



US 20210008161A1

(19) **United States**(12) **Patent Application Publication****D'Souza et al.**(10) **Pub. No.: US 2021/0008161 A1**(43) **Pub. Date: Jan. 14, 2021**(54) **METHODS AND COMPOSITIONS FOR
IMPROVED HOMOLOGY DIRECTED
REPAIR***C12N 5/0783* (2006.01)*A61K 35/17* (2006.01)*C12N 5/0789* (2006.01)*A61P 37/06* (2006.01)*A61P 7/06* (2006.01)(71) Applicant: **CRISPR THERAPEUTICS AG**, Zug
(CH)(52) **U.S. Cl.**(72) Inventors: **Sanjay D'Souza**, Cambridge, MA (US);
Jason West, Lexington, MA (US)CPC *A61K 38/1709* (2013.01); *A61K 48/00*
(2013.01); *C12N 7/00* (2013.01); *C12N 15/11*
(2013.01); *C12N 9/22* (2013.01); *C12N*
5/0636 (2013.01); *A61K 35/17* (2013.01);
C12N 5/0647 (2013.01); *A61P 37/06*
(2018.01); *A61P 7/06* (2018.01); *C12N*
2310/20 (2017.05); *C12N 2800/80* (2013.01);
C12N 2750/14143 (2013.01); *C12N 2501/145*
(2013.01); *C12N 2501/26* (2013.01); *C12N*
2501/125 (2013.01); *C12N 2501/2303*
(2013.01); *C12N 2500/02* (2013.01); *C12N*
15/86 (2013.01)(21) Appl. No.: **16/904,493**(22) Filed: **Jun. 17, 2020****Related U.S. Application Data**(60) Provisional application No. 62/862,539, filed on Jun.
17, 2019.**Publication Classification**(51) **Int. Cl.***A61K 38/17* (2006.01)*C12N 15/86* (2006.01)*C12N 7/00* (2006.01)*C12N 15/11* (2006.01)*C12N 9/22* (2006.01)

(57)

ABSTRACT

The disclosure features methods of increasing repair of a DNA double stranded break (DSB) in a target gene by the homology-directed repair (HDR) pathway. The disclosure also features compositions for use in the methods.

Specification includes a Sequence Listing.

FIG. 1A

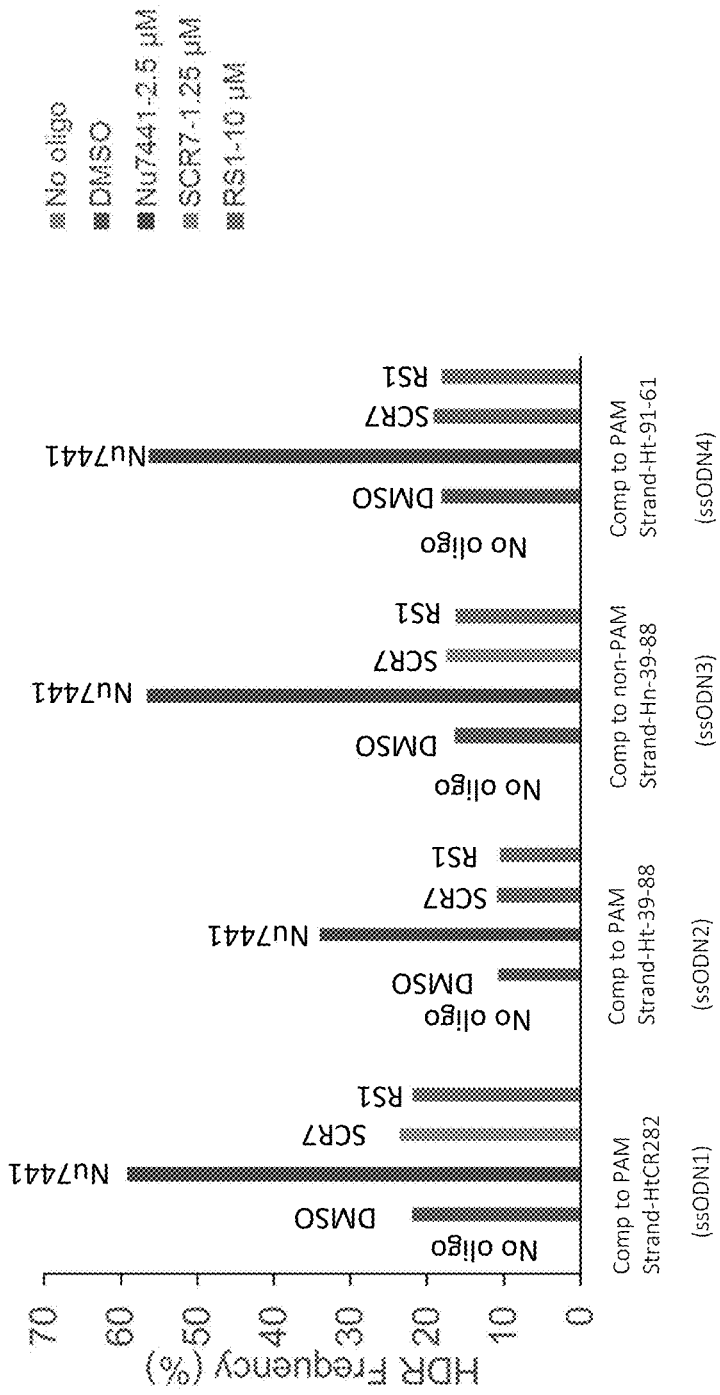
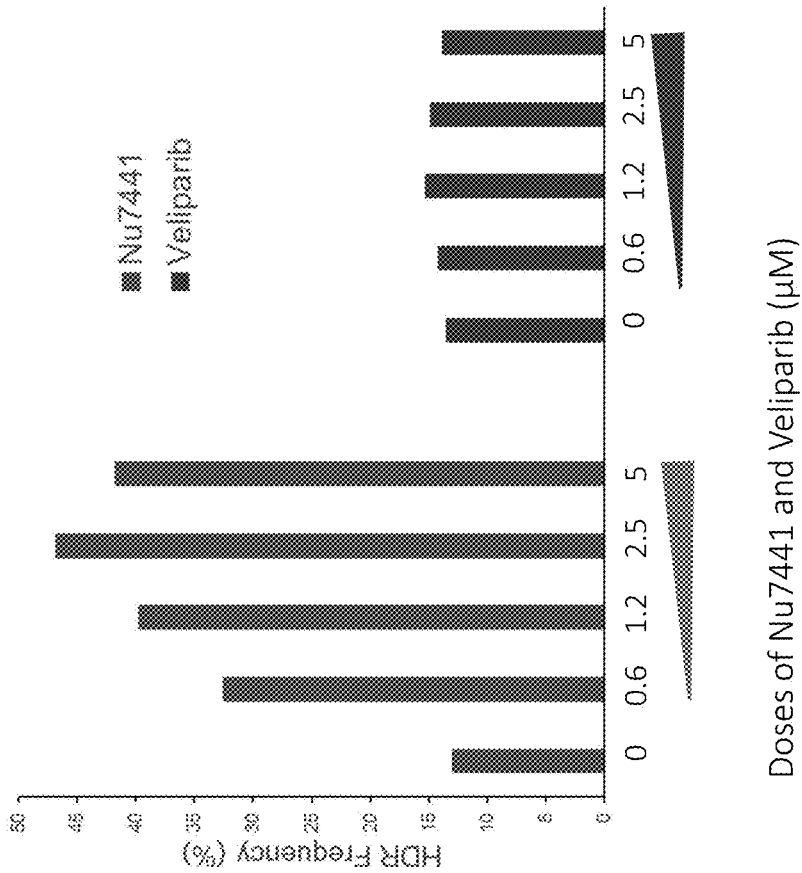


