUUUACUACUUGAAAAGU GGCACCGAGUCGGGCGUUU	123 mG*mA*mU*CGGGACUGGCACUUCGUUUUAGAm GmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU AAGGCUAGUCGUUAUCAmCmAmCmUmUmGmAmAm AmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	145
G000555	UGCAUCUGAGAACCCUUAGG GUUUUAGAGCUAGAAAAGC AAGUAAAUAAGCUAGUC CGUUUACUACUUGAAAAGU GGCACCGAGUCGGGCGUUU	124 mU*mG*mC*AUCUGAGAACCCUUAGGGUUUUAGAm GmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU AAGGCUAGUCGUUAUCAmCmAmCmUmUmGmAmAm AmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	146
G000666	CACCUUUGUCUGGGAAACA GUUUUAGAGCUAGAAAAGC AAGUAAAUAAGCUAGUC CGUUUACUACUUGAAAAGU GGCACCGAGUCGGGCGUUU	125 mC*mA*mC*UCUUGUCUGGGAAACAGUUUUAGAm GmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU AAGGCUAGUCGUUAUCAmCmAmCmUmUmGmAmAm AmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	147
G000667	AUCGUUACAGGAAAUCUGA GUUUUAGAGCUAGAAAAGC AAGUAAAUAAGCUAGUC CGUUUACUACUUGAAAAGU GGCACCGAGUCGGGCGUUU	126 mA*mU*mC*GUUACAGGAAAUCUGAGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU UAAGGCUAGUCGUUAUCAmCmUmUmGmAmAm mAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	148

TABLE 6 - continued

Mouse albumin sgRNAs and modification pattern			
Guide ID	Full Sequence	SEQ NO: Full Sequence Modified	SEQ ID NO:
G000668	GCAUCUUCAGGGAGUAGCUU GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	127 mG*mC*mA*UCUUUCAGGGAGUAGCUUGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	149
G000669	CAAUCUUAAAUAUGUUGUG GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	128 mC*mA*mA*UCUUUAAAUAUGUUGUGGUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	150
G000670	UCACUCUUGUCUGUGGAAAC GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	129 mU*mC*mA*CUCUUUCUGUGGAAACGUUUUAGA GmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAA AAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	151
G011722	UGCUUGUAUUUUUUCUAGUAA GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	130 mU*mG*mC*UUGUAUUUUUUCUAGUAGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	152
G011723	GUAAAUAUCUACUAAAGACAA GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	131 mG*mU*mA*AAUAUCUACUAAAGACAAGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	153
G011724	UUUUUCUAGUAAUGGAAGCC GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	132 mU*mU*mU*UUCUAGUAAUGGAAGCCGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	154
G011725	UUAAUAAAUGUAAUAAA GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	133 mU*mU*mA*UAUAAAUGUAAUAAAUGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	155
G011726	GCACAGAUAAAACACUUAA GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	134 mG*mC*mA*CAGAUAAAACACUUAAAGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	156
G011727	CACAGAUAAAACACUUAC GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	135 mC*mA*mC*AGAUAAAACACUUAAAGUUUUAGA GmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAA AAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	157
G011728	GGUUUUAAAUAUAUAUGU GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	136 mG*mG*mU*UUUUAAAUAUAUGGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	158
G011729	UCAGAUUUCCUGUAACGAU GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGGAGCUUUU	137 mU*mC*mA*GAUUUUCUGUAACGAUGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAA UAAGGCCUAGUCGGUUAUCAmAmCmUmUmGmAmA mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	159

TABLE 6 -continued

Mouse albumin sqRNAs and modification pattern					
Guide ID	Full Sequence	SEQ NO:	SEQ ID NO:	Full Sequence Modified	SEQ ID NO:
G011730	CAGAUUUUCUGUAACGAUC GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	138	160	mC*mA*mG*AUUUUCCUGUAACGAUCGUUUUAGAm GmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU AAGGCUGAGUCGUUAUCAmAmCmUmUmGmAmAm AmAmAmGmUmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	
G011731	CAAUGGUAAAUAAGAAAUA GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	139	161	mC*mA*mA*UGGUAAAUAAGAAAUAAGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU UAAGGCUGAGUCGUUAUCAmAmCmUmUmGmAmAm mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	
G013018	GGAAAAACUGAAGGGGGCAA GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	140	162	mG*mG*mA*AAACUGAAGGGGGCAAGUUUUAGA mGmCmUmAmGmAmAmUmAmGmCAAGUUAAAAAU UAAGGCUGAGUCGUUAUCAmAmCmUmUmGmAmAm mAmAmAmGmUmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	
G013019	GGCGAUACUCACUCUUGUCUG GUUUUAGAGCUAGAAAUGC AAGUUAAAUAAGGUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	141	163	mG*mG*mC*GAUCACACUUCUUGUCUGGUUUUAGAm GmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAAU AAGGCUGAGUCGUUAUCAmAmCmUmUmGmAmAm AmAmAmGmUmGmCmAmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU*mU	

TABLE 7

Cyno albumin guide RNA			
Guide ID	Guide Sequence	Cyno Genomic Coordinates (mf5)	SEQ ID NO:
G009844	GAGCAACCUCACUCUUGUCU	chr5:61198711-61198731	2
G009845	AGCAACCUCACUCUUGUCUG	chr5:61198712-61198732	165
G009846	ACCUCACUCUUGUCUGGGGA	chr5:61198716-61198736	166
G009847	CCUCACUCUUGUCUGGGGAA	chr5:61198717-61198737	167
G009848	CUCACUCUUGUCUGGGGAAG	chr5:61198718-61198738	168
G009849	GGGGAAAGGGGAGAAAAAAA	chr5:61198731-61198751	169
G009850	GGGAAGGGGAGAAAAAAA	chr5:61198732-61198752	170
G009851	AUGCAUUUGUUUCAAAAU	chr5:61198825-61198845	3
G009852	UGCAUUUGUUUCAAAAU	chr5:61198826-61198846	172
G009853	UGAUUCCUACAGAAAAAGUC	chr5:61198852-61198872	173
G009854	UACAGAAAAAGUCAGGAUA	chr5:61198859-61198879	174
G009855	UUUCUUCUGCCUUAAAACAG	chr5:61198889-61198909	175
G009856	UUAUAGUUUUAUUCAAC	chr5:61198957-61198977	176
G009857	AUUUAUGAGAACAGCAC	chr5:61199062-61199082	5
G009858	GAUCAACAGCACAGGUUUG	chr5:61199070-61199090	6
G009859	UUAAAUAAGCAUAGUGCAA	chr5:61199096-61199116	7
G009860	UAAAAGCAUAGUGCAAUGGAU	chr5:61199101-61199121	8
G009861	UAGUGCAAUGGAUAGGUUU	chr5:61199108-61199128	9

TABLE 7-continued

Cyno albumin guide RNA			
Guide ID	Guide Sequence	Cyno Genomic Coordinates (mf5)	SEQ ID NO:
G009862	AGUGCAAUGGAUAGGUUUUA	chr5:61199109-61199129	182
G009863	UUACUUUGCACUUUCCUUAG	chr5:61199186-61199206	183
G009864	UACUUUGCACUUUCCUUAGU	chr5:61199187-61199207	184
G009865	UCUGACCUUUUAAAAUACCU	chr5:61199238-61199258	185
G009866	UACUUAAAACUUUAAAAUACU	chr5:61199367-61199387	10
G009867	AAAGUUGAACAAUAGAAAAA	chr5:61199401-61199421	11
G009868	AAUGCAUAAUCUAAGUAAA	chr5:61198812-61198832	2
G009869	AUUAUCCUGACUUUUUCUGU	chr5:61198860-61198880	189
G009870	UGAAUUAUUUCUCUGUUUAA	chr5:61198901-61198921	190
G009871	UAAUUUUUUCCCCACUA	chr5:61199203-61199223	191
G009872	AAAAGGUCAGAAUUGUUUAG	chr5:61199229-61199249	192
G009873	AACAUCUAGGUAAAUAAAA	chr5:61199246-61199266	193
G009874	UAAAUAUUUCAACAUCCU	chr5:61199258-61199278	13
G009875	UUGUCAUGUAUUUCUAAAUA	chr5:61199322-61199342	195
G009876	UUUGUCAUGUAUUUCUAAAA	chr5:61199323-61199343	196

TABLE 8

Cyno sqRNA and modification patterns			
Guide ID	Full Sequence	SEQ ID NO:	SEQ ID NO:
		Full Sequence	Modified
G009844	GAGCAACCUCACUCUUGUCU GUUUUAGAGCUAGAAAAGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	34	mG*mA*mG*CAACCUCACUCUUGUCUGUUUUAG AmGmCmUmAmGmAmAmAmUmAmGmCAAGUUA AAAUAAGGCUAGUCGUUAUCAmCmUmUm GmAmAmAmAmGmUmGmGmCmAmCmGm AmGmUmCmGmUmGmCmUmU*mU*mU*mU
G009845	AGCAACCUCACUCUUGUCUG GUUUUAGAGCUAGAAAAGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	198	mA*mG*mC*AACCUCACUCUUGUCUGGUUUUAG AmGmCmUmAmGmAmAmAmUmAmGmCAAGUUA AAAUAAGGCUAGUCGUUAUCAmCmUmUm GmAmAmAmAmGmUmGmGmCmAmCmGm AmGmUmCmGmUmGmCmUmU*mU*mU
G009846	ACCUCACUCUUGUCUGGGGA GUUUUAGAGCUAGAAAAGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	199	mA*mC*mC*UCACUCUUGUCUGGGGAGUUUU AGAmGmCmUmAmGmAmAmAmUmAmGmCAA GUUAAAUAAGGCUAGUCGUUAUCAmCmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCm CmGmAmGmUmCmGmUmGmCmUmU*mU*mU
G009847	CCUCACUCUUGUCUGGGGA GUUUUAGAGCUAGAAAAGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	200	mC*mC*mU*CACUCUUGUCUGGGGAAGUUUU GAmGmCmUmAmGmAmAmAmUmAmGmCAAGU UAAAUAAGGCUAGUCGUUAUCAmCmAmCm UmGmAmAmAmAmGmUmGmGmCmAmCm GmAmGmUmCmGmUmGmCmUmU*mU*mU
G009848	CUCACUCUUGUCUGGGGAAG GUUUUAGAGCUAGAAAAGC AAGUUAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	201	mC*mU*mC*ACUCUUGUCUGGGGAAGGUUUU AGAmGmCmUmAmGmAmAmAmUmAmGmCAA GUUAAAUAAGGCUAGUCGUUAUCAmCmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCm CmGmAmGmUmCmGmUmGmCmUmU*mU*mU

TABLE 8 -continued

Cyno sgRNA and modification patterns			
Guide ID	Full Sequence	SEQ ID NO: Full Sequence Modified	SEQ ID NO:
G009849	GGGGAAAGGGGAGAAAAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	202 mG*mG*GAAGGGGAGAAAAAAAAGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	235
G009850	GGGAAGGGGAGAAAAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	203 mG*mG*GAAGGGGAGAAAAAAAAGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	236
G009851	AUGCAUUUGUUUCAAAAUU GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	35 mA*mU*mG*CAUUGUUUCAAAAUAGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	67
G009852	UGCAUUUGUUUCAAAAUU GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	36 mU*mG*mC*AUUUGUUUCAAAAUAGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	68
G009853	UGAUUCCUACAGAAAAAGUC GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	206 mU*mG*mA*UUCCUACAGAAAAAGUCGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	239
G009854	UACAGAAAAAGUCAGGAUAA GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	207 mU*mA*mC*AGAAAAAGUCAGGAUAGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	240
G009855	UUUCUUUCUGCCUUUAAA GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	208 mU*mU*mU*CUCUCGCCUUUAAAACAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	241
G009856	UUUAAGUUUUAUUCAAC GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	209 mU*mU*mA*UAGUUUUAUUCAACAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	242
G009857	AUUUAUGAGAUCAACAGCAC GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	37 mA*mU*mU*UAUGAGAUCAACAGCACGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	69
G009858	GAUCAACAGCACAGGUUUUG GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	38 mG*mA*mU*CAACAGCACAGGUUUUGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	70
G009859	UUAAAUAAGCAUAGUGCAA GUUUUAGAGCUAGAAAUAGC AAGUUAACAUAGGCUGAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	39 mU*mU*mA*AAUAAAGCAUAGUGCAAGUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUGAGCCGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	71

TABLE 8 -continued

Cyno sgRNA and modification patterns			
Guide ID	Full Sequence	SEQ ID NO: Full Sequence Modified	SEQ ID NO:
G009860	UAAAAGCAUAGUGCAAUGGAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	40 mU*mA*mA*AGCAUAGUGCAAUGGAUGUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	72
G009861	UAGUGCAAUGGAUAGGUUU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	41 mU*mA*mG*UGCAAUGGAUAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	73
G009862	AGUGCAAUGGAUAGGUUU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	215 mA*mG*mU*GCAAUGGAUAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	248
G009863	UUACUUUGCACUUUCCUUAG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	216 mU*mU*mA*CUUUGCACUUUCCUUAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	249
G009864	UACUUUGCACUUUCCUUAG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	217 mU*mA*mC*UUUGCACUUUCCUUAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	250
G009865	UCUGACCUUUUAUUUUACCU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	218 mU*mC*mU*GACCUUUUAUUUUACCUAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	251
G009866	UACUAAAACUUUUAUUUUACU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	42 mU*mA*mC*UAAAACUUUUAUUUUACGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	74
G009867	AAAGUUGAACAAUAGAAAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	43 mA*mA*mA*GUUGAACAAUAGAAAAAGUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	75
G009868	AAUGCAUAAUCUAAAGUAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	44 mA*mA*mU*GCAUAAUCUAAAGUAAAAGUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	76
G009869	AUUAUCCUGACUUUUUCUGU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	222 mA*mU*mU*AUCUGACUUUUUCUGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	255
G009870	UGAAAAUUUCCUCUGUUUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUJGAAAAGU GGCACCGAGUCGGUGCUUUU	223 mU*mG*mA*AUUAUCCUCUGUUUAGUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU	256

TABLE 8 -continued

Cyno sgRNA and modification patterns								
Guide ID	Full Sequence	SEQ NO:	SEQ ID	SEQ NO:	SEQ ID	SEQ NO:	SEQ ID	SEQ NO:
Full Sequence	Modified							
G009871	UAUUUUUCUUUUGCCCACUA GUUUUAGAGCUAGAAAUGC AAGUAAAAAUAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	224	mU*mA*mA*UUUUCUUUUGCCCACUAGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUm GmAmAmAmAmGmUmGmGmCmAmCmCmGm AmGmUmCmGmGmUmGmCmU*mU*mU*mU		257			
G009872	AAAAGGUAGAGAAUUGUUUAG GUUUUAGAGCUAGAAAUGC AAGUAAAAAUAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	225	mA*mA*mA*AGGUAGAGAAUUGUUUAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU		258			
G009873	AACAUCUAGGUAAAAUAAA GUUUUAGAGCUAGAAAUGC AAGUAAAAAUAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	226	mA*mA*mC*AUCCUAGGUAAAAUAGGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU		259			
G009874	UAUUAAAUCAACAUCCU GUUUUAGAGCUAGAAAUGC AAGUAAAAAUAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	45	mU*mA*mA*UAUUUUCAACAUCCUGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU		77			
G009875	UUGUCAUGUAAAUCUAAAA GUUUUAGAGCUAGAAAUGC AAGUAAAAAUAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	228	mU*mU*mG*UCAUGUAAAUCUAAAUGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU		261			
G009876	UUUGUCAUGUAAAUCUAAAA GUUUUAGAGCUAGAAAUGC AAGUAAAAAUAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	229	mU*mU*mU*GUCAUGUAAAUCUAAAAGUUUUAG AmGmCmUmAmGmAmAmUmAmGmCAAGUUAA AAUAAGGCUAGUCGGUUAUCAmAmCmUmUmGm AmAmAmAmGmUmGmGmCmAmCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*mU		262			

TABLE 9

Vector Components and Sequences								
Plasmid ID	5' ITR	1 st orientation			2 nd orientation			3' ITR
		Splice Acceptor	Transgene	Poly-A	Poly-A	Transgene	Splice Acceptor	
P00147	(SEQ ID NO: 263)	Mouse Splice (R338L)-Acceptor (SEQ ID NO: 265) NO: 264)	Human Factor IX (SEQ ID NO: 266)	SEQ (R338L)	SEQ 267	Human Factor IX (R338L) (SEQ ID NO: 268) NO: 269)	Mouse Splice (R338L) (SEQ ID NO: 270)	(SEQ ID NO: 270)
P00411	(SEQ ID NO: 263)	Human Factor IX Splice (R338L)-Acceptor HiBit (SEQ ID NO: 271) NO: 272)	Human Factor IX (R338L)-Acceptor HiBit (SEQ ID NO: 273) NO: 274)	Human Factor IX (R338L)-Acceptor HiBit (SEQ ID NO: 274) NO: 275)	Human Factor IX Splice (R338L)-Acceptor HiBit (SEQ ID NO: 275) NO: 276)	Human Factor IX Splice (R338L)-Acceptor HiBit (SEQ ID NO: 276) NO: 277)	Human Factor IX Splice (R338L)-Acceptor HiBit (SEQ ID NO: 277) NO: 278)	(SEQ ID NO: 278)

TABLE 9-continued

Plasmid ID	5' ITR	Vector Components and Sequences							
		1 st orientation				2 nd orientation			
		Splice Acceptor	Transgene	Poly-A	Poly-A	Splice Acceptor	Transgene	Poly-A	3' ITR
P00415	(SEQ ID NO: 263)	Mouse Albumin Splice Acceptor (SEQ ID NO: 275) (SEQ ID NO: 264)	Nluc-P2A- GFP (SEQ ID NO: 266)	SEQ ID NO: 266	SEQ ID NO: 267	Nluc-P2A- GFP (SEQ ID NO: 276)	Mouse Albumin Splice Acceptor (SEQ ID NO: 270) (SEQ ID NO: 269)	(SEQ ID NO: 270)	
P00418	(SEQ ID NO: 263)	Mouse Albumin Splice Acceptor (SEQ ID NO: 272)	Human Factor IX (R338L)- HiBit (SEQ ID NO: 273)	SEQ ID NO: 266	SEQ ID NO: 267	Human Factor IX (R338L)- HiBit (SEQ ID NO: 273)	Mouse Albumin Splice Acceptor (SEQ ID NO: 270) (SEQ ID NO: 269)	(SEQ ID NO: 270)	

5' ITR Sequence (SEQ ID NO: 263) :

TTGGCCACTCCCTCTCGCGCCTCGCTCACTGAGGCCGGGACCAAAGGTC

GCCCGACGCCGGGCTTGCCGGCGGCCTCAGTGAGCGAGCGCGCAGAGA

GGGAGTGGCCAATCCATCACTAGGGGTTCT

Mouse Albumin Splice Acceptor (1st orientation) (SEQ ID NO: 264) :

TAGGTCAGTGAAGAGAAGAACAAAAGCAGCATATTACAGTTAGTGTCTTCATCA

ATCTTTAAATATGTTGTGGTTCTCTCCCTGTTCCACAG

Human Factor IX (R338L), 1st Orientation (SEQ ID NO: 265) :

TTTCTTGATCATGAAACGCCAACAAAATTCTGAATCGGCAAAGAGGTATAATTCA

GGTAAATGGAAAGAGTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAA

GTGTAGTTTGAGAAGCACGAGAAGTTTGAAAACACTGAAAGAACAACTGAAT

TTTGGAAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTAAATGGCG

GCAGTTGCAAGGATGACATTAATTCTATGAATGTTGGTGTCCCTTGGATTGAAAG

GAAAAGACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAG

TTTGTTAAAATAGTGTGATAACAAGGTGGTTGCTCCTGTACTGAGGGATATCGA

CTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTCCATGTGGAAGAGTT

TCTGTTTACAAACTCTAAGCTCACCGTGCTGAGACTGTTTCTGATGTGGACT

ATGTAATTCTACTGAAGCTGAAACCATTGGATAACATCACTCAAAGCACCAAT

CATTAAATGACTCACTGGGTGTTGGAGAAGATGCCAACCAGGTCAATTCC

CTTGGCAGGTTGTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTA

ATGAAAAATGGATTGTAACTGCTGCCACTGTGTTGAAACTGGTGTAAAATTACAG

TTGTCGAGGTGAAACATAATATTGAGGAGACAGAACATACAGAGCAAAGCGAAAT

GTGATTGCAATTATTCCCTCACCAACATACAATGCACTTAAAGTACAACCAT

GACATTGCCCTCTGGAACGGACGAACCCCTAGTGCTAACAGCTACGTTACACCT

ATTTGCATTGCTGACAAGGAATACACGAACATCTTCTCAATTGGATCTGGCTAT

GTAAGTGGCTGGGAAGAGTCTCCACAAAGGGAGATCAGCTTAGTTCTCAGTAC

CTTAGAGTCCACTGTTGACCGAGCCACATGTCCTATCTACAAAGTTACCATCT

ATAACAAACATGTTCTGTGCTGGCTTCCATGAAGGGAGGTAGAGATTGTCAGGAG

- continued

ATAGTGGGGACCCATGTTACTGAAGTGGAAAGGGACCAGTTCTTAACGTGGAAATTA

TTAGCTGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTA

TCCCGGTATGTCAACTGGATTAAGGAAAAACAAAGCTCACTTAA

Poly-A (1st orientation) (SEQ ID NO: 266) :

CCTCGACTGTGCCCTAGTGCAGCCATCTGTTGCCCCCTCCCCGTGCCCTC

CTTGACCCCTGGAAGGTGCCACTCCCACTGTCCTTCTAATAAAATGAGGAATTGC

ATCGCATTGTCAGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGCAGGACAG

CAAGGGGGAGGATTGGAAGACAATAGCAGGCATGCTGGGATGCGGTGGCTA

TGGCTTCTGAGGCGGAAAGAACAGCTGGGCTCTAGGGGTATCCCC

Poly-A (2nd orientation) (SEQ ID NO: 267) :

AAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGT

TGTTAACCTGTTATTGCAGCTTAAATGGTACAAATAAGCAATAGCATCACAAA

TTTCACAAATAAACGATTTTCACTGCATTCTAGTTGTTGTCCAAACTCATC

AATGTATCTTATCATGTCTG

Human Factor IX (R338L), 2nd Orientation (SEQ ID NO: 268) :

TTAGGTGAGCTTAGTCTTTCTTTATCCAATTCACTGAGCGAGACCTTCGTATAG

ATGCCATATTCCTTCATCGCACATTCCCTCCCCAACTTATTATCCGGTCAAGA

AACTTGTCTCGACTTCAGTGACGTGTGGTCCACCTGAATCACCTGGCATGAGTC

GCGACCGCCCTCGTAAACCCAGCACAAACATGTTATTGAAATCGTAAATTCTG

GGACAGAAGACAGGTGCGCTATCGACCAACGGGACGCGCAAATATTGAGACGA

GGCTGATCGACCTTGTGGAAGACCCGCCCCCACCACCTCACATATCCGCTCCAA

ATTTCAAGAAGATATTGTATATTCTTATCGGCTATACAAATCGGGTAACATAGG

AGTTAAGTACGAGTGGCTCGTCCAGCTCCAGGAGGCTATATCATGGTTACTGT

TTATAGCGCATTAAATTGTGATGGGTATGATCCTGATAACATTCTTTCTGTC

AGTATGCTCAGTTCTTCAATGTTGTTGCCAGCCACGACGTAATCTAACCCCC

GTCTCGACACAGTGTGCCGGCTTACAATCCACTTCAATTGACTATGGAGCCCCA

CAAAACCGCTGACTTTCCGTTGAGCACCACTGCCATGAAATTGCCAGGTTA

GCGTCCTCGCCCCGACAACCCTAGTAAAGTCATTAAATGACTGTGGATTGT

ATATTATCAAGAATCGTTCGCTTCAGTAGAGTTAACGTAGTCCACATCGGAAAA

ACTGTCTCGCCCTGTCAACTTGATGTCAGTGGACACACTTACCGACCGCACGG

AAGGGCACCGCCGGTCACAGCTCTTGATTCAGCGAGCCGGTAGCCCTCAGTG

CAACTACACACAACCTTGTTGCGCGAATTTCACAGAATTGCTCGCATCGTCCA

TTTTAATGTTGAGGTGACGTCAACTCGCAGTTTCTCCTCAAAACCAAAAGGG

CACCAACACTCGTAGGAATTATCGTCTTACAACCTCCCCCATTGAGACATGGA

TTAGATTGCACTGGTCCCCATCGACATATTGCTTCCAGAACACTGAGTGGTCCGTTCTG

- continued

TATTCTCAAACACCTCGCGCCTTCTCAAAAAGTCATTTCCCTCCATAACACTCTCG

CTCCAAGTCCCTTGACGAATTCTCAAGCTTCCTGAGTTACCTTTAGGCCGG

TTAAGTATCTTATTCGCGTTTCGTGGTCCAGAAA

Mouse Albumin Splice Acceptor (2nd orientation) (SEQ ID NO: 269) :
CTGTGGAACAGGGAGAAAAACCACAAACATATTAAAGATTGATGAAGACAA

CTAACTGTAATATGCTTTGTTCTTCACTGACCTA

3' ITR Sequence (SEQ ID NO: 270) :

AGGAACCCCTAGTGATGGAGTGGCACTCCCTCTGCGCGCTCGCTCACTG

AGGCCGCCGGCAAAGCCCGGGCGTCGGCGACCTTGGTCGCCCGCCTCAGTG

AGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Human Factor IX Splice Acceptor (1st Orientation) (SEQ ID NO: 271) :
GATTATTTGGATTTAAACAAAGACTTCTTAAGAGATGTAAATTTCATGATGTT

TTCTTTTTGCTAAAATAAAGAATTATTCTTTACATTCAG

Human Factor IX (R338L)-HiBit (1st Orientation) (SEQ ID NO: 272) :
TTCTTGATCATGAAAAGCCAACAAATTCTGAATCGGCAAAGAGGTATAATTCA

GGTAAATTGAAAGAGTTGTTCAAGGGAACCTTGAGAGAGAATGTATGAAAGAAAA

GTGTAGTTGAAAGAACGACGAGAAGTTTGAAAACACTGAAAGAACAACTGAAT

TTTGGAACAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTAAATGGCG

GCAGTTGCAAGGATGACATTAATTCTATGAATGTTGGTGTCCCTTGATTGAAAG

GAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAG

TTTGTAAAAATAGTGCTGATAACAAGGTGGTTGCTCCTGACTGAGGGATATCGA

CTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTCCATGTGAAAGAGTT

TCTGTTCACAAACTCTAACGCTCACCGTGCTGAGACTGTTTCTGATGTGGACT

ATGTAATTCTACTGAAGCTGAAACCATTGGATAACATCACTCAAAGCACCAAT

CATTTAATGACTTCACTCGGGTTGTTGGAGAAGATGCCAACCCAGGTCAATTCC

CTTGGCAGGTTGTTGAAATGGAAAGTTGATGCATTCTGTGAGGGCTATCGTTA

ATGAAAATGGATTGTAACTGCTGCCACTGTGTTGAAACTGGTGTAAAATTACAG

TTGTCGAGGTGAAACATAATTGAGGAGACAGAACATACAGAGCAAAGCGAAAT

GTGATTGAAATTATTCTCACACAAACTACAATGCAGCTATTAAATAAGTACAACCAT

GACATTGCCCTCTGAACTGGACGAACCCCTAGTGCTAACAGCTACGTTACACCT

ATTTGCATTGCTGACAAGGAATACACGAACATCTTCTCAAAATTGGATCTGGCTAT

GTAAGTGGCTGGGGAGAGTCTCCACAAAGGGAGATCAGCTTAGTTCTCAGTAC

- continued

CTTAGAGTTCCACTTGTGACCGAGGCCACATGTCCTATCTACAAAGTTACCATCT

ATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTATGTCAAGGAG

ATAGTGGGGACCCATGTTACTGAAGTGGAAAGGGACCAGTTCTTAAGTGGAAATTA

TTAGCTGGGTGAAGAGTGTGCAATGAAGGAAATATGGAATATATACCAAGGTC

TCCCGGTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTGTCAGCGGATGGAG

ACTGTTCAAGAAGATCAGCTAA

Human Factor IX (R338L) -HiBit (2nd Orientation) (SEQ ID NO: 273) :
TTAGGAAATCTCTTAAACAGCCGCCAGCGCTCACGGTGAGCTAGTCTTCTTT

ATCCAATTTCACCGTAGCGAGAGACCTTCGTATAGATGCCATATTCCTTCATCGCA

CATTCCCTCCCCAACTTATTATCCGGTCAAGAAACTGTTCTCGACTTCAGTGA

CGTGTGGTCCACCTGAATCACCTGGCATGAGTCGCACCGCCCTCGTGAAACCCAG

CACAAAACATGTTATTGTAATCGTAAATTCTGTGGACAGAAAGACAGGTGCGCTAT

CGACCAACGGGACGCGCAAATATTGAGAACGAGGGCTGATGACCTTGTTGGAAG

ACCCGCCCTCACACTCACATATCCGCTCCAAATTCAAGAAGATATTGTATAT

TCTTATCGGCTATAACAAATCGGGTAACATAGGAGTTAAGTACGAGTGGCTCGTCC

AGCTCCAGGAGGGCTATATCATGGTTGACTTGTATAGCGGCATTATAATTGTGA

TGGGTATGATCTGATAACATTCTTTCTGTTCACTGCTAGTATGCTCAGTTCTTCAATGT

TGTGTTGCCAGCCACGACCGTAATCTAACCCCCGTCGACACAGTGTGCGGCCG

TTACAATCCACTTTCTATTGACTATGGAGCCCCACAAAACCGTCGACTTTCCGTT

GAGCACACCTGCCATGGAAATTGGCAGGTTAGCGTCTCGCCCCGACAACCC

AGTAAAGTCATTAATGACTGTGTTGTTATATTCAAGAATCGTTCGC

TTCAGTAGAGTTAACGTAGTCCACATGGGAAAAACTGTCTGGCCCTGTCAACTT

TGATGTCTGGCACACTTACCCGACCGCACGGGAAGGGCACCGCGTTACAGC

TCTTTGATTCTCAGCGAGCGGTAGCCCTAGTGCACACTACACAACACTTGTGTC

GGCGGAATTTACAGAATTGCTCGCATCGTCATTAAATGTTGAGGTGACGTCC

AACTCGCAGTTTCCTTCAAAACAAAAGGGCACCAACACTCGTAGGAATTATA

TCGTCTTACAACCTCCCCCATTCAAGACATGGATTAGATTGCACTGGTCCCCATCGA

CATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAACACCTCGCGCTTC

TTCAAAACTGCATTTCCTCCATACACTCTCGCTCCAAGTCCCTGCACGAATTCT

TCAAGCTTCCTGAGTTACCTTTAGGCCGGTTAAGTATCTTATTGCGTTTCTG

GGTCCAGAAA

Human Factor IX Splice Acceptor (2nd Orientation) (SEQ ID NO: 274) :
CTGAAATGTAAGAATAATTCTTTAGTTAGCAAAAAAGAAAACATCATGAAAA

TTTACATCTCTTAAGAAAGTCTTGTGTTAATCCAAATAATC

Nluc-P2A-GFP (1st Orientation) (SEQ ID NO: 275) :

TTCTTGATCATGAAACGCCAACAAATTCTGAATCGGCCAAGAGGTATAATTCA

GGTAAATTGGAAGAGTTGTTCAAGGGAACCTTGAGAGAGAAATGTATGGAAGAAA

GTGTAGTTTGAGAAGAACGAGTATTCACTTGGAGGACTTGTGGTACTGGAGGCA

AACCGCTGGTTATAATCTCGACCAAGTACTGGAACAGGGCGGGTAAGTCCCTCTT

TCAGAATTGGGTGTAAGCGTCACACCAATCCAGCGATTGTTGCTGGAGAGAA

- continued

```

CGGACTAAAATTGACATCCATGTTATCATTCCATATGAAGGTCTCAGTGGAGACCA
AATGGGGCAGATCGAGAAGATTTCAAGGTAGTTACCCAGTCGACGATCACCACTT
CAAAGTCATTCTCCACTATGGCACACTTGTATCGACGGAGTAACCTAATATGAT
TGATTACTTGGTCGCCGTATGAGGGCATCCAGTGTGATGGCAAAAGATCAC
CGTAACAGGAACGTTGTGGAATGGGAACAAGATAATCGACGAGAGATTGATAAAC
CAGACGGGTCACTCCTGTTCAAGGTTACAATTAAACGGCGTACAGGATGGAGACTCT
GTGAACGAATACTGGCCACAAATTTCACTCCTGAAGCAGGCCGGAGACGTGGAG
GAAAACCCAGGGCCCGT GAGCAAGGGCGAGGAGCTGTTCACCGGGTGGTGCCTCAT
CCTGGTCAGCTGGACGGCGACGTAAACGGCCACAAGTTCAAGCGTGTCCGGGAGG
GCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACCGC
AAGCTGCCGTGCCCTGGCCACCCCTCGT GACCACCCCTGACCTACGGCGTGCAGTGC
TTAGGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTCAAGTCCGCATGCC
GAAGGCTACGTCCAGGAGCGCACCATCTTCTCAAGGACGACGGCAACTACAAGAC
CCCGGCCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGG
GCATCGACTCAAGGAGGACGGCAACATCCTGGGGCACAGCTGGAGTACAACACTAC
AACAGCCACAACGTCTATATCATGGCGACAAGCAGAAGAACGGCATCAAGGTGAA
CTTCAGATCCGCCACAACATCGAGGACGGCGACGCGTGCAGCTGCCGACCACCTACC
AGCAGAACACCCCCATGGCGACGGCCCGTGTGCTGCCGACAACCACACTACCTG
AGCACCCCAGTCGCCCTGAGCAAAGACCCCAAGAGAACGGCGATCACATGGTCCT
GCTGGAGTCGTGACGCCGCCGGGATCACTCTGGCATGGACGAGCTGTACAAGG
GAGGAGGAAGCCCGAAGAAGAAGAGAAAGGTCTAA

```

Nluc-P2A-GFP (2nd Orientation) (SEQ ID NO: 276) :

```

TTACACCTTCCTCTTCTTCTGGGCTGCCCGCCCTGTACAGCTCGCCATGCC
AGGGTGATGCCGGCGCGGTCAAGAACTCCAGCAGCACCATGTGGTCCCTCTCG
TTGGGTCTTGCTCAGGGCGCTCTGGTGCTCAGGTAGTGGTGTGCGGAGCAGC
ACGGGGCGTCGCCGATGGGGTGTCTGCTGGTAGTGGTGTGCGCAGCTGCACGCTG
CCGTCCTCGATGTTGTGCTGATCTTGAAGTTCACCTGATGCCGTTCTGCTTGT
CGGCATGATGTACAGTTGTGCTGTTGAGTGTACTCCAGCTGTGGCCAGGA
TGTGCGTCCTCTTGAAGTCGATGCCCTCAGCTCGATCTGTCACCAGGGTGTG
GCCCTCGAACCTCACCTGGCCCTGGCTTGTAGTTGCCGTCGCTTGAAGAACGAT
GGTCCTCTCCTGCCACGTAGGCCCTGGGATGGCGCTCTTGAAGAACGCTGTGCTG
CATGTGGTCGGGTACCTGCTGAAGCACTGCACGCCGTAGGTAGGGTGGTCACCA
GGGTGGGCCAGGGCACGGGCAGCTGCCGGTGGTGCAGATGAACCTCAGGGTCA
TGCCGTTAGGGCGTCGCCCTGCCCTGCCCTGCCGTCACGCTGAACCTGTGGCGT
ACGTCGCCGTCAGCTCCACCAAGGATGGGCACGCCGGTGAACAGCTCCCGCC
CTTGCCTCACGGGCCGGGTTCTCCTCCACGTCGCCGGCTGCTTCAAGCAGGTGAA
GTTGGTGGCCAGGATCCTCTCGCACGCCCTGCCGTCAGGCCGTTGATGGTCAC
CCTGAAACAGCAGGCTGCCGTGGGTTGATCAGCCTCTCGTCGATGATCTGTTGCC
GTTCCACAGGGTGCCGGTCACGGTGTCTTCTGCCGTCGAACACGGCGATGCCCTC
GTAGGGCCTGCCGAAGTAGTCGATCATGTTGGGGTCAGGCCGTCGATCACCAGG

```

- continued

TGCCGTAGTGCAGGATCACCTGAAGTGGTGTGTCGACGGGGTACACCACCTGA
AAATCTTCTCGATCTGGCCCATCTGGTCGCCGCTCAGGCCCTCGTAGGGATGATCA
CGTGGATGTCGATCTCAGGCCGTTCTGCCGCTCAGCACGATCCTCTGGATGGGG
TCACGCTCACGCCAGGTTCTGGAACAGGCTGCTCACGCCGCTCAGCACCT
GGTCCAGGTTGAGCCGGCTGCCTCCAGTCGCCACGAAGTCCTCCAGGGTGA
ACACGGCCTCTCGAAGCTGCACCTCTCCATGCACTCCCTCCAGGTTGCCCTG
CACGAACCTCCAGCTTGCCGTGTTGACCTCTGGCCTGTTAGGATCTGTTG
GCGTTCTCGTGGTCCAGGAA

P00147 full sequence (from ITR to ITR) :

(SEQ ID NO: 277)

TGGGCCACTCCCTCTCGCGCCTCGCTCACTGAGGCCGGCAGCAAAGGTC
GCCCGACGCCGGCTTGCCGGCGGCCTCAGTGAGCGAGCGCGCAGAGA
GGGAGTGCCAACATCCATCACTAGGGTTCTAGATCTCTTAGGTCAGTGAAGAGA
AGAACAAAAGCAGCATATTACAGTTAGTTGTCATCAATTTAAATATGTTGT
GTGGTTTTCTCCCTGTTTACAGTTTCTTGATCATGAAAACGCCAACAAAAT
TCTGAATCGGCCAACAGGTATAATTCAAGGTAATTGGAAGAGTTGTCAGAGA
ACCTTGAGAGAGAATGTATGGAAGAAAAGTAGTTGAAAGAACGACGAGAAGTT
TTGAAAACACTGAAAGAACACTGAATTGGAAGCAGTAGTTGATGGAGATCA
GTGTGAGTCCAATCCATGTTAAATGGCGCAGTTGCAAGGATGACATTAATTCTA
TGAATGTTGGTGCCTTGGATTGAAAGGAAAGACTGTGAATTAGATGTAACATG
TAACATTAAGAATGGCAGATGCGAGCAGTTGTAAGGATAGTGTGATAACAAGG
TGGTTGCTCTGACTGAGGGATATCGACTTGCAGAAAACCAGAACGCTGTGAAAC
CACCACTGCCATTCCATGTGAAGAGTTCTGTTCACAAACTCTAACGCTCACCC
GTGCTGAGACTGTTTCTGTGACTATGTAATTCTACTGAAGCTAACCA
TTTGATAACACTCAAAGCACCAATCTTAATGACTCAGTTGCTGGTTGTTG
GTGGAGAAGATGCCAACCAGGTCATCCCTGGCAGGTTGTTGAATGGTAAAG
TTGATGCATTCTGGAGGCTATCGTAATGAAAAATGGATTGTAAGTGTGCCCC
ACTGTGTTGAAACTGGTTAAATTACAGTTGTCGAGGTAAACATAATTGAGG
AGACAGAACATACAGAGCAAAAGCGAAATGTGATTGAAATTCTCACCAAC
TACAATCGAGCTATTAAGTACAACCATGACATTGCCCTCTGGAACGGACGAA
CCCTTAGTGCCTAACAGCTACGTTACACCTATTGCAATTGCTGACAAGGAATACAG
AACATCTCTCAAATTGGATCTGGCTATGTAAGTGGCTGGGAAGAGTCTCCAC
AAAGGGAGATCAGCTTAGTTCTCAGTACCTAGAGTCCACTTGTGACCGAGCC
ACATGTCTTCTACAAAGTCACCATCTATAACAACATGTTCTGTGCTGGCTTCC
ATGAAGGAGGATGAGATTGTCAGGAGATAGTGGGGACCCATGTTACTGAA
GTGGAAAGGGACCAAGTTCTTAAGTGGAAATTAGCTGGGTGAAGAGTGTGCAAT
GAAAGGCAAATATGGAATATACCAAGGTATCCCGGTATGTCACGGATTAAGG
AAAAAACAAAGCTACTTAACCTCGACTGTCGCTTCTAGTTGCCAGCCATCTGTTGT
TTGCCCTCCCCGTGCCTTCTTGACCCCTGGAAGGGTGCACCTCCACTGTCCTTCC
TAATAAAATGAGGAAATTGCATCGCATTGTCAGTAGGTGTCATTCTATTCTGGGG