FIG. 1B



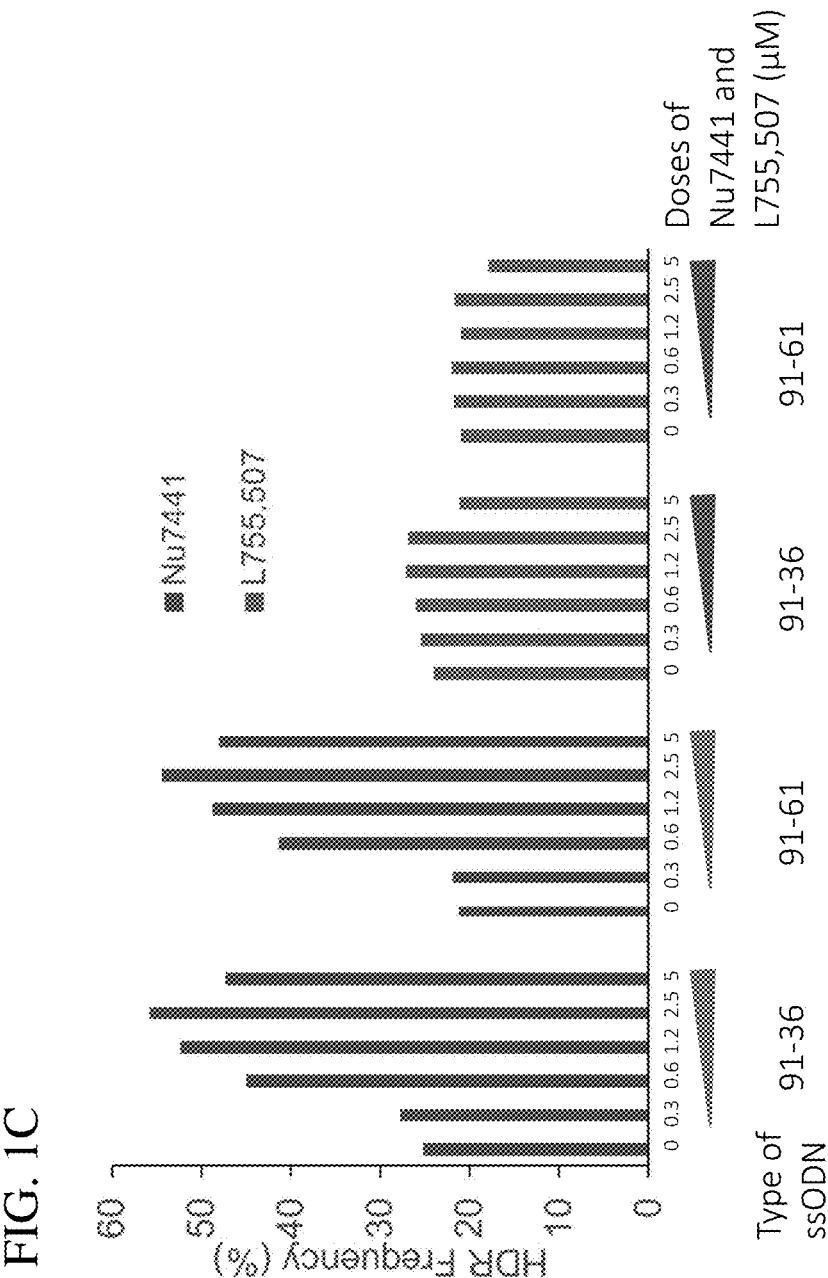


FIG. 1D

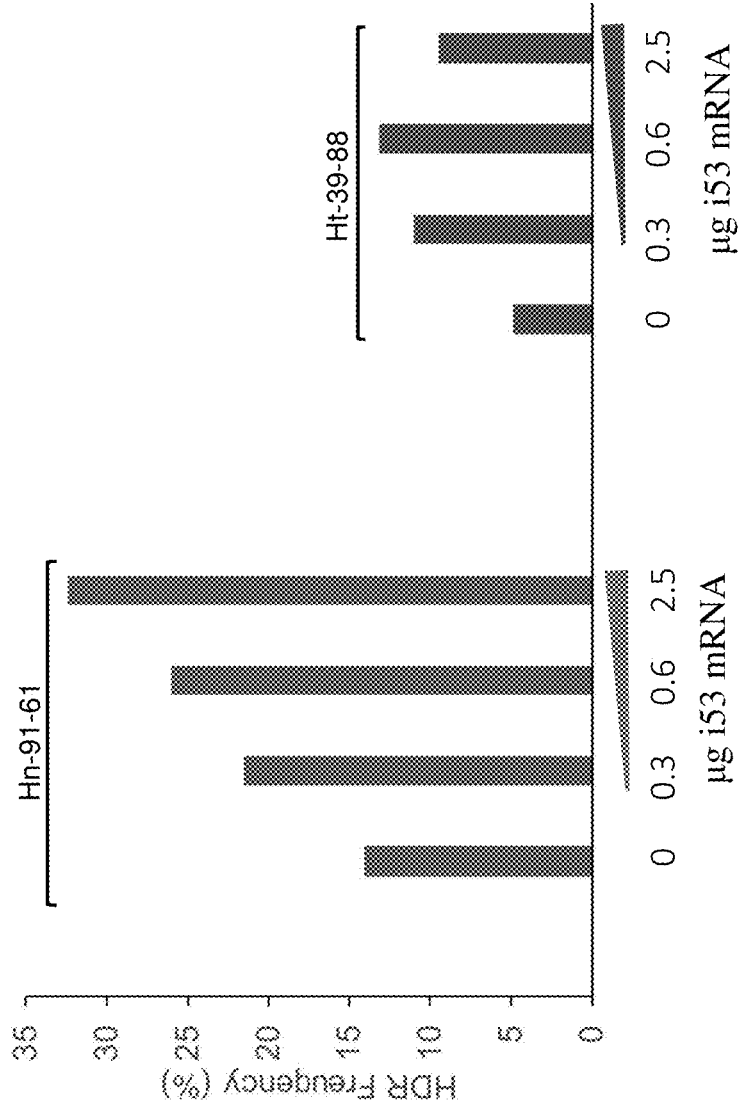
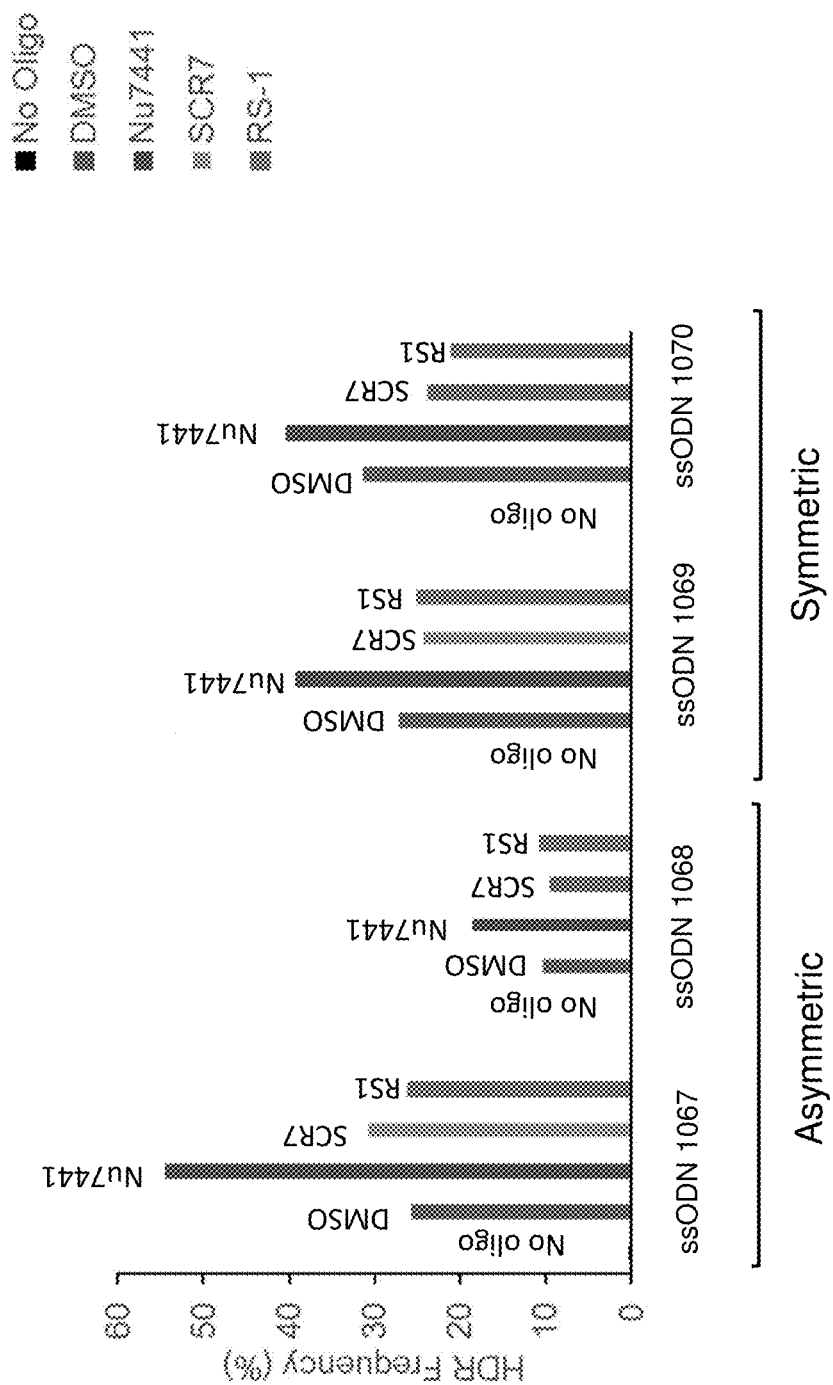
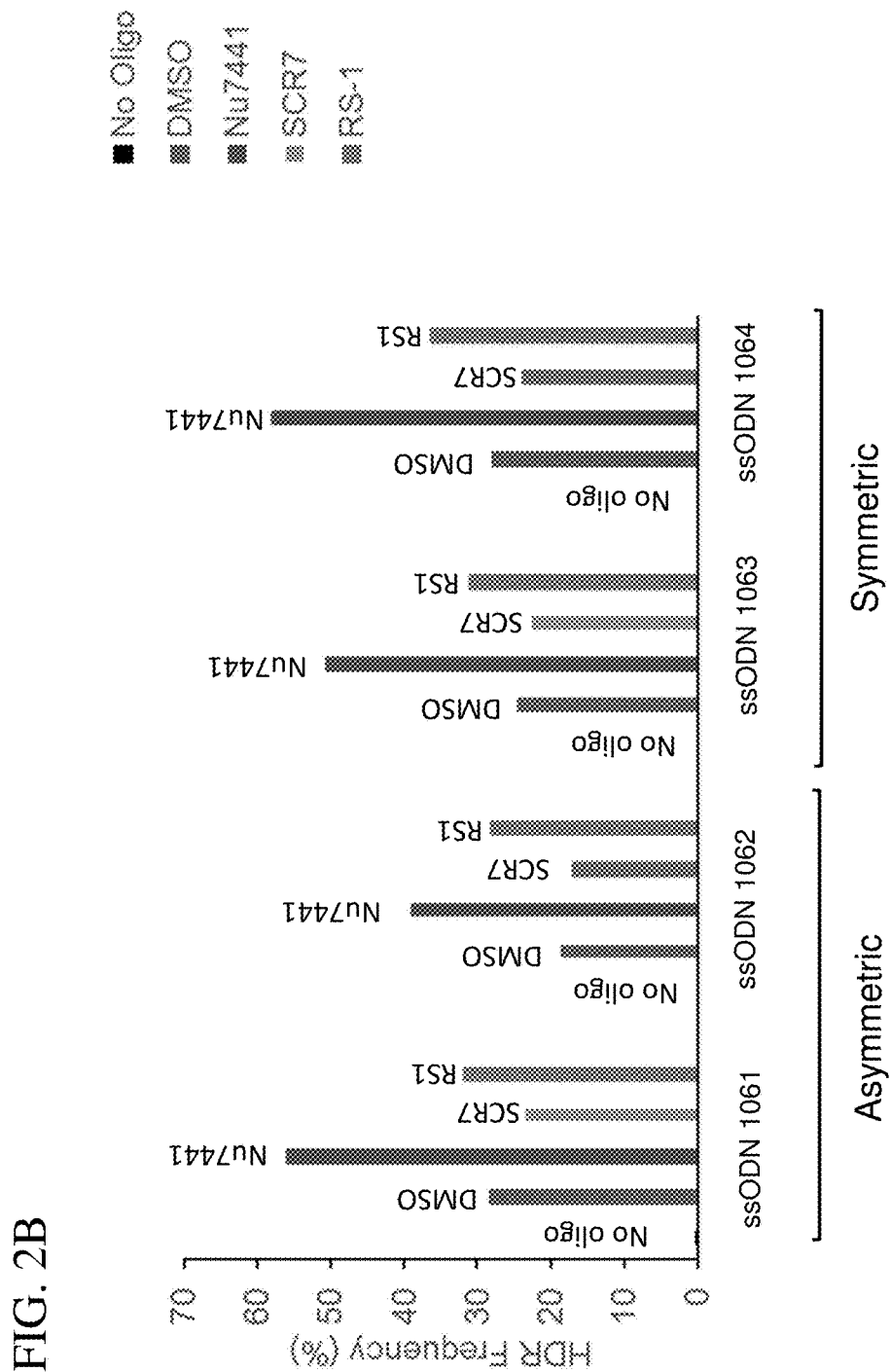