- continued

GGTGGGGTGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATG
 CTGGGGATGCGGTGGCTCATGGCTCTGAGGCGAAAGAACCGACTGGGCTCT
 AGGGGGTATCCCCAAAAACCTCCACACCTCCCCTGAACCTGAAACATAAAATG
 AATGCAATTGTTGTTAACTGTTATTGCAGCTTATAATGGTTACAAATAAGCA
 ATAGCATCACAAATTTCACAATAAAGCATTTTCACTGCATTCTAGTTGTGGTT
 GTCCAAAATCTCATCAATGTATCTTATCATGTCGTAGGTGAGCTTAGTCTTTCTTT
 ATCCAATTCACGTAGCGAGAGACCTCGTATAGATGCCATATTCCCTTCATCGCA
 CATTCCCTCCCCAACTTATTATCCCGTCAAGAAACTGTTCCCGACTTCAGTGA
 CGTGTGGTCCACCTGAATCACCTGGCATGAGTCGCACCGCCCTCGTGAACCCAG
 CACAAAACATGTTATTGTAATCGTAAATTCTGTGGACAGAACAGGTCGCTCTAT
 CGACCAACGGGACGCGAAATATTGCAGAACGAGGCTGATCGACCTTGTTGGAAG
 ACCCGCCCCCACCACACATATCCGCTCCAAATTCAAGAAGATATTGTATAT
 TCTTTATCGGCTATAAAATCGGGTAACATAGGAGTTAAGTACGAGTGGCTCGTCC
 AGCTCCAGGAGGGCTATATCATGGTTGACTTGTATAGCGGCATTATAATTGTGA
 TGGGGTATGATCCTGATAAACATTCTTTCTGTCAGTATGCTCAGTTCTCAATGT
 TGTGTTGCCAGCCACGACCGTAATCTAACCCCCGTCTGACACAGTGTGCGGCC
 TTACAATCCACTTTCACTGACTATGGAGCCCCACAAAACCGTCGACTTCCGTT
 GAGCACCACCTGCCATGAAATTGCCAGGTTAGCGTCCTGCCCGACAACCC
 AGTAAAGTCATTAATGACTGTGTTGATTGTATATTCAAGAACGTTGCGC
 TTCAGTAGAGTTAACGTAGTCCACATGGGAAAAACTGTCGGCCCTGTCAACTT
 TGATGTCGGCACACTTACCGACCGCACGGGAGGGCACCGCCGTTACAGC
 TCTTTGATTCTCAGCGAGCCCGTAGCCCTCACTGCAACTACACACAATTGTTGTC
 GGCGGAATTTCACAGAATTGCTCGCATCGTCATTGTTAATGTTGCAAGGTGACG
 AACTCGCAGTTTCCTCAAACACAAAAGGGCACCAACACTCGTAGGAATTATA
 TCGTCTTACAACCCCCCATTCAAGACATGGATTAGATTCGATTGGTCCCCATCGA
 CATATTGCTTCAGAACACTGAGTGGCCGTTCTGTATTCTCAAAACACCTCGCGC
 TTCAAAAACGCTTTCCATACACTCTCGCTCCAAGTCCCTGCAAGAATTCT
 TCAAGCTTCCTGAGTTACCTTTAGGCCGTTAAGTATCTTATTGCGTTTC
 GGTCCAGAAAATGTTGAAACAGGGAGAGAAAACCACACAAATTTAAAGA
 TTGATGAAGAACAACTAATGCTGCTTTGTTCTCTTCACTGACCTAA
 GAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTGCGCGCTCG
 CTCACTGAGGCCGCCGGCAAAGCCGGCGTGGCGACCTTGTCGCCCGC
 CTCAGTGAGCGAGCGAGCGCAGAGAGGGAGTGGCCAA
 P00411 full sequence (from ITR to ITR) :
 (SEQ ID NO: 278)
 TTGGCCACTCCCTCTGCGCGCTCGCTCACTGAGGCCGGGACCAAAGGTC
 GCCCGACGCCGGCTTGCCGGCGGCCTCAGTGAGCGAGCGCAGAGA
 GGGAGTGGCCAATCCATCACTAGGGGTTCTAGATCTCTGATTATTGATTAAA
 ACAAAAGACTTCTTAAGAGATGAAAATTTCATGATGTTCTTTGCTAAA
 AAAGAATTATTCTTTACATTCACTGAGTTCTGATCATGAAAACGCCAACAAATTC

- continued

TGAATCGCCAAAGAGGTATAATTCAAGTAAATTGGAAGAGTTGTCAGGGAAC
CTTGAGGAGAAATGTATGGAAGAAAAGTGTAGTTTGAGAAGCAGAGAAGTT
TGAAAACACTGAAAGAACAACTGAATTGGAGCAGTATGTTGATGGAGATCAGT
GTGAGTCCAATCATGTTAATGGCGCAGTTGCAAGGGATGACATTAATTCTATG
AATGTTGGTGTCCCTTGGATTGAGAAGAAAGACTGTGAATTAGATGTAACATGTA
ACATTAAGAATGGCAGATGCGAGCAGTTGTAAGGAAATAGTGTGATAACAAGGTG
GTTTGCTCCTGACTGAGGGATATCGACTTGAGAAAACCAGAAGTCCTGTGAACCA
GCAGTGCCATTCCATGTGGAAGAGTTCTGTTCACAAACTCTAACGTCACCCGT
GCTGAGACTGTTTCTGATGTTGACTATGTAATTCTACTGAAAGCTGAAACCA
TGGATAACATCACTAAAGCACCCAACTCATTTAATGACTTCACTCGGTTGTTG
GAGAAGATGCCAACCCAGGTCAATTCCCTGGCAGGTTGTTGAATGGTAAAGTTG
ATGCATTCTGTGGAGGCTCATCGTTAATGAAAAATGGATTGTAAGTGTGCTGCCACT
GTGTTGAAACTGGTGTAAAATTACAGTTGTCGCAGGTGAAACATAATTGAGGAGA
CAGAACATACAGAGCAAAGCGAAATGTGATTGAAATTCTCACCACAACTAC
AATGCAGCTATTAAATAAGTACAACCATGACATTGCCCTCTGGAACGGAAACCGAAC
TTAGTGCTAACAGCTACGTTACACCTATTGCTTGACAGGAATACACGAAC
ATCTCCCAAAATTGGATCTGGCTATGTAAGTGGCTGGGAAGAGTCTTCACAA
GGGAGATCAGCTTAGTTCTCAGTACCTAGAGTTCCACTGTTGACCGAGCCACA
TGTCTTATCTACAAAGTTCACCATCTATAACACATGTTCTGTGCTGGCTTCCATG
AAGGAGGTAGAGATTCAAGGAGATAGTGGGGACCCATGTTACTGAAGTG
GAAGGGACCAGTTCTTAACGGAATTATTAGCTGGGTGAAGAGTGTGCAATGAA
AGGCAAATATGAAATATACCAAGGTCTCCCGTATGTCACGGATTAAAGGAA
AAACAAAGCTCACTGTCAGCGGATGGGAGACTGTTCAAGAAGATCAGCTAACCTCGA
CTGTGCCTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCGTGCCTTCTTGAC
CCTGGAAAGGTGCCACTCCCACGTCCCTTCTAATAAAATGGGAATTGCAATCGCA
TTGCTGAGTAGGTGTCATTCTATTCTGGGGGTGGGTGGGCAGGACAGCAAGG
GGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATGCGGTGGCTATGGCT
TCTGAGGCGGAAAGAACAGCTGGGCTCTAGGGGTATCCCCAAAAACCTCCA
CACCTCCCCGTAACACTGAAACATAAAATGAATGCAATTGTTGTTAACTGTT
ATTGCACTTAAATGGTTACAAATAAGCAATAGCATCACAAATTTCACAAATAA
GCATTTTTCACTGCATTCTAGTTGTTGTCACAAACTCATCAATGTTACTTATC
ATGTCGTTAGGAAATCTCTTAAACAGCCGCCAGCGCTCACGGTGAGCTAGTCT
TTCTTTTATCCAATTACGTAGCGAGAGACCTTGTATAGATGCCATTTCCTT
CATCGCACATTCTCCCCCAACTTATTATCCGGTCAAGAAACTTGTCTCGACT
TCAGTGACGTGTTGGTCCACCTGAATCACCTGGCATGAGTCGCGACCGCCCTCGTGA
AACCCAGCACAAACATGTTATTGTAATGTAATTCTGTGGACAGAACAGAGT
CGCTCTATCGACCAACGGGACCGCAGCGAAATATTGCAAGAACGAGGGCTGATCGACCTT
TGTGGAAGACCCGCCCCACCCACTCACATATCCGCTCCAAATTCAAGAAGATAT
TTGTATATTCTTATCGGCTATACAAATCGGGTAACATAGGAGTTAAGTACGAGTG

- continued

```

GCTCGTCCAGCTCAGGAGGGTATATCATGGTGACTTGTATAGCGCATTAT
AATTGTGATGGGGTATGATCCTGATAACATTCTTTCTGTTAGTGTCACTTTC
TTCAATGTTGTTGCCAGCACGACCGTAATCTAACCCCCGTCGACACAGTG
TGCAGCCGTTACAATCCACTTTCAATTGACTATGGAGCCCCACAAAACGCGTCGAC
TTTCGTTGAGCACCACCTGCCATGGAAATTGGCCAGGTTAGCGTCCTGCC
GACAACCCTAGTAAAGTCATTAATGACTGTGGAATTGTGTATATTCAAGAAT
CGTTTGGCTTCAGTAGAGTTAACGTAGTCCACATGGGAAAAACTGTCTGCC
TGTCAACTTGATGTCGGACACACTAACCGACCGCACGGGAAGGGACCGCC
GTTCACAGCTCTTGATTCAGCGAGGCCAGCCAGTGCACACTACACAA
CTTGTGTCGGCGAATTTACAGAATTGTCGCATGTCATTTAATGTTGCA
GGTGAACGCCACTCGCAGTTTCCAAACAAAAGGGCACCAACACTCGTA
GGAATTATATCGTCTTACAACCTCCCCCATTCAAGACATGGATTAGATTGCAATTGG
TCCCCATCGACATATTGCTCCAGAACTCAGGGTCCGTTCTGTATTCTCAAACACCT
CGCGCGCTTCTCAAAACTGCATTTCCTCCATACACTCTCGCTCCAAGTCCC
CAGAATTCTCAAGCTTCTGAGTTACCTTTAGGCCGGTTAAGTATCTTATTC
GCGTTTCTGTCAGAAAAGTAAATGAAAAGAATAATTCTTAGTTAGCA
AAAAAGAAAACATCATGAAATTACATCTTAAAGAAAGCTTTAGTTAATCC
AAATAATCAGAGATCTAGGAACCCCTAGTGATGGAGTTGCCACTCCCTCTGCGC
GCTCGCTCGCTCACTGAGGCCGCCGGCAAGCCGGCGTCGGCGACCTTGG
TCGCCCGGCTCAGTGAGCGAGCGAGCGCAGAGAGGGAGTGGCAA

```

P00415 full sequence (from ITR to ITR) :

(SEQ ID NO: 279)

```

TTGCCACTCCCTCTCGCGCCTCGCTCGCTCACTGAGGCCGGCACAAAGTC
GCCGACGCCGGCTTGCCTGGCGCCTAGTGAGCGAGCGCAGAGA
GGGAGTGGCCAACCCATCACTAGGGTTCTAGATCTCTAGTCAGTGAAGAGA
AAACAAAGCAGCATATTACAGTTAGTTGTCATCAATTAAATATGTTG
GTGGTTTCTCCCTGTTCCACAGTTCTGATCATGAAACGCCAACAAAT
TCTGAATGCCAAAGAGGTATAATTAGGAAATTGGAAGAGTTGTCAGAGGA
ACCTTGAGAGAGATGTATGGAAGAAAAGTGTAGTTGAAGAAGCAGTATTCA
TTGGAGGACTTGTGCGTGACTGGAGGCAACCGCTGGTTATAATCTGACCAAGTA
CTGGAACAGGGGGGGTAAGTCCCTCTTCAGAATTGGGTAAAGCGTCACACCA
ATCCAGCGATTGTGTTCTGGAGAGAACGGACTCAAATTGACATCCATGTT
ATTCCATATGAAGGTCTCAGTGGAGACCAATGGGCAGATCGAGAAGATTTCAA
GGTAGTTACCCAGTCGACGATCACCACCTCAAAGTCATTCTCCACTATGGCACACT
TGTTATCGACGGAGTAACCTAAATGATTGATTACTTGGCGCCGTATGAGGG
CATCGCAGTGGATGGCAAAAGATCACCGTAACAGGAACGTTGTGGAATGGGA
ACAAGATAATCGACGAGAGATTGATAAATCCAGACGGTCACCTGTCAGGGTT
ACAATTAACGGCGTACAGGATGGAGACTCTGTGAACGAATACTGGCCACAAATT
TTCACTCTGAAGCAGGCCGGAGACGTGGAGGAAACCCAGGGCCCGTGAGCAAGG
GCGAGGAGCTGTTCACCGGGGTGGTGCCTGAGCTGGACGGCGACGTA

```

- continued

AACGGCCACAAGTTCA CGTGTCCGGCGAGGGCGATGCCACCTACGGCAA
GCTGACCCCTGAAGTCATCTGCACCACCGCAAGCTGCCGTGCCCTGCCAACCT
CGTGACCACCTGACCTACGGCGTGCAGTGCTCAGCGTACCCGACCACATGAA
GCAGCACGACTTCTCAAGTCCGCCATGCCGAAGGCTACGTCAGGAGCGCACCA
TCTTCTCAAGGACGACGGCAACTACAAGACCCGCGCGAGGTGAAGTTCGAGGG
GACACCCCTGGTGAAACCGCATCGAGCTGAAGGGCATCGACTCAAGGAGGACGGCAA
CATCCTGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCATGG
CCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAG
GACGGCGCGTGCAGCTGCCGACCACTACCGCAGAACACCCCCATGGCGACGG
CCCCGTGCTGCTGCCGACAACCAACTACCTGAGCACCCAGTCGCCCTGAGCAAAG
ACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCTGTGACCGCCGCCGG
ATCACTCTCGGCATGGACGAGCTGTACAAGGGAGGAGGAAGCCCGAAGAAGAAGA
GAAAGGTCTAACCTCGACTGTGCTTAGTTGCCAGCCATCTGTTGTTGCCCTCC
CCCGTGCTCTCTTGACCTGGAAGGTGCCACTCCACTGTCTTTCTAATAAAATG
AGGAAATTGCATCGCATTGTCAGTAGGTGTCATTCTATTCTGGGGGTGGGTGG
GGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATGC
GGTGGGCTCTATGGCTCTGAGGCGGAAAGAACCCAGCTGGGCTCTAGGGGTATC
CCCCAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGT
TGTTGTTAACATTGTTATTGCGAGCTTAAATGGTTACAAATAAGCAATAGCATCAC
AAATTTCACAAATAAGCATTTCACTGCATTCTAGTTGGTTGTCCAACACTC
ATCAATGTATCTTATCATGTCGTTACACCTTCTCTTCTGGGGCTGCCGCC
CCTTGACAGCTCGTCCATGCCAGGGTGATGCCGCCGCCGTACGAACCTCAGCA
GCACCATGTGGCCCTCTCTGCTGGGTCTCAGGGCGCTCTGGGTGCTCAG
GTAGTGGTTGTCGGGCAGCAGCACGGGCCGTCGCCGATGGGGTGTCTGCTGGT
AGTGGTCGGCCAGCTGCACGCTGCCGTCTCGATGTTGCTGATCTGAAGTTCA
CCTTGATGCCGTTCTCTGCTTGCGCATGATGTACACGTTGGCTGTTGTTAGTT
GTACTCCAGCTGTGGCCAGGATGTTGCCGTCTCTGTAAGTCGATGCCCTCAG
CTCGATCTGTTACCAGGGTGTGCCCTCGAACCTCACCTGGCCCTGGTCTGTAG
TTGCCGTCGTCCTTGAGAAGATGGCTCTCTGACGTAAGCCCTGGGATGGCG
CTCTTGAGAAGATGTCGTCGTTCATGTTGCTGGGTACCTGCTGAAGCACTGCACG
CCGTAAGGTGCGTCAACAGGGTGGGCCAGGGCACGGGAGCTTGGCTCGGGTGGT
GCAGATGAACCTCAGGGTCAGCTTGCGTAGGTGGCGTCGCCCTGCCCTGCC
CACGCTGAACATTGTCGGCGTTCACGTCGCCGTCCAGCTCCACCAAGGATGGCAC
GCCGGTGAACAGCTCCTGCCCTGCTCACGGGCCGGGTTCTCTGCCACGTCGCC
GCCCTGCTTCAGCAGGCTGAAGTTGGTGGCCAGGATCTCTCGCACAGCCTCAGCC
GGTCACGCCGTTGATGGTCACCTGAACAGCAGGCTGCCGTGGGTTGATCAGCCT
CTCGTCGATGATCTTGTGCCGTTCCACAGGGTGGCGTCACGGTGATCTTGTGCC
TCGAACACGGCGATGCCCTCGTAGGGCTGCCGAAGTAGTCGATCATGTTGGGGTC
ACGCCGTCGATCACCAAGGGTGGCGTAGTGCAGGATCACCTGAAGTGGTGGTC

- continued

ACGGGGTACACCACCTGAAAATCTTCGACTGCCCATCTGGTCGCCGCTCAGG
CCCTCGTAGGGATGATCACGTGGATGTCGATCTTCAGGCGTTCTCGCCGCTCAGC
ACGATCCTCTGGATGGGGTACGCTCACGCCAGGTTCTGGAACAGGCTGCTCAGC
CCGCCTGCTCCAGCACCTGGCCAGGTTGAGCCGGCGGTCTGCCTCAGTCGCC
ACGAAGTCCTCAGGGTAACACGCCCTCCTCGAAGCTGCACCTCTCCATGCAC
TCCCTCTCCAGGTTGCCCTGCAGAACCTCCAGCTGCCGTGTTGACCTCTTGG
GCCTGTTCAAGGATCTTGGTGGCGTTCTCGTGGTCCAGGAA

P00418 full sequence (from ITR to ITR) :

(SEQ ID NO: 280)

TTGGCCACTCCCTCTCGCGCCCTCGCTCACTGAGGCCGGCGACCAAAGGTC
GCCCGACGCCGGCTTGCCGGCGGCCTCAGTGAGCGAGCGCAGCAGA
GGGAGTGGCCAACTCCATCACTAGGGTTCTAGATCTCTTAGGTCAAGAGA
AGAACAAAAAGCAGCATATTACAGTTAGTTGTCATCAATTTAAATATGTTGT
GTGGTTTTCTCCCTGTTCCACAGTTTCTTGATCATGAAACGCCAACAAAT
TCTGAATCGGCCAAAGGGTATAATTCAAGGTAATTGGAAGAGTTGTTCAAGGGA
ACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTGAAAGAACGACGAGAAGTT
TTGAAAACACTGAAAGAACAACTGAATTGGAAAGCAGTATGTTGATGGAGATCA
GTGTGAGTCCAATCCATGTTAAATGGCGGCAGTTGCAAGGATGACATTAAATTCTA
TGAATGTTGGTGTCCCTTGGATTGAAAGGAAAGAAACTGTGAATTAGATGTAACATG
TAACATTAAGAATGGCAGATGCGAGCAGTTGTAAGGAAACTAGTGTGATAAACAGG
TGGTTTGCTCCTGACTGAGGGATATCGACTTCGAGAAAACCAGAACGTCCTGTGAA
CAGCAGTGCCATTCCATGTGGAAGAGTTCTGTTCACAAACTTCTAACGTCACCC
GTGCTGAGACTGTTTCTGATGTGACTATGTAATTCTACTGAAGCTGAAACCA
TTTGGAATAACACTCAAAGCACCCAACTATTAATGACTCAGTCACTCGGGTGTG
GTGGAGAAGATGCCAACCCAGGTCAATTCCCTGGCAGGTTGTTGAATGGTAAAG
TTGATGCATTCTGTGGAGGCTATCGTAATGAAAATGGATTGTAACTGCTGCC
ACTGTGTTGAAACTGGTGTAAAATTACAGTTGCGCAGGTAACTGAACTATTGAGG
AGACAGAACATACAGAGCAAAGCGAAATGTGATTGAAATTCTCACCACAA
TACAATGCAGCTATTAAAGTACAACCATGACATTGCCCTCTGGAACGGACGAA
CCCTTAGTGCTAACAGCTACGTTACACCTATTGCTTGCTGACAAGGAATACAG
AACATCTCCTCAAATTGGATCTGGCTATGTAAGTGGCTGGGAAGAGTCTTCCAC
AAAGGGAGATCAGCTTAGTTCTCAGTACCTAGAGTTCCACTTGTGACCGAGCC
ACATGTCTTCTATCTACAAAGTCACCATCTATAACACATGTTCTGTGCTGGCTTCC
ATGAAGGAGGTAGAGGATTGATGTCAAGGAGATAGTGGGGACCCATGTTACTGAA
GTGGAAGGGACCAAGTTCTTAACGAAATTAGCTGGGGTGAAGAGTGTGCAAT
GAAAGGCAAATATGGAATATACCAAGGTCTCCCGGTATGTCAACTGGATTAAAGG
AAAAAACAAAGCTACTGTCAGCGGATGGAGACTGTTCAAGAACGATCAGCTAACCT
CGACTGTGCCCTAGTTGCCAGGCCATCTGTTGCCCCCTCCCCGTGCTTCC
GACCCCTGGAAGGTGCCACTCCACTGTCCTTCTAATAAAATGAGGAAATTGCA
GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCA

- continued

AGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGCTCATG
GCTTCTGAGGCCAAGAACAGCTGGGCTCTAGGGGTATCCCCAAAAACCTC
CCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTAACTTG
TTTATTGCAGCTTATAATGGTACAAAATAAGCAATAGCATCACAAATTCACAAAT
AAAGCATTTTTCACTGCATTCTAGTTGTTGTCACACTCATCAATGTATCTT
ATCATGTCTGTTAGGAAATCTCTAAACAGCGCAGCCGCTCACGGTGAGCTTAG
TCTTTCTTTATCCAATTACGTAGCGAGAGACCTCGTATAGATGCCATATTC
CTTCATCGCACATTCCCTCCCCAACTTATTATCCCGTCAAGAAACTGTTCC
ACTTCAGTGACCTGTTGGTCCACCTGAATCACCTGGCATGAGTCGACCC
TGAAACCCAGCACAAAATGTTATTGTAATCGTAAATTCTGTGGACAGAAAGACA
GGTCGCTCTATCGACCAACGGACGCGCAAATATTGAGACAGGAGCTGATCGAC
CTTGTGGAAGACCCGCCCCACCCACTCACATATCCGCTCCAAATTCAAGAAGA
TATTGTATATTCTTATCGGTATACAAATGGGTAACATAGGAGTTAAGTACGA
GTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTTAGCGGAT
TATAATTGTGATGGGTATGATCCTGATAACATTCTTTCTGTTAGTGT
TTCTCAATGTTGTTGTCGCCAGCCACGACCGTAATCTAACCCCCGTC
GTGTGCGCCGTTACAATCCACTTTCTGACTATGGAGCCCCACAAAC
GACTTTCCGTTGAGCACCACCTGCCATGGAAATTGGCAGGTTAGCG
CCCGACAACCCCTAGTAAAGTCATTAATGACTGTGGATTGTGTTAATT
AAATCGTTCGGCTTCAGTAGAGTTAACGTTAGCTCACATCGGAAAAACT
CTTGTCAACTTGATGTCGAGACACACTTACCCGACCGCACGGGAAGGG
CGGTTCACAGCTTTGATTCTCAGCGAGCCGTAAGCTCAGTCAACTAC
AACTTTGTTGCGCGGAATTTCAGAATTGCTCGCATCGCCATT
CAGGTGACGTCAAACCGCAGTTCTTCCAAACCAAAAGGG
AACTCGTCAAGCTTCAAGCTTCTGAGTTACCTTTAGGCC
TAGGAATTATCGCTTACAACACTCCCCCATT
GGTCCCCATCGACATATTGCTTCAAGACTCAGTGGTCC
CTCGCGCGCTTCTC
TGCACGAATTCTCAAGCTTCTGAGTTACCTTTAGGCC
TCGCGTTTCGTTGAGAAACTGTGGAAACAGGGAGAGAAAAAC
ATATTTAAAGATTGATGAAGACA
CACTGTAATATGCTGCTTTGTTCTCT
CACTGACCTAAAGAGATCTAGGAACCC
GCGCTCGCTCGCTCACTGAGGCCGCCGG
GGTGC
P00123 full sequence (from ITR to ITR) :
(SEQ ID NO: 281)
GGCCACTCCCTCTGCGCCTCGCTCACTGAGGCCGGCGAC
CCGACGCCGGCTTGCCGGCGCTCAGTGAGCGAGCGCAGAGAG
GAGTGGCCA
GAGTGGCCA
TGAAGAGAAGAAC
TATGTTGTTGTTCTCC
TATGTTGTTGTTGATCATGAAACGCCA

- continued

ACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAAGGTAATTGGAAAGAGTTGTTC
 AAGGGAACCTTGAGAGAGAATGTATGGAAAGAAAAGTGTAGTTTGAAGAACACGA
 GAAAGTTTTGAAAACACTGAAAAACAACGTAAATTGGAAAGCAGTATGTTGATGG
 AGATCAGTGTGAGTCCAATCCATGTTAAATGGCGGCAGTGCAGGATGACATTAA
 TTCCCTATGAATGTTGGTCCCTTGGATTGAAGGAAAGAACGTGAATTAGATGT
 AACATGTAACATTAAGAATGGCAGATGCGAGCAGTTGTAAGGAAAGAACGTGATA
 ACAAGGTGGTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAACGTCT
 GTGAACCAGCAGTGCCATTCCATGTGGAAGAGTTCTGTTCACAAACTCTAAGC
 TCACCCGTGCTGAGACTGTTCCATGTGACTATGTAATTCTACTGAAGCTGA
 AACCATTTGGATAACATCACTCAAAGCACCCAACTATTAATGACTTCACTCGGGT
 TGTTGGTGGAGAAGATGCCAACCCAGGTCAATTCCCTGGCAGGTGTTGAATGG
 TAAAGTTGATGCAATTCTGTGGAGGCTCTATGTTAATGAAAAATGGATTGTAAGTGC
 TGCCCACGTGTTGAAACTGGTAAAATTACAGTTGTCGCAGGTGAACATAATAT
 TGAGGAGACAGAACATACAGAGCAGCGAAATGTGATTGAAATTCTCACC
 ACAACTACAATGCACTATTAAAGTACAACCATGACATTGCCCTCTGGAACACTGG
 ACGAACCCCTAGTGCTAACAGCTACGTTACACCTATTGCAATTGCTGACAAGGAAT
 ACACGAACATCTCCTCAAATTGGATCTGGCTATGTAAGTGGCTGGGAAGAGTCT
 TCCACAAAGGGAGATCAGCTTAGTTCTCAGTACCTTAGAGTTCCACTGTTGACC
 GAGCCACATGCTTCTATCTACAAAGTCACCATCTATAACACATGTTCTGCTG
 GCTTCCATGAAGGAGGTAGAGATTGATGTCAAGGAGATAGTGGGGACCCATGTT
 ACTGAAGTGGAGGGACCAAGTCTTAACGTTGAAATTAGTGGCTGGGTGAAGAGTG
 TCGCAATGAAAGGAAATATGGAAATATAACCAAGGTATCCCGTATGTCACGTTG
 TTAAGGAAAAAAACAAAGTCACCTAACCTCGACTGTCCTCTAGTTGCCAGCCATC
 TGTTGTTGCCCTCCCCGTGCCCTTGACCCCTGGAAGGTGCCACTCCACTGTC
 CTTTCCTAATAAAATGAGGAATTGATCGCATGTCATTGCTGAGTAGGTGTCATTCTATT
 TGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAG
 GCATGCTGGGGATGCGGTGGGCTCTATGGCTCTGAGGCGGAAAGAACAGCTGG
 GCTCTAGGGGTATCCCCACTAGTCCACTCCCTCTGCGCGCTCGCTCGCTACTG
 AGGCCGGCGACCAAAGTCGCCCCGACGCCCGGGCTTGCCGGCGGCCTCAGTG
 AGCGAGCGAGCGCCAGAGAGGGA

P00204 full sequence (from ITR to ITR) :

(SEQ ID NO: 282)

GGCCACTCCCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGCGACCAAAGGTGCG
 CGACGCCGGCTTGCCGGCGGCTCAGTGAGCGAGCGAGCGCGAGAGAGG
 GAGTGGCCAACCTCATCACTAGGGGTTCTGGAGGGGTGGAGTCGTGACCTAGGTC
 GTCTCCGGCTCTGCTTTTCCAGGGGTGTGTTCGCCGAGAACGCACTGTAAGAGTTT
 ATGTTTTTCATCTGCTTGATTTCTAGTAATGGAAGCCTGGTATTTAAAATA
 GTTAAATTTCTTCTAGTGCTGATTTCTAGATTATTACTGTTGTTGTTATTAT
 TGTCATTATTCATCTGAGAACCTAGGTCACTGAGAGAAGAACAAAAGCAGCAT
 ATTACAGTTAGTGTCTCATCAATCTTAAATATGTTGTTGTTCTCCCTGT

- continued

TTCCACAGTTTCTTGATCATGAAAACGCCAACAAATTCTGAATCGGCCAAGAG
 GTATAATTCAAGGAAATTGGAGAGAGTTGTTCAAGGGAACCTTGAGAGAGAAATGTA
 TGGAGAAAAGTGTAGTTGAAGAACGAGAAGTTTGAAAACACTGAAAGA
 ACAACTGAATTGGAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGT
 TAAATGGCGCAGTTGCAAGGATGACATTAATTCTATGAATGTTGGTGTCCCTT
 GGATTTGAAGGAAAGAACGTGAAATTAGATGTAACATGTAACATTAAGAATGGCAG
 ATGCGAGCAGTTGAAAAATAGTGTGATAACAAGGTGGTTGCTCCTGTACTGA
 GGGATATCGACTTGAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCAT
 GTGGAGAGTTCTGTTCACAAACTCTAAGCTCACCCGTGCTGAGACTGTTTCC
 TGATGTGGACTATGTAATTCTACTGAAGCTGAAACCATTGGATAACATCACTCA
 AACGCACCAATTTAATGACTTCACTCGGGTTGTTGGAGAGATGCCAAACC
 AGGTCAATTCCCTGGCAGGTTGTTGAATGGTAAAGTTGATGCATTCTGGAGG
 CTCTATCGTTAATGAAAATGGATTGTAACTGCTGCCACTGTGTTGAAACTGGTGT
 TAAAATTACAGTTGTCGAGGTGAAACATAATTGAGGAGACAGAACATACAGAGC
 AAAAGCGAAATGTGATTGAAATTCTCACACAAACTACAATGCAGCTATTAA
 AGTACAACCATGACATTGCCCTCTGAACTGGACGAACCCCTAGTGCTAACAGCT
 ACCTTACACCTATTGCAATTGCTGACAAGGAATACACGAACATCTCTCAAATTG
 GATCTGGCTATGTAAGTGGCTGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTA
 GTTCTTCAGTACCTTAGAGTTCCACTTGTGACCGAGCCACATGTTCTATCTACAA
 AGTTCAACCATCTATAACAAACATGTTCTGCTGGCTTCCATGAAGGGAGTAGAGATT
 CATGTCAGGAGATACTGGGGACCCATGTTACTGAAGTGGAGGGACAGTTCT
 TAACTGGAATTATTAGCTGGGTGAAGAGTGTGCAATGAAAGGAAATATGGAAT
 ATATACCAAGGTATCCGGTATGTCAACTGGATTAAAGGAAAAACAAAGCTCACTT
 AACCTCGACTGTGCCCTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCGTGCC
 TCCTTGACCTGGAAAGGTGCCACTCCACTGTCCTTCTAATAAAATGAGGAATT
 GCATCGCATTGTCGAGTAGGTGTCATTCTATTCTGGGGGGGGGGGGCAGGAC
 AGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATGCCGGGGCT
 CTATGGCTTCTGAGGCGGAAAGAACCCAGCTGGGCTCTAGGGGTATCCCCCTAG
 GTGGTTATTATTGATATTTGGTATCTTGATGACAATAATGGGGATTTGA
 AAAGCTTAGCTTAAATTCTTTAATTAAAAAAATGCTAGGCAGAATGACTCAA
 TTACGTTGGATACTGGTGAATTATTACGGCTCATAGGGCCCTGCTCGACCAT
 GCTATACTAAAAATTAAAAGTGTACTAGTCCACTCCCTCTGCGCGCTCGCTCG
 CACTGAGGCCGGCGACCAAGGTGCCCCGACGCCGGCTTGCCGGCGCCT
 CAGTGAGCGAGCGAGCGCGCAGAGAGGG

P00353 full sequence (from ITR to ITR) :

(SEQ ID NO: 283)

TTGGCCACTCCCTCTGCGCGCTGCTCGCTCACTGAGGCCGGCGACCAAAGGTC
 GCGCGACGCCGGCTTGCCGGCGCCCTAGTGAGCGAGCGAGCGCAGAGA
 GGGAGTGGCCAACCTCCATCACTAGGGGTTCTAGATCTGATTGAAAGCTTAGCT
 TAAATTCTTTAATTAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGAT

- continued

ACAGTTGAATTATTACGGCTCATAGGGCTGCCCTGCTCGACCATGCTATACTAAA
AATTAAAAGTGTGTTACTAATTTATAATGGAGTTCCATTATTTACCTTTA
TTCTTATTACCATTGTCTTAGATATTACAAACATGACAGAAACACTAAATCT
TGAGTTGAATGCACAGATATAAACACTTAACGGGTTTAAAAATAATGTTGGT
GAAAAAAATAACTTGAGTGTAGCAGAGAGAACATTGCCACCTTCAGATTTCC
TGTAACGATCGGAACTGGCATCTTCAGGGAGTAGCTTAGGTAGTGAAGAGAAGA
ACAAAAAGCAGCATATTACAGTTAGTTGTCTCATCAATCTTAAATATGTTGTG
GTTTTCTCTCCCTGTTCCACAGTTTCTTGATCATGAAAACGCCAACAAATTCT
GAATCGGCCAAGAGGTTCTTGATCATGAAAACGCCAACAAATTCTGAATCGGC
CAAAGAGGTATAATTCAAGGTTAGGAAAGAGTTGTTCAAGGGAACCTTGAGAGA
GAATGTATGGAAGAAAAGTGTAGTTTGAGAAGACGAGAAGAGTTTGAAAACAC
TGAAAGAACAACTGAATTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCA
ATCCATGTTAAATGGCGCAGTTGCAAGGATGACATTAATTCTATGAATGTTGGT
GTCCTTGGATTGAGGAAAGAAACTGTGAATTAGATGTAACATGTAACATTAAGA
ATGGCAGATCGCAGCTTTGTAAGGAAACTGCTGATAACAAGGTGGTTGCTCCT
GTACTGAGGGATATCGACTTGCAAGAACAGAAGTCCTGTGAACCAGCAGTGCCA
TTCCATGTGGAAGAGTTCTGTTCAAAACTCTAAGCTCACCGTGCTGAGACTG
TTTCCTGATGTGACTATGTAATTCTACTGAAGCTGAAACCATTGGATAACAT
CACTCAAAGCACCAACTATTAAATGACTTCACTCGGGTTGTTGGAGAAGATGC
CAAACCAGGTCAATTCCCTGGCAGGTTGTTGAATGGTAAAGTTGATGCATTCTG
TGGAGGCTCTATGTTAAATGGATTGTAAGTGTGCTGCCACTGTGTTGAAAC
TGGTGTAAATTACAGTTGTCGCAGGTGAACATAATTGAGGAGACAGAACATA
CAGAGCAAAGCAGGAAATGTGATTCGAATTCTCACCACAACATGCAAGCT
ATTAATAAGTACAACCATGACATTGCCCTCTGGAACGGACAACTTAGTGCTA
AACAGCTACGTTACACCTATTGCTGACAGGAAACACGAACATCTCCTC
AAATTGGATCTGGCTATGTAAGTGGCTGGGAAGAGCTTCCACAAAGGGAGATC
AGCTTTAGTTCTCAGTACCTAGAGTTCCACTTGTTGACCGAGCCACATGTTCTA
TCTACAAAGTTACCATCTATAACACATGTTCTGTGCTGGCTTCCATGAAGGAGGT
AGAGATTCTGTCAGGAGATAGTGGGGACCCATGTTACTGAAGTGGAAAGGGAC
CAGTTCTTAACGGATTATTAGCTGGGTGAAGAGTGTGCAATGAAAGGCAAAT
ATGGAATATATACCAAGGTATCCGGTATGCAACTGGATTAAAGGAAAAACAAAG
CTCACTTAACCTCGACTGTGCCCTCTAGTTGCCAGCCATCTGTTGCCCCCTCCC
CGTCCTCTGACCCCTGGAGGTGCCACTCCACTGCTTCTCTAAATAAAATGA
GGAAATTGCATCGCATTGCTGAGTAGGTGTCAATTCTATTCTGGGGGTGGGTGG
GCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATGCG
GTGGGCTCTATGGCTCTGAGGCGGAAAGAACAGCTGGGCTCTAGGGGTATCC
CCGTGAGATCGCCCATCGGTATAATGATTGGGAGAACACATTCAAAGGCTGTA
AGTTATAATGCTGAAAGCCCCTTAATTCTGGTAGTATTAGTTAAAGTTAAA
ACACCTTTCCACCTTGAGTGTGAGAATTGTAGAGCAGTGCTGTCCAGTAGAAATG

- continued

```
TGTGCATTGACAGAAAAGACTTGGATCTGTCTGACAATGGCAGGCCAGAGATC
ACAAGGCTATCAAGCACTTGACATGGCAAGTGTAACTGAGAAGCACACATTCAA
ATAATAGTTAATTAAATTGAATGTATCTAGCCATGTGTGGCTAGTAGCTCCCTTCCT
GGAGAGAGAATCTGGAGGCCACATCTAACTTGTAAAGTCTGGAATCTTATTTTAT
TTCTGGAAAGGTCTATGAACATAGTTGGGGCAGCTCAGTTACTAACTTTAAT
GCAATAAGAACATCTCATGGTATCTTGAGAACATTATTTGTCTTGTAGATCTAGGA
ACCCCTAGTGTGGAGTTGCCACTCCCTCTGCGCGCTCGCTCGCTCACTGAGGC
CGCCCGGGCAAAGCCCCGGCGTCGGCGACCTTGGTCGCCCGCCTCAGTGAGCG
AGCGAGCGCGAGAGAGGGAGTGGCAA
```

P00354 full sequence (from ITR to ITR) :

(SEQ ID NO: 284)

```
TTGGCCACTCCCTCTCGCGCGCTCGCTCACTGAGGCCAGCGAGA
GCCCGACGCCCGGGCTTGCCGGCGGCCTAGTGAGCGAGCGAGA
GGGAGTGGCCAECTCATCACTAGGGTTCTAGATCTTAGCCTGGCAAATGAA
GTGGGTAACCTTCTCCCTCTCGTCTCGGCTCTGCTTTCCAGGGGTGTG
TTCGCGAGAACGACGTAAGAGTTATGTTTCTAGCTCTGCTGTATTTCAG
TAATGGAAGCCTGGTATTTAAAATAGTTAAATTCCCTTAGTGCTGATTCTAGAT
TATTATTACTGTGTTGTTATTATTGTATTGTCATGAGAACCTTAGGTG
GTTTATTTGATATATTTGGTATTTGATGACAATAATGGGGATTTGAAAG
CTTAGCTTAAATTCTTTAATTAAAAAAATGCTAGGAGAACGACTCAAATTA
CGTTGGATACAGTTGAATTATTACGGTCTCATAGGGCTGCCCTGCTGACCAGCT
ATACTAAAAATAAAAGTGTGTTACTAATTATAATGGAGTTCCATTATATT
TACCTTATTCTTATTACCATTTGCTTAGAGATATTACAAACATGACAGAAACA
CTAAATCTGAGTTGAATGCACAGATATAAACACTTAACGGTTTAAAATAA
ATGTTGGTAAAAAATAACTTGAGTGTAGCAGAGAGGAACCTGCCACCTCA
GATTTCCTGAAACGATCGGGAACTGGCATCTCAGGGAGTAGCTTAGGTAGTGAA
GAGAAGAACAAAAAGCAGCATATTACAGTTAGTGCTTCATCAATTTAAATAG
TTGTGTGGTTTCTCTCCCTGTTCCACAGTTCTTGATCATGAAAACGCAACA
AAATTCTGAATCGGCCAAGAGGTATAATTCAAGGTAAATTGGAAGAGTTGTTCAA
GGAACCTTGAGAGAGAATGTGGAAGAAAAGTGTAGTTGAGAACGACGAGA
AGTTTTGAAAACACTGAAAGAACAACTGAATTGGAAAGCAGTATGTTGATGGAG
ATCAGTGTGAGTCCAATCCATGTTAAATGGGGCAGTTGCAAGGATGACATTAA
CCTATGAATGTTGGTGTCCCTTGGATTGAGGAAAGAACGACTGTGAATTAGATGAA
CATGTAACATTAAGAATGGCAGATGCGAGCAGTTGTAAGGAAACTGTGATAAAC
AAGGTGGTTGCTCCTGTAAGGGATATCGACTTGCAAGAAAACAGAACGACTG
GAACCAAGCAGTGCCATTCCATGTGGAAGAGTTCTGTTCAAAACTCTAAGCTC
ACCCGTGCTGAGACTGTTTCTGATGTGGACTATGTAATTCTACTGAGCTGAA
ACCATTTGGATAACATCACTCAAAGCACCAATCATTAAATGACTTCAGTGGTT
GTTGGTGGAGAAGATGCCAACCGGTCAATTCCCTGGCAGGTTGTTGAATGGT
AAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAATGGATTGTAAGCTG
```

- continued

GCCCACTGTGTTGAAACTGGTAAAATTACAGTTGTCGAGGTGAAACATAATATT
 GAGGAGACAGAACATACAGAGCAAAGCGAAATGTGATTGCAATTATTCCCTCACCA
 CAACTACAATGCAGCTATTAATAAGTACAACCAGCATTGCCCTCTGGAACCTGGA
 CGAACCTTAGTGCTAACAGCTACGTTACACCTATTGCTGACAAGGAATA
 CACGAACATCTCCTCAAATTGGATCTGGCTATGTAAGTGGCTGGGAAGAGTCTT
 CCACAAAGGGAGATCAGCTTAGTTCTCAGTACCTTAGAGTCCACTTGTGACCG
 AGCCACATGTCTTCTATCTACAAAGTCACCATCTATAACACATGTTCTGCTGGC
 TTCCATGAAGGAGGTAGAGATTCAATGTCAGGAGATAGTGGGGACCCATGTTAC
 TGAAGTGGAAAGGAGCAGTTCTTAACTGGAATTAGCTGGGTGAAGAGTGTG
 CAATGAAAGGCAAATATGGAATATACCAAGGTATCCGGTATGTCACACTGGATT
 AAGGAAAAAACAAGCTCACTAACCTCGACTGTGCCTCTAGTTGCCAGCCATCTG
 TTGTTGCCCTCCCCGTGCCCTTGACCTGGAAAGGTGCCACTCCACTGTCCT
 TTCCCTAATAAAATGAGGAATTGCATGCATTGCTGAGTAGGTGTCATTCTATTCT
 GGGGGTGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGG
 CATGCTGGGATGCGGTGGCTATGGCTTGAGGGCGAAAGAACCCAGCTGGGG
 CTCTAGGGGTATCCCCTGAGATGCCATCGGTATAATGATTGGAGAACACA
 TTCAAAAGGCTGTAAGTTATAATGCTGAAAGCCACTTAATATTCTGGTAGTATT
 AGTTAAAGTTAAACACCTTTCCACCTTGAGTAGTGAGAATTGAGACAGTGC
 TGCCAGTAGAAATGTCATTGACAGAAAGACTGTGGATCTGCTGAGCAATGT
 GGCAAGGAGATCACAAGGCTATCAAGCACTTGACATGGCAAGTGTAACTGAG
 AACACACATTCAAATAATGTTAATTGAAATGTATCTAGCCATGTGTGGCT
 AGTAGCTCTTCTGGAGAGAGAATCTGGAGGCCACATCTAACTGTTAAGTCTGG
 AATCTTATTTTATTCGGAAAGGTCTATGAACTATAGTTGGGGCAGCTCACT
 TACTAACTTTAATGCAATAAGAATCTCATGGTATCTTGAGAACATTATTTGTCT
 TTGTTAGTACTGAAACCTTACATGTAAGTAAGGGCTATGAACTTAAAGTCACATCT
 CCAACCTTAGTAATGTTAATGTTAGTAAAAAAATGAGTAATTAAATTATTTAGA
 AGGTCAATAGTATCATGTTACCTAAACAGAGGTATATGGTTAGAAAAAGAAC
 ATTCAAAGGACTTATATAATCTAGCCTGACAATGAATAATTAGAGAGTAGTT
 TGCCCTGTTGCCATGTTACATGACACATATGTGCTCTGCACCTCAGC
 ATGGTAGAGTCATATTCAAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCC
 TCTCTGCGCCTCGCTCGCTCACTGAGGCCGCCGGCAAAGCCGGCGTGGC
 GACCTTGGTCGCCCGCCTCAGTGAGCGAGCGAGCGCAGAGAGGGAGTGGCCA

A

P00350: The 300/600 bp HA F9 construct (for G551) (SEQ ID NO: 285)
 TTGGCCACTCCCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGCGACCAAAGGTC
 GCCCGACGCCGGCTTGCCGGCGGCCTCAGTGAGCGAGCGAGCGCAGAGA
 GGGAGTGGCCAATCCATCACTAGGGGTTCTAGATCTAAGTATATTAGAGCGAGTC
 TTTCTGACACAGATCACCTTCTATCAACCCACTAGCCTCTGGCAAAATGAAGT
 GGGTAACCTTCTCCTCCTCTCGTCTCGGCTCTGCTTTCCAGGGTGTGTTT

- continued

CGCCGAGAAGCACGTAAGAGTTTATGTTTTCATCTCTGCCTGTATTTCTAGTA
ATGGAAGCCTGGTATTTAAATAGTTAAATTCCCTTAGTGCTGATTTCTAGATTA
TTATTACTGTTGTTGTTATTATGTCATTATTCATCTGAGAACCTTTCTTGA
TCATGAAAACGCCAACAAAATTCTGAATCGGCCAACAGAGGTATAATTCAAGTAAAT
TGGAAAGAGTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGT
TTGAAAGAACGAGAACAGTTGAAAACACTGAAAGAACAACTGAATTGGAA
GCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTAAATGGCGCAGTTG
CAAGGATGACATTAATTCTATGAATGTTGGTCCCTTGATTTGAAGGAAAGAA
CTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGGCAGCAGTTGTAA
AAATAGTGTGATAACAAGGTGGTTGCTCCTGTACTGAGGGATATCGACTTGCAGA
AAACCAGAAGTCTGTGAACCAGCAGTGCCATTCCATGTGGAAGAGTTCTGTTTC
ACAAACTCTAACGCTCACCGTGTGAGACTGTTTCCATGTGACTATGTAAA
TTCTACTGAAGCTGAAACCATTGGATAACATCACTCAAAGCACCAATCATTAA
TGACTTCACTCGGGTTGTTGGAGAACGATGCCAACCCAGGTCAATTCCCTGGCA
GGTGTTTGAATGGAAAGTTGATGCCATTGTGAGGGCTATCGTTAATGAAAA
ATGGATTGTAACTGCTGCCACTGTGTTGAAACTGGTGTAAATTACAGTTGCGC
AGGTGAACATAATTGAGGAGACAGAACATACAGAGCAAAGCGAAATGTGATTC
GAATTATTCCCTACCACAACTACAATGCAGCTATTAAAGTACAACCATGACATTG
CCCTCTGGAACTGGACGAACCTTAGTGCTAACAGCTACGTTACACCTATTGCA
TTGCTGACAAGGAATACACGAACATCTCCTCAAATTGGATCTGGCTATGTAAGT
GCTGGGGAAAGAGTCTTCACAAAGGGAGATCAGCTTAGTTCTCAGTACCTAGAG
TTCCACTTGTGACCGAGCCACATGCTTCTATCTACAAAGTCACCATCTATAACAA
CATGTTCTGTGCTGGCTTCATGAAGGAGGTAGAGATTGTCAGGAGATAGTGG
GGGACCCATGTTACTGAAAGTGGAAAGGGACAGTTCTTAAGTGGATTAGCTG
GGGTGAAGAGTGTGCAATGAAAGGAAATATGGAATATACCAAGGTATCCCGT
ATGTCAACTGGATTAAGGAAAAACAAAGCTACTAACCTCGACTGTGCCTCTAG
TTGCCAGCCATCTGTTGTTGCCCTCCCCGTGCCTCCTTGACCCGTGAAGGTGCC
ACTCCCACGTGCTTCTTAATAAAATGAGGAAATTGATCGCATTGCTGAGTAGG
TGTCAATTCTATTCTGGGGGTGGGTGGGCAGGACAGCAAGGGGGAGGATTGGGA
AGACAATAGCAGGCATGCTGGGATGCGGTGGGCTATGGCTTGAGGCGAAA
GAACCAGCTGGGCTCTAGGGGTATCCCCCTAGGTGGTTATTGATATT
TTGGTATCTTGATGACAATAATGGGGATTGAAAGCTAGCTTAAATTCTT
TAATTAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGATACAGTTGAAT
TTATTACGGTCTCATAGGGCCTGCCTGCTGACCAGTGTACACTAAAAATTAAAAGT

- continued

GTGTGTTACTAATTATAAAATGGAGTTCCATTATTTACCTTATTCTTATT
CCATTGTCCTAGATATTACAAACATGACAGAACACTAAAGATCTAGGAACCC
CTAGTGATGGAGTTGCCACTCCCTCTGCGCCTCGCTCGCTACTGAGGCC
CGGGCAAAGCCCCGGCGTCGGCGACCTTGGTCGCCCGCCTAGTGAGCGAGCG
AGCGCGCAGAGAGGGAGTGGCAA

P00356: The 300/2000 bp HA F9 construct (for G551) (SEQ ID NO: 286)
TTGGCCACTCCCTCTGCGCCTCGCTCGCTACTGAGGCCGGCACAAAGGTC
GCCGCACGCCGGCTTGCCGGCGCCTCAGTGAGCGAGCGCAGAGA
GGGAGTGGCCAATCCATCACTAGGGTCCCTAGATCTAAGTATTAAGCGAGTC
TTTCTGCACACAGATCACCTTCCTATCAACCCCCTAGCCTGGCAAAATGAAGT
GGTAACCTTCCTCCTCTCGTCTCCGCTCTGCTTTCCAGGGGTGTGTT
CGCCGAGAAGCAGTAAGAGTTTATGTTTCTGCTCTGCTGTATTTCTAGTA
ATGGAAGCCTGGTATTTAAATAGTTAAATTTCTTAGTGCTGATTTCTAGATTA
TATTACTGTTGTTGTTATTATTGTCATTATTGCACTGAGAACCTTTCTTGA
TCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAAGTAAAT
TGGAAAGAGTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGT
TTGAAGAACGAGAAGTGGAAACACTGAAACAACGTGAATTGGAA
GCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTAAATGGCGCAGTTG
CAAGGATGACATTAATTCTATGAATGTTGGTCCCTTGATTGAAAGGAAAGAA
CTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGGCAGCTTGTAA
AAATAGTGTGATAACAAGGTGGTTGCTCTGACTGAGGGATATGACTTGCAGA
AAACCGAGTCCTGTGAACCAGCAGGCCATTCCATGTGAAAGAGTTCTGTT
ACAAACTCTAACGTCACCCGTCGACTGTTTCTGATGTGGACTATGTA
TTCTACTGAAGCTGAAACCATTGGATAACATCACTCAAAGCACCAATCTTAA
TGACTTCACTCGGGTTGGTGGAGAAGATGCCAACCCAGGTCAATTCCCTGGCA
GGTTGTTTGAATGGAAAGTTGATGCATTCTGTTGGAGGCTATGTTAATGAAAA
ATGGATTGTAACTGCTGCCACTGTGTTGAAACTGGTTAAATTACAGTTGTCGC
AGGTGAACATAATTGAGGAGACAGAACATACAGAGCAAAGCGAAATGTGATT
GAATTATTCTCACCACAACTACAATGCGACTTAAAGTACAACCATGACATTG
CCCTCTGGAACGGACGCCACCTTAGTGCTAAACAGCTACGTTACACCTATTGCA
TTGCTGACAAGGAATACACGAACATCTCCCAAATTGGATCTGGTATGTAAGTG
GCTGGGGAAAGAGTCTTCCACAAAGGGAGATCAGCTTAGTTCTCAGTACCTAGAG
TTCCACTGTTGACCGACGCCACATGCTTCTATCTACAAAGTCACCATCTATAACAA
CATGTTCTGCTGGCTTCCATGAAGGAGGTAGAGATTGTCAGGAGATAGTGG
GGGCCACCATGTTACTGAAGTGGAGGGACAGTTCTTAACGAAATTAGCTG
GGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATACCAAGGTATCCCGT
ATGTCACACTGGATTAAGGAAAAACAAAGCTACTTAACCTCGACTGTGCCTCTAG
TTGCCAGCCATCTGTTGCCCCTCCCCGTGCCCTCCTGACCCCTGGAAGGTGCC
ACTCCCACGTGCTTCTCTAATAAAATGAGGAAATTGCATCGCATTGTCAGTAGG

- continued

TGTCATTCTATTCTGGGGGTGGGTGGGCAGGACAGCAAGGGGGAGGATGGGA
AGACAATAGCAGGCATGCTGGGATGCGGTGGCTCATGGCTCTGAGGCGGAAA
GAACCAGCTGGGCTCTAGGGGTATCCCCCTAGGTGGTTATATTATGATATATT
TTTGGTATCTTGATGACAATAATGGGGATTTGAAAGCTTAGCTTAAATTCTTT
TAATTAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATACAGTTGAAT
TTATTACGGTCTCATAGGCCTGCCTGCGACCATGCTATACTAAAAATTAAAGT
GTGTGTTACTAATTATAAATGGAGTTCCATTATATTACCTTATTCTTATTAA
CATTGTCTTAGATATTACAAACATGACAGAACACTAAATCTTGAGTTGAA
TGCAAGATATAAACACTTAACGGGTTTAAAGATAATAATGTTGGTAAAAAATAT
AACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTCAGATTTCTGTAACGATC
GGGAACCTGCATCTCAGGGAGTAGCTTAGGTAGTCAGTGAAGAGAAGAACAAAAAGCA
GCATATTACAGTTAGTTGTCTTCATCAATCTTAAATATGTTGTGGTTTCTCTCC
CTGTTTACAGACAAGAGTGAGATGCCCATCGGTATAATGATTGGAGAACAA
CATTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCCTTAATATTCTGGTAGTA
TTAGTTAAAGTTTAAACACCTTTCCACCTTGAGTGTGAGAATTGAGAGCAGT
GCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGTGCTGAGCAAT
GTGGCAGCCAGAGATCACAGGCTATCAAGCCTTGCACATGGCAAGTGTAACTG
AGAACACACATTCAAATAATAGTTAATTGAATGTATCTAGCCATGTGTGG
CTAGTAGCTCCTTCCTGGAGAGAGAATCTGGAGCCACATCTAATTGTTAAGTCT
GGAATCTATTCTGAAAGGTCTATGAACTATAGTTGGGGCAGCTCA
CTTACTAATTGCAATAAGATCCATGGTATCTTGAGAACATTATTGTCTCT
TTGTAGTACTGAAACCTTACATGTGAAGTAAGGGCTATACTTAAGTCACATCT
CCAACCTTAGTAATGTTAATGTAGTAAAAAAATGAGTAATTAAATTATTAGA
AGGTCAATAGTATCATGTATTCAAATAACAGAGGTATATGGTTAGAAAAGAAC
ATTCAAAGGACTTATATAATCTAGCCTTGACAATGAATAAATTAGAGAGTAGTT
TGCCTGTTGCCTCATGTTCATAAATCTATTGACACATATGTGCATCTGCACCTCAGC
ATGGTAGAAGTCCATATCCTTGCTGGAAAGGCAGGTGTTCCCATTACGCCCTCAG
AGAATAGCTGACGGGAAGAGGCTTCTAGATAGTTGTATGAAAGATATAACAAATC
TCGCAGGTACACAGGCATGATTGCTGGTGGAGAGGCCACTGCCTCATACTGA
GGTTTTGTGCTGCTTTCAGAGTCCTGATTGCCTTCCCAGTATCTCAGAAATG
CTCATACGATGAGCATGCCAAATTAGTCAGGAAGTAACAGACTTGCAGAACAGT
GTGTTGCCGATGAGTCTGCCGCAACTGTGACAATCCCTGTGAGTACCTCTGAT
TTTGTGGATCTACTTCCTGCTTCTGAACTCTGTTCAAAGCCAATCATGACTCCA
TCACCTAAGGCCCGGGAACACTGTGGCAGAGGGCAGCAGAGAGATTGATAAAGCC
AGGGTGATGGGAATTCTGTGGACTCCATTCTAGTAATTGCAGAACGCTACAAT

- continued

ACACTCAAAAAGTCTACCACATGACTGCCAAATGGGAGCTTGACAGTGACAGTG
ACAGTAGATATGCCAAAGTGGATGAGGGAAAGACCACAAGAGCTAAACCTGTAAA
AAGAACTGTAGGCAACTAAGGAATGCAGAGAGAAAGATCTAGGAACCCCTAGTGAT
GGAGTTGGCCACTCCCTCTCGCGCTCGCTCGCTCACTGAGGCCGCCGGCAGA
GCCCGGGCGTCGGCGACCTTGGTCGCCCGCCTCAGTGAGCGAGCGAGCGCA
GAGAGGGAGTGCCAA
P00362: The 300/1500 bp HA F9 construct (for G551) (SEQ ID NO: 287)
TTGGCCACTCCCTCTCGCGCTCGCTCGCTCACTGAGGCCGCCGGCAGA
GCCCGACGCCCGGCTTGCCCGGCCCTCAGTGAGCGAGCGAGA
GGGAGTGGCCAECTCATCACTAGGGTTCTAGATCTAAGTATATTAGCGAGTC
TTTCTGCACACAGATCACCTTCCTATCAACCCACTAGCCTGGCAAATGAAGT
GGTAACCTTCTCCTCCTCGTCTCCGGCTTGCTTTCCAGGGTGTGTT
CGCCGAGAACGAGCTAAGAGTTTATGTTTCTGCTGTATTTCTAGTA
ATGGAAGCTGGTATTTAAAAATAGTTAAATTTCCTTAGTGCTGATTTCTAGATTA
TTATTACTGTTGTTGTTATTATTGTCATTATTGCACTGAGAACCTTTCTTGA
TCATGAAAACGCCAACAAAATTCTGAATCGGCAAAGAGGTATAATTCAAGTAAAT
TGGAGAGTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGT
TTGAGAACGAGAGTTGAAAACACTGAAAGAACAACTGAATTGGAA
GCAGTATGTTGAGATCAGTGTGAGTCAATCCATGTTAAATGGCGCAGTTG
CAAGGATGACATTAATTCTATGAATGTTGGTGTCCCTTGATTGAAAGGAAAGAA
CTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGGCAGCAGTTGTAA
AAATAGTGTGATAACAAGTGGTTGCTCCTGACTGAGGGATACTGACTTGCAGA
AAACACAGAACGAGCTGAAACCGCAGTGCCTTCAATGTGAAAGAGTTCTGTTTC
ACAAACTCTAACGCTACCCGTGCTGAGACTGTTTCTGATGTGACTATGTAAA
TTCTACTGAAGCTGAAACCATTGGATAACATCACTCAAAGCACCAATCATTAA
TGACTTCACTCGGGTTGTTGGAGAAGATGCCAACCCAGGTCAATTCCCTGGCA
GGTTGTTTGAATGGAAAGTTGATGCAATTCTGTGGAGGCTATGTTAAATGAAAA
ATGGATTGTAACTGCTGCCACTGTGTTGAAACTGGTGTAAAAATTACAGTTGCGC
AGGTGAACATAATTGAGGAGACAGAACATACAGAGCAGCGAAATGTGATTC
GAATTATTCCCTACCCACAACTACAATGCAAGCTATTAAAGTACAACCATGACATTG
CCCTCTGGAACTGGACGAACCTTAGTGCTAACAGCTACGTTACACCTATTGCA
TTGCTGACAAGGAATACACGAACATCTCCTCAAATTGGATCTGGCTATGTAAGTG
GCTGGGGAAAGAGTCTTCCACAAAGGGAGATCAGCTTAGTTCTCAGTACCTAGAG
TTCCACTTGTGACCGAGGCCACATGTCTTCTATCTACAAAGTCACCATCTATAACAA
CATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTGATGTCAAGGAGATAGTGG
GGGACCCCATGTTACTGAAGTGGAAAGGGACCAAGTTCTTAACGAAATTAGCTG
GGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATACCAAGGTATCCCGT
ATGTCAACTGGATTAAGGAAAAACAAAGCTCACTTAACCTCGACTGTGCCTCTAG
TTGCCAGCCATCTGTTGTTGCCCTCCCCGTGCCTTCCTTGACCTGGAAAGGTGCC

- continued

```

ACTCCCACGTCTTCTTAATAAAATGAGGAATTGCATCGCATTGTCTGAGTAGG
TGTCAATTCTATTCTGGGGGGTGGGGGGAGGACAGCAAGGGGGAGGATGGGA
AGACAATAGCAGGCATGCTGGGATGCGGTGGGCTATGGCTCTGAGGCCGAAA
GAACCACTGGGCTCTAGGGGTATCCCCCTTAGGTGGTTATATTATGATATATT
TTGGTATCTTGATGACAATAATGGGGATTTGAAAGCTAGCTTAAATTCTTT
TAATTAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATACAGTTGAAT
TTATTACGGTCTCATAGGGCCTGCCTGCGACCATGCTATACTAAAATTAAAAGT
GTGTGTTACTAATTATAAATGGAGTTCCATTATATTACCTTATTCTTATTAA
CCATTGTCTTAGATATTACAAACATGACAGAACACTAAATCTGAGTTGAA
TGCACAGATATAAACACTTAACGGGTTTAAAAATAATGTTGGTAAAAAAATAT
AACTTGAGTGAGCAGAGAGGAACCAATTGCCACCTTCAGATTTCGTAAAGATC
GGGAACGGCATCTTCAGGGAGTAGCTTAGGTAGTGAAGAGAAGAACAAAAGCA
GCATATTACAGTTAGTTGCTTCATCAATCTTAAATATGTTGTGGTTTCTCTCC
CTGTTCCACAGACAAGAGTGAGATGCCCATCGGTATAATGATTGGAGAACAA
CATTCAAGGCTGTAAGTTATAATGCTGAAAGCCCCTTAATTTCTGGTAGTA
TTAGTTAAAGTTAAACACCTTTCCACCTTGAGTGTGAGAATTGTAGAGCAGT
GCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGTGCTGAGCAAT
GTGGCAGCCAGAGATCACAAGGCTATCAAGCATTGACATGGCAAGTGTAACTG
AGAACACACATTCAAATAATAGTTAATTGAATGTATCTAGCCATGTGTGG
CTAGTAGCTCCTTCCTGGAGAGAGAATCTGGAGCCACATCTAATTGTTAAGTCT
GGAATCTTATTTTATTCTGGAAAGGTCTATGAACTATAGTTGGGGCAGCTCA
CTTACTAACTTTAATGCAATAAGATCCATGGTATCTGAGAACATTATTGTCTCT
TTGTAGTACTGAAACCTTATACATGTGAAGTAAGGGCTATACTTAAGTCACATCT
CCAACCTTAGTAATGTTTAATGTTAGTAAAAAAATGAGTAATTAAATTATTAGA
AGGTCAATAGTATCATGATTCCAATAACAGAGGTATATGGTTAGAAAAGAAACA
ATTCAAAGGACTTATAATATCTAGCCTTGACAATGAATAAATTAGAGAGTAGTT
TGCTGTTGCCTCATAAATCTATTGACACATATGTGCATCTGCACCTCAGC
ATGGTAGAAGTCCATATCCTTGCTGGAAAGGCAGGTGTTCCATTACGCCCTCAG
AGAATAGCTGACGGGAAGAGGCTTCTAGATAGTTGATGAAAGATATAACAAATC
TCGCAGGTACACAGGCATGATTGCTGGTGGAGAGGCCATTAGATCTAGGAAC
CCCTAGTAGGGAGTTGCCACTCCCTCTCGCGCGCTCGCTCGCTCACTGAGGCCG
CCGGGCAAAGGCCGGCGTCGGCGACCTTGGTCGCCCGCCTCAGTGAGCGAG
CGAGCGCGCAGAGAGGGAGTGGCAA

```

Factor IX R338L polypeptide encoded in P00147

(SEQ ID NO: 702)

YNSGKLEEFVQGNLERECMEEKCSFEEAREVFENTERTEFWKQYVDGDQCESNPCLNGSCK

DDINSYEWCPGFEGKNCELDVTCNIKNGRCEQFKNSADNKVVCSCTEGYRLAENQKSCEPA

VPFFPCGRVSVSQTSKLTRAETVFPDVYVNSTEATILDNITQSFNDFTRVVGGEDAKPGQPP

- continued

WQVVLNGKVDAPCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIETEHTEQKRNVIRIIPHNN
YNAAINKYNHDIALLELDEPLVLSNYSVTPICIAKEYTNIFLKFGSGYVSGWGRVFHKGRSALVL
QYLRVPLVDRATCLLSTKFTIYNNMFCAGFHEGGRDSCQGDGGPHVTEVEGTSFLTGIISWGEE
CAMKGKYGIYTKVSRYVNWIKEKTKLT

Cas9 ORF
(SEQ ID NO: 703)
ATGGATAAGAAGTACTCAATCGGGCTGGATATCGGAACTAATTCCGTGGTTGGC
AGTGATCACGGATGAATACAAGTGCCGTCCAAGAACGTTCAAGGTCTGGGAACA
CCGATAGACACAGCATCAAGAAAAATCTCATCGGAGCCCTGCTGTTGACTCCGC
GAAACCGCAGAAGCAGACCCGGCTCAAACGTACCGCAGGCGACGCTACACCCGGC
GAAGAACGCTCATCTGCTATCTGCAAGAGATCTTCGAAACGAAATGGCAAAGGTC
ACGACAGCTTCTCCACCGCTGGAAGAATCTTCTGGTGGAGGAGGACAAGAAG
CATGAACGGCATCTATCTTGAAACATCGTCGACGAAGTGGCGTACACGAAA
GTACCCGACCATCTACCATCTCGGAAGAAGATCTTCTGGTGGAGGAGGACAAGAAG
ACCTCAGATTGATCTACTTGGCCCTGCCCATATGATCAAATTCCGGGACACTTCC
TGATCGAAGCGATCTGAACCTGATAACTCCGACGTGGATAAGCTTTCATCAAC
TGGTGCAGACCTACAACCAACTGTTGAAAGAAAACCAATCAATGCTAGCGCGTC
GATGCCAAGGCCATCTGTCCGCCGGCTGTGCAAGTCGCGGCCCTGAAAACCT
GATCGCACAGCTGCCGGAGAGAAAAAGAACGGACTTTCGGCAACTTGATCGCTC
TCTCACTGGGACTCACTCCAAATTCAAGTCAAATTGACCTGGCGAGGACGCGA
AGCTGCAACTCTCAAAGGACACCTACGACGACGACTGGACAATTGCTGGCACAA
ATTGGCGATCAGTACGCGGATCTGTTCTGGCTAAGAACCTTGGACGCAACT
TTGCTGTCGATATCCTGCGCGTGAACACCGAAAATAACCAAGCGCCCTAGGCC
TCGATGATTAAGCGGTACGACGAGCATCACCAGGATCTCACGCTGCTCAAAGCG
CGTAGAGACAGCAACTGCGTGGAAAGTACAAGGAGATCTTCTTGACCGAGTCAAAG
ATGGGTACGCGGGTACATCGATGGAGGCGCTAGCCAGGAAGAGTTCTATAAGTTC
ATCAAGCCAATCTGGAAAAGATGGACGGAACCGAACGAAACTGCTGGTCAAGCTGAA
CAGGGAGGATCTGTCGGAAACAGAGAACCTTGACAAACGGATCATTCCCACC
AGATCCATCTGGGTGAGCTGCACGCCATCTTGGCGGCCAGGAGGACTTTACCCAT
TCCTCAAGGACAACCGGGAAAAGATCGAGAAAATTCTGACGTTCCGCATCCGTATT
ACGTGGGCCACTGGCGCGGCCATTGGCGCTTGGATGACTAGAAAATCA
GAGGAACCGATCTGTCGGAAATTGAGGAAGTGTGGATAAGGGAGCTTGGC
ACAAAGCTTCAACGAATGACCAACTCGACAAGAACCTCCAAACGAGAAG
TGCTTCTAAGCACAGCCTCTTACGAAACTTCACTGTCTACAACGAACGACT
AAAGTGAATAACGTTACTGAAGGAATGAGGAAGCCGGCTTCTGTCCGGAGAACAG
AAGAAAGCAATTGTCGATCTGCTGTTCAAGACCAACCGAACGGTACCGTCAAGCA
GCTTAAAGAGGACTACTTCAAGAAGATCGAGTGTGGACTCAGTGGAAATCAGCG
GGGTGGAGGACAGATTCAACCGCTCGCTGGGAACCTATCATGATCTCTGAAGATCA
TCAAGGACAAGGACTTCTGACAACGAGGAACGAGGACATCTGGAAAGATATC
GTCTGACCTTGACCCCTTCTGAGGATCGCGAGATGATCGAGGAGAGGCTTAAGACC

- continued

TACGCTCATCTTCGACGATAAGGTATGAAACAACCACTCAAGCGCCGGTACACT
GGTTGGGGCCGCCTCTCCGCAAGCTGATCAACGGTATTGCGATAAACAGAGCGG
TAAAACATCCTGGATTCTCAAATCGGATGGCTTCGCTAATCGTAACCTCATGCA
ATTGATCCACGACGACAGCCTGACCTTAAGGAGGACATCCAAAAGCACAAGTGT
CCGGACAGGGAGACTCACTCCATGAACACATCGGAATCTGCCGGTTCGCCGGCG
ATTAAGAAGGGAACTCTGCAAACGTGAAGGTGGTCGACGAGCTGGTGAAGGTCA
GGGACGGCACAAACCGGAGAATATCGTATTGAAATGGCCCGAGAAAACCAGACTA
CCCAGAAGGGCCAGAAAAACTCCCGCAAAGGATGAAGCGGATCGAAGAAGGAAT
CAAGGAGCTGGCAGCCAGATCCTGAAAGAGCACCCGGTGGAAAACACGCAGCTG
CAGAACGAGAAGCTACCTGTACTATTGCAAATGGACGGGACATGTACGTGGA
CCAAGAGCTGGACATCAATCGTTGCTGATTACGACGTGGACCACATCGTCCACA
GTCCTTCTGAAGGATGACTCGATCGATAACAAGGTGTTGACTCGCAGCGACAAGA
ACAGAGGGAAAGTCAGATAATGTGCCATCGGAGGAGGTGTAAGAAGATGAAGAA
TTACTGGCGCGACTCTGAATGCGAAGCTGATTACCCAGAGAAAGTTGACAATCT
CACTAAAGCCGAGCGCGCGGACTCTAGAGCTGGATAAGGCTGGATTCAAC
GGCAGCTGGTCGAGACTCGCAGATTACCAAGCACGTGGCGAGATCTTGACTCC
CGCATGAACACTAAATCGACGAGAACGATAAGCTCATCCGGAAAGTGAAGGTGAT
TACCCCTGAAAAGCAAACCTTGTGTCGGACTTCCGAAAGGACTTCAGTTTACAAAGT
GAGAGAAATCAACAACCTACCATCACGCGCATGACGCATACTCAACGCTGTGGTC
GTACCGCCCTGATCAAAAGTACCTAAACTTGAATCGGAGTTGTGTAACGGAGACT
ACAAGGTCTACGACGTGAGGAAGATGATAGCCAAGTCCGAACAGGAATCGGAA
ACGCAACTGCGAAATACTCTTACTCAAACATCATGAACTTTCAAGACTGAAAT
TACGCTGGCCAATGGAGAAATCAGGAAGAGGCCACTGATCGAAACTAACGGAGAA
ACGGGCGAAATCGTGTGGACAAGGGCAGGGACTTCGCAACTGTTGCAAAGTGCT
CTCTATGCCGCAAGTCATATTGTGAAAGAAAAGCAGTGCACCGCGGATTTTC
AAAGGAATCGATCCTCCAAAGAGAAATAGCGACAAGCTCATTGCACGCAAGAAAG
ACTGGGACCCGAAAGAAGTACGGAGGATTGATCGCCACTGTCGATACTCCGTC
CTCGTGGTGGCCAAGGTGGAGAAGGGAAAGAGCAAAAGCTCAAATCCGTCAAAG
AGCTGCTGGGATTACCATCATGGAACGATCCTCGTTGAGAAGAAGAACCCGATTGATT
TCCTCGAGGCAGGGTTACAAGGGAGGTGAAGAAGGATCTGATCATCAAACCTCCC
AAGTACTCACTGTCGAACCTGAAAATGGTCGGAAGCGCATGCTGGCTTCGCCGG
AGAACTCCAAAAGGAAATGAGCTGGCTTGCGTAGCAAGTACGTCAAACCTCTCA
TCTTGCTTCGCACTACGAAAACCTAAAGGTCAACCGGAAGATAACGAACAGAAC
AGCTTTTCGTGGAGCAGCACAGCATTATCTGGATGAAATCATCGAACAAATCTCCG

- continued

AGTTTCAAAGCGGTGATCCTCGCCACGCCAACCTCGACAAAGTCCTGTCGGCT
ACAATAGCATAGAGATAAGCCGATCAGAGAACAGGGCGAGAACATTATCCACTTG
TTCACCCCTGACTAACCTGGAGCCCCAGCCGCTTCAGTACTTCGATACTACTATC
GATCGAAAAGATAACACGTCCACCAAGGAAGTCTGGACGCGACCCTGATCCACCA
AAGCATCACTGGACTCTACGAAACTAGGATCGATCTGTCGAGCTGGTGGCGAT
U-dep Cas9 ORF
(SEQ ID NO: 704)
ATGGACAAGAAGTACAGCATCGGACTGGACATCGAACAAACAGCGTCGGATGGC
AGTCATCACAGACGAATACAAGGTCGGAGCAAGAAGTTCAAGGTCTGGAAACA
CAGACAGACACAGCATCAGAAGAACCTGATCGGAGCACTGCTGTCGACAGCGG
GAAACAGCAGAAGCAACAAGACTGAAGAGAACAGCAAGAAGAACATACAAGAA
GAAAGAACAGAATCTGCTACCTGCAGGAAATCTTCAGCAACGAAATGGCAAAGGTC
GACGACAGCTTCTCCACAGACTGGAAGAAAGCTTCTGGTCAAGAACAGAACAGAA
GCACGAAAGACACCCGATCTCGGAAACATCGTCGACGAGTCGACATACCACGAAA
AGTACCCGACATCTACACCTGAGAAAGAACGCTGGTCGACAGCACAGAACAGCA
GACCTGAGACTGATCTACCTGGCACTGGCACACATGATCAAGTTCAAGGAGACTTC
CTGATCGAAGGGAGACCTGAACCCGGACAACAGCGACGTCGACAAGCTGTTCATCCA
GCTGGTCAGACATACAACCCAGCTGTTGAAGAAAACCCGATCAACGCAAGCGGAG
TCGACGCAAAGGCAATCCTGAGCGCAAGACTGAGCAAGAGCAGAACAGACTGGAAA
CCTGATCGCACAGCTGCGGGAGAAAAGAAGAACGGACTGTTCGGAAACCTGATCG
CACTGAGCTGGGACTGACACCGAACTTCAGAGCAACTTCGACCTGGCAGAACAG
GCAAAGCTGCAAGGACACATACGACGACGACCTGGACAACACCTGAGCGAC
ACAGATCGGAGACCACTGAGCGACACCTGTTCTGGCAGCAAGAACCTGAGCGAC
CAATCCTGCTGAGCGACATCCTGAGAGTCACACAGAAAATCAAAGGCACCGCTG
AGCGCAAGCATGATCAAGAGATACTGAGAACACACCAGGACCTGACACTGCTGAA
GGCACTGGTCAGACAGCAGCTGCCGGAAAAGTACAAGGAAATCTTCGACCCAGA
GCAAGAACGGATACGAGGATACTGACGGAGGAGCAAGCCAGGAAGAACCTTA
CAAGTTCATCAAGCCGATCCTGGAAAAGATGGACGGAACAGAGAACACTGCTGGTCA
AGCTGAACAGAGAACGACTGCTGAGAAAGCAGAGAACATTGACAAACGGAAGCAT
CCCGCACAGATCCACCTGGGAGAACTGCACGCAATCCTGAGAACAGAGAACAG
TCTACCCGTTCTGAGGACAACAGAGAAAAGATGAAAAGATCCTGACATTCA
ATCCCCTGACTACGTCGGACCGCTGGCAAGAGGAAACAGCAGGATTCGACATTGAC
AAGAAAAGAGCGAAGAACAAATCACACCGTGAACCTCGAAGAACGTCGACAAAG
GGAGCAAGCGCACAGAGCTCATCGAAAGAACGACAAACTTCGACAAAGAACCTGCC
GAACGAAAAGGTCTGCCGAAGCACGACACGCTGCTGTACGAATACTCACAGTCTACA
ACGAACTGACAAAGGTCAAGTACGTACAGAACAGGAATGAGAACAGCCGGCATTCTG
AGCGGAGAACAGAACAGAACAGGCAATCGTCGACCTGCTGTTCAAGAACAAACAGAACAG
TCACAGTCAGCTGAAGGAAGACTACTTCAGAACAGATCGAACATGCTTCGACAGC
GTCGAAATCAGCGGAGTCGAAGAACAGATTCAACGCAAGCCTGGAACATACCACGA
CCTGCTGAAGATCATCAAGGACAAGGACTTCCTGGACAAACGAAGAAAACGAAGACA

- continued

TCCTGGAAGACATCGCCTGACACTGACACTGTTCGAAGACAGAGAAATGATCGAA
GAAAGACTGAAGACATACGCACACCTGTCGACGACAAGGTATGAAGCGACTGAA
GAGAAGAAGATAACAGGATGGGAAGACTGAGCAGAAAGCTGATCAACGGAATC
AGAGACAAGCAGAGCGGAAAGACAATCCTGGACTTCTGAAAGAGCGACGGATTGCG
AACACAGAAACTCATGCAGCTGATCCACGACAGCCTGACATTCAAGGAAGACA
TCCAGAAGGCACAGGTCAAGGTCAGCGGACAGGGAGACAGCCTGACGAACACATCGCAAA
CCTGGCAGGAAGCCCGCAATCAAGAAGGGATCCTGCAGACAGTCAGGTCGTCG
ACGAACGGTCAAGGTATGGGAAGACACAAGCGGAAAACATGTCATCGAAATG
GCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAATG
AAAGAGAATCGAAGAAGGAATCAAGGAACGGGACTGGGAGGCCAGATCCTGAAAGGAACACC
CGGTCGAAAACACACAGCTGCAAGCAGAAAAGCTGTACCTGTACTACCTGAGAAC
GGAAGAGACATGTACGTCGACAGGAACTGGGACATCAACAGACTGAGCGACTACGA
CGTCGACCACATCGTCCCGCAGAGCTTCTGAAGGACGACAGCATCGACAACAAGG
TCCGTACAAGAAGCACAAGAACAGAGGAAAGAGCGACAACGTCCGAGCGAAGA
AGTCGTCAAGAAGATGAAGAACTACTGGGAGACAGCTGCTGAACGCAAAGCTGATCA
CACAGAGAAAGTTGACAACCTGACAAGGACAGAGGAGAGAGGAGACTGAGCGAACT
GCACAAAGGACAGGATTCATCAAGAGACAGCTGGTCAAGGACAGCATCGACAACAAG
CACGTGCAAGATCCTGGACAGCAGAATGAACACAAAGTACGACGAAAAGACA
AGCTGATCAGAGAAGTCAAGGTATCACACTGAAGAGCAAGCTGGTCAGCGACTTC
AGAAAGGACTTCCAGTTCTACAAGGTCAAGAGAAATCAACAACTACCAACCACGCACA
CGACGCATACCTGAACGCACTGTCGGAACAGCACTGATCAAGAAGTACCCGAAGC
TGGAAAGCGAATTGCTACGGAGACTACAAGGTCTACGACGTCAAGAAAGATGATC
GCAAAGCGAACAGGAAATCGGAAAGGCAACAGCAAAGTACTTCTACAGCAA
CATCATGAACTTCTCAAGACAGAAATCACACTGGCAACCGGAGAAATCAGAAAAGA
GACCGCTGATCGAACACAGCGGAGAAACAGGAGAAATCGTCTGGACAAGGGAAAG
AGACTTCGCAACAGTCAGAAAGGTCTGAGCATGCCGCAAGTCACATCGTCAAGA
AGACAGAAAGTCCAGACAGGAGGATTCAAGCAAGGAAAGCATCTGCCGAGAGAAA
CAGCGACAAGCTGATCGAACAGGAAAGAAGGACTGGGACCGAAGAAGTACGGAGGA
TTCGACAGCCGACAGTCGACATACAGCGCTGGTCGCAAGGTCGAAAGGGAG
AAAGAGCAAGAAGCTGAAGAGCGTAAGGAACGGACTGCTGGGAAATCACAATCATGGAA
AGAAGCAGCTCGAAAAGAACCCGATCGACTTCTGGAAAGCAAAGGGATAACAAGGA
AGTCAAGAAGGACTGATCATCAAGCTGCCGAAAGTACAGCCTGTTGAACTGGAAA
ACCGAAGAAAGAGAAATGCTGGCAAGCGCAGGAGAACTGCAGAAGGGAAAAGACA
GGCACTGCCGAGCAAGTACGTCAACTTCTGTACCTGGCAAGCCACTACGAAAAGC
TGAAGGGAAAGCCCGGAGACAACGAAACAGAAGCAGCTGTTCGAACAGCACAA
GCACTACCTGGACGAAATCATCGAACAGATCAGCGAATTCAAGCAAGAGAGTCATCC
TGGCAGACGCAAACCTGGACAAGGTCTGAGCGCATACAACAAGCAGAGACAA
GCCGATCAAGAGAACAGGAGAAAACATCATCCACCTGTTCACACTGACAAACCTGG
GAGCACCAGGACATTCAAGTACTTCGACACAACAATCGACAGAAAGAGATAACACA

- continued

ACGCACAAAGGAAGTCTGGACCAACACTGATCCACCAGAGCATCACAGGACTGTA

CGAAACAAGAACGACCTGAGCCAGCTGGGAGGAGACGGAGGAGGAAGCCCCGAAG

AAGAAGAGAAAAGGTCTAG

mRNA comprising U dep Cas9

(SEQ ID NO: 705)