FIG. 2A





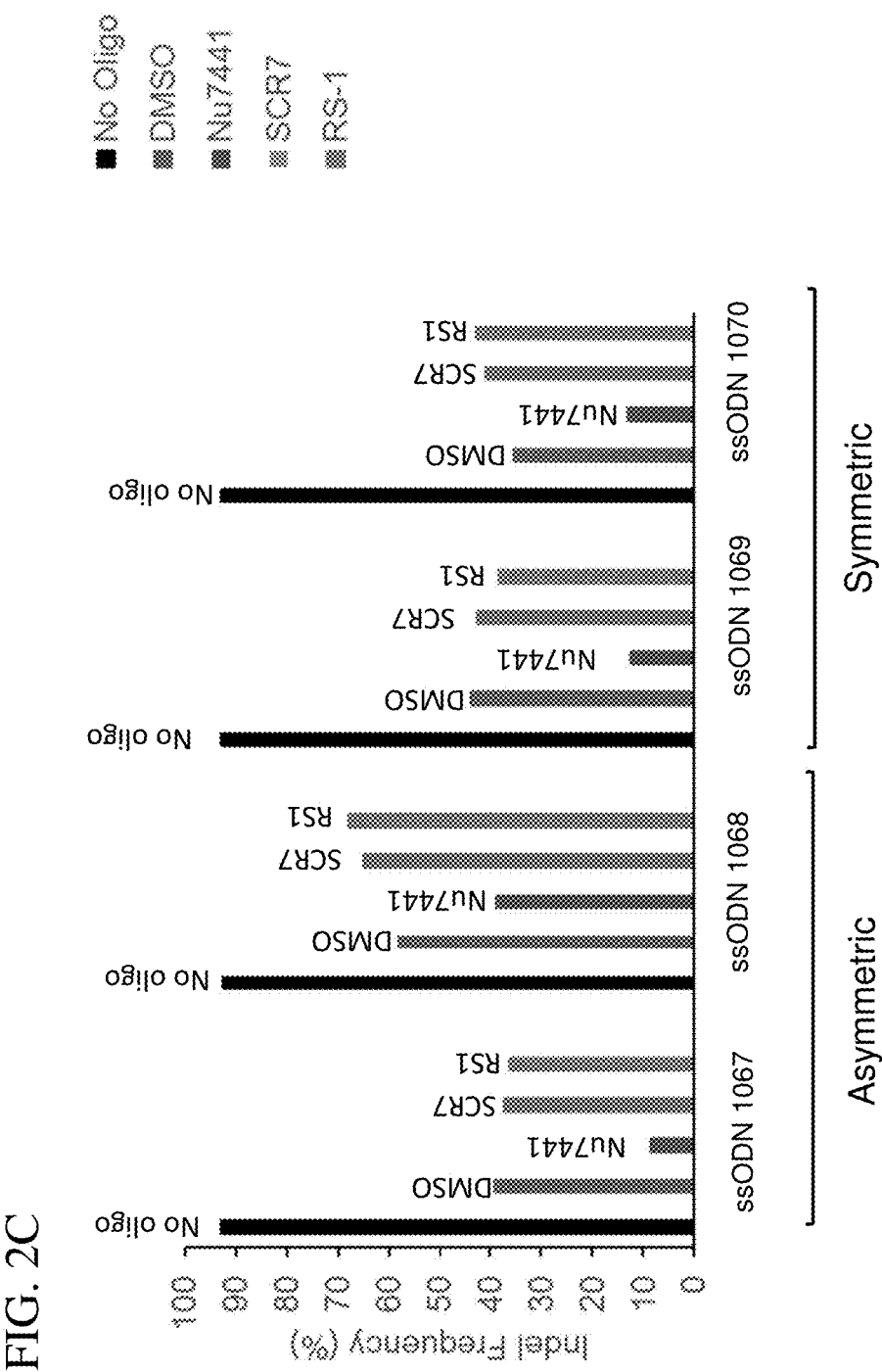


FIG. 2D

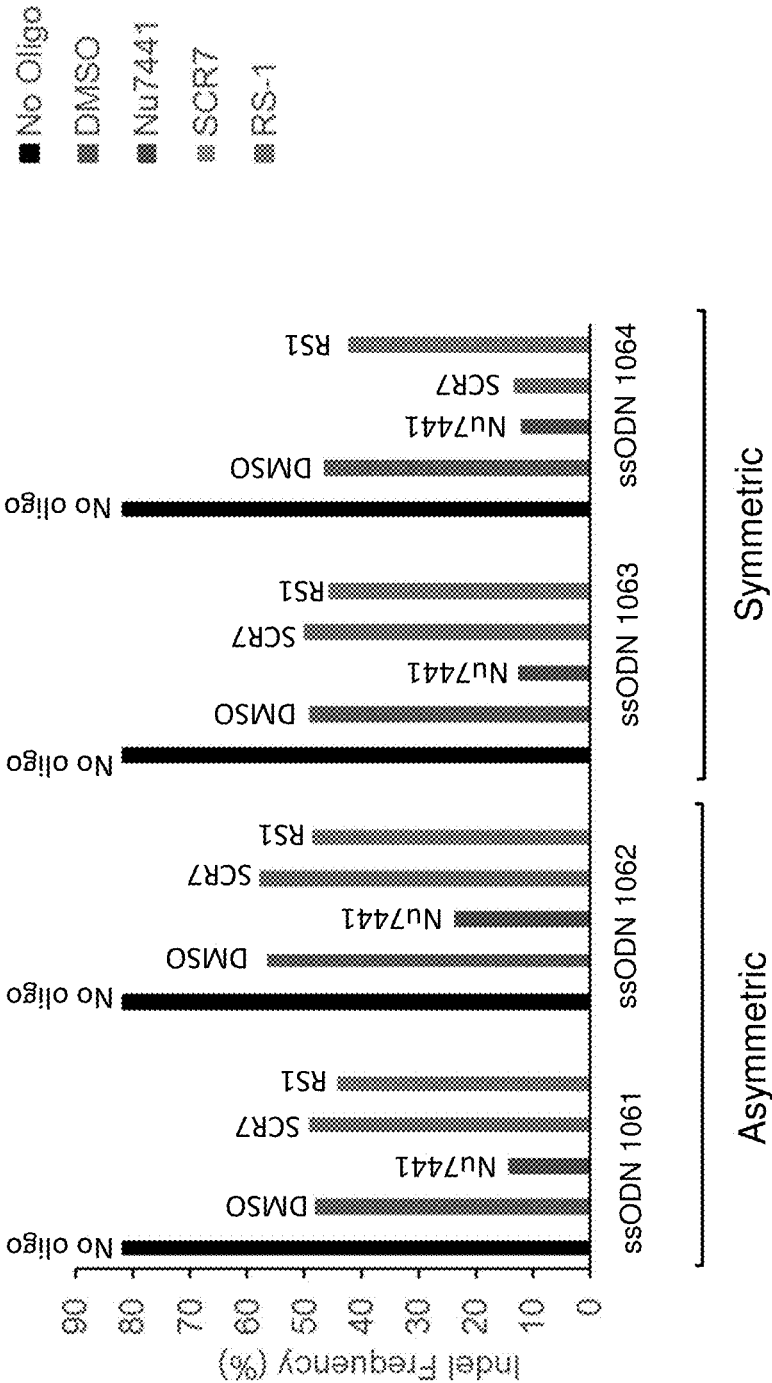


FIG. 3

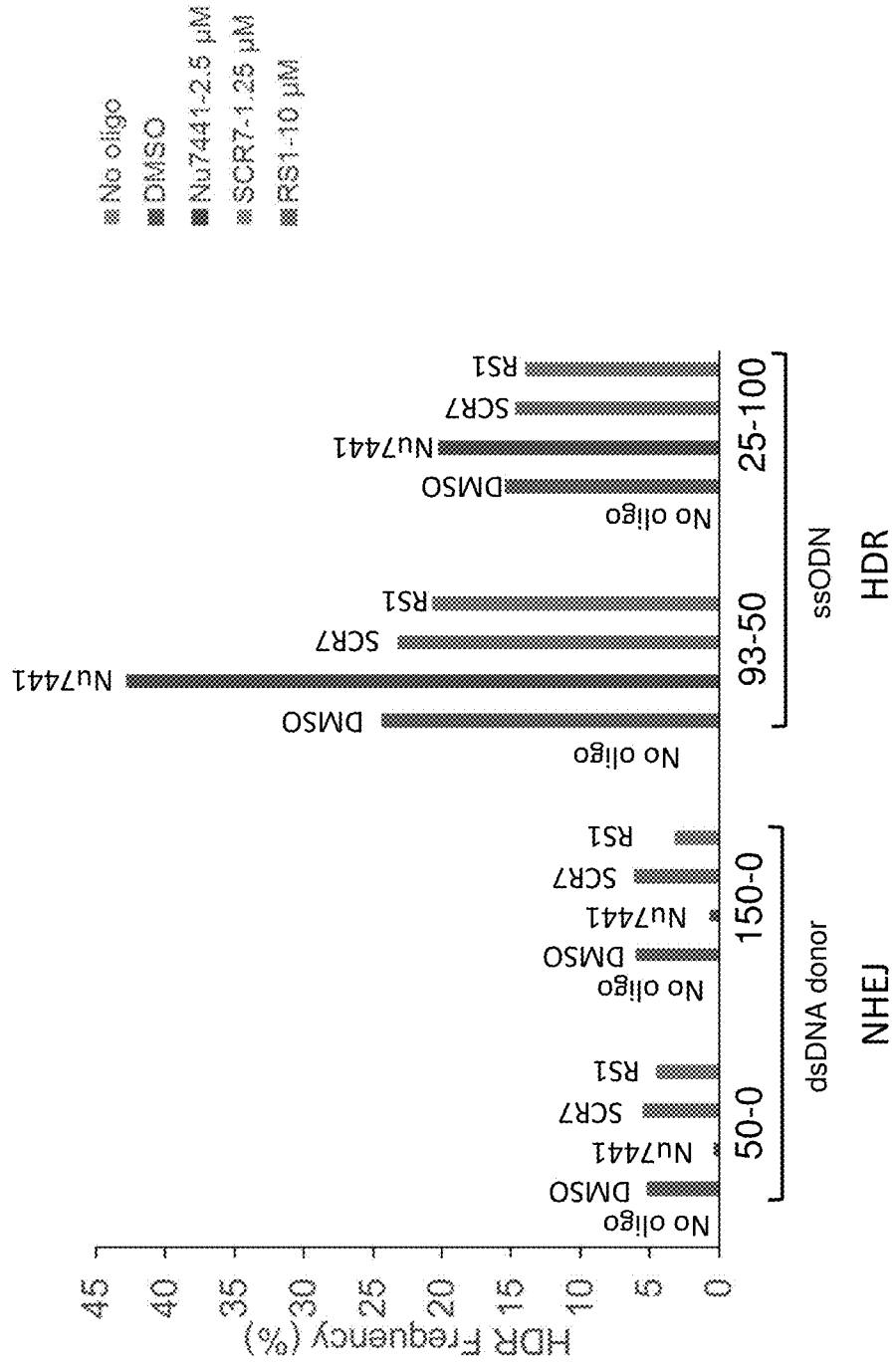


FIG. 4A

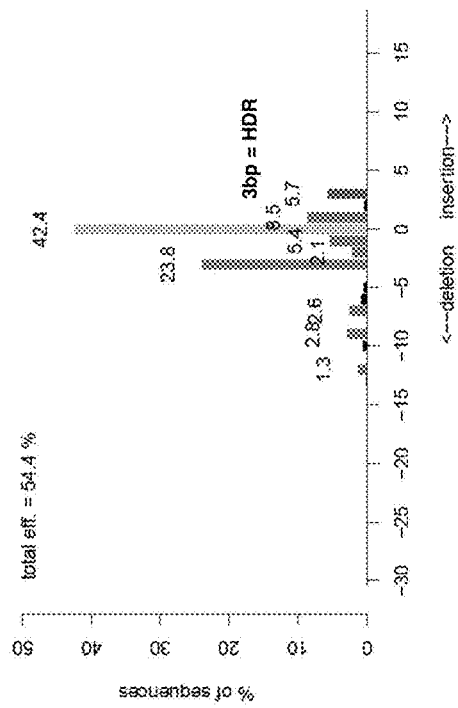
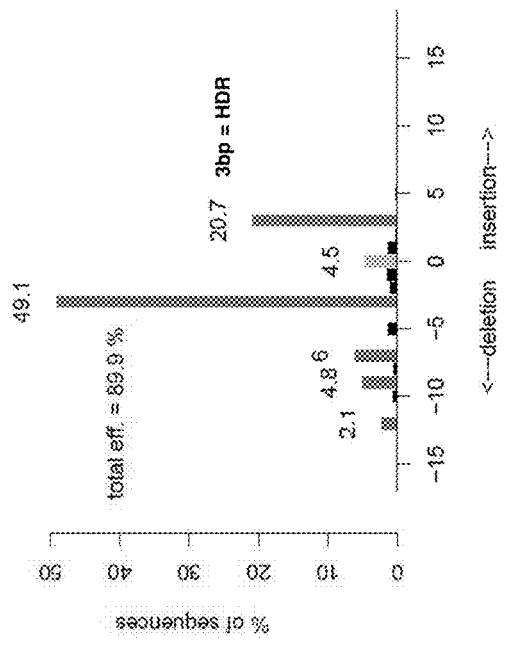


FIG. 4B



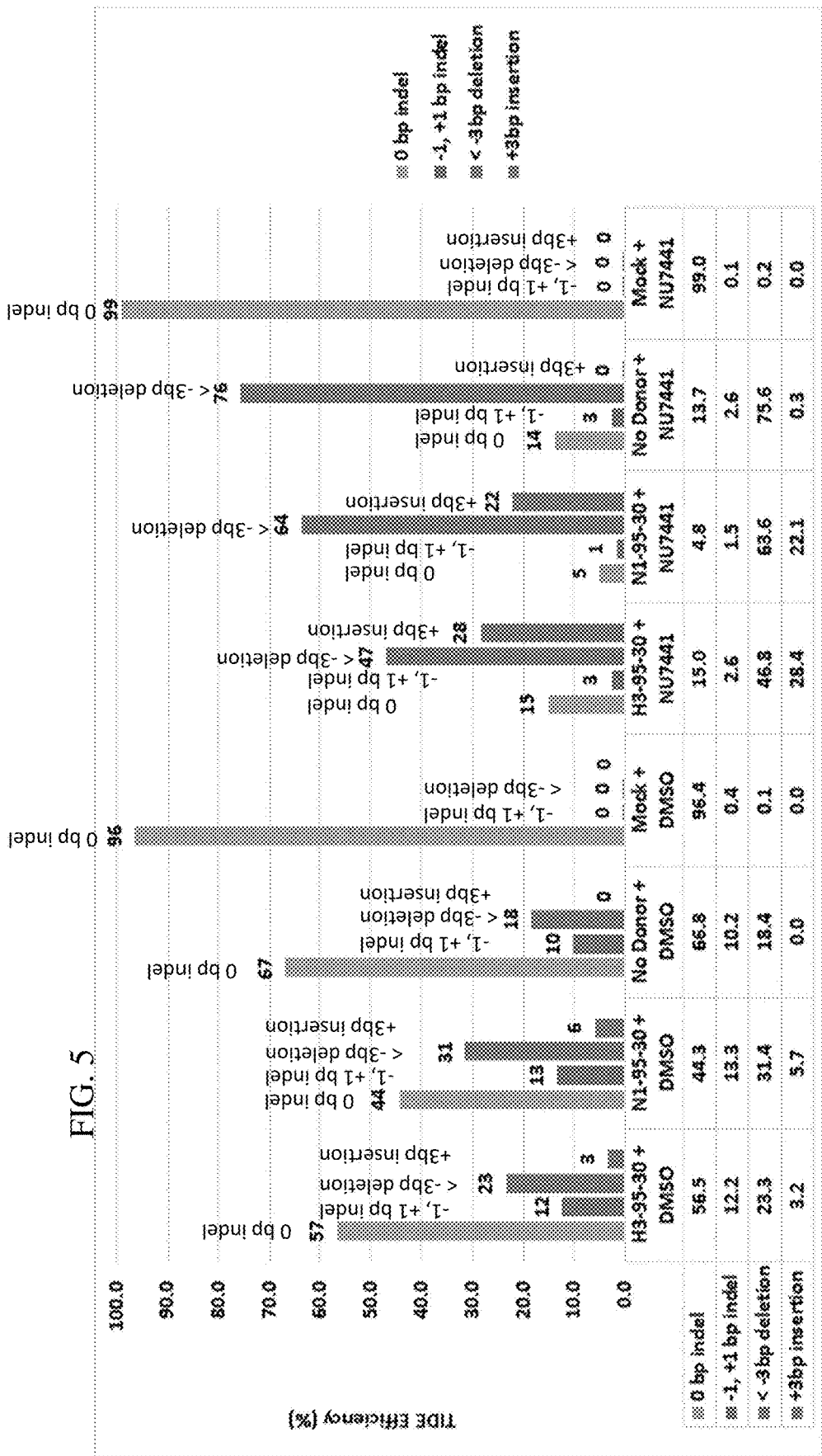


FIG. 6A

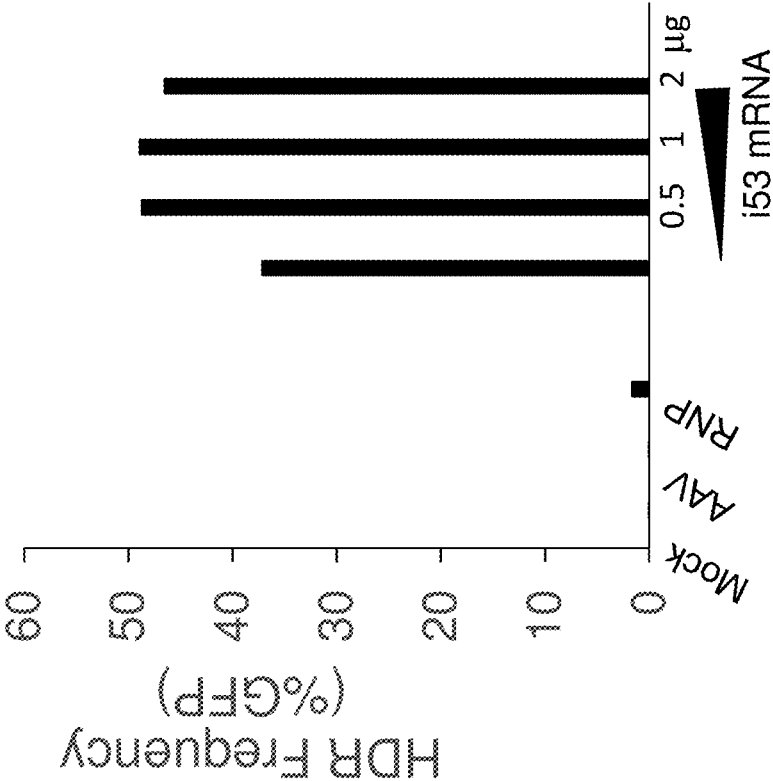


FIG. 6B

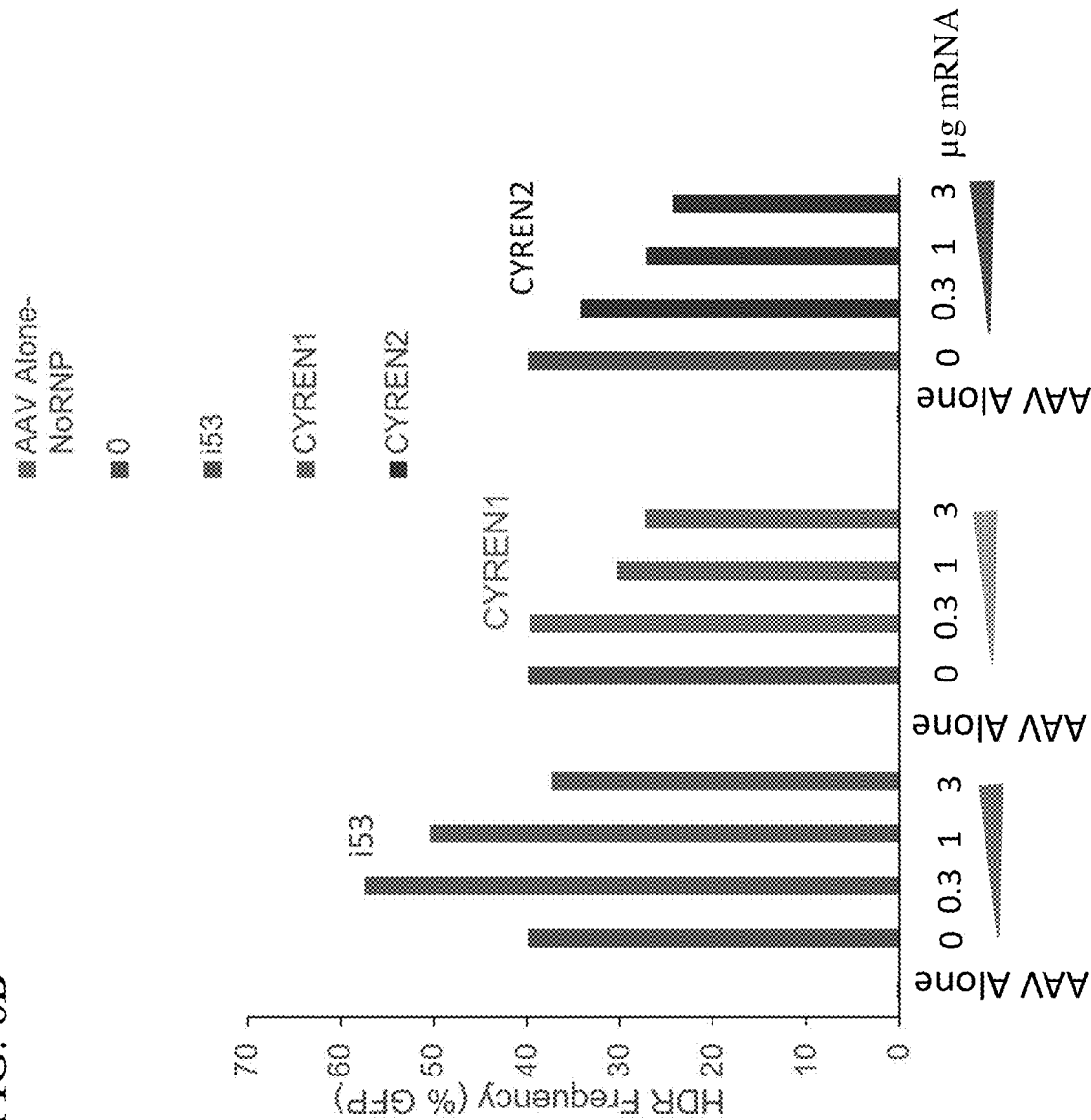


FIG. 6C