GGGUCCCCCAGUCGGCGUCCAGCGGCUUGCUUGUUCGUGUGUGUGUCGUUGCAG

GCCUUUUUCGGAUCCGCCACCAUAGGACAAGAACGUACAGCAUCGGACUGGACAUCG

GAACAAAACAGCGUCGGAUGGGAGUCAUCACAGACAGCAUCAAGGUCCCGAGCAA

GAAGGUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAACCCUGAU

CGGAGCACUGCGUUCGACAGCGAGAACAGCAGAACAGCAAGACUGAAGAG

AACAGCAAGAACAGAACACAGAACAGAACAGAACUUCGUACUGCAGGA

AAUCUUCAGCAACGAAAUGGCAAAGGUUCGACAGACAGCUUCUCCACAGACUGGAA

GAAAGCUUCCUGGUCGAAGAACAGAACAGAACAGCACCCGAUCUUCGGA

AAACAUUGUCGACGAUGUCGCAUACCACGAAAGUACCCGACAAUCAUCCUGA

GAAAAGCUGGUUCGACAGCACAGAACAGCACCCGAGACUGAUCUACCCUG

ACUGGCACACAUGAUCAAGUUCAGAGGACACUUCUGAUCGAAGGAGACUGAAC

CCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACC

AGCUGUUCGAAAGAAAACCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCU

GAGCGCAAGACUGAGCAAGAGCAGAACAGUGAAACCUUGAUCGCACAGCUGCCG

GGAGAAAAGAAGAACCGACUGUUCGGAAACCUUGAUCGCACUGAGCUGGGACUG

ACACCGAACUUAAGAGCAACUUCGACCUGGCAGAACAGCAGCAAAGCUGCAGCUGA

GCAAGGGACACAUACGACGACGACCUGGACAACCUGCUGGCCACAGAACGGAGCCA

GUACCGAGACCUGUUCCUGGCAGCAAAGAACCUUGAGCGACCCAAUCCUGCUGAGC

GACAUCUGAGAGUACACAGAAAACACAGGACCGCUGAGCGCAAGCAUGA

UCAAGAGAUACGACGAACACCACAGGACCUACACUGCUGAAGGCACUGGUCAG

ACAGCAGCUGCCGAAAAGUACAAAGGAAAUCUUCUUCGACCGAGAGCAAGAACGG

AUACGCAGGAUACAUUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAU

CAAGCCGAUCCUGGAAAAGAUGGACCGAACAGAACAGACUGCUGGUCAAGCUGAA

CAGAGAAGACCUGCUGAGAAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGAC

CAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAACAGACAGGAAGACUUCUACC

CGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCC

CGUACUACGUCCGGACCGCUGGCCAGAGGAAACAGCAGAUUCGCAUGGAUGACAA

GAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAACUGCUGCAGAACGG

GAGCAAGCGCACAGAGCUUCAUCGAAAGAACAGAACUUCGACAAGAACUGCC

GAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGCUAC

AACGAACUGACAAAGGUCAAGUACGUCACAGAACAGGAAGAACAGCCGGCAUUC

CUGAGCGGAGAACAGAACAGAACUGCUGACAGCAGAACAGAACAGAACAGAAC

AAGGUACAGUCAAGCAGCUGAAGGAAGACUACUCAAGAACAGAACUGAAUUC

GACAGCGUGCAAUCAGCGGAGUCGAGAACAGAACUACGCAAGCCUGGGAAACA

UACCACGACCUGCUGAAGAACAUCAAGGACAAGGACAUUCUGGACAACGAAGAAA

- continued

ACGAAGACAUCUGGAAGACAUCGUCCUGCACUGACACUGUUCGAAAGACAGAGA
AAUGAUCGAAGAAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAU
GAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGAAAGACUGAGCAGAAAGCU
GAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCUGGAACUCCUGAA
GAGCGACGGAUUCGCAAACAGAAAACUUCAUGCAGCUGAUCACGACGCCUG
ACAUUCAAGGAAGACAUCAGAGCACAGGUACAGGGACAGGGAGACAGCCUGC
ACGAACACAUCGCAAACCUGGCAGGAAGCCCAGCAAUCAAGAAGGGAAUCUGCA
GACAGUCAAGGUUCGACGAAUCUGGUCAAGGUCAUGGGAAAGACACAAGCCGGA
AAACAUUCGUCAUCGAAUUGGCAAGAGAAAACAGACAACACAGAAGGGACAGAA
GAACAGCAGAGAAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAG
CCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAGCUG
UACCUGUACUACUCUGCAGAACCGGAAGAGACAUGUACUGUCCAGGAACUGGACA
UCAACAGACUGAGCGACUACGACGUCCACACUGGUCCCGCAGACGUUCCUGAA
GGACGACAGCAUCGACAACAAGGUCCUGACAAGAACGACAAAGAACAGAGGAAA
GAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAAGAUGAACUACUGGAG
ACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAG
GCAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUCUCAAGAGACAG
CUGGUCAAGAACAGACAGAUCAACAAAGCACGUCCAGAACUCCUGGACAGCAGAA
UGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAAGUCAAGGUCAUCA
CACUGAAGAGCAAGCUGGUACGCGACUUCAGAAAAGGACUUCAGUUCUACAAGG
UCAGAGAAAUCACACUGGCAAACGGGAGAAAUCAGAAAAGAGACCGCUGAUCG
CGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUUACCG
AGACUACAGGUUCACGACGUCAAGAACUACUGCAGAACAGCGAACAGGAAA
CGGAAAGGCAACAGCAAAGUACUUCUUCACAGCAACAUCAUGAACUUCUUAAG
ACAGAAAUCACACUGGCAAACGGGAGAAAUCAGAAAAGAGACCGCUGAUCGAAACA
AACGGGAGAAACAGGAGAAAUCGUUCGGACAAGGGAGAGACUUCGCAACAGUC
AGAAAGGUCCUGAGCAUGCCGAGGUCAACAUUCGUCAAGAAGAGACAGGUCCAG
ACAGGAGGAUCAGCAAGGAAGCAUCUGCCGAAGAGAAAACAGCGACAAGCUG
AUCGCAAGAAAAGGACUGGGACCCGAAGAAGUACGGAGGAUCAGCCCG
ACAGUCGCCAUACAGCGUCCUGGUCCUGCGAACAGGCGAAAAGGGAAAGAGCAAG
AAGCUGAAGAGGUCAAGGAACUGCUGGGAAUCACAUCAUGGAAAGAAGCAGC
UUCGAAAAGAACCGAUCGACUUCUGGAAGCAAAGGGAUACAAGGAAGUCAAG
AAGGACCUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAAACUGGAAAACGGA
AGAAAGAGAAUCGUCCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAAUCUG
CUGCCGAGCAAGUACGUCAACUUCUGUACCCUGGCAAGCCACUACGAAAGCUGA
AGGGAAAGCCCGAAGACAACGAAACAGAAGCAGCUGUUCGACAGCACAAGCA
CUACCUUGGACGAAAUCUGAAGAACUGCAGCGAAUUCAGCAAGAGAGUCAUCC
GCAGACGCAAACCUUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGC
CGAUCAGAGAACAGGCAAGAAACAUCAUCCACCUUCACUGACAAACCUGG

- continued

AGCACCCGGCAGCAUUCAGAACUUUCGACACAACAAUCGACAGAGAAGAGAUACACA
AGCACAAAGGAAGGUCCUGGACGCAACACUGAUCACCAGAGCAUCACAGGACUGU
ACGAAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAGGAGGAAGCCCGA
AGAAGAAGAGAAAGGCUAGCUAGCCAUCACAUUUAAAAGCAUCUCAGCCUACC
AUGAGAAUAGAGAAAGAAAUGAAGAUCAAUCGUUAUCAUCUCUUUUUCUU
UUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUCUUUAUCA
UUUUGCCUCUUUCUGUGCUCAUUAAAUGGAAAGAACCUUCGAGA
AA
AA

SEQUENCE LISTING

Sequence total quantity: 1130

SEQ ID NO: 1 moltype = DNA length = 709
FEATURE Location/Qualifiers
source 1..709
mol_type = other DNA
organism = Homo sapiens

SEQUENCE: 1
gtaagaatc cattttctta ttgttcaact tttattctat tttccagta aaataaaggta 60
tttagaaact ctgcatctt aaagaattat ttggcattt atttctaaaa tggcatagta 120
ttttgtattt gtgaagtctt acaaggttat cttattaata aaattcaaac atcctaggta 180
aaaaaaaaaa aaggtcagaa ttgttagtg actgtat tcttttcgcg actaaggaaa 240
gtgcagaatgactttagatc actgaaactt cacagaatag ggttgaatg tgaatttcata 300
actatccccaa agacatccatc attgcactt gcttatttta aaaaccacaa aacctgtgt 360
gttgatctca taaatagaaac ttgttattt atttattttc atttttgtct gtcttcttgg 420
ttgctgttga tagacactaa aagagtat gatattatct aagtttgaat ataaggctat 480
aaatatttaa taatttttaa aatgtatc ttggtatgg aattttctt ctgtttaaag 540
gcagaagaaa taatttgaaca tcatcttgat tttttctgtt ggaatcagag cccaaatattt 600
tgaacaaat gcataatcta agtcaaatgg aaagaaatataaaaatgttac atttactt 660
cttggtttct tcagttttaa acaatccccc tttttcttcc ctggcccg 709

SEQ ID NO: 2 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 2
gagcaacctc actctttgtct 20

SEQ ID NO: 3 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 3
atgcattttt ttc当地at 20

SEQ ID NO: 4 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 4
tgcattttgtt tcaaaaatatt 20

SEQ ID NO: 5 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 5
atttatgaga tcaacagcac 20

SEQ ID NO: 6 moltype = RNA length = 20

-continued

FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 6	
gatcaacagc acagggtttg	20
SEQ ID NO: 7	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 7	
ttaaataaaag catagtgcaa	20
SEQ ID NO: 8	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 8	
taaagcatag tgcaatggat	20
SEQ ID NO: 9	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 9	
tagtgcaatg gataggtctt	20
SEQ ID NO: 10	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 10	
tactaaaact ttattttact	20
SEQ ID NO: 11	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 11	
aaaggttgaac aatagaaaaaa	20
SEQ ID NO: 12	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 12	
aatgcataat ctaagtcaaa	20
SEQ ID NO: 13	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 13	
taataaaaatt caaacatcct	20
SEQ ID NO: 14	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 14	
gcatctttaa agaattttt	20
SEQ ID NO: 15	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct

-continued

SEQUENCE: 15		
tttggcattt atttctaaaa		20
SEQ ID NO: 16	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 16		
tgtattttgtg aagtcttaca		20
SEQ ID NO: 17	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 17		
tccttagtaa aaaaaaaaaaa		20
SEQ ID NO: 18	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 18		
taattttctt ttgcgcacta		20
SEQ ID NO: 19	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 19		
tgactgaaac ttcacagaat		20
SEQ ID NO: 20	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 20		
gactgaaact tcacagaata		20
SEQ ID NO: 21	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 21		
ttcatttttag tctgtttct		20
SEQ ID NO: 22	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 22		
attatctaaat tttgaatata		20
SEQ ID NO: 23	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 23		
aatttttaaa atagttattct		20
SEQ ID NO: 24	moltype = RNA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 24		
tgaatttttc ttctgtttaa		20
SEQ ID NO: 25	moltype = RNA length = 20	

-continued

FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 25	
atcatcctga gtttttctgt	20
SEQ ID NO: 26	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 26	
ttaactaaac ttttatTTTAC	20
SEQ ID NO: 27	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 27	
acctttttt ttTTTTACCT	20
SEQ ID NO: 28	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 28	
agtgcataatgg ataggTCTTT	20
SEQ ID NO: 29	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 29	
tgattccTAC agaaaaaACTC	20
SEQ ID NO: 30	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 30	
tgggcaaggg aagaaaaaaaaa	20
SEQ ID NO: 31	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 31	
cctcaCTT GTCTGGGCAA	20
SEQ ID NO: 32	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 32	
acctcaCTC TGTCTGGGCA	20
SEQ ID NO: 33	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 33	
tGAGCACCT CACTTTGTC	20
SEQ ID NO: 34	moltype = RNA length = 100
FEATURE	Location/Qualifiers
source	1..100
	mol_type = other RNA
	organism = synthetic construct

-continued

```

SEQUENCE: 34
gagcaaccc actcttgctc gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 35      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 35
atgcatttgtt tccaaaatat gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 36      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 36
tgcatattttt tccaaaatatt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 37      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 37
atttatgaga tcaacagcac gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 38      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 38
gatcaaacagc acagggtttt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 39      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 39
ttaaataaag catagtgc当地 gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 40      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 40
taaaggcatag tgcaatggat gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 41      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 41
tagtgcaatg gataggctt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 42      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 42
tactaaaact ttatttact gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

```

-continued

```

SEQ ID NO: 43      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 43
aaagttgaac aatagaaaaaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 44      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 44
aatgcataat ctaagtcaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 45      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 45
taataaaaatt caaacatcct gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 46      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 46
gcatctttaa agaattttt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 47      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 47
tttgcattt atttctaaaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 48      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 48
tgtatttttg aagtcttaca gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 49      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 49
tccttagttaa aaaaaaaaaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 50      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 50
taattttctt ttgcgcacta gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 51      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA

```

-continued

```
SEQUENCE: 51          organism = synthetic construct
tgactgaaac ttccagaat gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 52          moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 52
gactgaaact tcacagaata gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cggttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 53          moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 53
ttcatttttag tctgtttct gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 54          moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 54
attatctaaat ttgttataata gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cggttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 55          moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 55
aatttttaaa atatgtttct gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 56          moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 56
tgaatttttc ttctgtttaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 57          moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 57
atcatctga gttttctgt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 58          moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 58
ttactaaac ttatatttac gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 59          moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 59
accttttttt tttttcacct gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100
```

-continued

```

SEQ ID NO: 60      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 60
atgtcaatgg ataggcttt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 61      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 61
tgattccatc agaaaaactc gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 62      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 62
tgggcaagg aaaaaaaaaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 63      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 63
cctcaactt gtctggcaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 64      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 64
acctcaactt tgtctggca gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 65      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 65
tgagcaacct cactttgtc gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 66      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 66
gagcaacctc actcttgtc gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 67      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 67
atgcatttgtt ttcaaaatat gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 68      moltype = RNA  length = 100
FEATURE
source
1..100

```

-continued

```

mol_type = other RNA
organism = synthetic construct
SEQUENCE: 68
tgcattttgtt tcaaaaattt gtttttagago tagaaatagc aagttaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 69      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 69
atttatgaga tcaacagcac gtttttagago tagaaatagc aagttaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 70      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 70
gatcaacagc acagggtttt gtttttagago tagaaatagc aagttaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 71      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 71
ttaataaaag catagtgc当地 gtttttagago tagaaatagc aagttaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 72      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 72
taaagcatag tgcaatggat gtttttagago tagaaatagc aagttaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 73      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 73
tagtgcaatg gataggcttt gtttttagago tagaaatagc aagttaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 74      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 74
tactaaaact ttatttact gtttttagago tagaaatagc aagttaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 75      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 75
aaagttgaac aatagaaaaa gtttttagago tagaaatagc aagttaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 76      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 76
aatgcataat ctaagtcaa gtttttagago tagaaatagc aagttaaaat aaggctagtc 60

```

-continued

```

cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100
SEQ ID NO: 77      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 77
taataaaaatt caaacatcct gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 78      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 78
gcacatcttaa agaatttattt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 79      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 79
tttggcattt atttctaaaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 80      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 80
tgtatatttg aagtcttaca gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 81      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 81
tccttagttaa aaaaaaaaaaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 82      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 82
taattttctt ttgcgcacta gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 83      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 83
tgactgaaac ttcacagaat gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 84      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 84
gactgaaact tcacagaata gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 85      moltype = RNA  length = 100
FEATURE           Location/Qualifiers

```

-continued

```

source          1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 85
ttcatttttag tctgttttct gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 86      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 86
attatctaag tttgaatata gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 87      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 87
aattttaaa atagttttct gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 88      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 88
tgattttttc ttctgtttaa gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 89      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 89
atcatctga gtttttctgt gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 90      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 90
ttactaaaac ttattttac gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 91      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 91
acctttttt ttttttacct gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 92      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 92
agtgcatagg ataggcttt gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 93      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 93

```

-continued

```

tgattcctac agaaaaactc gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 94      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 94
tgggcaagg aaaaaaaaaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 95      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 95
cctcacttct gtctggcaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 96      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 96
acctcactct tgcgtggca gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 97      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 97
ttagcaacct cactttgtc gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 98      moltype = RNA length = 20
FEATURE           Location/Qualifiers
source            1..20
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 98
atttgcatct gagaaccctt 20

SEQ ID NO: 99      moltype = RNA length = 20
FEATURE           Location/Qualifiers
source            1..20
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 99
atcgggaact ggcatttca 20

SEQ ID NO: 100     moltype = RNA length = 20
FEATURE           Location/Qualifiers
source            1..20
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 100
gttacagggaa aatctgaagg 20

SEQ ID NO: 101     moltype = RNA length = 20
FEATURE           Location/Qualifiers
source            1..20
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 101
gatcgggaac tggcatttc 20

SEQ ID NO: 102     moltype = RNA length = 20
FEATURE           Location/Qualifiers
source            1..20
                  mol_type = other RNA
                  organism = synthetic construct

```

-continued

SEQUENCE: 102 tgcatctgag aacccttagg	20
SEQ ID NO: 103 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 103 cactcttgtc tgtggaaaca	20
SEQ ID NO: 104 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 104 atcggtacag gaaaatctga	20
SEQ ID NO: 105 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 105 gcatcttcag ggagtagctt	20
SEQ ID NO: 106 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 106 caatcttaa atatgttgta	20
SEQ ID NO: 107 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 107 tcactcttgt ctgtggaaac	20
SEQ ID NO: 108 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 108 tgcttgatt tttctagtaa	20
SEQ ID NO: 109 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 109 gtaaatatct actaagacaa	20
SEQ ID NO: 110 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 110 tttttctagt aatggaagcc	20
SEQ ID NO: 111 FEATURE source moltype = RNA length = 20 Location/Qualifiers 1..20 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 111 tttatattttt gatataatttt	20
SEQ ID NO: 112 moltype = RNA length = 20	

-continued

FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 112	
gcacagatat aaacacttaa	20
SEQ ID NO: 113	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 113	
cacagatata aacacttaac	20
SEQ ID NO: 114	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 114	
ggttttaaaa ataataatgt	20
SEQ ID NO: 115	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 115	
tcagatttcc ctgttaacgat	20
SEQ ID NO: 116	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 116	
cagattttcc tgtaaacgatc	20
SEQ ID NO: 117	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 117	
caatggtaaa taagaataaa	20
SEQ ID NO: 118	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 118	
ggaaaatctg aagggtggcaa	20
SEQ ID NO: 119	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 119	
ggcgatctca ctcttgtctg	20
SEQ ID NO: 120	moltype = RNA length = 100
FEATURE	Location/Qualifiers
source	1..100
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 120	
atttgcacatct gagaaccctt gtttttagago tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt	100
SEQ ID NO: 121	moltype = RNA length = 100
FEATURE	Location/Qualifiers
source	1..100
	mol_type = other RNA

-continued

```

SEQUENCE: 121          organism = synthetic construct
atcgggaact ggcacattca gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 122          moltype = RNA   length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 122
gttacaggaa aatctgagg gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 123          moltype = RNA   length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 123
gatcggaaac tggcatcttc gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 124          moltype = RNA   length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 124
tgcatctgag aacccttagg gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 125          moltype = RNA   length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 125
cactcttgtc tgtggaaaca gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 126          moltype = RNA   length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 126
atcgttacag gaaaatctga gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 127          moltype = RNA   length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 127
gcatcttcag ggagtagctt gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 128          moltype = RNA   length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 128
caatcttaa atatgttgt gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 129          moltype = RNA   length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 129
tcactcttgt ctgtggaaac gtttttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

```

-continued

```

SEQ ID NO: 130      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 130
tgcttgatt tttcttagaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 131      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 131
gtaaatatct actaagacaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 132      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 132
tttttctagt aatggaaagcc gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 133      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 133
ttatattattt gatatatatttt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 134      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 134
gcacagatata aacacttaac gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 135      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 135
cacagatata aacacttaac gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 136      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 136
ggttttaaaa ataataatgt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 137      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 137
tcagatttc ctgtaacgt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 138      moltype = RNA  length = 100
FEATURE
source
1..100

```

-continued

```

mol_type = other RNA
organism = synthetic construct
SEQUENCE: 138
cagattttcc tgttaacgatc gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 139      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 139
caatggtaaa taagaataa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 140      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 140
ggaaaatctg aaggtgccaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 141      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 141
ggcogatctca ctcttgcctg gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 142      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 142
atttgcatct gagaaccctt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 143      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 143
atcgggaact ggcatctca gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 144      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 144
tttacaggaa aatctgaagg gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 145      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 145
gatcgggaac tggcatcttc gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 146      moltype = RNA  length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 146
tgcatctgag aacccttagg gtttagago tagaaatagc aagtaaaaat aaggctagtc 60

```

-continued

```

cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100
SEQ ID NO: 147      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 147
cactcttgc tgtggaaaca gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 148      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 148
atcggttacag gaaaatctga gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 149      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 149
gcacatcttag ggagtagctt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 150      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 150
caaatcttaa atatgttgt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 151      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 151
tcactcttgt ctgtggaaac gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 152      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 152
tgcttgatt tttctagtaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 153      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 153
gttaatatct actaagacaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 154      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 154
tttttctagt aatggaaagcc gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt          100

SEQ ID NO: 155      moltype = RNA  length = 100
FEATURE           Location/Qualifiers

```

-continued

```

source          1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 155
ttatattatt gatatatttt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 156      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 156
gcacagatat aaacacttaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 157      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 157
cacagatata aacacttaac gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 158      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 158
ggttttaaa ataataatgt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 159      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 159
tcagatttc ctgttaacgat gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 160      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 160
cagattttcc tgtaacgatc gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 161      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 161
caatggtaaa taagaataa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 162      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 162
ggaaaatctg aagggtggcaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                         100

SEQ ID NO: 163      moltype = RNA  length = 100
FEATURE           Location/Qualifiers
source            1..100
               mol_type = other RNA
               organism = synthetic construct
SEQUENCE: 163

```

-continued

```

ggcgatctca ctcttgtctg gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaagt ggcaccgagt cggtgccttt 100

SEQ ID NO: 164      moltype = length =
SEQUENCE: 164
000

SEQ ID NO: 165      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 165
agcaacctca ctcttgtctg 20

SEQ ID NO: 166      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 166
acctcaactc tgtctgggaa 20

SEQ ID NO: 167      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 167
cctcaactt gtctgggaa 20

SEQ ID NO: 168      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 168
ctcaacttg tctgggaaag 20

SEQ ID NO: 169      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 169
gggaagggg agaaaaaaaaa 20

SEQ ID NO: 170      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 170
gggaaggggg gaaaaaaaaa 20

SEQ ID NO: 171      moltype = length =
SEQUENCE: 171
000

SEQ ID NO: 172      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 172
tgcatttgtt tcaaaatatt 20

SEQ ID NO: 173      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 173
tgattcctac agaaaaagtc 20

SEQ ID NO: 174      moltype = RNA length = 20

```

-continued

FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 174	
tacagaaaaa gtcaggataa	20
SEQ ID NO: 175	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 175	
tttcttgc cttaaacag	20
SEQ ID NO: 176	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 176	
ttagatttt atattcaaac	20
SEQ ID NO: 177	moltype = length =
SEQUENCE: 177	
000	
SEQ ID NO: 178	moltype = length =
SEQUENCE: 178	
000	
SEQ ID NO: 179	moltype = length =
SEQUENCE: 179	
000	
SEQ ID NO: 180	moltype = length =
SEQUENCE: 180	
000	
SEQ ID NO: 181	moltype = length =
SEQUENCE: 181	
000	
SEQ ID NO: 182	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 182	
agtgcataatgg ataggcctta	20
SEQ ID NO: 183	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 183	
ttactttgca ctttccttag	20
SEQ ID NO: 184	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 184	
tactttgcac tttccttagt	20
SEQ ID NO: 185	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 185	
tctgacccctt tattttacct	20
SEQ ID NO: 186	moltype = length =

-continued

SEQUENCE: 186
000

SEQ ID NO: 187 moltype = length =
SEQUENCE: 187
000

SEQ ID NO: 188 moltype = length =
SEQUENCE: 188
000

SEQ ID NO: 189 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 189
attatcctga ctttttctgt 20

SEQ ID NO: 190 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 190
tgaatttttc ctctgtttaa 20

SEQ ID NO: 191 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 191
taattttctt ttgccccacta 20

SEQ ID NO: 192 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 192
aaaaggtagt aattgtttag 20

SEQ ID NO: 193 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 193
aacatccatgt gtaaaaataaa 20

SEQ ID NO: 194 moltype = length =
SEQUENCE: 194
000

SEQ ID NO: 195 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 195
tttgtcatgtat tttctaaaaat 20

SEQ ID NO: 196 moltype = RNA length = 20
FEATURE Location/Qualifiers
source 1..20
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 196
tttgtcatgtat tttctaaaaa 20

SEQ ID NO: 197 moltype = length =
SEQUENCE: 197
000

SEQ ID NO: 198 moltype = RNA length = 100

-continued

```

FEATURE          Location/Qualifiers
source           1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 198
agaaacctca ctcttgtctg gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 199      moltype = RNA length = 100
FEATURE          Location/Qualifiers
source           1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 199
acactcactct tgcgtgggaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 200      moltype = RNA length = 100
FEATURE          Location/Qualifiers
source           1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 200
cctcactctt gtctggggaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 201      moltype = RNA length = 100
FEATURE          Location/Qualifiers
source           1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 201
ctcactcttgcgtggggaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 202      moltype = RNA length = 100
FEATURE          Location/Qualifiers
source           1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 202
gggaagggg agaaaaaaaaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 203      moltype = RNA length = 100
FEATURE          Location/Qualifiers
source           1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 203
gggaagggg agaaaaaaaaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 204      moltype = length =
SEQUENCE: 204
000

SEQ ID NO: 205      moltype = length =
SEQUENCE: 205
000

SEQ ID NO: 206      moltype = RNA length = 100
FEATURE          Location/Qualifiers
source           1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 206
tgattccatc agaaaaaagt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt                100

SEQ ID NO: 207      moltype = RNA length = 100
FEATURE          Location/Qualifiers
source           1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 207

```

-continued

```

tacagaaaaa gtcaggataa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 208      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 208
ttcttcgtc cttaaacag gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 209      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 209
ttatagttt atattcaaac gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 210      moltype = length =
SEQUENCE: 210
000

SEQ ID NO: 211      moltype = length =
SEQUENCE: 211
000

SEQ ID NO: 212      moltype = length =
SEQUENCE: 212
000

SEQ ID NO: 213      moltype = length =
SEQUENCE: 213
000

SEQ ID NO: 214      moltype = length =
SEQUENCE: 214
000

SEQ ID NO: 215      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 215
atgtcaatgg ataggctta gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 216      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 216
ttacttgca ctttccttag gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 217      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 217
taccttgac tttccttagt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 218      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 218
tctgacctt tatttacct gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

```

-continued

```
SEQ ID NO: 219      moltype = length =
SEQUENCE: 219
000

SEQ ID NO: 220      moltype = length =
SEQUENCE: 220
000

SEQ ID NO: 221      moltype = length =
SEQUENCE: 221
000

SEQ ID NO: 222      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 222
attatcctga cttttctgt gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 223      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 223
tgaaattttt ctctgtttaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 224      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 224
taatttttctt ttgcccaacta gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 225      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 225
aaaagggttag aattgttttag gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 226      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 226
aacatccttag gtaaaaataaa gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 227      moltype = length =
SEQUENCE: 227
000

SEQ ID NO: 228      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 228
ttgtcatgtat tttctaaat gtttagago tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 229      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
```

-continued

```

SEQUENCE: 229
tttgcatgt atttctaaaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 230      moltype = length =
SEQUENCE: 230
000

SEQ ID NO: 231      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 231
acgaacctca ctcttgctg gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 232      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 232
acctcactct tgcgtgggaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 233      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 233
ccctcactctt gtctggggaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 234      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 234
ctcaactcttgcgtggggaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 235      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 235
ggggaaaggaa agaaaaaaaaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 236      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 236
ggggaaaggaa agaaaaaaaaa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 237      moltype = length =
SEQUENCE: 237
000

SEQ ID NO: 238      moltype = length =
SEQUENCE: 238
000

SEQ ID NO: 239      moltype = RNA length = 100
FEATURE
source
1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 239

```

-continued

```

tgattcctac agaaaaagtc gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 240      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 240
tacagaaaaa gtcaggataa gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 241      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 241
tttcttctgc cttaaacag gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 242      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 242
ttatagttt atattcaac gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 243      moltype = length =
SEQUENCE: 243
000

SEQ ID NO: 244      moltype = length =
SEQUENCE: 244
000

SEQ ID NO: 245      moltype = length =
SEQUENCE: 245
000

SEQ ID NO: 246      moltype = length =
SEQUENCE: 246
000

SEQ ID NO: 247      moltype = length =
SEQUENCE: 247
000

SEQ ID NO: 248      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 248
agtgcataatgg ataggctta gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 249      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 249
ttactttgca ctttccttag gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 250      moltype = RNA length = 100
FEATURE           Location/Qualifiers
source            1..100
                  mol_type = other RNA
                  organism = synthetic construct
SEQUENCE: 250
tactttgcac tttccttagt gtttagago tagaaatagc aagtaaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

```

-continued

SEQ ID NO: 251 moltype = RNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 251 tctgaccc ttatttaccc gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 252 moltype = length =
SEQUENCE: 252 000

SEQ ID NO: 253 moltype = length =
SEQUENCE: 253 000

SEQ ID NO: 254 moltype = length =
SEQUENCE: 254 000

SEQ ID NO: 255 moltype = RNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 255 attatccctga cttttctgt gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 256 moltype = RNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 256 tgaattttc ctctgtttaa gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 257 moltype = RNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 257 taattttctt ttgcccacta gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 258 moltype = RNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 258 aaaaggtcag aattgttttag gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 259 moltype = RNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 259 aacatccctag gtaaaaataaa gtttttagago tagaaatagc aagttaaaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgcttt 100

SEQ ID NO: 260 moltype = length =
SEQUENCE: 260 000

SEQ ID NO: 261 moltype = RNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other RNA
organism = synthetic construct

-continued

SEQUENCE: 261
ttgtcatgt tttctaaaat gtttagagc tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcctt 100

SEQ ID NO: 262 moltype = RNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other RNA
organism = synthetic construct

SEQUENCE: 262
ttgtcatgt atttctaaaat gtttagagc tagaaatagc aagttaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgcctt 100

SEQ ID NO: 263 moltype = DNA length = 145
FEATURE Location/Qualifiers
source 1..145
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 263
ttggccactc cctctctgcg cgctcgctcg ctcaactgagg ccggcgacc aaaggtcgcc 60
cgacccggc gcttgcccg ggccgcctca gtgacgagc gagcgcgcag agagggagtg 120
gccaactcca tcacttagggg ttct 145

SEQ ID NO: 264 moltype = DNA length = 100
FEATURE Location/Qualifiers
source 1..100
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 264
taggtcaatg aagagaagaa caaaaagcacatattacag tttagtgtct tcatcaatct 60
ttaaatatgt tttgtgtt ttctctccct gttccacag 100

SEQ ID NO: 265 moltype = DNA length = 1296
FEATURE Location/Qualifiers
source 1..1296
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 265
tttcttgatc atgaaaacgc caacaaaatt ctgaatccgc caaagaggta taattcagg 60
aaatttggaaag agttttgttca agggAACCTT gagagagaat gtatggaaaga aaagtgttagt 120
tttggaaagac cacggaaatgt ttttgaaac actgaagaa caactgtttaa ttggaaacgg 180
tatgttgcgat gagatcgatc tgatgtccaaat cctatgtttaa atggccgcg ttggaaaggat 240
gacattaaatt cctatgtatc ttggatgtccc ttggatgttggaaaggaaat ctgtgat 300
gtatgttaatc gtaacattaa gaatggcaga tgccggcgtt tttgttggaaat tagtgtctgt 360
aacaagggtgg tttgtccctg tactggggata tttgttggaaat cttgttggaaat 420
aaaccggcgc tgccatgttcc atgtggaaat gtttggatgtt cacaacttcc taatgttcc 480
cgtgtgttgcatc ttggatgttcc tgatgtggac tatgttggaaat ctactggaaatc tggatgttcc 540
ttggataaca tcactcaaaat caccatca tttaaatgtact tcactccggg tttgttggaa 600
gaagatgttca aaccagggtca atcccttgg cttgttggaaat tttgttggaaat agttgttgc 660
ttctgtgttgc gcttctatgt tttgttggaaat tttgttggaaat ctgttccccc tttgttggaa 720
actgtgttgc tttgttggaaat tttgttggaaat cttgttggaaat cttgttggaaat 780
gagccatgttca gaaatgttgcatc ttggatgttgc tatgttggaaat cttgttggaaat 840
aagtacaacc atgacatttc cttttgttgc ttggacggaa ctggacggaa acatgttgc 900
tttacacca tttgttggaaat ttggatgttgc ttggatgttgc ttggatgttgc 960
gttacatgttca gtttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc 1020
tttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc 1080
tataacaaca tttgttggaaat ttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc 1140
atgtggggac cccatgttca tttgttggaaat ttggatgttgc ttggatgttgc ttggatgttgc 1200
ttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc 1260
gttacatgttca gtttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc ttggatgttgc 1296

SEQ ID NO: 266 moltype = DNA length = 276
FEATURE Location/Qualifiers
source 1..276
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 266
cctcgactgt gccttcttagt tgccggccat ctgtttttt cccctcccccc gtgccttcc 60
tgacccttgc aggtgtccat cccactgtcc ttctcttataa aatgtggggaa attgttgc 120
attgttgcgtt taggtgtcat ttatcttgg ggggttggggat gggccggggc agcaagggg 180
aggatgtggaa agacaatagc aggcatgtcc gggatgttgc gggctctatg gtttcttgc 240
cgaaaagaaac cagctggggc tttttttttt atcccc 276

SEQ ID NO: 267 moltype = DNA length = 192
FEATURE Location/Qualifiers
source 1..192

-continued

-continued

gatgtacat	gtaacattaa	gaatggcaga	tgcgagcagt	ttttaaaaaa	tagtgctgat	360
aacaaggctgg	tttgcctctg	tactggggaa	tatcgacttg	cagaaaaacc	gaagtcctgt	420
gaaccacag	tgccattttcc	atgtggaaa	gttctgttt	cacaaacttc	taagtcacc	480
cgtgctgaga	ctgttttcc	tgatgtggac	tatgtaaat	ctactga	tgaaaccatt	540
ttggataaca	tcactcaaag	caccaatca	ttaatgact	tcactcggtt	tggtgggtaa	600
gaagatgoca	aaccaggta	atccccttg	caggtgttt	tgaatggtaa	agtgtatgca	660
ttctgtggag	gtctatcgt	taatggaaaa	tggattgtta	ctgctgccca	ctgtgttgaa	720
actctgtgtt	aaattacagt	tgtcgaggt	gaacataata	ttagggagac	agaacataaca	780
gagccaaagc	gaaatgtat	togaatttt	cctcaccaca	actacaatgc	agctattaat	840
aagtacaacc	atgacattgc	ccttctggaa	ctgacgacaa	ccttagtgc	aaacagctac	900
gttacaccta	tttgcattgc	tgacaaaggaa	tacacgaca	tcttcctaa	atttggatct	960
ggctatgtaa	gtggctgggg	aagagtctc	cacaaaggaa	gatcagott	agtcttcag	1020
taccttagag	ttccacttgt	tgaccgagcc	acatgtctt	tatctacaaa	gttcaccatc	1080
tataacaaca	tggttctgtt	tggttccat	gaaggaggaa	gatggatcg	tcaaggagat	1140
agtgggggac	cccatgttac	tgaatggaa	gggaccagg	tcttaactgg	aattatttagc	1200
tggggtaa	agtgtcaat	gaaaaggaa	tatggatat	ataccaagtc	ctcccgat	1260
gtcaactgga	ttaaggaaaa	aacaaagctc	actgtcagcg	gatggagact	gttcaagaag	1320
atcagctaa						1329

SEQ ID NO: 273	moltype = DNA	length = 1329				
FEATURE	Location/Qualifiers					
source	1..1329					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 273						
tttagaaatc	ttcttaaaca	ggccgcacgc	gctcacgggt	agcttagtct	tttcttttat	60
ccaaatccacg	tagcgagaga	ccttcgtata	gtgcacat	ttccccctca	tcgcacat	120
ctcccccccaa	cttattatcc	ccgtcaagaa	acttgcctt	tcgacttcag	tgacgtgtgg	180
tccacactgaa	tcaccttggc	atgagtcgcg	accgcctcg	tgaacccag	cacaaacat	240
gttattgtaa	atcgtaaaat	tcgtggacag	aagacaggtc	gctctatcga	ccaacggac	300
gcccacat	atgcacat	gggtgtatcg	acctttgtgg	aagacegcgc	cccacccact	360
cacatatacg	ttcccaaaat	tgcagaacga	gggtgtatcg	ttttatcgg	ctatacaat	420
cggggtaaca	taggagttaa	gtacggatgg	ctcgccacgc	tccaggaggg	ctatcatg	480
gttgtacttg	ttttagatgcgg	cattataa	gtgtgggg	atgatcctga	taacattcct	540
tttctgttca	tgatgtctcg	tttcttcaat	gttgtgttcg	ccgccccaga	ccgttaatctt	600
aacccccctgc	tgcacacat	gtgcccgcgt	tacaatccac	ttttcatgtt	ctatggagac	660
ccccacaaaac	ggtgtcgactt	ttccgtttag	caccacgtc	catggaaatt	ggccagggtt	720
agcgccatcg	cccccgacaa	ccctagtaaa	gtcattaaat	gactgtgtgg	attgtgttat	780
attatcaaga	atcggttcgg	cttcgtatgt	gttacgttgc	tccacatccg	aaaaaactgt	840
ctcgccgcctt	gtcaacttttgc	atgttggaa	cacacttacc	cgacgcacg	ggaagggcac	900
cgcccggttca	cagcttttttgc	gatcttcgc	gagccggtag	ccctcgtgc	aactacacac	960
aactttgttgc	tcggcggaaat	ttttacagaa	ttgtcgcgt	cgccatttttgc	taatgttgc	1020
gggtgacgtcc	aactcgacgt	ttttcccttc	aaaacccaaaa	gggcacccaa	actcgtagga	1080
atttatatcg	tctttacaa	tcccccatt	cagacatggaa	tttagattcgc	attggcccc	1140
atcgacat	tgcttcaga	actcgtgttgc	ccgttctgttgc	ttctcaacaa	cctcgccgcgc	1200
ttcttcacaa	ctgcatttttgc	cctccatata	ctctcgctc	aagttcccttgc	acgacatttc	1260
ttcaagctt	ctcgagttat	acctttttaggc	ccggtaatgtt	atcttattcgc	cgttttcggt	1320
gtccagaaaa						1329

SEQ ID NO: 274	moltype = DNA	length = 100				
FEATURE	Location/Qualifiers					
source	1..100					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 274						
ctgaaatgt	aaagaataat	tcttttagttt	tagaaaaaaa	gaaaacatca	tgaaaatttt	60
acatctctta	agaaaagtctt	tgttttaat	ccaaataatc			100

SEQ ID NO: 275	moltype = DNA	length = 1446				
FEATURE	Location/Qualifiers					
source	1..1446					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 275						
tttcttgc	atgaaaacgc	caacaaaattt	ctgaatcgcc	caaaggagta	taattcgat	60
aaatttggaa	agtttgcgtca	agggaaacctt	gagagagaat	gtatggaaa	aaagtgtat	120
tttggaaag	cagtatttc	tttggaggac	tttgcgtgttgc	actggaggca	aaccgctgttgc	180
tataatctcg	accaacttac	ggaacagggttgc	ggggtaatgtt	ccctcttca	gaatttgggt	240
gtaaacgtca	caccaatcc	cgccgttgc	ttgtctgttgc	agaacggact	aaaaatttgc	300
atccatgttca	tcattccata	tgaagggttgc	atgtggatcc	aaatggggca	gtcgagaa	360
attttcaagg	tagtttaccc	agtcgtatgttgc	caccacttca	aagtcatatttgc	ccactatggc	420
acacttgc	tcgacggatgtt	aactctaat	atgattgttgc	actttgggtgc	cccgatgtatgttgc	480
ggccatcg	tgtttgttgc	caaaaagatc	accgttaacag	gaacgtgttgc	gaatggaa	540
aagataatcg	acgagatgtt	gataaatcc	gacgggtcac	tcctgttgc	gttacattt	600
aacggcgtca	caggatggatgtt	cgtatgttgc	cgtatgttgc	ccacaaatttgc	ttcactcctg	660
aaggcaggcc	gagacgttgc	ggaaaaccca	ggggccgttgc	gcaaggccgttgc	ggagctgttgc	720

-continued

accgggggtgg	tgcctatcct	ggtcgagctg	gacggcgacg	taaaggccca	caagttagc	780
gtgtccggc	aggcggaggg	cgtggccaco	tacggcaac	tgaccctgaa	gttcatctgc	840
accacccggca	agetcggcgt	gcccggccc	accctcgta	ccaccctgac	ctacggcg	900
cagtgcgttca	gocgttaccc	cgaccatcg	aaggcagcag	acttcttcaa	gtccggcat	960
ccccaaaggct	acgttccagga	gcccacccat	ttcttcaagg	acgacggca	ctacaagacc	1020
cggcgccgagg	tgaagttcg	ggggacac	cttgtgaac	gcatcgact	gaagggcac	1080
gacttcaagg	aggacggca	cattctggg	cacaagctgg	agtacaacta	caacagccac	1140
aacgtctata	tcatggccg	caacagaag	aacggcat	aggtaac	caagatccgc	1200
cacaacatcg	aggacggcag	cgtgcagtc	gcccggact	accagcaga	cacccatc	1260
ggcgacggcc	ccgtgtgtct	gcccggacaa	cactacttgc	gacacccatc	cgccctgac	1320
aaagacccca	acgagaagcg	cgatcacat	gtctgtgt	agttctgtac	cgccggccgg	1380
atcaactctcg	gatggacga	gctgtacaag	ggaggaggaa	gcccgaaga	gaagagaaag	1440
gtctaa						1446

SEQ ID NO: 276 moltype = DNA length = 1446
 FEATURE Location/Qualifiers
 source 1..1446
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 276

ttacacccctc	ctcttttct	tggggctgco	gcccggctt	taacatcg	ccatggccag	60
ggtgatcg	gggggggtc	cgaactcc	cggcaccat	tggtccct	tctcggttgg	120
gtcttgc	aggggcgtct	gggtgtcg	gtatgtgtt	tggggcaga	gcacggggcc	180
gtcgccgt	gggggtgtt	gtgtgtgt	gtggcggac	tgcacgtgc	cgtctcgat	240
gttgtgc	atcttgaat	tcacottgt	gcccgtt	tgttgttgc	ccatgtatgt	300
cacgttgc	ctgtgtgt	tgtatcc	cttggggcc	aggatgtgc	cgtctctt	360
gaagtgc	ccttcaat	cgatctgt	caccagggt	tgccttca	acttccatc	420
ggccctgg	tgttagtgc	cgtcgtt	gaagaatg	gttccctt	gcacgtagcc	480
ctcggttgc	cgcttcttgc	agaatgt	ctgtttat	tgggtgggt	acctgtgtaa	540
gcactgc	cgttaggtca	gggtgttac	cagggtggc	cagggtcagg	gcagcttgc	600
gtgtgtgc	acgggttgc	gggttgcgt	gcccgttgc	ggtgcgttgc	cgccctcgcc	660
gttcacgt	aacttgc	cggttacgt	gcccgttgc	tccaccagg	tgggcacac	720
gcccgttgc	acttgc	ggggccggg	tttcccttca	cgtcgccggc	780	
ctgttgc	agggtgt	tgggtggc	gatcccttc	cacagcc	agccggcgtac	840
gcccgttgc	gtcaccc	acagcagg	gcccgttgc	tttgcgttgc	tctcgat	900
gatcttgc	cggttccaca	gggttgcgt	cacgggtatc	tttgcgttgc	cgaacacggc	960
gatgcctc	tagggctgc	cgaatgt	gatcatgtt	gggggtc	cgtcgat	1020
cagggtgc	tagtgc	tcacccgt	gtgtgttgc	tccacgggt	acaccac	1080
aaaaatctc	tccatgttgc	ccatgttgc	gcccgttgc	ccctcgat	ggatgtatc	1140
gtggatgt	atcttgc	cggttgc	gctcgtatc	atccctgttgc	tgggggtc	1200
gttcacgc	agggttgc	acaggctgt	cacccgc	tgctccagca	cctgggtc	1260
gttgtgc	cggttgc	tccagtc	cacgaatgt	tccagggt	acacggc	1320
ctcgaatgt	cacttctt	ccatgc	cctctcc	ttggcc	cgaaactc	1380
cagttgc	ctgtgttac	tettggc	gttcaggat	ttgttgg	gttcgttgc	1440
caggaa						1446

SEQ ID NO: 277 moltype = DNA length = 3570
 FEATURE Location/Qualifiers
 source 1..3570
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 277

ttggccactc	cctctctgc	cgctcg	ctcaatcg	ccgggggac	aaagggtgc	60
cgtacccgg	gttttgc	ggccggct	gtgagcg	gagcgc	agaggag	120
ccaaactca	tcaatcg	ttccat	tctttaggt	gtgaag	aaacaaaag	180
cagcatata	cgttagt	tcttcat	tctttaat	tgttgttgc	ttttcttc	240
cctgttca	cgttttct	tgtat	aaacggca	aaatcttgc	tccggca	300
aggtataatt	caggtaatt	ggaagat	tttcaagg	accttgc	agaatgtat	360
gaagaaaaat	gtatgttgc	agaacac	gaatgtttt	aaaacat	agaacaact	420
gaatgttgc	aggatgtat	tgatgttgc	cgtgtgt	ccaatccat	ttttatggc	480
ggcgtgtc	aggatgtat	taatccat	gaatgttgc	gtcccttgg	atttgaagg	540
aagaactgt	aattatgt	aacatgt	aatagaat	gcatgtgc	gcagttgt	600
aaaaatatgt	ctgataacaa	gggtgttgc	tccatgt	agggatatc	acttgc	660
aaccacaaat	cctgttgc	aggatgtgc	tttccatgt	gaagatgttgc	tgtttcaca	720
acttctaa	tcacccgt	tgatgttgc	tttccatgt	tggactatgt	aaattctact	780
gaagcttgc	aaatgttgc	taatcat	aaacggcc	aatctttaa	tgacttca	840
cgggtgttgc	gtggagaaga	tgccaaacca	ggtcaatttgc	cttgc	gggtgttgc	900
ggtaaagtgc	atgcattct	tggaggct	atcgttata	aaaaatgtt	tgtactgt	960
gcccactgt	ttgaaaactgg	acagtgtc	cagggt	taatatttgc	tttttttt	1020
gagacacaaat	atacagac	aaagcgaaat	tgatgttgc	ttatcttca	ccacaactac	1080
atgcgtatc	ttaatgttgc	caaccat	atggcccttc	tggacttgc	cgaaacctt	1140
gtgttgc	gttacgttac	acattttgc	attgtgttgc	agaataac	gacatctt	1200
ctcaatgttgc	gtatgttgc	tgttgc	ttggaaag	tcttccacaa	agggat	1260
gttttagtgc	ttcgttac	taggttca	tttgcgttgc	gagccacat	tcttctat	1320
acaatgttca	ccatgttgc	caacatgtt	tgtgttgc	tccatgttgc	aggtagag	1380
tcatgttgc	gagatgttgc	ggggcc	tttactgttgc	tttgcgttgc	tttgcgttgc	1440

-continued

actggaaatta	ttagctgggg	tgaagagtgt	gcaatgaaag	gcaaatatgg	aatatatacc	1500
aaggatcc	ggtatgtcaa	ctggattaa	aaaaaaacaa	agtcactta	acctcgactg	1560
tgcccttgc	ttgcgcggca	tctgtttt	ccccccccc	cgtgcctcc	ttgaccctgg	1620
aagggtccac	tcccaactgtc	ctttcttaat	aaaaatggag	aattgcatcg	cattgtctgg	1680
qtaggtgtca	ttctatttcg	gggggtgggg	tggggcaggag	cgaacagggg	qaaggatggg	1740
aagacaataq	caggcatgt	ggggatgggg	tgggctctat	ggcttcttag	gcccggaaagaa	1800
ccagctgggg	ctcttaggggg	tatccccaaa	aaacccccc	caccctcccc	tgaacactgaa	1860
acataaaaat	aatgcaatgt	tttgtttaa	cttggtttt	gcagcttata	atggttacaa	1920
ataaaagcaat	agcatcacaa	atttcacaaa	taaaggactt	ttttcaatgc	attctagttg	1980
tggtttgtcc	aaactcatca	atgtatctta	tcatgtctgt	taggtgaget	tagtctttt	2040
ttttatccaa	ttcacgttagc	gagagacctt	cgtatagatg	ccatattttc	ccttcatcg	2100
acatctcc	cccccaactta	ttatcccggt	caagaacatt	gttccttcga	cttcagtgac	2160
gtgtgttcca	cctgaaatcac	cttggcatga	gtcgccggcc	ccctcggtaa	acccagcaca	2220
aaacatgtta	ttgttaaatcg	taaatttgc	ggcagacaga	caggtcgttc	tatcgaccaa	2280
cgggacgcgc	aaatattgc	gaacggggc	tgatcgaccc	ttgtggaaga	ccggccccc	2340
cccaactcaca	tatccgctcc	caaatttca	gaagatattt	gtatatttt	tatccggtat	2400
acaatccggg	gtacatcgat	taatggat	gagtggctcg	tccagtc	ggaggggct	2460
atcatgggtt	tacttggttt	tagcgccat	ataatgtg	tggggatgt	tccgttataac	2520
attcccttcc	tgttcaatgt	gctcgtt	ttcaatgtt	tgggtccag	ccacgaccgt	2580
aatctaacc	cccgcttcga	cacagtgtc	ggccgttaca	atccacttt	cattgtactat	2640
ggggccccca	caaaacgggt	cgactttcc	gttgacccac	acccgttgc	gaatttggcc	2700
aggttttagcg	tcctcgcccc	cgacaaaccc	agtaaagtca	ttaaatgact	gtgtggatgg	2760
ttgttattata	tcaagaaatcg	ttccgggttc	agtagagtt	acgttagtca	catcggggaa	2820
aactgtctcg	gcccttgc	actttgtat	ctgggacaca	tttacccgac	cgccacggaa	2880
ggggcacccgg	gggttccacgc	tctttgtatt	ctcggcggc	cgtgtacccc	cagtgcata	2940
acacaacaat	tttgtgtcc	cggaattttt	acagaatttc	tgcgtatgtc	catttttaat	3000
tttgcgttgc	acgttcaact	cgcaatttt	tccttc	aaaaaggcc	accaacaccc	3060
gttaggaattt	atatcgctt	tacaactcc	ccatttcaga	catggattag	attcgat	3120
gtccccatcg	acatattgtt	tccagaaact	agtggttccgt	tctgttattt	caaacaccc	3180
geggcgcttc	ttaaaactgtc	attttcttc	catacatct	cgtccatgt	tcccttgcac	3240
gaatttctca	agttttcttc	agtttatactt	ttttagggccg	ttaaatgtat	tatttcgtt	3300
ttcgtgttcc	agaaaaaaactg	tggaaacagg	gagagaaaaa	ccacacaaca	tattttaaaga	3360
ttgatgaaga	caactaactg	taatatgtc	ctttttgtt	ttcttccatc	tgacctaaga	3420
gatcttagaa	cccccttagtg	tggagtggc	cactccct	ctgcccgttc	gtccgtctac	3480
ttggggccgc	cgggccaaac	ccggggcgtt	ggcgcaccc	gggtccccgg	cctcaatgt	3540
cgacgcacgc	cgccggaggg	gagtggccaa				3570

SEQ ID NO: 278 moltype = DNA length = 3636
FEATURE Location/Qualifiers
source 1..3636
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 278 Organism - Synthetic Construct

ttggccactc	cctctcgcg	cgctcgatcg	ctcaactgagg	ccggggcacc	aaaggtcgcc	60
cgacgccccg	gctttcccg	ggggccatca	gtgaaggcagc	gaggcgcag	agaggggatg	120
gccaactcc	tcactagggg	ttcttagatac	tctgtatatt	tggataaaa	acaagaactt	180
tcttaaaga	tgtaaaaaatt	tcatatgtat	ttcttttttgc	ctaaactaa	aaggattttc	240
ttttacattt	cagtttttt	tgatcatgaa	aacgocaaaca	aaattctgaa	tcggccaaag	300
aggttataatt	caggtaaatt	ggaagaggtt	gtcaaggga	acccctggag	agaatgtatg	360
gaagaaaatg	gtatgtttga	aaagacacga	gaagtttttg	aaaacacta	aaagacaact	420
gaattttgc	agcgtatgt	tgatggatg	cagtgtgatg	ccaatccatg	tttaaatggc	480
ggcagttgc	aggatgcacat	taatccat	gaatgtgtt	gtcccccttgg	attnaaaggaa	540
aagaactgt	aatttagatgt	aacatgtaac	attaagaatg	gcagatgcga	gcagggtttgt	600
aaaaaaatgt	ctgtataacaa	ggtgttttgc	tcctgtacty	aggggatatgc	acttgcagaa	660
aaccagaatg	cctgtgaacc	agcagtgc	tttccatgtg	gaagagtttgc	tgtttcacaa	720
acttcaagc	tcaccggcgtc	tgagactgtt	tttccatgtg	tggactatgt	aaattctact	780
gaagctgaaa	ccatTTTgga	taacatca	caaagcaccc	aatcattaa	tgacttca	840
cggggttgtt	gtggagaaga	tgccaaacca	ggtaatccc	cttggcagg	tgttttgaat	900
ggtaaagtgt	atgcatttc	tgagggtct	atcgtaatg	aaaaaaatggat	tgttaactgt	960
gcccactgtg	tttggaaacttgc	tggtttaaat	acatgtgtcg	cagggtacaa	taatattgttgc	1020
gagacagaaac	atacagagca	aaagcgaat	gtgattcgaa	tttattctca	ccacaactac	1080
aatgcagcta	ttaataaagta	caaccatgc	attgccttc	tggacttgg	cgaaccctta	1140
gtgctaaacaa	gctacgttac	accttattgtc	attgtgtaca	aggataacata	gaacatcttc	1200
ctcaaaatttg	gatctggcta	tgtaaatgtc	ttggggaaag	tcttccacaa	aggggatata	1260
gttttagttc	ttcagttact	tagatttcc	cttggtgacc	gagccacatg	tctttatctt	1320
acaaaagtca	ccatctataa	caacatgttc	tgtgtttgt	tccatgtagg	aggttagagat	1380
tcatgtcaag	gagatagtttgc	ggggacccat	gttactgttgc	tggaaaggac	cagggttctt	1440
actggaaat	tttagctgggg	tgaagagtgt	gcaatgaaat	gcaatatacg	aatatatacc	1500
aaggcttccc	ggatgttca	tcgttataag	ggaaaaaaacaa	agctctactgt	cagggtatgg	1560
agactgttca	agaagatcatg	ctaaatccgtc	ctgtgtcttc	tagttggccat	ccatctgttgc	1620
tttgccttc	ccccgtgtc	tccttgacc	tggaaaggatgc	cactccact	gtcccttctt	1680
ataaaaatgt	ggaaatttgc	tcgcattgtc	tgagtaggttgc	tcatttttatt	ctgggggggttgc	1740
gggtggggca	ggacagcga	ggggaggatt	gggaagacaa	tagcaggat	gttggggatg	1800
cgggtgggtt	tatgttttttgc	ggggggggaa	gaaccatgt	gggtcttgc	gggtatcccc	1860
aaaaaaatcc	ccacatcc	ccctgttgc	gaaacataaa	atgtatgttgc	ttgtgtttgt	1920
taacttgttt	attgcgtt	ataatgttgc	caataaaagc	ataatgtatca	caatatttcac	1980

-continued

aataaaaagca	ttttttcac	tgcattctag	ttgtgggtag	tccaaactca	tcaatgtatc	2040
ttatcatgtc	tgtagggaaa	tcttcttaaa	cagocgcac	ccgcacccg	tgagcttagt	2100
cttttcttt	atccaattca	cgtagcgaga	gaccttcgta	tagatgcac	atttcccctt	2160
catccgacat	tctcccccc	aacttattat	cccggtcaag	aaacttgttc	cttcgacttc	2220
agtgacgtgt	ggtcacactg	aatcacctt	gcatgagtgc	cgaccgcct	cgtgaacccc	2280
agcacaaaaac	atgttattgt	aaatcgtaaa	tttcgtggac	agaagacagg	tcgctctatc	2340
gaccaacggg	acgcgcacat	atgcagaa	gagggtgtat	gacccctt	ggaagacccg	2400
ccccccacca	ctcacatc	cgctccaaa	tttcaagaag	atatttgat	attttttatc	2460
ggctatacaa	atcggggtaa	cataggagtt	aagtacgagt	ggctgtc	gctccaggag	2520
ggttatatac	tggttgtact	tggtttagt	ggcattataa	ttgtgtatgg	gtatgtatct	2580
gataacatc	ttttttctgtt	cagatgtc	agtttctca	atgttgcgtt	cgccagccac	2640
gaccgtaaatc	ttaaccccc	tctcgacaca	gtgtcgccgc	gttacaatcc	acttttctatt	2700
gactatggg	ccccccacaa	acgegtcgac	tttccgttgc	agcaccac	gccatggaaa	2760
ttggccaggt	ttagegtctc	cgcccccc	aaccctgtat	aagtatccaa	atgactgtgt	2820
ggattgtgtt	attatataa	gaatcggtt	gggttcgat	gagttaaatc	agtccacatc	2880
ggggaaaaact	gtctcgcccc	ttgtcaactt	tgatgtctg	gacacacta	cccgaccgca	2940
cgggaaaggc	accgcgggtt	cacagcttt	ttgatttca	gegagccgt	agccctcagt	3000
gcaactacac	acaatcttgc	tgteggcgga	attttacag	aattgtcgc	atcgctcatt	3060
ttaatgttgc	cagggtacgt	ccaaactcgca	gtttttccct	tcaaaaacca	aagggcacca	3120
acactcgtag	gaatttat	cgtttata	actccccca	ttcagacatc	gatttagattc	3180
gcattgttcc	ccatcgacat	attgttca	gaactcagtg	gtccgttctg	tatttcaaa	3240
cacctcgccg	gttttctca	aactgcatt	ttcttccata	cactctcg	ccaagttccc	3300
ttgcacaaat	tcttcacat	ttcttgcgtt	atacccttta	ggccgggtta	gtatcttatt	3360
cgcgtttcc	ttgtccagaa	aaactgtat	gtttaaaaat	aattttttag	tttttagcaaa	3420
aaagaaaaca	tcatgaaaat	tttacatctc	ttaagaaatg	ctttgtttt	aatccaaata	3480
atcagagatc	taggaacccc	tagtgcgtt	gttggccact	ccctctctgc	gchgctcgctc	3540
gtcactcgag	ggcgccccgg	caaagcccg	gggtcgccgg	acctttggc	gccccggctc	3600
agtgagcgcag	cgagcgccca	gagaggggat	ggccaa			3636

SEQ ID NO: 279 moltype = DNA length = 3615
 FEATURE Location/Qualifiers
 source 1 .. 3615
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 279

ttggccactc	cctctctgcg	cgctcgctcg	ctcaactgagg	ccggggcgacc	aaagggtcgcc	60
cgacgcgggg	gttttgcgg	ggccggctca	gtgagcgacg	gagcgcgcag	agagggagtg	120
gccaactcca	tcactagggg	ttccttagtgc	tcttaggtca	gtgaagagaa	gaacaaaaag	180
caagcatatta	cagtttagtt	tcttcatca	tctttaaaata	tggttgcgtt	tttttctctc	240
cctgtttcca	cagggttttc	tgatcatca	aaacgcacca	aaatctgaa	tcggccaaag	300
agggtataatt	caggttataat	ggaagagtt	gttcaaggaa	accttgcgag	agaatgtatc	360
gaagaaaatgt	gtatgtttga	agaagcgtat	ttcaacttgg	aggacttgc	cggtgactgg	420
aggcacaaccg	ctgggtataa	tctcgaccaa	gtactggaa	agggcgccgg	aagtccctc	480
tttcagaattt	ttgggttgc	cgtcacacaa	atccagcgga	ttgtgttgc	tggagagaac	540
ggactcaaaa	tttgcacatca	tgttattc	ccatataa	gtctcgttgc	agaccaatgc	600
gggcagatcg	agaagatttt	caaggtagtt	tacccagtc	acgatcacca	cttcaagtc	660
atttctccat	atggcactact	ttgttgc	ggatgtactat	ctaataatgt	tgattatctt	720
gttgcggccgt	atggggccat	cgcaatgtt	atggccaaa	agatcacatc	aacaggaa	780
ttgtggaaatg	ggaacaagat	aatcgacgag	agatgtat	atccagacgg	gtcactctg	840
ttcagggttca	caattaacgg	cgteacacga	tggagactct	gtgaacgat	actggccaca	900
aatttttcata	tcctcgaagca	ggccggagac	gtggaggaaa	acccaggggc	cgtgaccaag	960
ggcgaggagc	ttttccacgg	gggtggcc	atccctgtcg	agttggacgg	cgacgataaac	1020
ggccacaatgt	tcagcgctgc	cgccggggc	gagggtcgat	ccacccatcg	caagctgacc	1080
ctgaagtca	tctgcaccac	cggaagctg	ccgcgtccct	ggccacccct	cgtgaccacc	1140
ctgacccatcg	gggtcgactg	tttcgcgc	taccccgac	acatgaa	gcacgacttc	1200
ttcaagtcgc	ccatgcggca	aggtcaatgc	caggagcgca	ccatctt	caaggacgac	1260
ggcaactaca	agacccgcgc	cgagggtaa	ttcgaggccg	acacccatgt	gaaccgcac	1320
gagctgaaagg	gcatecgactt	caaggaggac	ggcaacatcc	tggggcaca	gctggagatc	1380
aacttcaaca	gcacaaacgt	ctatcatcg	gcccacaac	agaagaa	catcaagg	1440
aacttcaaga	tccggccacaa	catecgaggac	ggcagcgatgc	agtcgcgc	ccactaccag	1500
cagaacaccc	ccatcgccgca	cgccccgt	ctgtgcggcc	acaacccatc	cctgagacacc	1560
cagtccgcgc	tgagcaaaaga	cccccaacgc	aagcgcgatc	acatgttgc	gctggagatc	1620
gtgaccgcgc	cggggatcac	tctcgatcg	gacgagatgt	acaaggggg	aggaagcccg	1680
aaagaaaggaa	gaaagggtcta	acccgtactg	tgcttctca	ttggccacca	tctgttgc	1740
ccccctccccc	ctgtgccttc	ttggccatgg	ttccactgtc	tttttctaa		1800
aaaatggagaa	aatttgcata	cattgtctga	gtatgttgc	ttcttatttc	gggggtgggg	1860
tggggcaggg	cagcaagggg	gaggattggg	aagacaatag	caggcatgt	ggggatgcgg	1920
tgggtcttat	ggcttctgtt	ggcggaaagaa	ccagctgggg	ctcttagggg	tatccccaaa	1980
aaacctccca	cacccccc	tgaacactgaa	acataaaaat	aatgcacat	ttgttgcgtt	2040
cttgcgttatt	cgacttata	atggttcaaa	ataaagcaat	agcatcacaa	atttcacaaa	2100
taaaagcattt	ttttcactgc	atttgcatttgc	ttgtttgc	aaactcatca	atgtatctt	2160
tcatgtgtgt	tacacccatc	tcttttttctt	ggggctgc	ccggccctgt	acagctcg	2220
catgcccagg	gtgtatgcgg	cgccggc	gaactccac	agcaccatgt	ggccctctt	2280
ctcggtgggg	tccttgcgtca	gggcgcctcg	gggtgc	ttgtgttgc	cgccgcac	2340
cacggggccgc	tcgcgcgttgc	gggtgttgc	ctgtgtatgg	tcggccatgt	gcacgctgc	2400
gttgcgtatgt	ttgtgcgttgc	tcttgcgtat	cacccatgt	ccgttcttct	gttgcgttgc	2460

-continued

catgatgtac	acgttgtggc	tgttgttagtt	gtactccagc	ttgtggccca	ggatgttgcc	2520
gtctccctcg	aagtctcgatgc	ccttcaggctc	gatccctgttc	accagggtgt	cgccctcgaa	2580
cttcaccccg	gcctcggtct	tgttagttgc	gtcgtccttg	aagaagatgg	tcctctctcg	2640
cacgtagccc	tccggcatgg	cgctttgaa	gaagtcgtgc	tgcttcatgt	ggtcggggta	2700
cctgtcttgg	cactgcacgc	cgttaggtcag	ggtgtcacc	agggtggggc	agggcacggg	2760
caaccttgcgc	gtgggtcgaga	tgaacttcag	ggtcagcttg	ccgttaggtgg	cgtegcctc	2820
gcctctcgcc	ctcacgtga	atttggggc	gttcacgtcg	ccgtccagct	ccaccaggat	2880
gggcacccacg	ccgggtaaaca	gctcttcgc	cttgcacag	gggcgggggt	tctcttcac	2940
gtcgcggccc	tgcttcagca	ggctgaagtt	ggtgccagg	atctctcgc	acagctcca	3000
ggcggtacgg	ccgggtatgg	tcccttcgtt	cacggaggctc	ccgtcggggt	tgatcggct	3060
ctcgctcgat	atcttgcgt	cgtttttttc	gggtccacag	gggtccggtc	acgggtat	3120
gaacacccgg	atgccttcgt	agggtctgc	gaagtagtgc	atcatgttgg	gggtcacgcc	3180
gtcgatcacc	agggtgcctgt	agtgcaggat	caccttgaag	ttgtgttgc	ccacggggta	3240
caccacccgt	aaaatcttc	cgatctggcc	catctgtcg	ccgtccagcc	cctctgtgg	3300
gatgtatcgc	ttggatgtcg	ttttcgggg	gttctcgcc	ctcagcacga	tcctctggat	3360
gggggtacgg	cttcaccccc	gggtttggaa	caggctgtc	acggccctc	gtccacagcac	3420
ctgggtccagg	ttgttagccgg	cggtctgcct	ccagtcgc	acgaagtctt	ccagggtgaa	3480
cacggccctcc	tccaaagctgc	actttccctc	catgcactcc	ctctccagg	tgcctgcac	3540
gaactcttcc	agcttgcgc	tgttgtacct	cttggcctg	ttcaggatct	tgttggcggt	3600
ctcgtgttcc	aggaa					3615

SEQ ID NO: 280	moltype = DNA	length = 3636				
FEATURE	Location/Qualifiers					
source	1..3636					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 280						
ttggccactc	cctcttcgtcg	cgctcgctcg	ctcaactgagg	ccggggcgacc	aaagggtcgcc	60
cgtacggccgg	gttttgcgg	ggccggctca	gtgagcgcgc	gagcgcggcg	agaggaggatg	120
gccaactcca	tcaacttaggg	ttcttagatgc	tcttaggtca	gtgaagagaa	gaacaaaaag	180
cagcatata	cagttagtt	tcttcatcaa	tctttaaata	ttgtgtgtgg	ttttcttc	240
cctgttttca	cagttttct	tgtatcataa	aaaccccaaca	aaattctgaa	tcggccaaag	300
aggtaatttt	caggtaattt	ggaagagtt	gttcaaggga	accttgcgg	agaatgtatg	360
gaagaaaaagt	gtatgttga	agaagcacga	gaagtttttgc	aaaactctga	aagaacaact	420
gaattttgg	aggcgtatgt	tgtatggat	cagtgtgatg	ccaaatccatg	tttaaatggc	480
ggcagtttgc	aggatgatc	taatttccat	gaatgttgc	gtcccttgg	atttgaagga	540
aagaactgtg	aattatgt	aacatgtaa	attaagaatg	gcagatgcga	gcagtttgc	600
aaaaatatgt	ctgataacaa	ggtgggttgc	tccctgtactg	agggatatacg	acttgcggaa	660
aaccacggat	cctgtgaaacc	agcgtgc	tttccatgtgc	gaagagtttc	tgtttcacaa	720
acttctaagg	tcacccgtgc	tgagactgtt	tttccatgtgc	tggtactatgt	aaatttctact	780
gaagctgaaa	ccattttgg	taacatact	caaaagcacc	aatcatttta	tgacttca	840
cgggttgc	gtggagaaga	tgccaaacca	gttcaattcc	cttggcaggt	tgtttgaat	900
gtgaaatgt	atgcattctg	tggaggctc	atcgtaatgt	aaaatggat	tgttaactgt	960
gcccactgt	ttgtttttgc	tttttttttt	acagttgtcg	cagggttgc	aatattttgg	1020
gagacacaa	atacagacg	aaagcggaaat	gtgattcgaa	ttatcttca	ccacaactac	1080
aatgcagtc	ttaataagta	caaccatgc	atttgccttc	ttggacttgc	cgaaccctta	1140
gtgtctaaac	gatcgatgt	accttttgc	atgtgtacca	aggaatatac	gaacatcttc	1200
ctcaatattt	gtatggctt	tgttaatgtt	tttttttttt	tttttttttt	tttttttttt	1260
gtttagtttgc	tttgcgtt	tagatgttca	tttttttttt	tttttttttt	tttttttttt	1320
acaaatgttca	ccatctataa	caacatgtt	tgtgtgttgc	ttcatgttgc	tttttttttt	1380
tcatgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1440
actggatgtt	ttatgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1500
aaggctcc	gttatgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1560
agactgttca	agaagatcag	ctaacctcg	ctgtgccttc	tagttgcgg	ccatcttttg	1620
tttgccttc	ccccgtgc	ttcttgc	tttttttttt	tttttttttt	tttttttttt	1680
aataatgtt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1740
gggtggggca	ggacacgca	ggggaggatt	ggggaggacaa	tagcggccat	gtctgggtat	1800
cggtgggtc	tatgttct	gaggcggaaa	gaaccagctg	gggtcttgg	tttttttttt	1860
aaaaaaaccc	ccacaccc	ccctgttgc	gaaacatataa	ttgtgtgtgt	tttttttttt	1920
taacttgg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1980
aaatataatgt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2040
ttatcatgtc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2100
tttttttttt	atccaaattca	ctgtacgttgc	tttttttttt	tttttttttt	tttttttttt	2160
catcgacat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2220
agtgcacgtt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2280
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2340
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2400
ccccccacca	ctcacatatac	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2460
ggctatcaa	atcggttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2520
ggctatcaa	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2580
gataacattc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2640
gaccgtatac	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2700
gactatgg	ccccccacca	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2760
ttggccagg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2820
ggatttgtt	atattatcaa	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2880
ggaaaaaaact	gtctcgcc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2940

-continued

cgggaaagggc	accggccggtt	cacagctt	ttgattctca	gcgagccggt	agccctcagt	3000
gcactcac	acaacttgt	tgtcgccga	attttacag	aattgtcgc	atcgccatt	3060
ttaatgtt	cagggtacgt	ccaaactcgca	gtttttccct	tcaaacc	aaggccacca	3120
acactcgtag	gaatttat	cgtcttaca	actcccccca	tccagacat	gatttagattc	3180
gcattgttc	ccatcgacat	attgttcca	gaactcagtg	gtccgttctg	tattctcaaa	3240
cacctcgcc	gotttttca	aactgcattt	ttctccata	cacttcgt	ccaaagtccc	3300
ttgcacgaat	tcttcaagct	ttctcgatgt	atacccttta	ggccggtaa	gtatcttatt	3360
cgcgtttcc	ttgtccagaa	aaactgtgga	aacaggaga	aaaaaccac	acaacatatt	3420
taaagattga	tgaagacaac	taactgtaat	atgctgttt	ttgttcttct	cttcaactgac	3480
ctaagagat	taggacccc	tagtgcgtt	gttggccact	ccctcttcgc	gchgctcgctc	3540
getcaactgag	gcccgggg	caaaggcc	gctcgccgg	acctttggc	gccccggctc	3600
agtgagcg	cgagcgca	gagaggagt	ggccaa			3636

SEQ ID NO: 281 moltype = DNA length = 1954
 FEATURE Location/Qualifiers
 source 1..1954
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 281

ggccactccc	tctctgcgcg	ctcgctcgct	cactgaggcc	gggcgaccaa	aggtcgcccg	60
acgccccggc	tttgcgggg	cggccctcgt	gagcgagcga	gchgctcg	aggaggatggc	120
caactccatc	actagggtt	cctggagggg	tggagtcgt	ataggtcagt	gaagagaaga	180
acaaaaaggc	gcatattaca	gttagtgc	tccatcaatc	tttaatat	ttgtgtgtt	240
tttctctccc	ttgtttccaca	gtttttctt	atcatgaaa	ccgcacaaa	attctgaatc	300
ggccaaagag	gtataattca	ggtaaattgg	aagagttgt	tcaagggaa	ctttagagag	360
aatgtatgga	agaaaatgt	agtttggaa	aagcacgaga	agttttggaa	aacactgaaa	420
gaacaactga	attttggaa	cagtatgtt	atggagatca	gtgtgatcc	aatccatgtt	480
taaatggcg	cagttcga	gtatgcattt	atccatgtat	atgttgggt	ccctttggat	540
ttgaaggaaa	gaactgtgaa	tttagtgtaa	catgttaat	taagaatggc	agatgcgagc	600
agttttgtaa	aaatagtgt	gataacaagg	tggtttgc	ctgtactgag	ggatatcgac	660
tttcacaaac	ccagaatgtt	tgtgacccat	tccatgtgga	agatttctg	720	
tttcacaaac	ttcttaagtc	accgtcgct	agactgttt	ttctgtatg	gactatgtaa	780
attctactga	agctgaaacc	attttggata	acatcactca	aagcacccaa	tcatttaatg	840
acttcactcg	ggttgttgg	ggagaagat	ccaaaccagg	tcaattccct	tggcgagg	900
ttttgaatgg	taaaatgtat	gtcattctgt	ggggctctat	cgatgtatggaa	aaatgtgattt	960
taactgctgc	ccactgtgtt	gaaactgtgtt	tttttttttt	agttgtcgca	ggtgacata	1020
atattgagga	gacagaaat	acagacaaa	aggaaatgt	gattcgattt	attctcacc	1080
acaactacaa	tgcagttt	aataagtaca	accatgacat	tgccttcgt	gaactggacg	1140
aacccttagt	gtataacacg	tacgttacat	ctatgtcat	tgctgacaa	gaatacacga	1200
acatcttcct	caaatttgg	tctggatgt	taatgtggc	ggggagatgc	ttccacaaag	1260
ggagatcgc	tttagtctt	cagtaccta	gagttccact	tgtgacccga	gccacatgc	1320
ttctatctac	aaagtttacc	atctataaca	acatgttctg	tgtgttgc	catgaaggag	1380
gtagagatc	atgtcaagga	gatagtgggg	gacccatgt	tactgtgg	gaaggagcc	1440
gttttcttac	tggaaattt	agctgggggt	aaaggtgtc	aatgtggaa	aaatgtggaa	1500
atatataacaa	ggatcccg	tatgtcaat	ggatgtggaa	aaaaacaa	ctcacttaac	1560
ctcgactgt	ccttctatgt	gcccgcattt	tgtgttgc	ccctcccc	tgccttc	1620
gaccctggaa	gggtggactt	ccactgttct	ttcttataata	aatgtggaaa	ttgatcgac	1680
ttgtctgtt	agggtgtt	ctatgttttt	gggtgggggt	gggcaggaca	gcaaggggg	1740
ggatggggaa	gacaatagca	ggcatgtgg	ggatgtgg	ggctctatgg	ttctgtgg	1800
ggaagaagaa	actgtgggtt	ctagggggta	tccacttct	ctctgcgc	1860	
tcgctcgctc	actgaggcc	ggcgcacaaa	ggtgcggcc	cgccgggtt	ttgcccggc	1920
ggcctcgtt	agcgagcg	cgcgcagaga	ggga			1954

SEQ ID NO: 282 moltype = DNA length = 2359
 FEATURE Location/Qualifiers
 source 1..2359
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 282

ggccactccc	tctctgcgcg	ctcgctcgct	cactgaggcc	gggcgaccaa	aggtcgcccg	60
acgccccggc	tttgcgggg	cggccctcgt	gagcgagcga	gchgctcg	aggaggatggc	120
caactccatc	actagggtt	cctggagggg	tggagtcgt	accttaggtc	tctccggctc	180
tgtttttttt	aggggtgtgt	ttcgcggata	agcactgttt	ttttttatgt	tttttcatct	240
ctgctgtgtt	tttttctatgt	atggaaatgtt	gttattttaa	tttttttttt	300	
gtgtctgtt	ctagatttt	attactgtt	ttgttgcatt	atttgcatt	360	
gagaactagg	tcaatgtt	gaagaacaaa	aagcagat	ttacgttgc	ttgtttttat	420
caatctttaa	atatgtgtt	ttgtttttt	ctccctgttt	ccacgtttt	tcttgcatt	480
gaaaacgc	acaaaatttct	gaatcgccca	aagaggata	attcaggtaa	atttggaaag	540
ttttgttca	ggaaccttgc	gagagaatgt	atggaaagaaa	agtgtgtttt	tgaagaagca	600
cgagaatgtt	ttgaaaacac	tgaaagaaca	actgtttttt	ggaaggatgt	tgttgcatt	660
gtatgtgtt	ttgttgcatt	ttgttgcatt	ggaaagaact	gtgttgcatt	ttgttgcatt	720
ttatgtatgtt	ttgttgcatt	ttgttgcatt	ggaaagaact	gtgttgcatt	ttgttgcatt	780
aacattaaga	atggcgatgt	cgagcgttt	ttgttttttt	gtgttgcatt	ttgttgcatt	840
tgctccgtt	ctgaggata	tcgacttgc	gaaaacc	agtccctgtt	accaggatgt	900
ccatcttccat	gttggaaatgt	ttctgtttca	caaacttctt	agtcaccc	tgctgttt	960
gtttttctgtt	atgttgcatt	ttgttttttt	actgtttttt	aaaccatttt	ggataacatc	1020

-continued

actcaaaggca	ccaaatcatt	taatgacttc	actcggttg	ttgggttggaga	agatgc	aaa	1080
ccaggtaat	tcccttgcca	gggtgttttg	aatggtaaag	ttgatgcatt	ctgtggaggc		1140
tctatcgta	atgaaaaatg	gattgtact	gctggccact	gttgttggaaac	tttgttggaa		1200
attacagtgtt	tcgcaggta	acataatatt	gaggagacag	aacataatggca	gcaaaaggca		1260
aatgtgatcc	gaattatccc	tcaccacac	tacaatgcg	cttataataa	gtacacccat		1320
gacattgccc	ttcttggact	ggacgaaacc	tttagtgc	taacacccat	tacacccatt		1380
tgcattgtctg	acaaggaaata	cacgaacatc	ttcttccaaat	ttggatctgg	ctatgttaagt		1440
ggctggggggaa	gagtttccca	caaaaggaga	tcagtttag	ttcttcgtca	ctttagatgg		1500
ccacttgttgc	accgaggccac	atgtttctca	tctacaatgt	tcacccatcta	taacaaacatg		1560
ttctgtgtctg	gcttcatcga	aggaggtaga	gattcatgtc	aaggagatag	ttgggggacc		1620
catgttactg	aagtggaaagg	gaccaggatc	ttaactggaa	ttattagtcg	gggtgaagag		1680
tgtgtcaatga	aaggcaataa	tggatataat	accaggatgt	ccccgtatgt	caatggatt		1740
aaggaaaaaa	caaagctcac	ttaacactcg	ctgtgccttc	tagttcccg	ccatctgtgg		1800
tttgccttc	cccccgtctg	tccttgaccc	tggaaagggtc	cattccctact	gtctttctt		1860
aataaaatga	ggaaattgc	tcgcattgtc	tgagtaggtg	tcatttttatt	ctgggggggt		1920
gggtggggca	ggacagcgaag	ggggaggatt	gggaaggacaa	tagcaggat	gtctggggatg		1980
cggtggggctc	tatgggttct	gagggggaaa	gaaccagctg	gggtctttag	gggtatcccc		2040
cttaggtgtt	tatattattt	atataattttt	ggatcttttg	atgacataaa	ttgggggattt		2100
tgtaaagctta	gttttaattt	tcttttaatt	aaaaaaaaat	gcttaggcaga	atgactcaaa		2160
ttacgttgg	tacagtgtt	tttattacgg	tctcataggg	cctgcgtct	cgaccatgt		2220
atactaaaaa	ttaaaaatgt	actagtccac	tccttcctcg	cgccgcgtct	cgctcaactga		2280
ggccggccga	ccaaaggctc	ccccacggcc	gggtttgc	cgggggccct	cagtgagcga		2340
ggcggccgc	agagggga						2359

-continued

```
ttttgtctct ttgttagatct aggaacccct agtgatggag ttggccactc cctctctgcg 2820
cgctcgctcg ctcaactgagg ccggccgggc aaagccccgg cgctggggca ccttggctcg 2880
cccgccctca gtgagcgcagc gagegcgcagc agagggagtg gccaa 2925
```

```
SEQ ID NO: 284 moltype = DNA length = 3477
FEATURE Location/Qualifiers
source 1..3477
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 284
ttggccactc cctctctgcg cgctcgctcg ctcaactgagg ccgggggagacc aaaggtcgcc 60
cgacgcgggg gttttggcccg ggccggccctca gtgagcgcagc gagegcgcagc agagggagtg 120
gccaactccca tcactaggggg ttcttagatc tttagctctcg gcaaaatgaa gtggggtaacc 180
tttctctcc tccttcgtcg ctccggctct gttttttccca ggggtgtgtt tcggccagaa 240
gcacgtaaaga gttttatgtt ttcttagatc tgctgttattt ttcttagaa tggaaagcctg 300
gtatTTAAAT atagttaaat ttctcttag tgctgatttc tagattata ttactgtgt 360
tgttggattt attgtcatta ttgcattctg agaaccctta ggtgggttata ttattgtat 420
atttttggta ttcttgatga caataatggg ggattttggaa agcttagctt taaatttctt 480
ttaattaaaaaaa aaaaatgtca ggcagaatac ctcaaattac gttggataca gttgaattta 540
ttacggcttc ataggggctcg cctgtcgacat cgtatatac taaaattaa aagtgtgtgt 600
tactaattttt ataaatggag ttccatTTA tatttacctt tattttttat ttaccatgt 660
cttagatgat atttacaaac atgacagaaaa cactaaatct tgagtttggaa tgccacagata 720
taaacactta acgggggttaaaaataataa tgttgggtgaa aaaatataac tttaggtgt 780
gcagaggggaa accatggccca ccttcagattt ttctgtacat gatccggaaac tggcatcttc 840
agggagtagc tttaggtcgt gaagagaaga acaaaaagca gcatattaca gtttagttgtc 900
ttcatcaatc tttaaatat ttgtgtgtt tttctctccc tgtttccaca gtttttctt 960
atcatgaaaaa cggccaaacaaat ttctgtatc ggcacaaagat gtataattca ggttaatattgg 1020
aagagttttt tcaagggaaac cttagagagaa aatgtatggaa agaaaaatgtt agttttggaa 1080
aagcacgaga agttttggaa aacactgaaa gaacaactgaa attttggaaag cagttatgtt 1140
atggagatca gtgtgagttcc aatccatgtt taaaatggcg gtagttgcaag gatgacattta 1200
attccatgtat atgttgggtt cccttggat tttgaaggaaa gaactgtgaa tttagatgtaa 1260
catgtacatc taagaatggc agatcgagc agttttgtaa aatagtgtt gataacaagg 1320
tggtttgcctt ctgtacttggag ggatattcgac ttgcagaaaaa ccagaatgtcc tggtaaccag 1380
cagtgcattt tccatgtggaa agatgttctg tttcacaaac ttctaaagctc accccgtctg 1440
agactgtttt tccatgtgtg gactatgtaa attctactgtaa agctgaaacc attttggata 1500
acatcaatc aagcacccaa tcattatgtt acttcactcg ggttgggtt gggaaatgt 1560
ccaaaccagg tcaattccct tggcagggtt tttgaatggaa taaatgttgc gcatctgt 1620
gaggcttatcgat ttccatgtggaa aatggatttggtaactgtgc ccactgtgtt gaaactgtgt 1680
ttaaaatattt acgttgcgcg ggttgcacata atattggatggaa gacagaacat acagaccaaa 1740
agcgaaatgtt gattcaattt attccttacc aacaatggaa tgcagtttattt aataatgtaca 1800
accatgacat tggcccttcgtt gaaactggggcc aacccttagt gctaaacacgc tacgttacac 1860
ctatttgcattt tgctgacaaatc gatatacaca gatattttccca tctggctatg 1920
taagtggctt gggaaatggc ttccacaaatggggcc gggatgttgc tttttttttt cttttttttt 1980
gagttccatc tttttttttt gggatgttgc gggatgttgc tttttttttt cttttttttt 2040
acatgttccatc tttttttttt gggatgttgc tttttttttt cttttttttt 2100
gaccatgtt gttttttttt gggatgttgc tttttttttt cttttttttt 2160
aaagatgtgtc aatggggggcc aatggggggcc tttttttttt cttttttttt 2220
ggatggatggaa aatggggggcc tttttttttt cttttttttt 2280
ttgttggatggaa aatggggggcc tttttttttt cttttttttt 2340
ttccatataaa aatggggggcc tttttttttt cttttttttt 2400
gggttgggggtt gggatgttgc tttttttttt cttttttttt 2460
ggatgttgc tttttttttt gggatgttgc tttttttttt 2520
tccctgtatc tttttttttt cttttttttt 2580
agttataatggc tttttttttt cttttttttt 2640
cctttttttt cttttttttt gggatgttgc tttttttttt 2700
tttgcacatggc tttttttttt gggatgttgc tttttttttt 2760
caaggactt tttttttttt gggatgttgc tttttttttt 2820
taatggatggc tttttttttt gggatgttgc tttttttttt 2880
ggccacatc tttttttttt gggatgttgc tttttttttt 2940
ctatagttt gggatgttgc tttttttttt gggatgttgc tttttttttt 3000
tttgcacatggc tttttttttt gggatgttgc tttttttttt 3060
ctatacttac tttttttttt gggatgttgc tttttttttt 3120
ttatTTTTT tttttttttt gggatgttgc tttttttttt 3180
tttgcacatggc tttttttttt gggatgttgc tttttttttt 3240
agggatgttgc tttttttttt gggatgttgc tttttttttt 3300
cttcacatggc tttttttttt gggatgttgc tttttttttt 3360
tccctgtatc tttttttttt gggatgttgc tttttttttt 3420
gacccatggc tttttttttt gggatgttgc tttttttttt 3477
```

```
SEQ ID NO: 285 moltype = DNA length = 2476
FEATURE Location/Qualifiers
source 1..2476
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 285
ttggccactc cctctctgcg cgctcgctcg ctcaactgagg ccgggggagacc aaaggtcgcc 60
```