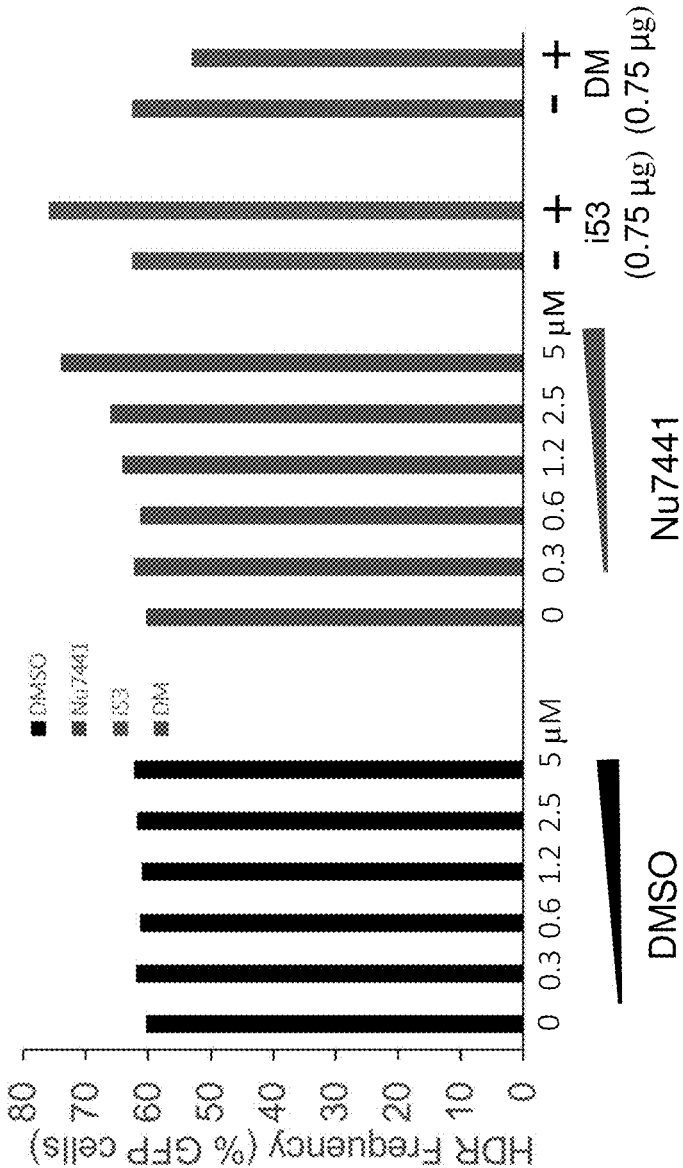


FIG. 7

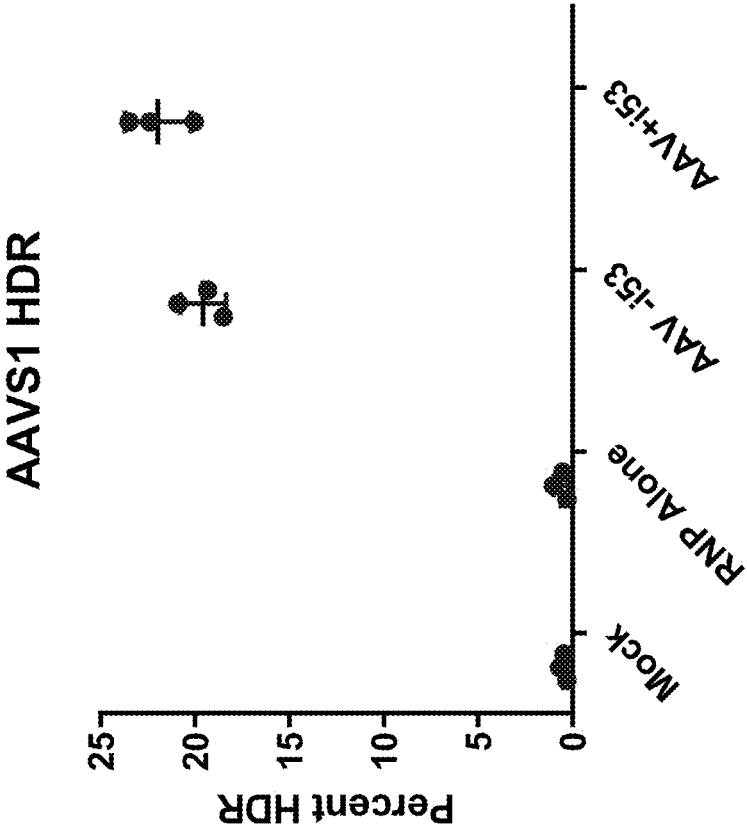


FIG. 8

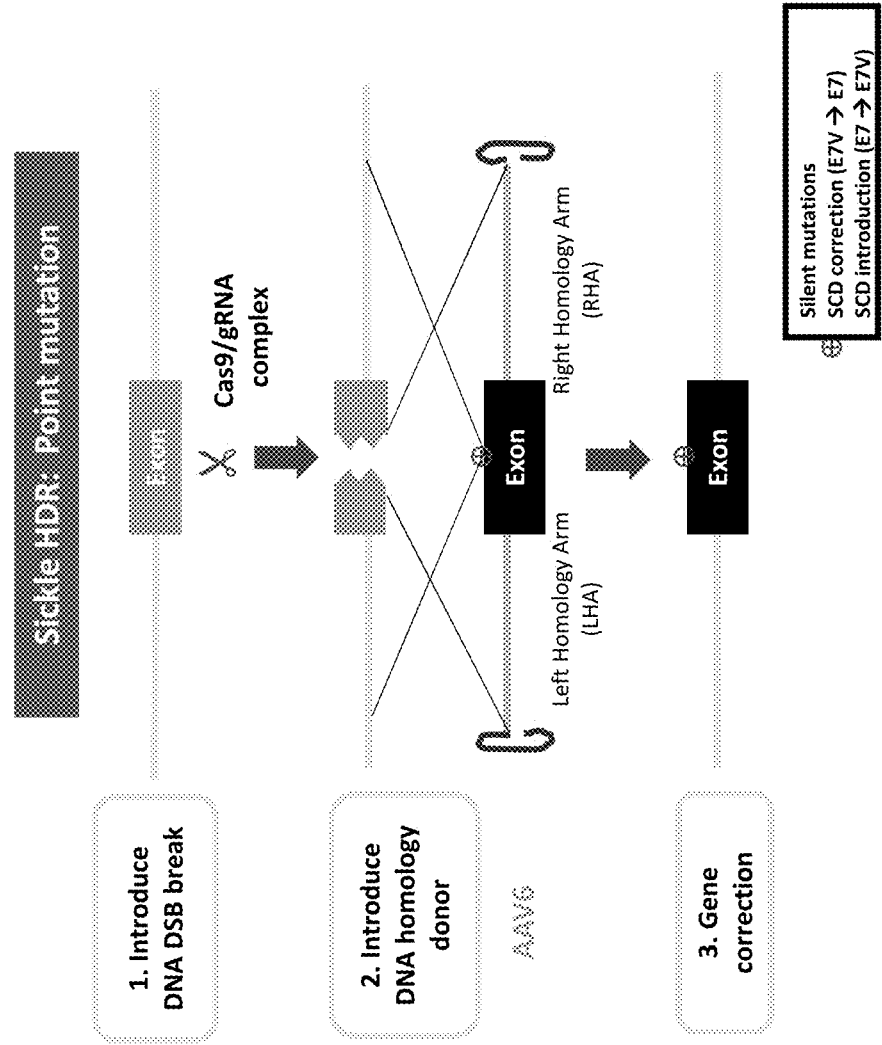


FIG. 9