-continued

gcacggcccg gcttgcggc ggcggccatca gtgagccgac gagcgcgcag agagggagt 120
 gccaactcca tcacttaggg ttccatgatc taagtatatt agagcgcgtc ttctgcaca 180
 cagatcacct ttccatcaa cccactagc ctctggccaa atgaatggg taaccttt 240
 ctcccttcctc ttctgcctcg gctctgtt ccctagggt gtgttgcgc gagaacgcg 300
 taagagttt atgttttgc atctctgtt gtattttct agtaatggaa gcctggatt 360
 taaaatagt taaaatttcc tttagtgctg attttctatgatttactatgttgcgtt 420
 ttatattgtt cattatggc atctgagac ctttttcttg atcatggaaaa cgccaaacaaa 480
 attctgaatc ggcccaaaagg gtataatcctt ggtaatattg aaagggttgt tcacggaaac 540
 cttagagagaa aatgtatgga agaaaaagtgt agtttgcgg aagcgcgaga agttttgaa 600
 aacactgaaa gaacaactgaa attttgcgg cagttatgtt atggagatca gtgtggact 660
 aatccatgt taaaatgggg cagttgcgg gatgacatatttactatgttgcgtt 720
 ccctttggat ttgaaggaaa acatgttgcg tttagatgttgc catgttacat taagaatggc 780
 agatcgagac agttttgtaa aatagtgct gataacaagg tggttgcgc ctgtactgag 840
 ggatatcgac ttgcagaaaa ccagaatgtcc tgtaaccagg cagttgcgtt tccatgtgg 900
 agagtttgc ttccacaaaat ttctatgtcc acccgttgcg agactgtttt tcctgtatgt 960
 gactatgtaa attctactgaa agcttgcgg aatttggatc acatcatca aacggcccaa 1020
 tcatttaatg acttcactcg ggttgggtt ggagaagatg ccaaccagg tcaattccct 1080
 tggcagggtt ttttgaatgg taaagtgtat gcatttcgtt gaggctatcg ctgttatgaa 1140
 aaatggatgtt taactgttgc ccactgtttt gaaactgtgtt taaaatttac agtttgcga 1200
 ggttgcacata atatttgcgg gacacaaatc acaggacaaa acggcaatgtt gatccggatt 1260
 attccctacc acaactacaa tgcagatatt aataagtaca accatgacat tgcccttgc 1320
 gaactggacg aacccttagt gctaaacacg tacgttacac ctatttgcat tgctgacaa 1380
 gaatcacacg acatcttcctt cccatgttgc tctggctatg taatggctg gggaaagatc 1440
 ttcccaaaagg ggagatcgc tttagttttt ctagtacccat gatgttccat tggttgcgg 1500
 gcccacatgtc ttctatctac aaagtgcacc atctatataaca acatgttgc tgctgggtt 1560
 catgaaggaaat gtagagatc atgttgcgg gatagtgggg gacccatgt tactgaatgt 1620
 gaaggggacca gtttcttaac tggaaatatt agctgggggtt aaggtgtgc aatggaaaggc 1680
 aataatggaa tataatccaaat ggtatcccg gttgtcaatc ggatataaggaa aaaaacaaacaa 1740
 ctcaacttaac ctgcactgtt ccccttagt gccacccatc tggttgcgc cccctccccc 1800
 tgccatcttcctt gaccctggaa ggttgcaccat ccactgttcc ttctatataaa aatggggaaa 1860
 ttgcatcgatc ttgttgcgtt aggtgttccat ctattttggg ggggtgggg gggcaggacaa 1920
 gcaaggggggg ggttggggaa gacataatc ggcacatgttgc ggtatgttgc ggtatgttgc 1980
 ttcttcgttgc ggaaaggaaacc agctggggctt ctatggggat tcccccttagt tggttataat 2040
 tattgtatata tttttgtat ctttgcgttca aataatgggg gattttgaaa gtttgcgtt 2100
 aaatattttttaat taaaataaa aaaaatgttgc ggcacatgttgc tcaaaatattc tggtatgtat 2160
 ttgaatatttac tgcgttccat tagggcgttgc ctgttgcacc atgttactatc aaaaatattt 2220
 agtgtgtgtt actaatattttaat taaaatgttgc ttccatttttatttacatttattt 2280
 taccattgtc tttagtagata tttaaaaaaacta tgacagaaaac actaaagatc taggaacccc 2340
 tagtgtatgg gttggccactt ccctctctgc ggcgttgcgtt gtttgcgttgc gcccggccgg 2400
 caaaggcccg ggcgttgcgtt acccttgcgtt gccggccctc agtgcggcgg cggccggccgtc 2460
 gagaggaggat ggcacaa 2476

```

SEQ ID NO: 286 moltype = DNA length = 4175
FEATURE Location/Qualifiers
source 1..4175
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 286
ttggccactt cctcttcgtcg cgctcgtcg ctcactgagg ccgggcgacc a
cgacgcccgg gctttgcccc ggccgcctta gtgagcgagc gagcgcgcag a
gccaatccca tcaactagggg ttccatagatc taatgtatatt agacgagtc t
cagatcacct ttccatcaa ccccaactagc ctctggccaa atgaagtggg t
cctcccttc ttcgtctccg gctctgcctt ttccagggggt gtgttgcgc g
taaggatgtt atgttttttc atctctgcgtt gtatttttct agtaatggaa g
ttaaaatagt taaaattttc tttagtgcgtt atttcttagat tattttact g
ttatttttgtt cattatttgc atctcgaaac cttttttcttg atcatgaaaa c
atctcgaaatc ggccaaagag gtataatcca ggttaaattgg aagagttgt t
cttggagagag aatgtatgga agaaaaatgtt agttttgtt aagcgcgaga a
aacatggaa gaacaactgg aattttggaaag cagttatgtt atggagatca g
atccatgtt taaaatggccgg cagttgcgaatc gatgcacatca attctatca a
ccctttggat ttgaaggaaa gaactgtgaa tttagatgtt catgtaaatc t
agatgcgagc agttttgtt aaatagtgtt gataacaagg tggtttgc t
ggatatccgac ttgcagaaaa ccgcggatcc tggaaacccgg cagttgcctt t
agatgttcc ttccacaacatcc ttcataatgc accccgtgtc agactgtttt t
gactatgttta attctactgg agctgaaatcc atttttggata acatcatca a
tcattttatgtt acttcactgg ggttgggtt ggagaagatg cccaaaccagg t
tggcagggtt ttttgaatgg taaagtgttgcattctgtt gaggctctat c
aatggatgtt taactgtgtc ccactgtgtt gaaactgtgtt taaaattac a
ggtaacatca atatttgaggaa gacagaacat acagagcaaa agcgaatgtt g
atccctacc aacaatccaa tgccatgtt aataagtaca accatcatca t
gaactggacg aacccttagt gctaaacacg tacgttacac ctatttgc t
gaatacacgca acatcttcctt caaatttggaa tctggatcatg taatggctg t
ttccacaacgg ggagatcggc ttttagttctt cagttccatca gagttccact t
gccacatgtt ttctatcttca aaaaatgttccacc atcttataaca acatgttcc t
catgaaggag gttagatgtt acatgtcaaggaa gatagtgggg gacccatgtt q
qaaqqqacca qtttcttaac tqqaattttt aqctgggtt qaaqtgtqc a

```

-continued

aaatatatggaa	tataataccaa	ggtatcccg	tatgtcaact	ggattaaagga	aaaaacaaaag	1740
ctcaactaac	ctcgactgtg	ccttcttagtt	gccagccatc	tgttgttgc	ccctcccccg	1800
tgccttcctt	gaccctggaa	ggtgcacactc	ccactgtcct	ttccataataa	aataggaaaa	1860
ttgcacatcga	ttgtctgagt	agggtcatt	ctattctggg	gggtgggggt	gggcaggaca	1920
gcaagggggaa	ggattgggaa	gacaatagca	ggcatgtcg	ggatgegggt	ggctctatgg	1980
cttctgagggc	ggaaaagaacc	agctggggct	ctagggggta	cccccttag	gtggttatat	2040
tattgtatata	tttttggat	ctttgatgac	aataatgggg	gatttgaaa	gcttagctt	2100
aaatttcctt	taattaaaaaa	aaaatgttag	gcagaatgac	tcaaaatcag	ttggatacag	2160
ttgaattttt	tacggtctca	tagggctgc	ctgctcgacc	atgtataact	aaaattaaaa	2220
agtgtgttgtt	actaattttt	taaatggat	ttccattttt	atttacattt	atttcttatt	2280
taccattgtc	tttagtagata	tttacaaaac	tgacagaaaac	actaaatctt	gagtttgaat	2340
gcacagatata	aaacactttaa	cggttttaa	aaataataat	gttgggtaaa	aaatataact	2400
tttagtgttag	cagagggaa	ccattggcc	cttcagattt	tcctgtaa	atcgggaaact	2460
ggcatcttca	ggggtaggt	taggtcgatg	aagagaagaa	caaaaacgcag	catattacag	2520
tttagttgtca	tcatcaatct	ttaaatatgt	tgtgtgggtt	ttctctcc	gtttccacag	2580
acaaggatgt	gatcgccat	cggtataatg	atttgggaga	acaacattt	aaaggccctgt	2640
aagttataat	gctgaaagcc	cacttaatat	ttctggtagt	attagttaaa	gtttaaaac	2700
acottttttcc	accttggatgt	tgagaattgt	agagcagtgc	tgtccagtag	aaatgtgtgc	2760
attgacatgg	agactgtgg	tctgtgtca	gcaatgtgg	agccagagat	cacaaggcta	2820
tcacagcaact	tgcacatggc	aaatgtact	gagaaggaca	cattcaata	atagttatt	2880
ttaatttgaat	gtatctagcc	atgtgtggct	agtagctct	ttccggaga	gagaatctgg	2940
agccccacatc	taacttggta	agtttggat	tttattttt	atttctggaa	aggctctatga	3000
actatagttt	ttggggcgc	tcacttacta	acttttaatg	caataagatc	catgttatct	3060
tgagaacattt	attttgtctc	ttttgttagat	tgaacccat	tatcatgtgaa	gttaggggtc	3120
tatacttaag	tcacatctcc	aaccttagta	atgttttaat	gtatgtaaaa	aatgatgtat	3180
taattttttt	tttagaaggta	aatagtatca	tgatattccaa	ataacagagg	tatatggta	3240
gaaaagaaac	atttcaaaagg	acttataata	tatctatgt	tgacaatgaa	taaattttaga	3300
gagtagtttgc	ctgtttgtcc	tcatgttcat	aaatcttattg	acacatatgt	gcatctgcac	3360
ttcagcatgg	tagaagtcc	tatcccttgc	cttggaaagg	cagggtgtcc	cattacgcct	3420
cagagaatag	ctgacgggaa	gagggtttt	atagtagtgc	atgaaaagata	tacaaaatct	3480
cgcaggatata	cacaggcatg	atttgcgtt	tggggagcgc	acttgcctca	tactgagggt	3540
ttttgtgtct	cttttgcag	tcctgttgc	ctttccccc	tatctccaga	aatgtctata	3600
cgtatgacat	gcacaaattt	tgcaaggatg	aaacagacttt	gcaaaagacgt	gtgttgcgcg	3660
tgagtctgc	gcacaaactgt	acaaaatccct	tgtgagttac	ttctgtat	gtggatctac	3720
ttccatgtt	tctggatctt	tgttccaaag	ccaaatcatg	ctccatca	taaggccccg	3780
ggaacacatgt	ggcagaggc	acagcagaga	ttgataaaac	cagggtgtat	ggaattttt	3840
gtgggactcc	atttcatatgt	aatttgcagaa	gctacaatac	actcaaaaag	tctcaccacaca	3900
tgactgccc	aatggggatct	tgacagtgc	agtgcacgt	gatatgc	agtggatgag	3960
ggaaaagatcca	caagagatca	accctgtaaa	aaqaaactgt	ggcaactaa	gaatgcagag	4020
agaaagatct	aggaaacccct	agtgtggag	ttggccactc	cctctctgc	cgctcgctcg	4080
ctcaactgagg	ccggccgggg	aaagccggg	cgtcgccgca	cctttggctc	cccgccctca	4140
gtgagcgagc	gagcgccg	agaggaggatg	gccaa			4175

SEQ ID NO: 287 moltype = DNA length = 3675
 FEATURE Location/Qualifiers
 source 1..3675
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 287

ttggccactc	cctctctgcg	cgctcgctcg	ctcaactgagg	ccggccgacc	aaaggtcgcc	60
cgacgcgggg	gttttgcggc	ggggccctca	gtgagcgacg	gagcgcggcag	agagggagtg	120
gccaacttca	tcactatgggg	ttccatagatc	taatgtatata	agagcgatc	tttctgcaca	180
caagatcacct	ttccatcaa	ccccatcgtc	ctctggcaaa	atgaatgtgg	taacctttt	240
cctccctctc	ttcgctccg	gctctgtt	ttccagggtt	gtgtttcgcc	gagaagcagc	300
taagagtttt	atgtttttc	atctctgtt	gtatttttct	agtaatggaa	gcctgttatt	360
ttaaaatatgt	taaattttcc	tttagtgcgt	atttcttagat	tatttattat	gttgggttt	420
tttattttgt	cattatgtc	atctgagaa	ctttttcttg	atcatgaaaa	cgcaacacaa	480
attctgtatc	ggccaaagag	gtataattca	ggtaaattgg	aagagtttt	tcaagggaa	540
cttgagagag	aatgtatgg	agaaaaatgt	agttttgt	agacgcggaga	agttttgaa	600
aacactgtaa	gaacaactgt	attttggatg	cagtatgttt	atggagatca	gtgtgagtc	660
atccatgtt	taatatggcg	cagtggcaag	gtatgcattat	atccatgtat	atgtgtgt	720
ccctttggat	ttgaaggaaa	gaactgtgaa	ttagatgtaa	catgtaacat	taagaatggc	780
agatgcgagc	agttttgtaa	aaatagtgt	gataacaagg	tgggtttgtc	ctgtactgt	840
ggatatacgac	ttgcagaaaa	ccaaatgtcc	tgtgaaccag	cagtgcattt	tcctatgtgg	900
agagtttctg	ttccacaaaac	ttctaaatgtc	accgtgtctg	agactgttt	tcctgtat	960
gactatgtaa	atttctactgt	agctgtaaa	attttggata	acatcaatca	aagcacccaa	1020
tcatttaatgt	acttctactgt	gggtgttgg	ggagaagatg	ccaaacccagg	tcaattccct	1080
tggcaggatgt	ttttgtatgg	taaagtgtat	gcattctgt	gagggtctat	cgttaatgtaa	1140
aaatggatgt	taactgtctg	ccactgtgtt	gaaaatgtgt	ttaaaaattac	agtgtcgca	1200
ggtgaacata	atatttggat	gacagaacat	acagagcaaa	agcgtaaatgt	gatcgtat	1260
attccttacc	acaactacaa	tgcgtat	aataagtaca	accatgcatt	tgcccttctg	1320
gaactggacg	aacccttagt	gctaaacago	tacgttacac	ctatgtat	tgctgacaa	1380
gaatacacgca	acatcttcc	caaatttgg	tctggctat	taatgtgg	ggggagatgc	1440
ttccacaaaag	ggagatc	tttagttt	cagtaccc	gatgttccact	tgttgacccg	1500
gcccacatgtc	ttctatctac	aaatgttac	atctataaca	acatgttctg	tgctggctc	1560
catgaaggag	gtagagatc	atgtcaagga	gatgttgggg	gaccccatgt	tactgtgt	1620

-continued

gaagggacca	gtttttaac	tggaattatt	agctgggtg	aagagtgtc	aatgaaaggc	1680
aaatatggaa	tataaccaa	ggatatcccg	tatgtcaact	ggattaaggaa	aaaaacaag	1740
ctcaactaac	ctcgactgt	ccttctatgt	gccagccatc	tgtgtttgc	ccctcccccg	1800
tgcccttc	gaccctggaa	ggtgcactc	ccactgtct	ttcctaataa	aataggaaaa	1860
ttgcacatcgca	ttgtctgagt	agggtcatt	ctattctggg	gggtgggtg	gggcaggaca	1920
gcaaggggga	ggatttggaa	gacaatagca	ggcatgtgg	ggatgegggtg	ggctctatgg	1980
cttctgagcc	ggaaaagaacc	agctgggtc	ctagggggta	cccccttag	gtggttatat	2040
tattgtatata	tttttgttat	ctttgtatgc	aataatgggg	gattttgaaa	gttttagctt	2100
aaattttttt	taattaaaaaa	aaaatgttag	gcagaatgac	tcaaaattacg	ttggatacacag	2160
ttgaattttt	tacggctcta	tagggcctgc	ctgctcgacc	atgtataact	aaaaattaaa	2220
agtgtgtgtt	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	2280
taccatgtc	tttagtagata	tttacaaaaca	tgacagaaaac	actaaatctt	gagtttgaat	2340
gcacagatata	aaacacttaa	cgggttttaa	aaataataat	gttggtaaaa	aaataataact	2400
ttttagttag	ttttagttag	ttttagttag	ttttagttag	ttttagttag	ttttagttag	2460
ggcatcttc	ggggatgtct	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	2520
tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	2580
acaagagtga	gtatcccatt	cggtataatg	atttgggaga	acaacattc	aaaggctgt	2640
aaatgtataat	gttggaaagcc	cacttaataat	ttctggtagt	attatgtaaa	gtttttaaaac	2700
acctttttcc	accttggatgt	ttttagttagt	ttttagttagt	ttttagttagt	ttttagttagt	2760
ttttagttagt	ttttagttagt	ttttagttagt	ttttagttagt	ttttagttagt	ttttagttagt	2820
tcaagcactt	tgcacatggc	aagtgttaact	gagaaggcaca	cattcaaaa	atatgttaatt	2880
tttaattgtat	gtatctagcc	atgtgtgggt	atgtgtccct	ttccctggaga	gagaatctgg	2940
agccccatc	taacttggta	atgtgtggta	ttttagttagt	ttttagttagt	ttttagttagt	3000
actataatgt	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3060
tgagaacatt	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3120
tataacttaag	tcacatctcc	aaccttagta	atgttttaat	gttagaaaaaa	aatgatgtat	3180
taattttttt	ttttagaaggc	aatatgtata	ttttagttagt	ttttagttagt	ttttagttagt	3240
gaaaaaaaac	aatttcaaaaagg	atcttataaa	tatttttttttta	ttttagttagt	ttttagttagt	3300
gagtagttt	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3360
ttttagttagt	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3420
cagagaatag	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3480
cggcggatata	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3540
tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3600
aaagccccgg	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3660
tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	tttttttttttta	3675

SEQ ID NO: 288 moltype = length =
 SEQUENCE: 288
 000

SEQ ID NO: 289 moltype = length =
 SEQUENCE: 289
 000

SEQ ID NO: 290 moltype = length =
 SEQUENCE: 290
 000

SEQ ID NO: 291 moltype = length =
 SEQUENCE: 291
 000

SEQ ID NO: 292 moltype = length =
 SEQUENCE: 292
 000

SEQ ID NO: 293 moltype = length =
 SEQUENCE: 293
 000

SEQ ID NO: 294 moltype = length =
 SEQUENCE: 294
 000

SEQ ID NO: 295 moltype = length =
 SEQUENCE: 295
 000

SEQ ID NO: 296 moltype = length =
 SEQUENCE: 296
 000

SEQ ID NO: 297 moltype = length =
 SEQUENCE: 297
 000

-continued

SEQ ID NO: 298 moltype = length =
SEQUENCE: 298
000

SEQ ID NO: 299 moltype = length =
SEQUENCE: 299
000

SEQ ID NO: 300 moltype = RNA length = 100
FEATURE Location/Qualifiers
misc_difference 1..20
note = n=a, c, u, g, unknown or other
source 1..100
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 300
nnnnnnnnnn nnnnnnnnnn gtttagago tagaaatagg aagttaaat aaggctagtc 60
cgtttatcaac ttgaaaaagt ggcaccgagt cggtgccttt 100

SEQ ID NO: 301 moltype = length =
SEQUENCE: 301
000

SEQ ID NO: 302 moltype = length =
SEQUENCE: 302
000

SEQ ID NO: 303 moltype = length =
SEQUENCE: 303
000

SEQ ID NO: 304 moltype = length =
SEQUENCE: 304
000

SEQ ID NO: 305 moltype = length =
SEQUENCE: 305
000

SEQ ID NO: 306 moltype = length =
SEQUENCE: 306
000

SEQ ID NO: 307 moltype = length =
SEQUENCE: 307
000

SEQ ID NO: 308 moltype = length =
SEQUENCE: 308
000

SEQ ID NO: 309 moltype = length =
SEQUENCE: 309
000

SEQ ID NO: 310 moltype = length =
SEQUENCE: 310
000

SEQ ID NO: 311 moltype = length =
SEQUENCE: 311
000

SEQ ID NO: 312 moltype = length =
SEQUENCE: 312
000

SEQ ID NO: 313 moltype = length =
SEQUENCE: 313
000

SEQ ID NO: 314 moltype = length =
SEQUENCE: 314
000

SEQ ID NO: 315 moltype = length =

-continued

SEQUENCE: 315
000

SEQ_ID NO: 316 moltype = length =
SEQUENCE: 316
000

SEQ_ID NO: 317 moltype = length =
SEQUENCE: 317
000

SEQ_ID NO: 318 moltype = length =
SEQUENCE: 318
000

SEQ_ID NO: 319 moltype = length =
SEQUENCE: 319
000

SEQ_ID NO: 320 moltype = length =
SEQUENCE: 320
000

SEQ_ID NO: 321 moltype = length =
SEQUENCE: 321
000

SEQ_ID NO: 322 moltype = length =
SEQUENCE: 322
000

SEQ_ID NO: 323 moltype = length =
SEQUENCE: 323
000

SEQ_ID NO: 324 moltype = length =
SEQUENCE: 324
000

SEQ_ID NO: 325 moltype = length =
SEQUENCE: 325
000

SEQ_ID NO: 326 moltype = length =
SEQUENCE: 326
000

SEQ_ID NO: 327 moltype = length =
SEQUENCE: 327
000

SEQ_ID NO: 328 moltype = length =
SEQUENCE: 328
000

SEQ_ID NO: 329 moltype = length =
SEQUENCE: 329
000

SEQ_ID NO: 330 moltype = length =
SEQUENCE: 330
000

SEQ_ID NO: 331 moltype = length =
SEQUENCE: 331
000

SEQ_ID NO: 332 moltype = length =
SEQUENCE: 332
000

SEQ_ID NO: 333 moltype = length =
SEQUENCE: 333
000

SEQ_ID NO: 334 moltype = length =

-continued

```
SEQUENCE: 334
000

SEQ_ID NO: 335      moltype =  length =
SEQUENCE: 335
000

SEQ_ID NO: 336      moltype =  length =
SEQUENCE: 336
000

SEQ_ID NO: 337      moltype =  length =
SEQUENCE: 337
000

SEQ_ID NO: 338      moltype =  length =
SEQUENCE: 338
000

SEQ_ID NO: 339      moltype =  length =
SEQUENCE: 339
000

SEQ_ID NO: 340      moltype =  length =
SEQUENCE: 340
000

SEQ_ID NO: 341      moltype =  length =
SEQUENCE: 341
000

SEQ_ID NO: 342      moltype =  length =
SEQUENCE: 342
000

SEQ_ID NO: 343      moltype =  length =
SEQUENCE: 343
000

SEQ_ID NO: 344      moltype =  length =
SEQUENCE: 344
000

SEQ_ID NO: 345      moltype =  length =
SEQUENCE: 345
000

SEQ_ID NO: 346      moltype =  length =
SEQUENCE: 346
000

SEQ_ID NO: 347      moltype =  length =
SEQUENCE: 347
000

SEQ_ID NO: 348      moltype =  length =
SEQUENCE: 348
000

SEQ_ID NO: 349      moltype =  length =
SEQUENCE: 349
000

SEQ_ID NO: 350      moltype =  length =
SEQUENCE: 350
000

SEQ_ID NO: 351      moltype =  length =
SEQUENCE: 351
000

SEQ_ID NO: 352      moltype =  length =
SEQUENCE: 352
000

SEQ_ID NO: 353      moltype =  length =
```

-continued

SEQUENCE: 353
000

SEQ_ID NO: 354 moltype = length =
SEQUENCE: 354
000

SEQ_ID NO: 355 moltype = length =
SEQUENCE: 355
000

SEQ_ID NO: 356 moltype = length =
SEQUENCE: 356
000

SEQ_ID NO: 357 moltype = length =
SEQUENCE: 357
000

SEQ_ID NO: 358 moltype = length =
SEQUENCE: 358
000

SEQ_ID NO: 359 moltype = length =
SEQUENCE: 359
000

SEQ_ID NO: 360 moltype = length =
SEQUENCE: 360
000

SEQ_ID NO: 361 moltype = length =
SEQUENCE: 361
000

SEQ_ID NO: 362 moltype = length =
SEQUENCE: 362
000

SEQ_ID NO: 363 moltype = length =
SEQUENCE: 363
000

SEQ_ID NO: 364 moltype = length =
SEQUENCE: 364
000

SEQ_ID NO: 365 moltype = length =
SEQUENCE: 365
000

SEQ_ID NO: 366 moltype = length =
SEQUENCE: 366
000

SEQ_ID NO: 367 moltype = length =
SEQUENCE: 367
000

SEQ_ID NO: 368 moltype = length =
SEQUENCE: 368
000

SEQ_ID NO: 369 moltype = length =
SEQUENCE: 369
000

SEQ_ID NO: 370 moltype = length =
SEQUENCE: 370
000

SEQ_ID NO: 371 moltype = length =
SEQUENCE: 371
000

SEQ_ID NO: 372 moltype = length =

-continued

SEQUENCE: 372
000

SEQ_ID NO: 373 moltype = length =
SEQUENCE: 373
000

SEQ_ID NO: 374 moltype = length =
SEQUENCE: 374
000

SEQ_ID NO: 375 moltype = length =
SEQUENCE: 375
000

SEQ_ID NO: 376 moltype = length =
SEQUENCE: 376
000

SEQ_ID NO: 377 moltype = length =
SEQUENCE: 377
000

SEQ_ID NO: 378 moltype = length =
SEQUENCE: 378
000

SEQ_ID NO: 379 moltype = length =
SEQUENCE: 379
000

SEQ_ID NO: 380 moltype = length =
SEQUENCE: 380
000

SEQ_ID NO: 381 moltype = length =
SEQUENCE: 381
000

SEQ_ID NO: 382 moltype = length =
SEQUENCE: 382
000

SEQ_ID NO: 383 moltype = length =
SEQUENCE: 383
000

SEQ_ID NO: 384 moltype = length =
SEQUENCE: 384
000

SEQ_ID NO: 385 moltype = length =
SEQUENCE: 385
000

SEQ_ID NO: 386 moltype = length =
SEQUENCE: 386
000

SEQ_ID NO: 387 moltype = length =
SEQUENCE: 387
000

SEQ_ID NO: 388 moltype = length =
SEQUENCE: 388
000

SEQ_ID NO: 389 moltype = length =
SEQUENCE: 389
000

SEQ_ID NO: 390 moltype = length =
SEQUENCE: 390
000

SEQ_ID NO: 391 moltype = length =

-continued

SEQUENCE: 391
000

SEQ_ID NO: 392 moltype = length =
SEQUENCE: 392
000

SEQ_ID NO: 393 moltype = length =
SEQUENCE: 393
000

SEQ_ID NO: 394 moltype = length =
SEQUENCE: 394
000

SEQ_ID NO: 395 moltype = length =
SEQUENCE: 395
000

SEQ_ID NO: 396 moltype = length =
SEQUENCE: 396
000

SEQ_ID NO: 397 moltype = length =
SEQUENCE: 397
000

SEQ_ID NO: 398 moltype = length =
SEQUENCE: 398
000

SEQ_ID NO: 399 moltype = length =
SEQUENCE: 399
000

SEQ_ID NO: 400 moltype = RNA length = 22
FEATURE
source
1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 400
gttttagagc tatgctgttt tg 22

SEQ_ID NO: 401 moltype = RNA length = 80
FEATURE
source
1..80
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 401
gttttagagc tagaaatagc aagttaaaat aaggctagtc cgtttatcaac ttgaaaaagt 60
ggcaccgagt cggtgctttt 80

SEQ_ID NO: 402 moltype = RNA length = 76
FEATURE
source
1..76
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 402
gttttagagc tagaaatagc aagttaaaat aaggctagtc cgtttatcaac ttgaaaaagt 60
ggcaccgagt cggtg 76

SEQ_ID NO: 403 moltype = length =
SEQUENCE: 403
000

SEQ_ID NO: 404 moltype = length =
SEQUENCE: 404
000

SEQ_ID NO: 405 moltype = length =
SEQUENCE: 405
000

SEQ_ID NO: 406 moltype = length =
SEQUENCE: 406
000

-continued

```
SEQ ID NO: 407      moltype = length =
SEQUENCE: 407
000

SEQ ID NO: 408      moltype = length =
SEQUENCE: 408
000

SEQ ID NO: 409      moltype = length =
SEQUENCE: 409
000

SEQ ID NO: 410      moltype = length =
SEQUENCE: 410
000

SEQ ID NO: 411      moltype = length =
SEQUENCE: 411
000

SEQ ID NO: 412      moltype = length =
SEQUENCE: 412
000

SEQ ID NO: 413      moltype = length =
SEQUENCE: 413
000

SEQ ID NO: 414      moltype = length =
SEQUENCE: 414
000

SEQ ID NO: 415      moltype = length =
SEQUENCE: 415
000

SEQ ID NO: 416      moltype = length =
SEQUENCE: 416
000

SEQ ID NO: 417      moltype = length =
SEQUENCE: 417
000

SEQ ID NO: 418      moltype = length =
SEQUENCE: 418
000

SEQ ID NO: 419      moltype = length =
SEQUENCE: 419
000

SEQ ID NO: 420      moltype = length =
SEQUENCE: 420
000

SEQ ID NO: 421      moltype = length =
SEQUENCE: 421
000

SEQ ID NO: 422      moltype = length =
SEQUENCE: 422
000

SEQ ID NO: 423      moltype = length =
SEQUENCE: 423
000

SEQ ID NO: 424      moltype = length =
SEQUENCE: 424
000

SEQ ID NO: 425      moltype = length =
SEQUENCE: 425
000
```

-continued

```
SEQ ID NO: 426      moltype = length =
SEQUENCE: 426
000

SEQ ID NO: 427      moltype = length =
SEQUENCE: 427
000

SEQ ID NO: 428      moltype = length =
SEQUENCE: 428
000

SEQ ID NO: 429      moltype = length =
SEQUENCE: 429
000

SEQ ID NO: 430      moltype = length =
SEQUENCE: 430
000

SEQ ID NO: 431      moltype = length =
SEQUENCE: 431
000

SEQ ID NO: 432      moltype = length =
SEQUENCE: 432
000

SEQ ID NO: 433      moltype = length =
SEQUENCE: 433
000

SEQ ID NO: 434      moltype = length =
SEQUENCE: 434
000

SEQ ID NO: 435      moltype = length =
SEQUENCE: 435
000

SEQ ID NO: 436      moltype = length =
SEQUENCE: 436
000

SEQ ID NO: 437      moltype = length =
SEQUENCE: 437
000

SEQ ID NO: 438      moltype = length =
SEQUENCE: 438
000

SEQ ID NO: 439      moltype = length =
SEQUENCE: 439
000

SEQ ID NO: 440      moltype = length =
SEQUENCE: 440
000

SEQ ID NO: 441      moltype = length =
SEQUENCE: 441
000

SEQ ID NO: 442      moltype = length =
SEQUENCE: 442
000

SEQ ID NO: 443      moltype = length =
SEQUENCE: 443
000

SEQ ID NO: 444      moltype = length =
SEQUENCE: 444
000
```

-continued

SEQ ID NO: 445 moltype = length =
SEQUENCE: 445
000

SEQ ID NO: 446 moltype = length =
SEQUENCE: 446
000

SEQ ID NO: 447 moltype = length =
SEQUENCE: 447
000

SEQ ID NO: 448 moltype = length =
SEQUENCE: 448
000

SEQ ID NO: 449 moltype = length =
SEQUENCE: 449
000

SEQ ID NO: 450 moltype = length =
SEQUENCE: 450
000

SEQ ID NO: 451 moltype = length =
SEQUENCE: 451
000

SEQ ID NO: 452 moltype = length =
SEQUENCE: 452
000

SEQ ID NO: 453 moltype = length =
SEQUENCE: 453
000

SEQ ID NO: 454 moltype = length =
SEQUENCE: 454
000

SEQ ID NO: 455 moltype = length =
SEQUENCE: 455
000

SEQ ID NO: 456 moltype = length =
SEQUENCE: 456
000

SEQ ID NO: 457 moltype = length =
SEQUENCE: 457
000

SEQ ID NO: 458 moltype = length =
SEQUENCE: 458
000

SEQ ID NO: 459 moltype = length =
SEQUENCE: 459
000

SEQ ID NO: 460 moltype = length =
SEQUENCE: 460
000

SEQ ID NO: 461 moltype = length =
SEQUENCE: 461
000

SEQ ID NO: 462 moltype = length =
SEQUENCE: 462
000

SEQ ID NO: 463 moltype = length =
SEQUENCE: 463
000

-continued

```
SEQ ID NO: 464      moltype = length =
SEQUENCE: 464
000

SEQ ID NO: 465      moltype = length =
SEQUENCE: 465
000

SEQ ID NO: 466      moltype = length =
SEQUENCE: 466
000

SEQ ID NO: 467      moltype = length =
SEQUENCE: 467
000

SEQ ID NO: 468      moltype = length =
SEQUENCE: 468
000

SEQ ID NO: 469      moltype = length =
SEQUENCE: 469
000

SEQ ID NO: 470      moltype = length =
SEQUENCE: 470
000

SEQ ID NO: 471      moltype = length =
SEQUENCE: 471
000

SEQ ID NO: 472      moltype = length =
SEQUENCE: 472
000

SEQ ID NO: 473      moltype = length =
SEQUENCE: 473
000

SEQ ID NO: 474      moltype = length =
SEQUENCE: 474
000

SEQ ID NO: 475      moltype = length =
SEQUENCE: 475
000

SEQ ID NO: 476      moltype = length =
SEQUENCE: 476
000

SEQ ID NO: 477      moltype = length =
SEQUENCE: 477
000

SEQ ID NO: 478      moltype = length =
SEQUENCE: 478
000

SEQ ID NO: 479      moltype = length =
SEQUENCE: 479
000

SEQ ID NO: 480      moltype = length =
SEQUENCE: 480
000

SEQ ID NO: 481      moltype = length =
SEQUENCE: 481
000

SEQ ID NO: 482      moltype = length =
SEQUENCE: 482
000
```

-continued

SEQ ID NO: 483 moltype = length =
SEQUENCE: 483
000

SEQ ID NO: 484 moltype = length =
SEQUENCE: 484
000

SEQ ID NO: 485 moltype = length =
SEQUENCE: 485
000

SEQ ID NO: 486 moltype = length =
SEQUENCE: 486
000

SEQ ID NO: 487 moltype = length =
SEQUENCE: 487
000

SEQ ID NO: 488 moltype = length =
SEQUENCE: 488
000

SEQ ID NO: 489 moltype = length =
SEQUENCE: 489
000

SEQ ID NO: 490 moltype = length =
SEQUENCE: 490
000

SEQ ID NO: 491 moltype = length =
SEQUENCE: 491
000

SEQ ID NO: 492 moltype = length =
SEQUENCE: 492
000

SEQ ID NO: 493 moltype = length =
SEQUENCE: 493
000

SEQ ID NO: 494 moltype = length =
SEQUENCE: 494
000

SEQ ID NO: 495 moltype = length =
SEQUENCE: 495
000

SEQ ID NO: 496 moltype = length =
SEQUENCE: 496
000

SEQ ID NO: 497 moltype = length =
SEQUENCE: 497
000

SEQ ID NO: 498 moltype = length =
SEQUENCE: 498
000

SEQ ID NO: 499 moltype = length =
SEQUENCE: 499
000

SEQ ID NO: 500 moltype = length =
SEQUENCE: 500
000

SEQ ID NO: 501 moltype = length =
SEQUENCE: 501
000

-continued

```
SEQ ID NO: 502      moltype =  length =
SEQUENCE: 502
000

SEQ ID NO: 503      moltype =  length =
SEQUENCE: 503
000

SEQ ID NO: 504      moltype =  length =
SEQUENCE: 504
000

SEQ ID NO: 505      moltype =  length =
SEQUENCE: 505
000

SEQ ID NO: 506      moltype =  length =
SEQUENCE: 506
000

SEQ ID NO: 507      moltype =  length =
SEQUENCE: 507
000

SEQ ID NO: 508      moltype =  length =
SEQUENCE: 508
000

SEQ ID NO: 509      moltype =  length =
SEQUENCE: 509
000

SEQ ID NO: 510      moltype =  length =
SEQUENCE: 510
000

SEQ ID NO: 511      moltype =  length =
SEQUENCE: 511
000

SEQ ID NO: 512      moltype =  length =
SEQUENCE: 512
000

SEQ ID NO: 513      moltype =  length =
SEQUENCE: 513
000

SEQ ID NO: 514      moltype =  length =
SEQUENCE: 514
000

SEQ ID NO: 515      moltype =  length =
SEQUENCE: 515
000

SEQ ID NO: 516      moltype =  length =
SEQUENCE: 516
000

SEQ ID NO: 517      moltype =  length =
SEQUENCE: 517
000

SEQ ID NO: 518      moltype =  length =
SEQUENCE: 518
000

SEQ ID NO: 519      moltype =  length =
SEQUENCE: 519
000

SEQ ID NO: 520      moltype =  length =
SEQUENCE: 520
000
```

-continued

```
SEQ ID NO: 521      moltype = length =
SEQUENCE: 521
000

SEQ ID NO: 522      moltype = length =
SEQUENCE: 522
000

SEQ ID NO: 523      moltype = length =
SEQUENCE: 523
000

SEQ ID NO: 524      moltype = length =
SEQUENCE: 524
000

SEQ ID NO: 525      moltype = length =
SEQUENCE: 525
000

SEQ ID NO: 526      moltype = length =
SEQUENCE: 526
000

SEQ ID NO: 527      moltype = length =
SEQUENCE: 527
000

SEQ ID NO: 528      moltype = length =
SEQUENCE: 528
000

SEQ ID NO: 529      moltype = length =
SEQUENCE: 529
000

SEQ ID NO: 530      moltype = length =
SEQUENCE: 530
000

SEQ ID NO: 531      moltype = length =
SEQUENCE: 531
000

SEQ ID NO: 532      moltype = length =
SEQUENCE: 532
000

SEQ ID NO: 533      moltype = length =
SEQUENCE: 533
000

SEQ ID NO: 534      moltype = length =
SEQUENCE: 534
000

SEQ ID NO: 535      moltype = length =
SEQUENCE: 535
000

SEQ ID NO: 536      moltype = length =
SEQUENCE: 536
000

SEQ ID NO: 537      moltype = length =
SEQUENCE: 537
000

SEQ ID NO: 538      moltype = length =
SEQUENCE: 538
000

SEQ ID NO: 539      moltype = length =
SEQUENCE: 539
000
```

-continued

```
SEQ ID NO: 540      moltype =  length =
SEQUENCE: 540
000

SEQ ID NO: 541      moltype =  length =
SEQUENCE: 541
000

SEQ ID NO: 542      moltype =  length =
SEQUENCE: 542
000

SEQ ID NO: 543      moltype =  length =
SEQUENCE: 543
000

SEQ ID NO: 544      moltype =  length =
SEQUENCE: 544
000

SEQ ID NO: 545      moltype =  length =
SEQUENCE: 545
000

SEQ ID NO: 546      moltype =  length =
SEQUENCE: 546
000

SEQ ID NO: 547      moltype =  length =
SEQUENCE: 547
000

SEQ ID NO: 548      moltype =  length =
SEQUENCE: 548
000

SEQ ID NO: 549      moltype =  length =
SEQUENCE: 549
000

SEQ ID NO: 550      moltype =  length =
SEQUENCE: 550
000

SEQ ID NO: 551      moltype =  length =
SEQUENCE: 551
000

SEQ ID NO: 552      moltype =  length =
SEQUENCE: 552
000

SEQ ID NO: 553      moltype =  length =
SEQUENCE: 553
000

SEQ ID NO: 554      moltype =  length =
SEQUENCE: 554
000

SEQ ID NO: 555      moltype =  length =
SEQUENCE: 555
000

SEQ ID NO: 556      moltype =  length =
SEQUENCE: 556
000

SEQ ID NO: 557      moltype =  length =
SEQUENCE: 557
000

SEQ ID NO: 558      moltype =  length =
SEQUENCE: 558
000
```

-continued

```
SEQ ID NO: 559      moltype = length =
SEQUENCE: 559
000

SEQ ID NO: 560      moltype = length =
SEQUENCE: 560
000

SEQ ID NO: 561      moltype = length =
SEQUENCE: 561
000

SEQ ID NO: 562      moltype = length =
SEQUENCE: 562
000

SEQ ID NO: 563      moltype = length =
SEQUENCE: 563
000

SEQ ID NO: 564      moltype = length =
SEQUENCE: 564
000

SEQ ID NO: 565      moltype = length =
SEQUENCE: 565
000

SEQ ID NO: 566      moltype = length =
SEQUENCE: 566
000

SEQ ID NO: 567      moltype = length =
SEQUENCE: 567
000

SEQ ID NO: 568      moltype = length =
SEQUENCE: 568
000

SEQ ID NO: 569      moltype = length =
SEQUENCE: 569
000

SEQ ID NO: 570      moltype = length =
SEQUENCE: 570
000

SEQ ID NO: 571      moltype = length =
SEQUENCE: 571
000

SEQ ID NO: 572      moltype = length =
SEQUENCE: 572
000

SEQ ID NO: 573      moltype = length =
SEQUENCE: 573
000

SEQ ID NO: 574      moltype = length =
SEQUENCE: 574
000

SEQ ID NO: 575      moltype = length =
SEQUENCE: 575
000

SEQ ID NO: 576      moltype = length =
SEQUENCE: 576
000

SEQ ID NO: 577      moltype = length =
SEQUENCE: 577
000
```

-continued

```
SEQ ID NO: 578      moltype = length =
SEQUENCE: 578
000

SEQ ID NO: 579      moltype = length =
SEQUENCE: 579
000

SEQ ID NO: 580      moltype = length =
SEQUENCE: 580
000

SEQ ID NO: 581      moltype = length =
SEQUENCE: 581
000

SEQ ID NO: 582      moltype = length =
SEQUENCE: 582
000

SEQ ID NO: 583      moltype = length =
SEQUENCE: 583
000

SEQ ID NO: 584      moltype = length =
SEQUENCE: 584
000

SEQ ID NO: 585      moltype = length =
SEQUENCE: 585
000

SEQ ID NO: 586      moltype = length =
SEQUENCE: 586
000

SEQ ID NO: 587      moltype = length =
SEQUENCE: 587
000

SEQ ID NO: 588      moltype = length =
SEQUENCE: 588
000

SEQ ID NO: 589      moltype = length =
SEQUENCE: 589
000

SEQ ID NO: 590      moltype = length =
SEQUENCE: 590
000

SEQ ID NO: 591      moltype = length =
SEQUENCE: 591
000

SEQ ID NO: 592      moltype = length =
SEQUENCE: 592
000

SEQ ID NO: 593      moltype = length =
SEQUENCE: 593
000

SEQ ID NO: 594      moltype = length =
SEQUENCE: 594
000

SEQ ID NO: 595      moltype = length =
SEQUENCE: 595
000

SEQ ID NO: 596      moltype = length =
SEQUENCE: 596
000
```

-continued

SEQ ID NO: 597 moltype = length =
SEQUENCE: 597
000

SEQ ID NO: 598 moltype = length =
SEQUENCE: 598
000

SEQ ID NO: 599 moltype = length =
SEQUENCE: 599
000

SEQ ID NO: 600 moltype = AA length = 7
FEATURE Location/Qualifiers
source 1..7
mol_type = protein
organism = synthetic construct
REGION 1..7
note = Simian virus 40 NLS
SEQUENCE: 600
PKKKRKV 7

SEQ ID NO: 601 moltype = AA length = 7
FEATURE Location/Qualifiers
source 1..7
mol_type = protein
organism = synthetic construct
REGION 1..7
note = Simian virus 40 NLS
SEQUENCE: 601
PKKKRRV 7

SEQ ID NO: 602 moltype = AA length = 16
FEATURE Location/Qualifiers
REGION 1..16
note = Nucleoplasmin bipartite NLS sequence
source 1..16
mol_type = protein
organism = synthetic construct
SEQUENCE: 602
KRPAATKKAG QAKKKK 16

SEQ ID NO: 603 moltype = length =
SEQUENCE: 603
000

SEQ ID NO: 604 moltype = length =
SEQUENCE: 604
000

SEQ ID NO: 605 moltype = length =
SEQUENCE: 605
000

SEQ ID NO: 606 moltype = length =
SEQUENCE: 606
000

SEQ ID NO: 607 moltype = length =
SEQUENCE: 607
000

SEQ ID NO: 608 moltype = length =
SEQUENCE: 608
000

SEQ ID NO: 609 moltype = length =
SEQUENCE: 609
000

SEQ ID NO: 610 moltype = length =
SEQUENCE: 610
000

SEQ ID NO: 611 moltype = length =

-continued

SEQUENCE: 611
000

SEQ_ID NO: 612 moltype = length =
SEQUENCE: 612
000

SEQ_ID NO: 613 moltype = length =
SEQUENCE: 613
000

SEQ_ID NO: 614 moltype = length =
SEQUENCE: 614
000

SEQ_ID NO: 615 moltype = length =
SEQUENCE: 615
000

SEQ_ID NO: 616 moltype = length =
SEQUENCE: 616
000

SEQ_ID NO: 617 moltype = length =
SEQUENCE: 617
000

SEQ_ID NO: 618 moltype = length =
SEQUENCE: 618
000

SEQ_ID NO: 619 moltype = length =
SEQUENCE: 619
000

SEQ_ID NO: 620 moltype = length =
SEQUENCE: 620
000

SEQ_ID NO: 621 moltype = length =
SEQUENCE: 621
000

SEQ_ID NO: 622 moltype = length =
SEQUENCE: 622
000

SEQ_ID NO: 623 moltype = length =
SEQUENCE: 623
000

SEQ_ID NO: 624 moltype = length =
SEQUENCE: 624
000

SEQ_ID NO: 625 moltype = length =
SEQUENCE: 625
000

SEQ_ID NO: 626 moltype = length =
SEQUENCE: 626
000

SEQ_ID NO: 627 moltype = length =
SEQUENCE: 627
000

SEQ_ID NO: 628 moltype = length =
SEQUENCE: 628
000

SEQ_ID NO: 629 moltype = length =
SEQUENCE: 629
000

SEQ_ID NO: 630 moltype = length =

-continued

SEQUENCE: 630
000

SEQ_ID NO: 631 moltype = length =
SEQUENCE: 631
000

SEQ_ID NO: 632 moltype = length =
SEQUENCE: 632
000

SEQ_ID NO: 633 moltype = length =
SEQUENCE: 633
000

SEQ_ID NO: 634 moltype = length =
SEQUENCE: 634
000

SEQ_ID NO: 635 moltype = length =
SEQUENCE: 635
000

SEQ_ID NO: 636 moltype = length =
SEQUENCE: 636
000

SEQ_ID NO: 637 moltype = length =
SEQUENCE: 637
000

SEQ_ID NO: 638 moltype = length =
SEQUENCE: 638
000

SEQ_ID NO: 639 moltype = length =
SEQUENCE: 639
000

SEQ_ID NO: 640 moltype = length =
SEQUENCE: 640
000

SEQ_ID NO: 641 moltype = length =
SEQUENCE: 641
000

SEQ_ID NO: 642 moltype = length =
SEQUENCE: 642
000

SEQ_ID NO: 643 moltype = length =
SEQUENCE: 643
000

SEQ_ID NO: 644 moltype = length =
SEQUENCE: 644
000

SEQ_ID NO: 645 moltype = length =
SEQUENCE: 645
000

SEQ_ID NO: 646 moltype = length =
SEQUENCE: 646
000

SEQ_ID NO: 647 moltype = length =
SEQUENCE: 647
000

SEQ_ID NO: 648 moltype = length =
SEQUENCE: 648
000

SEQ_ID NO: 649 moltype = length =

-continued

SEQUENCE: 649
000

SEQ_ID NO: 650 moltype = length =
SEQUENCE: 650
000

SEQ_ID NO: 651 moltype = length =
SEQUENCE: 651
000

SEQ_ID NO: 652 moltype = length =
SEQUENCE: 652
000

SEQ_ID NO: 653 moltype = length =
SEQUENCE: 653
000

SEQ_ID NO: 654 moltype = length =
SEQUENCE: 654
000

SEQ_ID NO: 655 moltype = length =
SEQUENCE: 655
000

SEQ_ID NO: 656 moltype = length =
SEQUENCE: 656
000

SEQ_ID NO: 657 moltype = length =
SEQUENCE: 657
000

SEQ_ID NO: 658 moltype = length =
SEQUENCE: 658
000

SEQ_ID NO: 659 moltype = length =
SEQUENCE: 659
000

SEQ_ID NO: 660 moltype = length =
SEQUENCE: 660
000

SEQ_ID NO: 661 moltype = length =
SEQUENCE: 661
000

SEQ_ID NO: 662 moltype = length =
SEQUENCE: 662
000

SEQ_ID NO: 663 moltype = length =
SEQUENCE: 663
000

SEQ_ID NO: 664 moltype = length =
SEQUENCE: 664
000

SEQ_ID NO: 665 moltype = length =
SEQUENCE: 665
000

SEQ_ID NO: 666 moltype = length =
SEQUENCE: 666
000

SEQ_ID NO: 667 moltype = length =
SEQUENCE: 667
000

SEQ_ID NO: 668 moltype = length =

-continued

SEQUENCE: 668
000

SEQ_ID NO: 669 moltype = length =
SEQUENCE: 669
000

SEQ_ID NO: 670 moltype = length =
SEQUENCE: 670
000

SEQ_ID NO: 671 moltype = length =
SEQUENCE: 671
000

SEQ_ID NO: 672 moltype = length =
SEQUENCE: 672
000

SEQ_ID NO: 673 moltype = length =
SEQUENCE: 673
000

SEQ_ID NO: 674 moltype = length =
SEQUENCE: 674
000

SEQ_ID NO: 675 moltype = length =
SEQUENCE: 675
000

SEQ_ID NO: 676 moltype = length =
SEQUENCE: 676
000

SEQ_ID NO: 677 moltype = length =
SEQUENCE: 677
000

SEQ_ID NO: 678 moltype = length =
SEQUENCE: 678
000

SEQ_ID NO: 679 moltype = length =
SEQUENCE: 679
000

SEQ_ID NO: 680 moltype = length =
SEQUENCE: 680
000

SEQ_ID NO: 681 moltype = length =
SEQUENCE: 681
000

SEQ_ID NO: 682 moltype = length =
SEQUENCE: 682
000

SEQ_ID NO: 683 moltype = length =
SEQUENCE: 683
000

SEQ_ID NO: 684 moltype = length =
SEQUENCE: 684
000

SEQ_ID NO: 685 moltype = length =
SEQUENCE: 685
000

SEQ_ID NO: 686 moltype = length =
SEQUENCE: 686
000

SEQ_ID NO: 687 moltype = length =

-continued

```

SEQUENCE: 687
000

SEQ ID NO: 688      moltype = length =
SEQUENCE: 688
000

SEQ ID NO: 689      moltype = length =
SEQUENCE: 689
000

SEQ ID NO: 690      moltype = length =
SEQUENCE: 690
000

SEQ ID NO: 691      moltype = length =
SEQUENCE: 691
000

SEQ ID NO: 692      moltype = length =
SEQUENCE: 692
000

SEQ ID NO: 693      moltype = length =
SEQUENCE: 693
000

SEQ ID NO: 694      moltype = length =
SEQUENCE: 694
000

SEQ ID NO: 695      moltype = length =
SEQUENCE: 695
000

SEQ ID NO: 696      moltype = length =
SEQUENCE: 696
000

SEQ ID NO: 697      moltype = length =
SEQUENCE: 697
000

SEQ ID NO: 698      moltype = length =
SEQUENCE: 698
000

SEQ ID NO: 699      moltype = length =
SEQUENCE: 699
000

SEQ ID NO: 700      moltype = AA length = 461
FEATURE
source          Location/Qualifiers
1..461
mol_type = protein
organism = Homo sapiens
SEQUENCE: 700
MQRVNMIMAE SPGLITICLL GYLLSAECTV FLDHENANKI LNRPKRYNSG KLEEFVQGNL 60
ERECMEEKCS FEEAREVFN TERTTEFWKQ YVDGDQCESN PCLNGGSCKD DINSYECWCP 120
FGFEGKNCEL VTDCNIKNGR CEQFCCKNSAD NKVVCSTEG YRLAENQKSC EPAPVFPCGR 180
VSVSQTSLT RAETVFPDVR YVNSTEATI LDNITQSTQS FNDFTRVGG EDAKPGQQFPW 240
QVVLNGKVA FCAGGSIVNEK WIVTAAHCV EHNIEETEHT EQKRNVIRII 300
PHINYNNAAIN KYNHDIALLE LDEPLVLNSY VTPICIADKE YTNIFLKFGS GYVSGWGRVF 360
HKGRSALVHQ YLRVPLVDRA TCLRSTKFTI YNNMFCAGFH EGGRDSCQGD SGGPHVTEVE 420
GTSFLTGIIS WGEECAMKGK YGIYTKVSRV VNWIKEKTKL T 461

SEQ ID NO: 701      moltype = AA length = 415
FEATURE
source          Location/Qualifiers
1..415
mol_type = protein
organism = Homo sapiens
SEQUENCE: 701
YNSGKLEEFV QGNLRECME EKCSFEEARE VFENTERTTE FWKQYVGDQ CESNPCLNGG 60
SCKDDINSYE CWCPFGFEGK NCEDVTCNI KNGRCEQFCK NSADNKVVCS CTEGYRLAEN 120
QKSCEPAVPF PCGRVSVSQT SKLTRAETVF PDVDYVNSTE AETILDNTQ STQSFNDFTF 180
VVGGEDAKPG QFPWQVVLNG KVDAFCGGSI VNEKWIVTAA HCVENTGVKIT VVAGEHNL 240

```

-continued

TEHTEQKRVN	IRIIPHYN	AAINKYNHD	ALLELDEPLV	LNSYVTPI	ADKEYTNIFL	300
KFGSGYVSGW	GRVFHKGRSA	LVLQYLRVPL	VDRATCLRST	KFTIYNNMFC	AGFHBRGRDS	360
CQGDSGGPHV	TEVEGTSFLT			GIISWGEECA	MKGKYGIYTK	415

SEQ ID NO: 702	moltype = AA length = 415					
FEATURE	Location/Qualifiers					
source	1..415					
	mol_type = protein					
	organism = synthetic construct					
SEQUENCE: 702						
YNSGKLEEFV	QGNLRECME	EKCSFEEARE	VFENTERTTE	FWKQYVDGDQ	CESNPCLNG	60
SCKDDINSYE	CWCPFGFEGH	NCELDVTNCI	KNGRCEQFCK	NSADNKVVC	CTEGYRLAEN	120
QKSCEPAVPF	PCGRVSVSQT	SKLTRAETVF	PDVYVNSTE	AETILDNTQ	STQSFNDFT	180
VVAGGEDAKPG	QFPWQVVLNG	KVDAGCGGS	VNEKWIVITA	HCVETGVKIT	VVAGEHNIEE	240
TEHTEQKRVN	IRIIPHYN	AAINKYNHD	ALLELDEPLV	LNSYVTPI	ADKEYTNIFL	300
KFGSGYVSGW	GRVFHKGRSA	LVLQYLRVPL	VDRATCLLST	KFTIYNNMFC	AGFHBRGRDS	360
CQGDSGGPHV	TEVEGTSFLT			GIISWGEECA	MKGKYGIYTK	415
SEQ ID NO: 703	moltype = DNA length = 4104					
FEATURE	Location/Qualifiers					
source	1..4104					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 703						
atggataaga	agtactcaat	cgggtggat	atcgaaacta	attccgtggg	ttggcagtg	60
atcacggat	aatacaaagt	ggctccaaag	aagttaaagg	tcctggggaa	caccgataga	120
cacagcatca	agaaaaaatct	catecgagcc	ctgctgttt	actccggcga	aaccgcagaa	180
gcgcacccgc	tcaaacgtac	cgcgaggcga	cgctacacc	ggccggaaaga	tcgcacatcgc	240
tatctgcaag	agatctttc	ggacaaatg	gcaaaaggctc	acgcacagctt	cttccaccc	300
cttggaaagaat	ctttctgtgt	ggaggaggac	aagaagatc	aacggcatcc	tatctttgg	360
aacatcgctc	acgaaatggc	gttacccacg	ccatctacca	tctcgccaa	tttgcggaaag	420
aatgttgggt	actactgtca	caaggccgc	ctcgatgt	tctacttggc	cctcgcccat	480
atgatcaat	ttccgcgaca	tttctgtatc	gaaggcgatc	tgaaccttga	taactccgc	540
gtggataa	gtttcattca	actgttgcag	acctaaca	actgttgc	agaaaaccca	600
atcaatgtca	ggccgcgtca	tgccaaaggcc	atctgttgc	cccggctgtc	gaagtcgcgg	660
cgcctcgca	acctgtatgc	acagctgcgc	ggagaaaaaa	agaacggact	tttgcgcac	720
ttgatcgctc	tctactgtgg	actactccc	aatttcaatg	ccaattttga	cctggccgag	780
gacgcga	tgcaactctc	aaaggacacc	tacgacgc	acttggacaa	tttgcgcga	840
caaaatggcc	atcaatgc	ggatctgtt	cttgcgcgt	agaaccc	ggacgcac	900
ttgtgttgc	atatacttgc	cgtgaacacc	aaataacc	aagccgcgt	tagcgcctc	960
atgattaa	ggtacgacg	gcatacc	gtctcaac	tgctcaaa	gctcggtg	1020
cagcaactgc	ctgaaaatg	caaggagatc	ttcttcgacc	agtccaagaa	tgggtacgca	1080
gggtacatcg	atggaggcgc	tagccagaa	gagtttata	agttcatca	gcacatctc	1140
gaaaatgg	acggaaatcg	agaatgtcg	gtcaagctg	acagggaga	tctgtc	1200
aaacagagaa	cttttgacaa	cgatccat	ccccccagg	tccatctgg	tgatgtc	1260
gccatcttgc	ggccgcagg	ggactttac	ccatctctca	aggacaac	ggaaaagatc	1320
gagaaaat	tgacgttcc	catecgat	tactggggc	cactggcgc	cgcaatctc	1380
cgttgcgt	gatgtatcg	aaaatcag	gaaaccatc	tttgcggaa	tttgcggaa	1440
gttgtggata	aggaggctc	ggcacaa	tccatcgac	gaatgacaa	tttcgacaa	1500
aatctccaa	acgagaaggt	gttccctaa	cacagcctc	tttacgata	tttactgtc	1560
tacaacaa	tgactaaatg	gaaatcgt	actgaaggaa	tgaggaac	ggcctttct	1620
tccggagaac	agaagaaatgc	aatttgc	ctgtgttca	agaccaac	caaggtgacc	1680
gtcaagc	ttaaaagg	ctacttca	aaatgtc	gttgcact	agtggaaatc	1740
agcgggttgg	aggacagatt	caacgctcg	ctgggac	atcatgtatc	cctgaagatc	1800
atcaaggaca	aggacttcc	tgcaac	gagaacgg	acatcttgc	agatatctc	1860
ctgacacttgc	cccttttgc	ggatcg	atgtatcg	agaggctt	gacctacgt	1920
catctcttc	aggatcaat	cttgc	ccggcgttgc	ccggcgttgc	ttttgggc	1980
cgcctctcc	gcaatgtat	caacgg	cgacataac	agacgcgt	aactatctc	2040
gatttctca	aatcgatgg	cttcgttca	cgtaacttca	tgcaatttgc	ccacgcac	2100
agectgc	tttggggat	catccaaaa	gcacaatgt	ccggcagg	agactcactc	2160
catgaacaca	tccgcgttgc	ggccgttgc	ccggcgttgc	agaaggaaat	tctgcac	2220
gtgaaggatgg	tcgacgtatc	ggtgaagg	atgggacgg	acaaacgg	aatatct	2280
attgaaatgg	cccgagaaaa	ccagactacc	cagaaggccc	agaaaaactc	ccgcgaaagg	2340
atgaaggcga	tccaaatgg	aatcaagg	ctgggcac	agatcttgc	agagcaccc	2400
gtggaaa	ccgcgttgc	gaacgaga	cttgcgttgc	actatggc	aatatggc	2460
gacatgtac	tggacca	gttgcgttgc	aatcttgc	tttgcgttgc	ccggac	2520
atcggttcc	actcttttgc	gaaggatgc	tcgtatc	acaagggtt	gactcg	2580
gacaagaaca	gagggaaatgc	agataatgt	ccatcg	aggtgttac	cctgaa	2640
aattacttgc	ggcgttcc	gtatgttgc	ctgttac	agagaaaat	tgacaatctc	2700
actaaaggcc	agcgcggcgg	actcttgc	ctggatagg	ctggattt	caaacggc	2760
ctgtgtc	ctccggcgttgc	taccaac	gttgcgttgc	tcttgcgtt	ccgcgttgc	2820
actaaat	acgagaac	taatgtc	ccggaaatgt	aggtgttac	cctgaa	2880
aaacttgcgt	ggacttcc	gaaggactt	cagtttaca	aaatgtg	gaga	2940
taccatc	cgcatgac	atacc	gttgcgttgc	gttgcgttgc	gatcaaaa	3000
taccctaa	ttgtatcg	gttgcgttgc	ggagacta	aggttac	cgtgagg	3060
atgatagcc	agtcgac	ggaaatcg	aaagcaat	cttttact	ca	3120

-continued

aacatcatga	acttttcaa	gactgaaatt	acgctggcca	atggagaaaat	caggaagagg	3180
ccactgtac	aaactaacgg	agaaaaacgggc	gaaatcgtgc	gggacaaggg	cagggacttc	3240
gcaactgttc	gcaaaagtgc	ctctatcgcc	caagtcaata	ttgtagaaga	aaccgaagtg	3300
caaacccggc	gattttcaa	ggaatcgatc	ctccccaaaga	gaaatagcga	caagcttatt	3360
gcacgcaga	aagactggg	cccgaagaag	tacggaggat	tcgattcgcc	gactgtcgca	3420
tactccgtcc	tcgtgggtggc	caagggtggag	aaggggaaaag	gcaaaaaaagct	caaattccgtc	3480
aaagagactgc	ttgggattac	catcatggaa	cgatccctcg	tcgagaagaa	cccgatttgat	3540
ttctctcgagg	cgaaagggtta	caaggagggt	aaaggaggatc	tgatcatcaa	actccccaaag	3600
tactcactgt	tcgaactgg	aaatggtcgg	aagcgcatgc	tggttccggc	cgggagaactc	3660
caaaaaggaa	atagatggc	cttgcctact	aagtacgtca	acttccctta	tcttgcgtcg	3720
cactacgaa	aactcaaagg	gtcacccgg	gataacgaaac	agaaggcagct	tttctgtggag	3780
cagcacaagg	attatctgg	tgaaatcatc	gaacaatct	ccgagtttc	aaagcgcgtg	3840
atcctcgccg	acgccaacct	cgacaaggatc	ctgtcgcc	acaataagca	tagagataag	3900
ccgatcagag	acacggccg	gaacattatc	cacttgtca	ccctgactaa	cctgggagcc	3960
ccagccgcct	tcaagtaactt	cgatactact	atcgatcgca	aaagatacac	gtccaccaag	4020
gaagttctgg	acggcaccct	gatccaccaa	agcatctact	gactctacga	aactaggatc	4080
gatctgtcg	agctgggtgg	cgat				4104

SEQ ID NO: 704 moltype = DNA length = 4140
 FEATURE Location/Qualifiers
 source 1..4140
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 704

atggacaaga	agtacagcat	cggactggac	atcgaaacaa	acagcgtcg	atgggcagtc	60
atcacagacg	aatacaagg	cccgagacag	aagttaacgg	tcctggaaa	cacagacaga	120
cacagcatca	agaagaac	gatcgagac	ctgctgtcg	acagcgaga	aacagcagaa	180
gcaacaacag	tgaaagaaac	agcaacaa	agatacaca	aaagaaaaaa	cagaatctgc	240
tacactgcagg	aaatcttcag	caacggaaatg	gcaaaagggtc	acgacacgtt	cttccacaga	300
ctggaagaaa	gttctctgg	cgaagaagac	aagaagcagc	aaagacaccc	gatcttcgga	360
aacatctgcg	acgaaggatc	atccacgac	aagtacatc	ccatctacca	ccttgagaaag	420
aaagctgtcg	acagacacaga	aaaggcagac	ctgagactga	tctacctggc	actggcacac	480
atgatcaatg	ttagggac	cttctgtatc	gaaggagacc	tgaacccgg	caacagcgac	540
gtcgacaagg	tggttcatcca	gctgggtccag	acatacaacc	agctgtcg	agaaaaacccg	600
atcaacacgg	ggggaggatc	cgcaacagg	atctcgatc	caagactgag	caagagcaga	660
agactggaaa	acctgtatgc	acagctggc	ggagaaaaaa	agaacggact	gttcggaaac	720
ctgatcgac	tgagcttggg	actgacacgg	aacttcaaga	gcaacttgc	cctggcagaa	780
gacgcaaaacg	tgcagctgag	caaggacaca	tacgacgacg	acctggacaa	cctgctggca	840
caqatcgagg	accatgtacg	agacactgtt	ctggcagca	agaacctgg	cgacgcatac	900
ctgctgtatgc	acatcttcag	agtcaacaca	gaaatcaca	aggcaccgct	gagcgcaga	960
atgatcaaga	gatacgtacg	acacccac	gacctgacac	tgctgaaggc	actgttca	1020
cagcagctgc	cgaaaaagta	caaggaaatc	ttcttcgacc	agagcaagaa	cggtacac	1080
ggatacatcg	acggaggagc	aaggcaggaa	gaattctaca	agttcatcaa	gccgatctcg	1140
gaaaagatgg	acggcaacaga	agaactgtg	gtcaagctg	acagagaaga	cttgcgtgaga	1200
aagcagagaa	cattcgacaa	cggaacatc	ccgcacccaga	tccacctgg	agaactgcac	1260
gcaatctcta	gaagacagga	agacttctac	ccgttctcg	aggacaacag	agaaaaagatc	1320
gaaaagatcc	tgacatctcg	aatccctgt	taatcgccg	cgctggcaag	aggaaacacg	1380
agatctgcgat	gatgtacatc	aaagagcggaa	gaaacaaatc	caccgtggaa	cttcgaagaa	1440
gtcgatcgaca	aggggacaa	cgccacagac	tccatcgaa	gaatgacaaa	cttcgacaa	1500
aacctcgccg	acgaaaaggat	cctgcgca	cacagctgc	tgtacgata	cttcacagtc	1560
tacaacacg	tgacatctcg	caatgtatgc	acagaaggaa	tgagaaaaggc	ggcattctcg	1620
agcggagaaac	agaagaac	aatctgtatc	ctgtcgatc	tgatcgatc	aaaggttcaca	1680
gtcaagcgc	tgaaaggaa	ctacttca	aaatgtcaat	gatccgacac	ctgtcgaaatc	1740
agcggatcg	aaagacatgt	caacgcac	ctggacat	accacgact	gctgaagatc	1800
atcaaggaca	aggacttctc	ggaaacacgg	gaaaacacgg	acatctcg	agacatctgc	1860
ctgacactga	cactgttgc	agacacatc	atgatcgatc	aaagactgaa	gacatcgac	1920
cacccgttgc	accacaaagg	catgaacatc	tgcaagatc	aaagatcaca	aggatgggaa	1980
agactgtatc	gaaagatgtat	caacggatc	agagacaacg	agagcggaaa	gacaatctcg	2040
gacttctcg	agagcgacgg	atccgacaa	tgcaatcgat	ccacgcac	ccacgcac	2100
agccgtatcg	tcaagatcg	catcgacaa	gcacaggatc	ggcggagg	agacagctg	2160
caccaacaca	tcgcaacac	ggcggacaa	ccggcaatc	agaaggaaat	cctgcacaca	2220
gtcaaggatcg	tcgacatcg	ggtcaaggatc	atgggaacac	acaacccgg	aaacatctgc	2280
atcgatctcg	caagagaaaa	ccagacaaca	cagaaggac	agaagaac	cagaga	2340
atgaagatcg	tccgaaagg	aatcaaggaa	ctggaaac	atggatctcg	ggaaacacccg	2400
gtcgaaaaca	cacatgtcg	gaacggaaa	ctgtacatcg	actatctcg	gaacggaa	2460
gacatgtatcg	tccgacccgg	actggacatc	aaacactgt	tgcaatcgat	ggatcgac	2520
atcgatccgc	agagcttctc	gaaggac	agcatcgatc	acaaggatct	gacaaga	2580
gacaagaaca	gaggaaagag	cgacaacatc	ccggac	aaatgtatc	ggatcgac	2640
aactactcg	gacatgtatcg	gaacggacaa	ctgtatcac	aaagaaaaat	cgacaacatc	2700
acaaggac	acggaggagg	actggac	ctggac	caggattat	caagac	2760
ctgtcgaaa	caagacat	cacaacgg	gtcgac	tccatcgat	ccgtcgac	2820
acaaggatcg	acgaaaacg	caatgtatcg	agagaatcg	atgtatcgatc	actgtatcg	2880
aagctgtatcg	tgcaatcgatc	aaaggactt	cgttctac	aggatcgatc	actgtatcg	2940
taccaccacg	cacacgac	atccatcgatc	gcagatcgatc	ggatcgatc	atgtatcgatc	3000
tacccgaaacg	tggaaaacg	attcgatcgatc	ggatcgatc	aggatcgatc	ctgtcgacaa	3060
atgatcgacaa	agagcgacaa	ggaaatcgatc	aaggac	aaatgtatcgatc	tttctac	3120

-continued

aacatcatga	acttcttcaa	gacagaaaatc	acaactggcaa	acgggaaaaat	cagaagaaaga	3180
ccgcgtatcg	aaacaacacgg	agaaaaacgg	gaaatcgtct	gggacaaggg	aagagacttc	3240
gcaacagtca	gaaaggatctt	gagcatccg	caggtaaca	tgcgtcaagaa	gacagaagtc	3300
cagacaggag	gattcagcaa	ggaaacgttc	ctgcccga	gaaacacgca	caagctgtac	3360
gcaagaaaga	aggactggg	cccgaagaag	tacggaggat	tcgacagccc	gacagtgcga	3420
taacagctcc	ttggtegtcgc	aaagggtcgaa	aaaggaaaaa	gcaagaagct	gaagagcgtc	3480
aaggaaactgc	ttggaaatcac	aatcatggaa	agaagcagct	tgcggaaa	cccgatcgac	3540
ttctctggaa	caaaggatata	caagggatc	agaaggacc	tgatcatcaa	gctggcgaa	3600
taacagctgt	tcgaaactgg	aaacggaaaga	aagagaatgc	ttggcaagcgc	aggagaactg	3660
cagaaggggaa	acgaaactggc	actggcgag	aagtacgtca	acttctgtt	cctggcaagc	3720
cactacgaaa	agctggagg	aagccccgg	gacaacgaa	agaagcagct	tttcgtcgaa	3780
cagcacaacg	actacctgg	cgaaatcatc	gaacatcg	cgcaatttcg	caagagagtc	3840
atcctggcg	acgcaaacct	ggacaaggtc	ctgagcgat	acaacaagca	cagagacaag	3900
ccgatcatcg	acacggcga	aaacatcatc	cacctgttc	cactgacaaa	cctggggagca	3960
ccggcagcat	tcaagtactt	cgacacaa	atcgacagaa	agagatacac	aagcacaaaag	4020
gaagtcttgg	acgcaacact	gatccccag	agcatcac	gactgtacg	aacaagaatc	4080
gacctgagcc	agctggagg	agacggagga	ggaagccccg	agaagaagag	aaaggcttag	4140

SEQ ID NO: 705 moltype = RNA length = 4501
 FEATURE
 source Location/Qualifiers
 1..4501
 mol_type = other RNA
 organism = synthetic construct

SEQUENCE: 705

gggtcccgca	gtcgccgtcc	agccgctctg	cttggtcgtg	tgttgtctg	tgccaggcctt	60
atccggatcc	gccccatgg	acaagaatg	cagcatcgaa	ctggacatcg	gaacaaacag	120
cgtccggatgg	gcagtcatc	cagacgata	caagggtcccg	agaacaaatg	tcaaggctct	180
ggggaaacaca	gacagacaca	gcatacggaa	aaacctgtatc	ggagactgtc	ttttcgacag	240
cgagaaaaaca	cgagaaggaa	caagactgaa	gagaacacgca	agaagaagat	acacaagaag	300
aaagaacaga	atctgtacc	tgccggaaa	cttcagcaac	gaaatggca	aggtcgacga	360
cagcttcttc	cacagactgg	aagaagactt	cctggcgaa	gaaagaca	agcagggaaag	420
acacccgatc	ttcgaaaca	tcgtcgacg	agtcgtatc	cacggaaaatg	acccgacaaat	480
ctaccacccgt	agaaaaaagc	ttggcgacag	cacagacaa	gcagacccgt	gactgtatca	540
cctggcactg	gcacacatg	tcaagttcg	aggacacttc	ctgatcgaa	gagacctgaa	600
cccgacaaac	acgcgacgtcg	acaaggctgt	catccagctg	gtcccgacat	acaacccagct	660
gttccggaaa	aacccgatc	acgcaaggcg	agtcgacgca	aaggcaatcc	tgagcgcaag	720
actggagac	acgcaaaagac	ttggaaaatc	gatccgacag	ctggccggag	aaaagaagaa	780
cggaactgttc	ggaaacactg	tcgcactgag	cctgggactg	acacccgact	tcaagagacaa	840
cttcgcactg	gcagaacagc	caaaggccca	gtctggacg	gacacatacg	acgcacgac	900
ggacaacactg	ttggccacaga	tcggagacca	gtacgcacag	ctgttccctgg	cagcaaaagaa	960
cctggagccg	gcaatcttc	tgaggcgatc	cctggagatc	aacacagaaa	tcacaaaggc	1020
accgctgacg	gcaagcatg	tcaagagata	cgacgaacac	caccaggacc	tgacactgtc	1080
gaaggcaactg	gtcagacacg	agtcgacccg	aaagtacaa	gaaatcttct	tcgaccagg	1140
caagaacacg	tacgcggat	acatcgacgg	aggagcaacg	caggaaatg	tctacaatgt	1200
catcaacgc	atccctggaa	agatggacgg	aaacaaatg	ctgctgtca	agctgtacag	1260
agaagacactg	ctgagaaatc	agagaacatt	cgacaaacgg	agcatccgc	accagatcca	1320
cctggagaa	ctgcacgcac	tcctggagaa	acaggaaatg	ttcttaccctg	tcctgtacgg	1380
caacagacaaa	aaatctgtac	attccggat	ccgtactact	tcggaccgct	1440	
ggcaagaggg	aacacgacat	tcgtcgat	gacaagaa	agcgaagaaa	caatcacacc	1500
gttggacttc	gaagaatgtcg	tcgacaaatgg	agcaacgc	cagatcttc	tcgaaagaaat	1560
gacaaatcttc	gacaaggaa	tcggccaa	aaagggtctt	ccggacacaa	gcctgtctg	1620
cgaatacttc	acatgttac	acgaaatgc	aaagggtca	tacgttac	aggaaatgt	1680
aaagccggca	ttccctgacg	gaaacacaaa	gaaaggcaatc	gtcgaccctg	ttttcaagac	1740
aaacagaaag	gtcacatgt	agcagctgaa	ggaagactac	ttcaaga	tcgaatgtt	1800
cgacacgttc	gaaatcgc	gagtgtcgaa	catatccaa	gcaaccccttgg	gaacatacc	1860
cgacccgt	aaatgtatc	aggacaaatg	tttccctgg	aacaaatgg	acgaagacat	1920
cctggagac	atcgtcttc	cactggact	gttccggac	agagaaatgt	tcgaaagaaag	1980
actgaagaca	taacgcac	tgttcgacg	caaggatcg	aagcagctg	agagaagaaag	2040
atacacacg	ttggggaa	tgacacgaaa	gctgtatc	ggaatcagag	acaacggac	2100
cgggaaagaca	atccctggact	tcctgtacgg	cgacggat	gcaaaatgg	acttcatgc	2160
gtgtatccac	gacgacac	tcgtatc	ggaaatgtatc	cagaaggccac	aggtcgacgg	2220
acaggggac	acgcgtcgac	aacacatcg	aaacctggc	ggaageccgg	caatcaagaa	2280
gggaatcttg	cacacatcg	aggtcgatcg	cgaactgtt	agggtatcg	gaagacacaa	2340
ggccggaaac	atcgatcg	aaatgtcgaa	agaaaaacccg	acaacacaga	agggacacaa	2400
gaacacgaca	gaaagaaatg	agagaatcg	agaaggatcg	aaaggactgg	gaagccgat	2460
cctggagaa	caccccgatcg	aaacacacaa	gtctgtatc	ggaaatgtatc	ttctacaatgt	2520
cctggatc	ggaagagaca	tgtacgtcg	ccaggatcg	gacatcaaca	gactgtacg	2580
ctacgcac	gaccacatcg	tccctcgatcg	tttccctgg	gacgacac	tcgacaaacaa	2640
gggtccatcg	agaacatcg	agaacatcg	aaaggtatcg	ttctacaatgt	gaagacacaa	2700
cgtcaagaatcg	atgaaatcg	actggatcg	gctgtatcg	gcaaaatgtcg	tcacacacg	2760
aaagttcgac	aacatcgatc	aggacacatcg	aggaggactg	acgcgtatcg	acaaggccgg	2820
attcatcg	agacacgtcg	tcgaaatcg	acagatcaca	aagcacgtcg	cacagatctt	2880
ggacacgaca	atgaaatcg	agtgacacg	aaacacatcg	ttttcaatgt	ttttcaatgt	2940
catcacactg	aaagacatcg	ttggtcacg	cttcagaaatg	gacttccatgt	tctacaatgt	3000
cagagaaaatc	aacaactacc	accacgcaca	cgacgcatac	ctgaaatcg	tcgtcgaaac	3060
agcactgtatc	aagaatgtatc	cgaaatgtatc	aaaggtatcg	ttttcaatgt	ttttcaatgt	3120

-continued

ctacgacgtc	agaaagatga	tcgcaaagag	cgaacaggaa	atcgaaaagg	caacagcaaa	3180
gtacttcttc	tcacgaca	tcatgaaactt	cttcaagacaa	gaaatcacac	tggcaaacgg	3240
agaatcaga	aagagccgc	tgatcgaaac	aaacggagaa	acaggagaaa	tcgtctggga	3300
caaggaaaga	gacttcgaa	cagttagaaa	ggtcctgac	atgcccgg	tcaacatcg	3360
caagaagaca	gaagttccaga	caggaggatt	cagcaaggaa	agcatctgc	cgaagagaaa	3420
cagcgacaag	ctgatcgcaa	gaaagaaggaa	ctgggacccc	aagaagtacg	gaggattcga	3480
cagccccgaca	gtcgcatata	ggtgtctgg	cgtcgcaaa	gtcgaaaagg	gaaagagcaa	3540
gaagctqaag	agcgtcaagg	aactgctggg	aatcacaatc	atggaaaaggaa	gcagcttcga	3600
aaagaacccg	atcgacttcc	tggaaacaaa	gggatataca	gaagtcaaga	aggacctgtat	3660
catcaagctt	ccgaaqtaca	gcctgttca	actggaaaac	ggaagaaaaa	gaatgtcg	3720
aagcgcggaa	gaactcgaga	agggacggaa	actggactc	ccgagcaat	acgtcaactt	3780
cctgtacctt	gcaagccact	acggaaaatgt	gaagggaa	ccggaa	acgaacagaa	3840
gcagctgttc	gtcgacagc	acaaggacta	cctggacgaa	atcatcgaa	agatcagcga	3900
attcagcag	agagtcatc	tggcagacgc	aaacctggac	aagggtcttga	gcgcatacaa	3960
caagcaca	gacaagccga	tcagaaaca	ggcagaaaac	atcatccacc	tgttcacact	4020
gacaaacccgt	ggagcaccgg	cagcatca	gtacttcgac	acaacaatcg	acagaaagag	4080
atacacaacg	acaaaaggaa	tcctggacgc	aaacactgtatc	caccagagca	tcacaggact	4140
gtacgaaaaa	agaatcgacc	tgagccatgt	ggggagggac	ggggagggaa	gccccaa	4200
gaagagaaaa	gtcttagctag	ccatcacatt	taaaagcatc	tcagcttacc	atgagaataa	4260
gagaaagaaaa	atagaatca	atgttattt	catctttttt	tcttttttgt	ttgtgttaag	4320
ccaacacccct	gtctaaaaaa	cataaatttc	ttaatcatt	ttgccttctt	tctctgtgt	4380
taattaataa	aaaaatggaa	agaacctca	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	4440
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	4500
a						4501