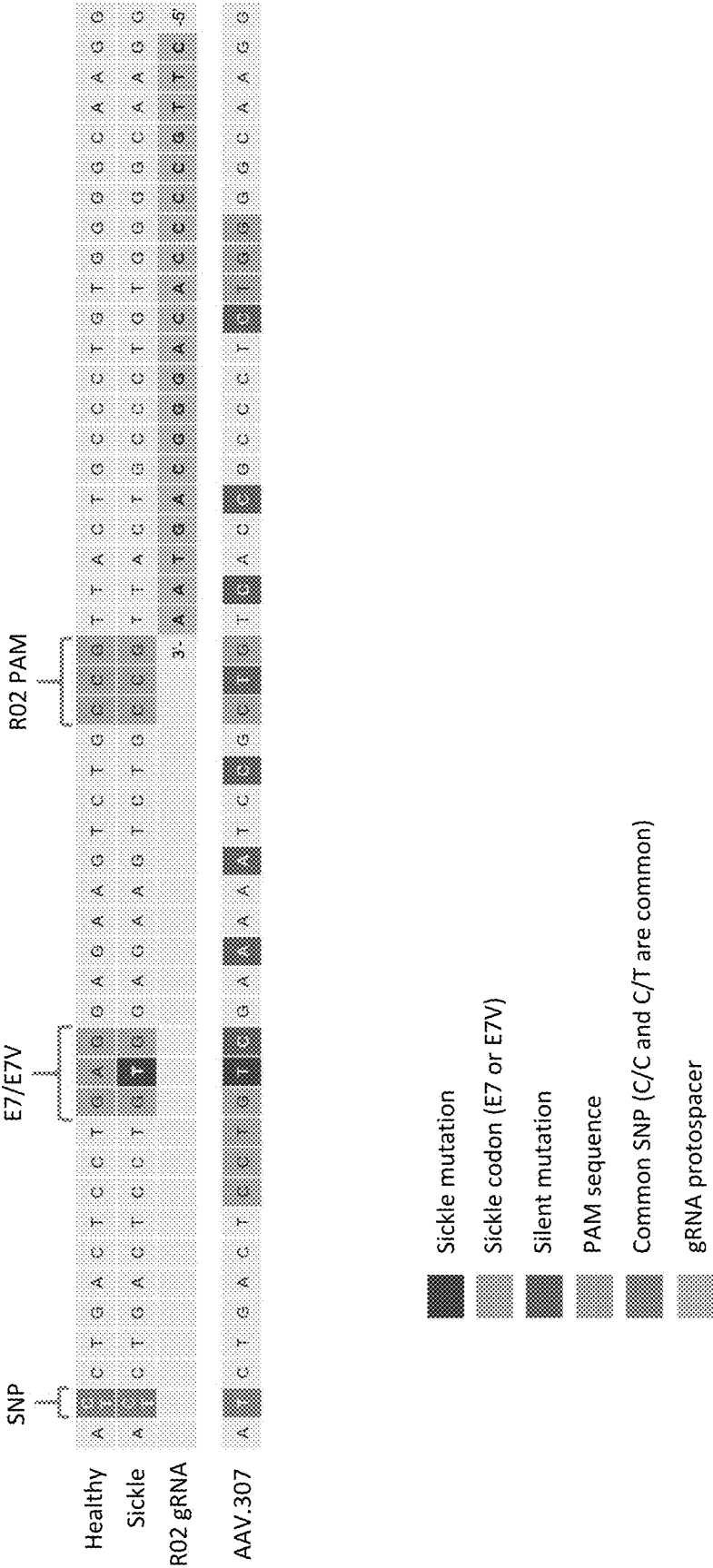


FIG. 10

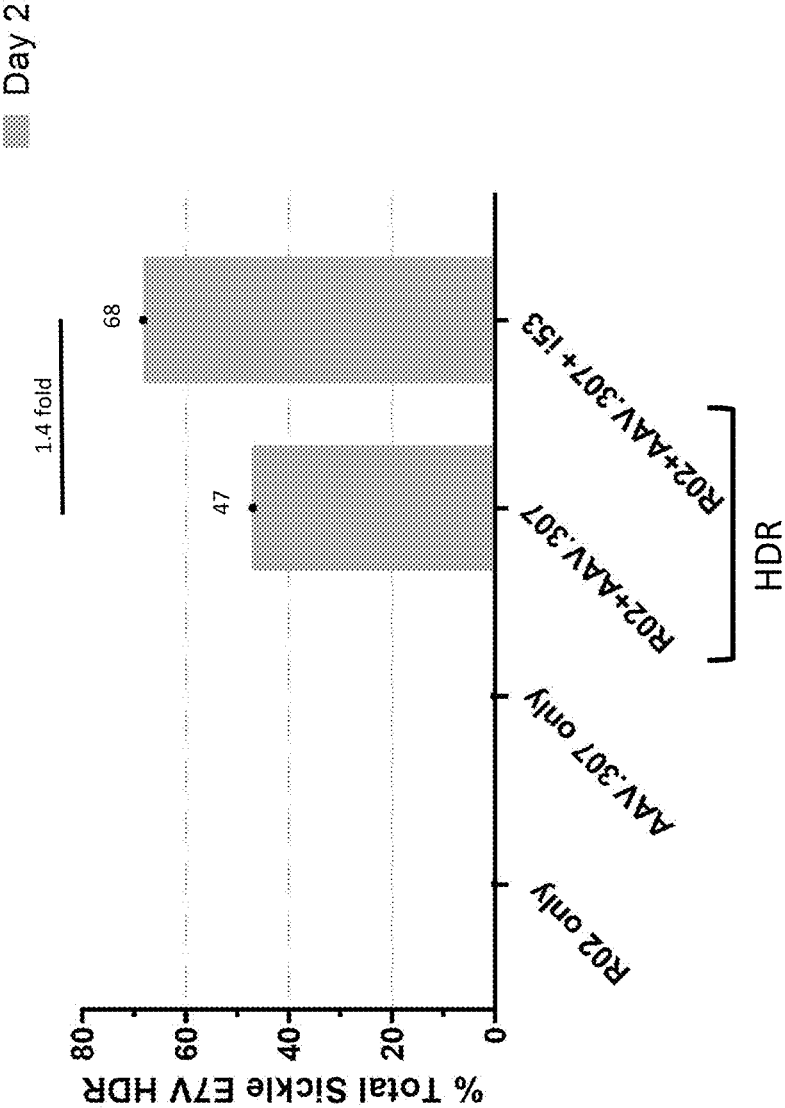


FIG. 11

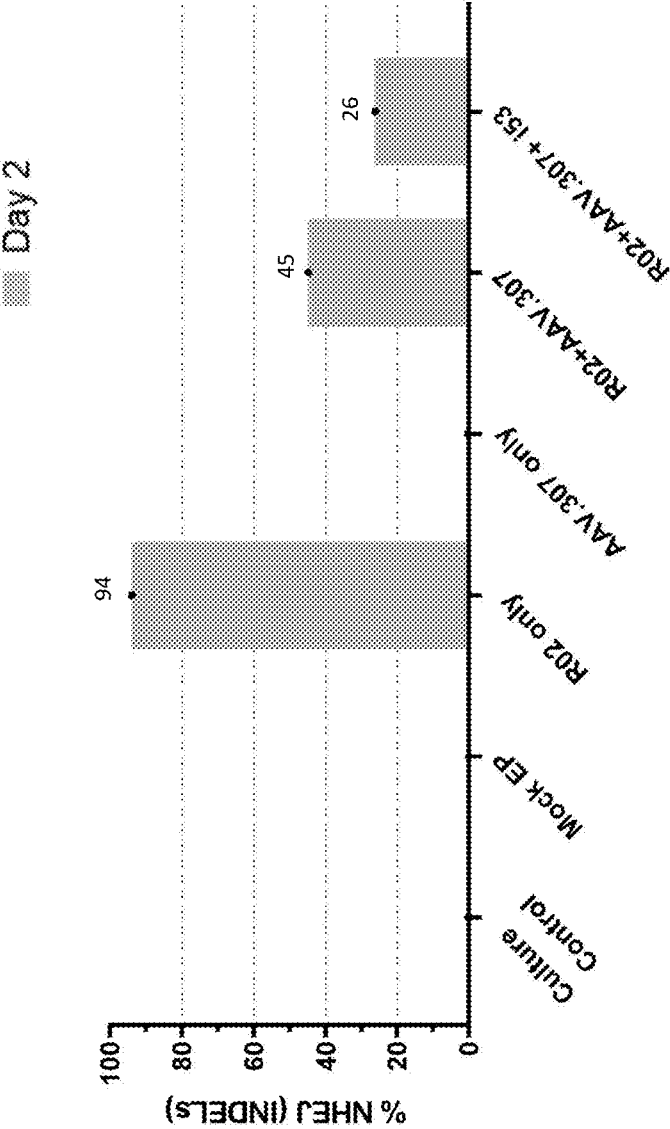


FIG. 12A

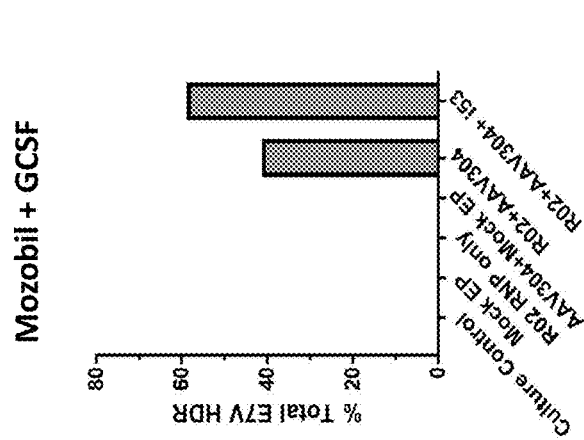


FIG. 12B

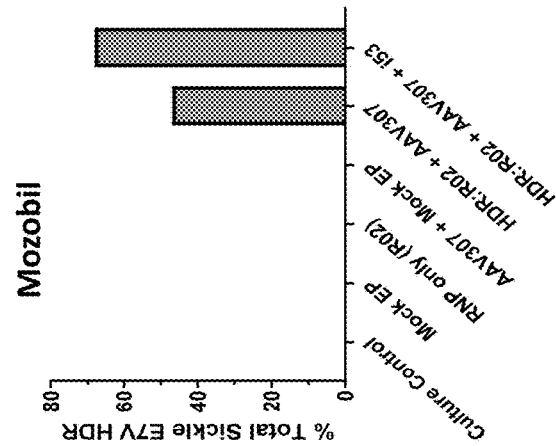


FIG. 13

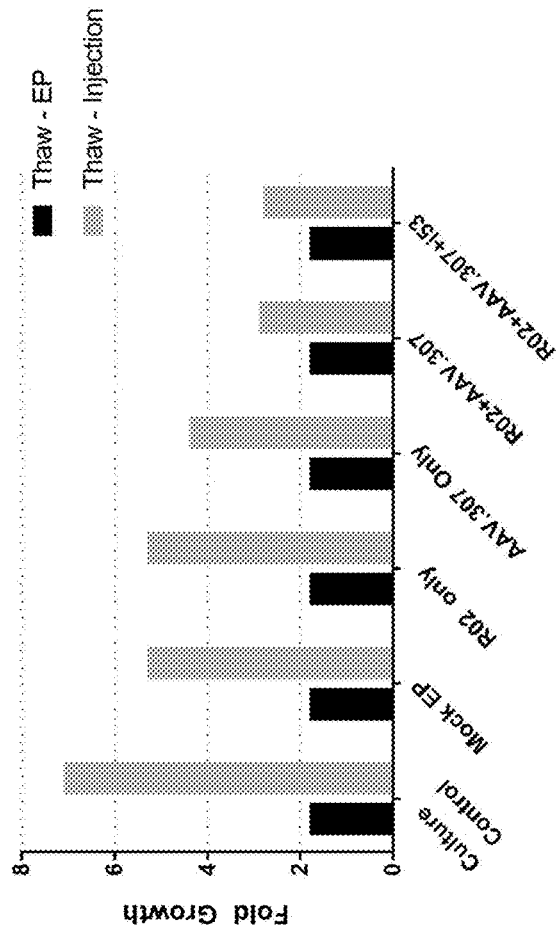


FIG. 14

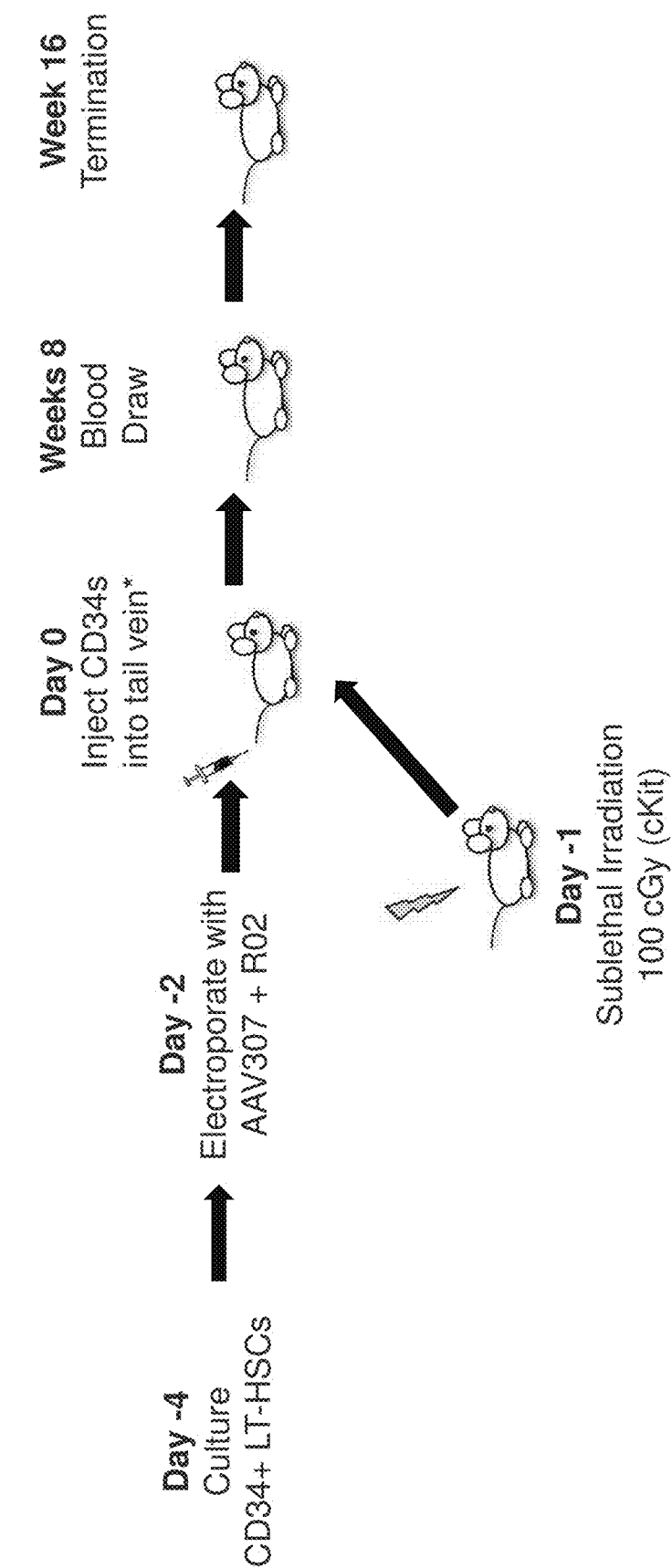