SEQ ID NO: 706 moltype = DNA length = 2802
 FEATURE Location/Qualifiers
 source 1..2802
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 706
 accacttca caatctgcta gcaagggtt a tgacgcgcgt gaacatgtatc atggcagaat 60
 caccaggcct catcaccatc tgccttttag gatatctact cagtgtgaa tgcgtttt 120
 ttcttgcata tggaaaacgaa aacaaaatttca tgaatcgcc aaagaggat aattcaggta 180
 aatttggaa gtttgcataa gggaaacccgt agagagaatg tatggaaaggaa aagtgtatgt 240
 ttgaaagacg acgagaatgtt tttggaaaaca ctgaaagacaa aactgtattt tggtggact 300
 atgttgcata gatgtatgtt ggttgcatac catgtttaa tggcggactg tgcaaggatg 360
 acattaaatc ctatgtatgtt tggtgtccct ttggatttga agggaaagaa tgcgttttt 420
 atgttgcata ttaacatcaag aatggccatgtt ggcggactgtt ttgtttttt aatgtgtata 480
 acaagggtgtt ttgccttgcata gttgtttttt atcgttgcatac aggttttttgcgtt 540
 aaccaggactt gocatttca tttttttttt acaaaacttca aatgttgcatac 600
 gtgtgtggac tttttttttt gatgtggact atgttttttca tactgtttttt gaaaccat 660
 tggataacatc cactcaacatc ttaatgtatc cacttcgggtt gttgtgtgg 720
 aagatgcacca acaggccatca tttttttttt gatgtgtttt gatgtgtttt gttgtgtcat 780
 tctgtggagg ctctatcgat tttttttttt aatgtttttt gatgtgtttt gttgtgtttt 840
 ctgtgtgtttt aatttgcata gttgtgtttt gatgtgtttt gatgtgtttt gttgtgtttt 900
 agcaaaacggc aatgtgtttt gatgtgtttt gatgtgtttt gatgtgtttt gatgtgtttt 960
 agtacaacatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1020
 ttacacat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1080
 gctatgtatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1140
 acctttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1200
 ataacaacatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1260
 gtgtgtgtttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1320
 ggggtgtttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1380
 tcaactgtatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1440
 tcattttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1500
 agatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1560
 attttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1620
 aatttacatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1680
 ctgttccat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1740
 tagcagttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1800
 agttctgtatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1860
 tgaagaaaga acacaggactt tttttttttt tttttttttt tttttttttt tttttttttt 1920
 cttcttccat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1980
 ctctttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2040
 catctttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2100
 cgttagtggggat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2160
 ggaaaagttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2220
 taatataaaataa tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2280
 acacatataa tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2340
 aggcattttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2400
 cccagacataa tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2460
 ccgttccat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2520
 cagcagtgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2580
 atgttcttctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2640
 agtcatttca tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2700

-continued

tttgattata gttaatccctt ctatctgaa tcttcttagag agttgtgac caactgacgt 2760
atgtttccct ttgtgaatta ataaactgggt gttctgggtc at 2802

SEQ_ID NO: 707 moltype = length =
SEQUENCE: 707
000

SEQ_ID NO: 708 moltype = length =
SEQUENCE: 708
000

SEQ_ID NO: 709 moltype = length =
SEQUENCE: 709
000

SEQ_ID NO: 710 moltype = length =
SEQUENCE: 710
000

SEQ_ID NO: 711 moltype = length =
SEQUENCE: 711
000

SEQ_ID NO: 712 moltype = length =
SEQUENCE: 712
000

SEQ_ID NO: 713 moltype = length =
SEQUENCE: 713
000

SEQ_ID NO: 714 moltype = length =
SEQUENCE: 714
000

SEQ_ID NO: 715 moltype = length =
SEQUENCE: 715
000

SEQ_ID NO: 716 moltype = length =
SEQUENCE: 716
000

SEQ_ID NO: 717 moltype = length =
SEQUENCE: 717
000

SEQ_ID NO: 718 moltype = length =
SEQUENCE: 718
000

SEQ_ID NO: 719 moltype = length =
SEQUENCE: 719
000

SEQ_ID NO: 720 moltype = length =
SEQUENCE: 720
000

SEQ_ID NO: 721 moltype = length =
SEQUENCE: 721
000

SEQ_ID NO: 722 moltype = length =
SEQUENCE: 722
000

SEQ_ID NO: 723 moltype = length =
SEQUENCE: 723
000

SEQ_ID NO: 724 moltype = length =
SEQUENCE: 724
000

SEQ_ID NO: 725 moltype = length =

-continued

SEQUENCE: 725
000

SEQ_ID NO: 726 moltype = length =
SEQUENCE: 726
000

SEQ_ID NO: 727 moltype = length =
SEQUENCE: 727
000

SEQ_ID NO: 728 moltype = length =
SEQUENCE: 728
000

SEQ_ID NO: 729 moltype = length =
SEQUENCE: 729
000

SEQ_ID NO: 730 moltype = length =
SEQUENCE: 730
000

SEQ_ID NO: 731 moltype = length =
SEQUENCE: 731
000

SEQ_ID NO: 732 moltype = length =
SEQUENCE: 732
000

SEQ_ID NO: 733 moltype = length =
SEQUENCE: 733
000

SEQ_ID NO: 734 moltype = length =
SEQUENCE: 734
000

SEQ_ID NO: 735 moltype = length =
SEQUENCE: 735
000

SEQ_ID NO: 736 moltype = length =
SEQUENCE: 736
000

SEQ_ID NO: 737 moltype = length =
SEQUENCE: 737
000

SEQ_ID NO: 738 moltype = length =
SEQUENCE: 738
000

SEQ_ID NO: 739 moltype = length =
SEQUENCE: 739
000

SEQ_ID NO: 740 moltype = length =
SEQUENCE: 740
000

SEQ_ID NO: 741 moltype = length =
SEQUENCE: 741
000

SEQ_ID NO: 742 moltype = length =
SEQUENCE: 742
000

SEQ_ID NO: 743 moltype = length =
SEQUENCE: 743
000

SEQ_ID NO: 744 moltype = length =

-continued

SEQUENCE: 744
000

SEQ_ID NO: 745 moltype = length =
SEQUENCE: 745
000

SEQ_ID NO: 746 moltype = length =
SEQUENCE: 746
000

SEQ_ID NO: 747 moltype = length =
SEQUENCE: 747
000

SEQ_ID NO: 748 moltype = length =
SEQUENCE: 748
000

SEQ_ID NO: 749 moltype = length =
SEQUENCE: 749
000

SEQ_ID NO: 750 moltype = length =
SEQUENCE: 750
000

SEQ_ID NO: 751 moltype = length =
SEQUENCE: 751
000

SEQ_ID NO: 752 moltype = length =
SEQUENCE: 752
000

SEQ_ID NO: 753 moltype = length =
SEQUENCE: 753
000

SEQ_ID NO: 754 moltype = length =
SEQUENCE: 754
000

SEQ_ID NO: 755 moltype = length =
SEQUENCE: 755
000

SEQ_ID NO: 756 moltype = length =
SEQUENCE: 756
000

SEQ_ID NO: 757 moltype = length =
SEQUENCE: 757
000

SEQ_ID NO: 758 moltype = length =
SEQUENCE: 758
000

SEQ_ID NO: 759 moltype = length =
SEQUENCE: 759
000

SEQ_ID NO: 760 moltype = length =
SEQUENCE: 760
000

SEQ_ID NO: 761 moltype = length =
SEQUENCE: 761
000

SEQ_ID NO: 762 moltype = length =
SEQUENCE: 762
000

SEQ_ID NO: 763 moltype = length =

-continued

SEQUENCE: 763
000

SEQ_ID NO: 764 moltype = length =
SEQUENCE: 764
000

SEQ_ID NO: 765 moltype = length =
SEQUENCE: 765
000

SEQ_ID NO: 766 moltype = length =
SEQUENCE: 766
000

SEQ_ID NO: 767 moltype = length =
SEQUENCE: 767
000

SEQ_ID NO: 768 moltype = length =
SEQUENCE: 768
000

SEQ_ID NO: 769 moltype = length =
SEQUENCE: 769
000

SEQ_ID NO: 770 moltype = length =
SEQUENCE: 770
000

SEQ_ID NO: 771 moltype = length =
SEQUENCE: 771
000

SEQ_ID NO: 772 moltype = length =
SEQUENCE: 772
000

SEQ_ID NO: 773 moltype = length =
SEQUENCE: 773
000

SEQ_ID NO: 774 moltype = length =
SEQUENCE: 774
000

SEQ_ID NO: 775 moltype = length =
SEQUENCE: 775
000

SEQ_ID NO: 776 moltype = length =
SEQUENCE: 776
000

SEQ_ID NO: 777 moltype = length =
SEQUENCE: 777
000

SEQ_ID NO: 778 moltype = length =
SEQUENCE: 778
000

SEQ_ID NO: 779 moltype = length =
SEQUENCE: 779
000

SEQ_ID NO: 780 moltype = length =
SEQUENCE: 780
000

SEQ_ID NO: 781 moltype = length =
SEQUENCE: 781
000

SEQ_ID NO: 782 moltype = length =

-continued

SEQUENCE: 782
000

SEQ_ID NO: 783 moltype = length =
SEQUENCE: 783
000

SEQ_ID NO: 784 moltype = length =
SEQUENCE: 784
000

SEQ_ID NO: 785 moltype = length =
SEQUENCE: 785
000

SEQ_ID NO: 786 moltype = length =
SEQUENCE: 786
000

SEQ_ID NO: 787 moltype = length =
SEQUENCE: 787
000

SEQ_ID NO: 788 moltype = length =
SEQUENCE: 788
000

SEQ_ID NO: 789 moltype = length =
SEQUENCE: 789
000

SEQ_ID NO: 790 moltype = length =
SEQUENCE: 790
000

SEQ_ID NO: 791 moltype = length =
SEQUENCE: 791
000

SEQ_ID NO: 792 moltype = length =
SEQUENCE: 792
000

SEQ_ID NO: 793 moltype = length =
SEQUENCE: 793
000

SEQ_ID NO: 794 moltype = length =
SEQUENCE: 794
000

SEQ_ID NO: 795 moltype = length =
SEQUENCE: 795
000

SEQ_ID NO: 796 moltype = length =
SEQUENCE: 796
000

SEQ_ID NO: 797 moltype = length =
SEQUENCE: 797
000

SEQ_ID NO: 798 moltype = length =
SEQUENCE: 798
000

SEQ_ID NO: 799 moltype = length =
SEQUENCE: 799
000

SEQ_ID NO: 800 moltype = length =
SEQUENCE: 800
000

SEQ_ID NO: 801 moltype = length =

-continued

SEQUENCE: 801
000

SEQ_ID NO: 802 moltype = length =
SEQUENCE: 802
000

SEQ_ID NO: 803 moltype = length =
SEQUENCE: 803
000

SEQ_ID NO: 804 moltype = length =
SEQUENCE: 804
000

SEQ_ID NO: 805 moltype = length =
SEQUENCE: 805
000

SEQ_ID NO: 806 moltype = length =
SEQUENCE: 806
000

SEQ_ID NO: 807 moltype = length =
SEQUENCE: 807
000

SEQ_ID NO: 808 moltype = length =
SEQUENCE: 808
000

SEQ_ID NO: 809 moltype = length =
SEQUENCE: 809
000

SEQ_ID NO: 810 moltype = length =
SEQUENCE: 810
000

SEQ_ID NO: 811 moltype = length =
SEQUENCE: 811
000

SEQ_ID NO: 812 moltype = length =
SEQUENCE: 812
000

SEQ_ID NO: 813 moltype = length =
SEQUENCE: 813
000

SEQ_ID NO: 814 moltype = length =
SEQUENCE: 814
000

SEQ_ID NO: 815 moltype = length =
SEQUENCE: 815
000

SEQ_ID NO: 816 moltype = length =
SEQUENCE: 816
000

SEQ_ID NO: 817 moltype = length =
SEQUENCE: 817
000

SEQ_ID NO: 818 moltype = length =
SEQUENCE: 818
000

SEQ_ID NO: 819 moltype = length =
SEQUENCE: 819
000

SEQ_ID NO: 820 moltype = length =

-continued

SEQUENCE: 820
000

SEQ_ID NO: 821 moltype = length =
SEQUENCE: 821
000

SEQ_ID NO: 822 moltype = length =
SEQUENCE: 822
000

SEQ_ID NO: 823 moltype = length =
SEQUENCE: 823
000

SEQ_ID NO: 824 moltype = length =
SEQUENCE: 824
000

SEQ_ID NO: 825 moltype = length =
SEQUENCE: 825
000

SEQ_ID NO: 826 moltype = length =
SEQUENCE: 826
000

SEQ_ID NO: 827 moltype = length =
SEQUENCE: 827
000

SEQ_ID NO: 828 moltype = length =
SEQUENCE: 828
000

SEQ_ID NO: 829 moltype = length =
SEQUENCE: 829
000

SEQ_ID NO: 830 moltype = length =
SEQUENCE: 830
000

SEQ_ID NO: 831 moltype = length =
SEQUENCE: 831
000

SEQ_ID NO: 832 moltype = length =
SEQUENCE: 832
000

SEQ_ID NO: 833 moltype = length =
SEQUENCE: 833
000

SEQ_ID NO: 834 moltype = length =
SEQUENCE: 834
000

SEQ_ID NO: 835 moltype = length =
SEQUENCE: 835
000

SEQ_ID NO: 836 moltype = length =
SEQUENCE: 836
000

SEQ_ID NO: 837 moltype = length =
SEQUENCE: 837
000

SEQ_ID NO: 838 moltype = length =
SEQUENCE: 838
000

SEQ_ID NO: 839 moltype = length =

-continued

SEQUENCE: 839
000

SEQ_ID NO: 840 moltype = length =
SEQUENCE: 840
000

SEQ_ID NO: 841 moltype = length =
SEQUENCE: 841
000

SEQ_ID NO: 842 moltype = length =
SEQUENCE: 842
000

SEQ_ID NO: 843 moltype = length =
SEQUENCE: 843
000

SEQ_ID NO: 844 moltype = length =
SEQUENCE: 844
000

SEQ_ID NO: 845 moltype = length =
SEQUENCE: 845
000

SEQ_ID NO: 846 moltype = length =
SEQUENCE: 846
000

SEQ_ID NO: 847 moltype = length =
SEQUENCE: 847
000

SEQ_ID NO: 848 moltype = length =
SEQUENCE: 848
000

SEQ_ID NO: 849 moltype = length =
SEQUENCE: 849
000

SEQ_ID NO: 850 moltype = length =
SEQUENCE: 850
000

SEQ_ID NO: 851 moltype = length =
SEQUENCE: 851
000

SEQ_ID NO: 852 moltype = length =
SEQUENCE: 852
000

SEQ_ID NO: 853 moltype = length =
SEQUENCE: 853
000

SEQ_ID NO: 854 moltype = length =
SEQUENCE: 854
000

SEQ_ID NO: 855 moltype = length =
SEQUENCE: 855
000

SEQ_ID NO: 856 moltype = length =
SEQUENCE: 856
000

SEQ_ID NO: 857 moltype = length =
SEQUENCE: 857
000

SEQ_ID NO: 858 moltype = length =

-continued

SEQUENCE: 858
000

SEQ_ID NO: 859 moltype = length =
SEQUENCE: 859
000

SEQ_ID NO: 860 moltype = length =
SEQUENCE: 860
000

SEQ_ID NO: 861 moltype = length =
SEQUENCE: 861
000

SEQ_ID NO: 862 moltype = length =
SEQUENCE: 862
000

SEQ_ID NO: 863 moltype = length =
SEQUENCE: 863
000

SEQ_ID NO: 864 moltype = length =
SEQUENCE: 864
000

SEQ_ID NO: 865 moltype = length =
SEQUENCE: 865
000

SEQ_ID NO: 866 moltype = length =
SEQUENCE: 866
000

SEQ_ID NO: 867 moltype = length =
SEQUENCE: 867
000

SEQ_ID NO: 868 moltype = length =
SEQUENCE: 868
000

SEQ_ID NO: 869 moltype = length =
SEQUENCE: 869
000

SEQ_ID NO: 870 moltype = length =
SEQUENCE: 870
000

SEQ_ID NO: 871 moltype = length =
SEQUENCE: 871
000

SEQ_ID NO: 872 moltype = length =
SEQUENCE: 872
000

SEQ_ID NO: 873 moltype = length =
SEQUENCE: 873
000

SEQ_ID NO: 874 moltype = length =
SEQUENCE: 874
000

SEQ_ID NO: 875 moltype = length =
SEQUENCE: 875
000

SEQ_ID NO: 876 moltype = length =
SEQUENCE: 876
000

SEQ_ID NO: 877 moltype = length =

-continued

SEQUENCE: 877
000

SEQ_ID NO: 878 moltype = length =
SEQUENCE: 878
000

SEQ_ID NO: 879 moltype = length =
SEQUENCE: 879
000

SEQ_ID NO: 880 moltype = length =
SEQUENCE: 880
000

SEQ_ID NO: 881 moltype = length =
SEQUENCE: 881
000

SEQ_ID NO: 882 moltype = length =
SEQUENCE: 882
000

SEQ_ID NO: 883 moltype = length =
SEQUENCE: 883
000

SEQ_ID NO: 884 moltype = length =
SEQUENCE: 884
000

SEQ_ID NO: 885 moltype = length =
SEQUENCE: 885
000

SEQ_ID NO: 886 moltype = length =
SEQUENCE: 886
000

SEQ_ID NO: 887 moltype = length =
SEQUENCE: 887
000

SEQ_ID NO: 888 moltype = length =
SEQUENCE: 888
000

SEQ_ID NO: 889 moltype = length =
SEQUENCE: 889
000

SEQ_ID NO: 890 moltype = length =
SEQUENCE: 890
000

SEQ_ID NO: 891 moltype = length =
SEQUENCE: 891
000

SEQ_ID NO: 892 moltype = length =
SEQUENCE: 892
000

SEQ_ID NO: 893 moltype = length =
SEQUENCE: 893
000

SEQ_ID NO: 894 moltype = length =
SEQUENCE: 894
000

SEQ_ID NO: 895 moltype = length =
SEQUENCE: 895
000

SEQ_ID NO: 896 moltype = length =

-continued

SEQUENCE: 896
000

SEQ_ID NO: 897 moltype = length =
SEQUENCE: 897
000

SEQ_ID NO: 898 moltype = length =
SEQUENCE: 898
000

SEQ_ID NO: 899 moltype = length =
SEQUENCE: 899
000

SEQ_ID NO: 900 moltype = length =
SEQUENCE: 900
000

SEQ_ID NO: 901 moltype = length =
SEQUENCE: 901
000

SEQ_ID NO: 902 moltype = length =
SEQUENCE: 902
000

SEQ_ID NO: 903 moltype = length =
SEQUENCE: 903
000

SEQ_ID NO: 904 moltype = length =
SEQUENCE: 904
000

SEQ_ID NO: 905 moltype = length =
SEQUENCE: 905
000

SEQ_ID NO: 906 moltype = length =
SEQUENCE: 906
000

SEQ_ID NO: 907 moltype = length =
SEQUENCE: 907
000

SEQ_ID NO: 908 moltype = length =
SEQUENCE: 908
000

SEQ_ID NO: 909 moltype = length =
SEQUENCE: 909
000

SEQ_ID NO: 910 moltype = length =
SEQUENCE: 910
000

SEQ_ID NO: 911 moltype = length =
SEQUENCE: 911
000

SEQ_ID NO: 912 moltype = length =
SEQUENCE: 912
000

SEQ_ID NO: 913 moltype = length =
SEQUENCE: 913
000

SEQ_ID NO: 914 moltype = length =
SEQUENCE: 914
000

SEQ_ID NO: 915 moltype = length =

-continued

SEQUENCE: 915
000

SEQ_ID NO: 916 moltype = length =
SEQUENCE: 916
000

SEQ_ID NO: 917 moltype = length =
SEQUENCE: 917
000

SEQ_ID NO: 918 moltype = length =
SEQUENCE: 918
000

SEQ_ID NO: 919 moltype = length =
SEQUENCE: 919
000

SEQ_ID NO: 920 moltype = length =
SEQUENCE: 920
000

SEQ_ID NO: 921 moltype = length =
SEQUENCE: 921
000

SEQ_ID NO: 922 moltype = length =
SEQUENCE: 922
000

SEQ_ID NO: 923 moltype = length =
SEQUENCE: 923
000

SEQ_ID NO: 924 moltype = length =
SEQUENCE: 924
000

SEQ_ID NO: 925 moltype = length =
SEQUENCE: 925
000

SEQ_ID NO: 926 moltype = length =
SEQUENCE: 926
000

SEQ_ID NO: 927 moltype = length =
SEQUENCE: 927
000

SEQ_ID NO: 928 moltype = length =
SEQUENCE: 928
000

SEQ_ID NO: 929 moltype = length =
SEQUENCE: 929
000

SEQ_ID NO: 930 moltype = length =
SEQUENCE: 930
000

SEQ_ID NO: 931 moltype = length =
SEQUENCE: 931
000

SEQ_ID NO: 932 moltype = length =
SEQUENCE: 932
000

SEQ_ID NO: 933 moltype = length =
SEQUENCE: 933
000

SEQ_ID NO: 934 moltype = length =

-continued

SEQUENCE: 934
000

SEQ_ID NO: 935 moltype = length =
SEQUENCE: 935
000

SEQ_ID NO: 936 moltype = length =
SEQUENCE: 936
000

SEQ_ID NO: 937 moltype = length =
SEQUENCE: 937
000

SEQ_ID NO: 938 moltype = length =
SEQUENCE: 938
000

SEQ_ID NO: 939 moltype = length =
SEQUENCE: 939
000

SEQ_ID NO: 940 moltype = length =
SEQUENCE: 940
000

SEQ_ID NO: 941 moltype = length =
SEQUENCE: 941
000

SEQ_ID NO: 942 moltype = length =
SEQUENCE: 942
000

SEQ_ID NO: 943 moltype = length =
SEQUENCE: 943
000

SEQ_ID NO: 944 moltype = length =
SEQUENCE: 944
000

SEQ_ID NO: 945 moltype = length =
SEQUENCE: 945
000

SEQ_ID NO: 946 moltype = length =
SEQUENCE: 946
000

SEQ_ID NO: 947 moltype = length =
SEQUENCE: 947
000

SEQ_ID NO: 948 moltype = length =
SEQUENCE: 948
000

SEQ_ID NO: 949 moltype = length =
SEQUENCE: 949
000

SEQ_ID NO: 950 moltype = length =
SEQUENCE: 950
000

SEQ_ID NO: 951 moltype = length =
SEQUENCE: 951
000

SEQ_ID NO: 952 moltype = length =
SEQUENCE: 952
000

SEQ_ID NO: 953 moltype = length =

-continued

SEQUENCE: 953
000

SEQ_ID NO: 954 moltype = length =
SEQUENCE: 954
000

SEQ_ID NO: 955 moltype = length =
SEQUENCE: 955
000

SEQ_ID NO: 956 moltype = length =
SEQUENCE: 956
000

SEQ_ID NO: 957 moltype = length =
SEQUENCE: 957
000

SEQ_ID NO: 958 moltype = length =
SEQUENCE: 958
000

SEQ_ID NO: 959 moltype = length =
SEQUENCE: 959
000

SEQ_ID NO: 960 moltype = length =
SEQUENCE: 960
000

SEQ_ID NO: 961 moltype = length =
SEQUENCE: 961
000

SEQ_ID NO: 962 moltype = length =
SEQUENCE: 962
000

SEQ_ID NO: 963 moltype = length =
SEQUENCE: 963
000

SEQ_ID NO: 964 moltype = length =
SEQUENCE: 964
000

SEQ_ID NO: 965 moltype = length =
SEQUENCE: 965
000

SEQ_ID NO: 966 moltype = length =
SEQUENCE: 966
000

SEQ_ID NO: 967 moltype = length =
SEQUENCE: 967
000

SEQ_ID NO: 968 moltype = length =
SEQUENCE: 968
000

SEQ_ID NO: 969 moltype = length =
SEQUENCE: 969
000

SEQ_ID NO: 970 moltype = length =
SEQUENCE: 970
000

SEQ_ID NO: 971 moltype = length =
SEQUENCE: 971
000

SEQ_ID NO: 972 moltype = length =

-continued

SEQUENCE: 972
000

SEQ_ID NO: 973 moltype = length =
SEQUENCE: 973
000

SEQ_ID NO: 974 moltype = length =
SEQUENCE: 974
000

SEQ_ID NO: 975 moltype = length =
SEQUENCE: 975
000

SEQ_ID NO: 976 moltype = length =
SEQUENCE: 976
000

SEQ_ID NO: 977 moltype = length =
SEQUENCE: 977
000

SEQ_ID NO: 978 moltype = length =
SEQUENCE: 978
000

SEQ_ID NO: 979 moltype = length =
SEQUENCE: 979
000

SEQ_ID NO: 980 moltype = length =
SEQUENCE: 980
000

SEQ_ID NO: 981 moltype = length =
SEQUENCE: 981
000

SEQ_ID NO: 982 moltype = length =
SEQUENCE: 982
000

SEQ_ID NO: 983 moltype = length =
SEQUENCE: 983
000

SEQ_ID NO: 984 moltype = length =
SEQUENCE: 984
000

SEQ_ID NO: 985 moltype = length =
SEQUENCE: 985
000

SEQ_ID NO: 986 moltype = length =
SEQUENCE: 986
000

SEQ_ID NO: 987 moltype = length =
SEQUENCE: 987
000

SEQ_ID NO: 988 moltype = length =
SEQUENCE: 988
000

SEQ_ID NO: 989 moltype = length =
SEQUENCE: 989
000

SEQ_ID NO: 990 moltype = length =
SEQUENCE: 990
000

SEQ_ID NO: 991 moltype = length =

-continued

SEQUENCE: 991
000

SEQ_ID NO: 992 moltype = length =
SEQUENCE: 992
000

SEQ_ID NO: 993 moltype = length =
SEQUENCE: 993
000

SEQ_ID NO: 994 moltype = length =
SEQUENCE: 994
000

SEQ_ID NO: 995 moltype = length =
SEQUENCE: 995
000

SEQ_ID NO: 996 moltype = length =
SEQUENCE: 996
000

SEQ_ID NO: 997 moltype = length =
SEQUENCE: 997
000

SEQ_ID NO: 998 moltype = length =
SEQUENCE: 998
000

SEQ_ID NO: 999 moltype = length =
SEQUENCE: 999
000

SEQ_ID NO: 1000 moltype = length =
SEQUENCE: 1000
000

SEQ_ID NO: 1001 moltype = length =
SEQUENCE: 1001
000

SEQ_ID NO: 1002 moltype = length =
SEQUENCE: 1002
000

SEQ_ID NO: 1003 moltype = length =
SEQUENCE: 1003
000

SEQ_ID NO: 1004 moltype = length =
SEQUENCE: 1004
000

SEQ_ID NO: 1005 moltype = length =
SEQUENCE: 1005
000

SEQ_ID NO: 1006 moltype = length =
SEQUENCE: 1006
000

SEQ_ID NO: 1007 moltype = length =
SEQUENCE: 1007
000

SEQ_ID NO: 1008 moltype = length =
SEQUENCE: 1008
000

SEQ_ID NO: 1009 moltype = length =
SEQUENCE: 1009
000

SEQ_ID NO: 1010 moltype = length =

-continued

SEQUENCE: 1010
000

SEQ_ID NO: 1011 moltype = length =
SEQUENCE: 1011
000

SEQ_ID NO: 1012 moltype = length =
SEQUENCE: 1012
000

SEQ_ID NO: 1013 moltype = length =
SEQUENCE: 1013
000

SEQ_ID NO: 1014 moltype = length =
SEQUENCE: 1014
000

SEQ_ID NO: 1015 moltype = length =
SEQUENCE: 1015
000

SEQ_ID NO: 1016 moltype = length =
SEQUENCE: 1016
000

SEQ_ID NO: 1017 moltype = length =
SEQUENCE: 1017
000

SEQ_ID NO: 1018 moltype = length =
SEQUENCE: 1018
000

SEQ_ID NO: 1019 moltype = length =
SEQUENCE: 1019
000

SEQ_ID NO: 1020 moltype = length =
SEQUENCE: 1020
000

SEQ_ID NO: 1021 moltype = length =
SEQUENCE: 1021
000

SEQ_ID NO: 1022 moltype = length =
SEQUENCE: 1022
000

SEQ_ID NO: 1023 moltype = length =
SEQUENCE: 1023
000

SEQ_ID NO: 1024 moltype = length =
SEQUENCE: 1024
000

SEQ_ID NO: 1025 moltype = length =
SEQUENCE: 1025
000

SEQ_ID NO: 1026 moltype = length =
SEQUENCE: 1026
000

SEQ_ID NO: 1027 moltype = length =
SEQUENCE: 1027
000

SEQ_ID NO: 1028 moltype = length =
SEQUENCE: 1028
000

SEQ_ID NO: 1029 moltype = length =

-continued

SEQUENCE: 1029
000

SEQ_ID NO: 1030 moltype = length =
SEQUENCE: 1030
000

SEQ_ID NO: 1031 moltype = length =
SEQUENCE: 1031
000

SEQ_ID NO: 1032 moltype = length =
SEQUENCE: 1032
000

SEQ_ID NO: 1033 moltype = length =
SEQUENCE: 1033
000

SEQ_ID NO: 1034 moltype = length =
SEQUENCE: 1034
000

SEQ_ID NO: 1035 moltype = length =
SEQUENCE: 1035
000

SEQ_ID NO: 1036 moltype = length =
SEQUENCE: 1036
000

SEQ_ID NO: 1037 moltype = length =
SEQUENCE: 1037
000

SEQ_ID NO: 1038 moltype = length =
SEQUENCE: 1038
000

SEQ_ID NO: 1039 moltype = length =
SEQUENCE: 1039
000

SEQ_ID NO: 1040 moltype = length =
SEQUENCE: 1040
000

SEQ_ID NO: 1041 moltype = length =
SEQUENCE: 1041
000

SEQ_ID NO: 1042 moltype = length =
SEQUENCE: 1042
000

SEQ_ID NO: 1043 moltype = length =
SEQUENCE: 1043
000

SEQ_ID NO: 1044 moltype = length =
SEQUENCE: 1044
000

SEQ_ID NO: 1045 moltype = length =
SEQUENCE: 1045
000

SEQ_ID NO: 1046 moltype = length =
SEQUENCE: 1046
000

SEQ_ID NO: 1047 moltype = length =
SEQUENCE: 1047
000

SEQ_ID NO: 1048 moltype = length =

-continued

SEQUENCE: 1048
000

SEQ_ID NO: 1049 moltype = length =
SEQUENCE: 1049
000

SEQ_ID NO: 1050 moltype = length =
SEQUENCE: 1050
000

SEQ_ID NO: 1051 moltype = length =
SEQUENCE: 1051
000

SEQ_ID NO: 1052 moltype = length =
SEQUENCE: 1052
000

SEQ_ID NO: 1053 moltype = length =
SEQUENCE: 1053
000

SEQ_ID NO: 1054 moltype = length =
SEQUENCE: 1054
000

SEQ_ID NO: 1055 moltype = length =
SEQUENCE: 1055
000

SEQ_ID NO: 1056 moltype = length =
SEQUENCE: 1056
000

SEQ_ID NO: 1057 moltype = length =
SEQUENCE: 1057
000

SEQ_ID NO: 1058 moltype = length =
SEQUENCE: 1058
000

SEQ_ID NO: 1059 moltype = length =
SEQUENCE: 1059
000

SEQ_ID NO: 1060 moltype = length =
SEQUENCE: 1060
000

SEQ_ID NO: 1061 moltype = length =
SEQUENCE: 1061
000

SEQ_ID NO: 1062 moltype = length =
SEQUENCE: 1062
000

SEQ_ID NO: 1063 moltype = length =
SEQUENCE: 1063
000

SEQ_ID NO: 1064 moltype = length =
SEQUENCE: 1064
000

SEQ_ID NO: 1065 moltype = length =
SEQUENCE: 1065
000

SEQ_ID NO: 1066 moltype = length =
SEQUENCE: 1066
000

SEQ_ID NO: 1067 moltype = length =

-continued

SEQUENCE: 1067
000

SEQ_ID NO: 1068 moltype = length =
SEQUENCE: 1068
000

SEQ_ID NO: 1069 moltype = length =
SEQUENCE: 1069
000

SEQ_ID NO: 1070 moltype = length =
SEQUENCE: 1070
000

SEQ_ID NO: 1071 moltype = length =
SEQUENCE: 1071
000

SEQ_ID NO: 1072 moltype = length =
SEQUENCE: 1072
000

SEQ_ID NO: 1073 moltype = length =
SEQUENCE: 1073
000

SEQ_ID NO: 1074 moltype = length =
SEQUENCE: 1074
000

SEQ_ID NO: 1075 moltype = length =
SEQUENCE: 1075
000

SEQ_ID NO: 1076 moltype = length =
SEQUENCE: 1076
000

SEQ_ID NO: 1077 moltype = length =
SEQUENCE: 1077
000

SEQ_ID NO: 1078 moltype = length =
SEQUENCE: 1078
000

SEQ_ID NO: 1079 moltype = length =
SEQUENCE: 1079
000

SEQ_ID NO: 1080 moltype = length =
SEQUENCE: 1080
000

SEQ_ID NO: 1081 moltype = length =
SEQUENCE: 1081
000

SEQ_ID NO: 1082 moltype = length =
SEQUENCE: 1082
000

SEQ_ID NO: 1083 moltype = length =
SEQUENCE: 1083
000

SEQ_ID NO: 1084 moltype = length =
SEQUENCE: 1084
000

SEQ_ID NO: 1085 moltype = length =
SEQUENCE: 1085
000

SEQ_ID NO: 1086 moltype = length =

-continued

SEQUENCE: 1086
000

SEQ_ID NO: 1087 moltype = length =
SEQUENCE: 1087
000

SEQ_ID NO: 1088 moltype = length =
SEQUENCE: 1088
000

SEQ_ID NO: 1089 moltype = length =
SEQUENCE: 1089
000

SEQ_ID NO: 1090 moltype = length =
SEQUENCE: 1090
000

SEQ_ID NO: 1091 moltype = length =
SEQUENCE: 1091
000

SEQ_ID NO: 1092 moltype = length =
SEQUENCE: 1092
000

SEQ_ID NO: 1093 moltype = length =
SEQUENCE: 1093
000

SEQ_ID NO: 1094 moltype = length =
SEQUENCE: 1094
000

SEQ_ID NO: 1095 moltype = length =
SEQUENCE: 1095
000

SEQ_ID NO: 1096 moltype = length =
SEQUENCE: 1096
000

SEQ_ID NO: 1097 moltype = length =
SEQUENCE: 1097
000

SEQ_ID NO: 1098 moltype = length =
SEQUENCE: 1098
000

SEQ_ID NO: 1099 moltype = length =
SEQUENCE: 1099
000

SEQ_ID NO: 1100 moltype = length =
SEQUENCE: 1100
000

SEQ_ID NO: 1101 moltype = length =
SEQUENCE: 1101
000

SEQ_ID NO: 1102 moltype = length =
SEQUENCE: 1102
000

SEQ_ID NO: 1103 moltype = length =
SEQUENCE: 1103
000

SEQ_ID NO: 1104 moltype = length =
SEQUENCE: 1104
000

SEQ_ID NO: 1105 moltype = length =

-continued

SEQUENCE: 1105
000

SEQ_ID NO: 1106 moltype = length =
SEQUENCE: 1106
000

SEQ_ID NO: 1107 moltype = length =
SEQUENCE: 1107
000

SEQ_ID NO: 1108 moltype = length =
SEQUENCE: 1108
000

SEQ_ID NO: 1109 moltype = length =
SEQUENCE: 1109
000

SEQ_ID NO: 1110 moltype = length =
SEQUENCE: 1110
000

SEQ_ID NO: 1111 moltype = length =
SEQUENCE: 1111
000

SEQ_ID NO: 1112 moltype = length =
SEQUENCE: 1112
000

SEQ_ID NO: 1113 moltype = length =
SEQUENCE: 1113
000

SEQ_ID NO: 1114 moltype = length =
SEQUENCE: 1114
000

SEQ_ID NO: 1115 moltype = length =
SEQUENCE: 1115
000

SEQ_ID NO: 1116 moltype = length =
SEQUENCE: 1116
000

SEQ_ID NO: 1117 moltype = length =
SEQUENCE: 1117
000

SEQ_ID NO: 1118 moltype = length =
SEQUENCE: 1118
000

SEQ_ID NO: 1119 moltype = length =
SEQUENCE: 1119
000

SEQ_ID NO: 1120 moltype = length =
SEQUENCE: 1120
000

SEQ_ID NO: 1121 moltype = length =
SEQUENCE: 1121
000

SEQ_ID NO: 1122 moltype = length =
SEQUENCE: 1122
000

SEQ_ID NO: 1123 moltype = length =
SEQUENCE: 1123
000

SEQ_ID NO: 1124 moltype = length =

-continued

```

SEQUENCE: 1124
000

SEQ_ID NO: 1125      moltype = length =
SEQUENCE: 1125
000

SEQ_ID NO: 1126      moltype = length =
SEQUENCE: 1126
000

SEQ_ID NO: 1127      moltype = length =
SEQUENCE: 1127
000

SEQ_ID NO: 1128      moltype = length =
SEQUENCE: 1128
000

SEQ_ID NO: 1129      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 1129
gagtccgagc agaagaagaa                                         20

SEQ_ID NO: 1130      moltype = RNA length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 1130
gacccccctcc accccgcctc                                         20

```

1.-143. (canceled)

144. A method of expressing Factor IX in a liver cell or population of liver cells, comprising administering:

- (i) a nucleic acid construct comprising a Factor IX protein coding sequence;
 - (ii) an RNA-guided DNA binding agent or a nucleic acid encoding the RNA-guided DNA binding agent; and
 - (iii) a guide RNA (gRNA) comprising a sequence targeting intron 1 of an albumin locus,
- thereby expressing Factor IX in the liver cell or population of liver cells.

145. The method of claim 144, wherein the RNA-guided DNA binding agent is Cas9.

146. The method of claim 145, wherein the method comprises administering an mRNA encoding the Cas9.

147. The method of claim 146, wherein the gRNA and the mRNA encoding the Cas9 are administered in a lipid nanoparticle.

148. The method of claim 144, wherein the nucleic acid construct is administered in a viral vector.

149. The method of claim 148, wherein the viral vector is an adeno-associated viral (AAV) vector.

150. The method of claim 144, wherein the nucleic acid construct is a bidirectional nucleic acid construct comprising: (a) a first segment comprising a first coding sequence for Factor IX and (b) a second segment comprising a reverse complement of a second coding sequence for Factor IX.

151. The method of claim 150, wherein the bidirectional nucleic acid construct does not comprise a homology arm, does not comprise a promoter that drives expression of the first coding sequence, and does not comprise a promoter that drives expression of the second coding sequence.

152. The method of claim 144, wherein the nucleic acid construct is administered in an AAV vector, the RNA-guided DNA binding agent or the nucleic acid encoding the RNA-guided DNA binding agent and the gRNA are administered in a lipid nanoparticle, and the AAV vector and the lipid nanoparticle are administered simultaneously or sequentially, in any order.

153. The method of claim 144, wherein the liver cell or population of liver cells is a human liver cell or population of human liver cells.

154. A method of treating a Factor IX deficiency, comprising administering to an individual with the Factor IX deficiency:

- (i) a nucleic acid construct comprising a Factor IX protein coding sequence;
 - (ii) an RNA-guided DNA binding agent or a nucleic acid encoding the RNA-guided DNA binding agent; and
 - (iii) a guide RNA (gRNA) comprising a sequence targeting intron 1 of an albumin locus,
- thereby expressing Factor IX in the individual.

155. The method of claim 154, wherein the RNA-guided DNA binding agent is Cas9.

156. The method of claim 155, wherein the method comprises administering an mRNA encoding the Cas9.

157. The method of claim 156, wherein the gRNA and the mRNA encoding the Cas9 are administered in a lipid nanoparticle.

158. The method of claim 154, wherein the nucleic acid construct is administered in a viral vector.

159. The method of claim 158, wherein the viral vector is an adeno-associated viral (AAV) vector.

160. The method of claim 154, wherein the nucleic acid construct is a bidirectional nucleic acid construct compris-

ing: (a) a first segment comprising a first coding sequence for Factor IX and (b) a second segment comprising a reverse complement of a second coding sequence for Factor IX.

161. The method of claim **160**, wherein the bidirectional nucleic acid construct does not comprise a homology arm, does not comprise a promoter that drives expression of the first coding sequence, and does not comprise a promoter that drives expression of the second coding sequence.

162. The method of claim **154**, wherein the nucleic acid construct is administered in an AAV vector, the RNA-guided DNA binding agent or the nucleic acid encoding the RNA-guided DNA binding agent and the gRNA are administered in a lipid nanoparticle, and the AAV vector and the lipid nanoparticle are administered simultaneously or sequentially, in any order.

163. The method of claim **154**, wherein the individual is a human.

* * * * *