FIG. 15

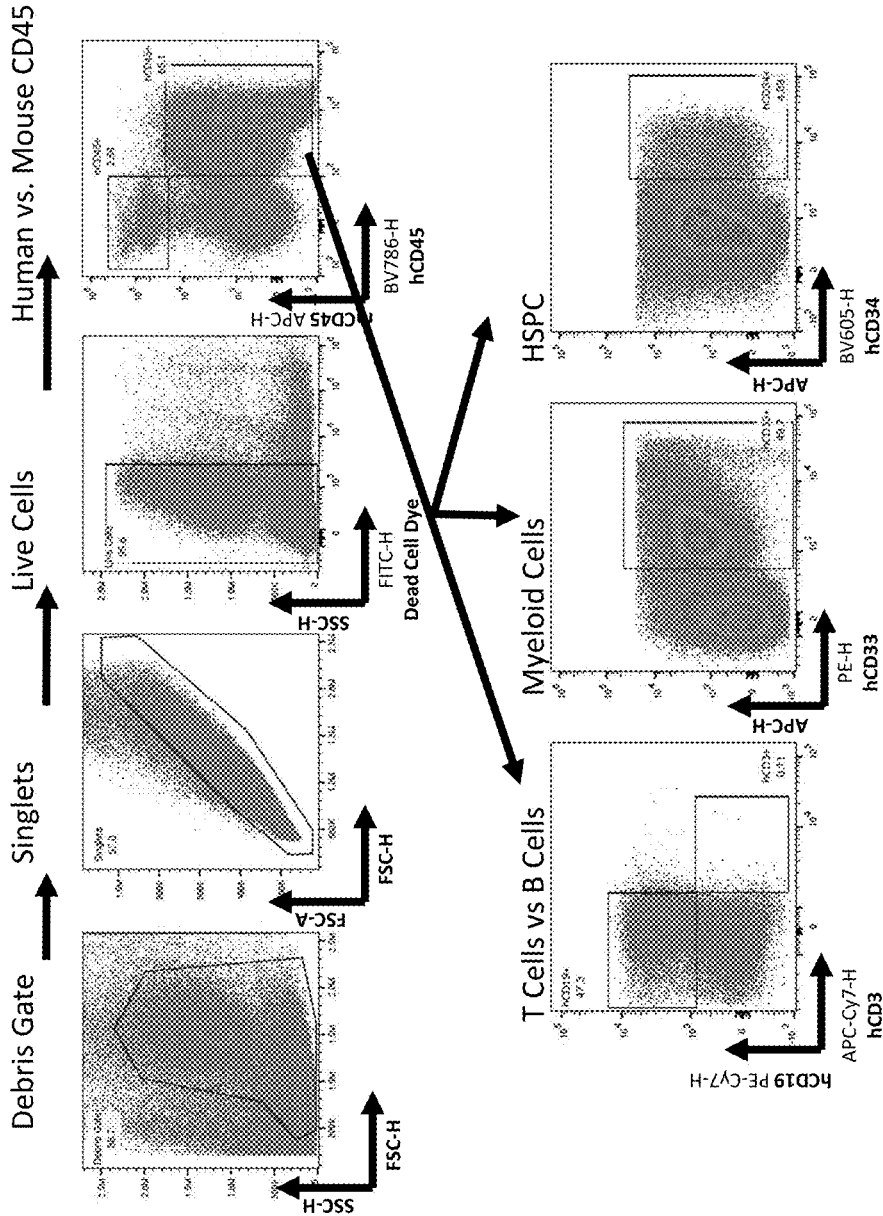


FIG. 16

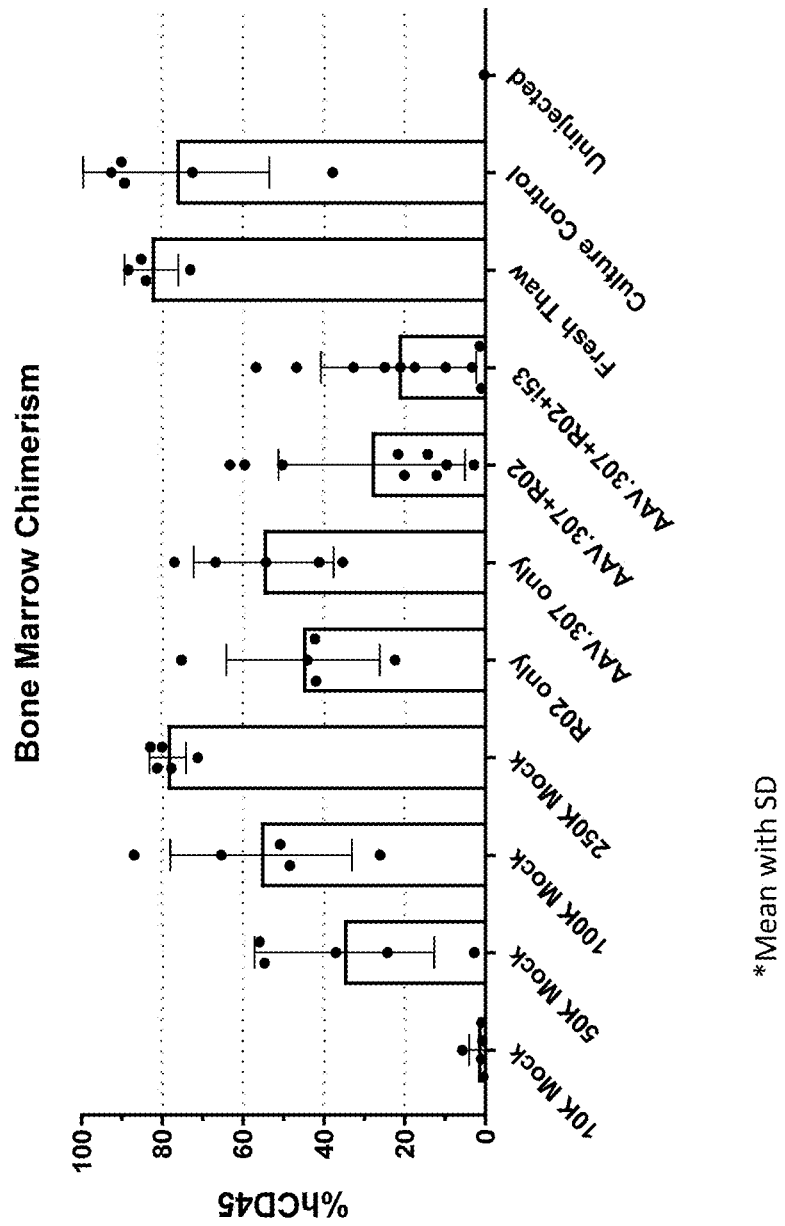


FIG. 17

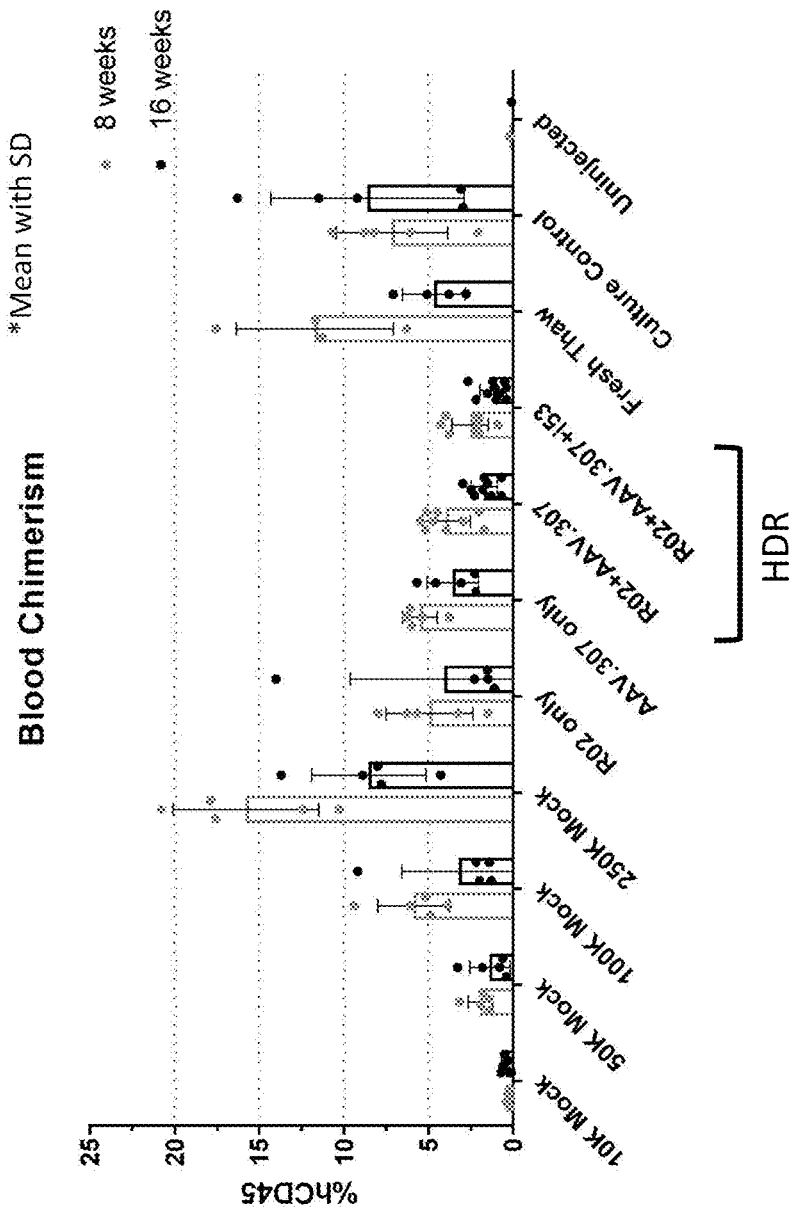


FIG. 18A

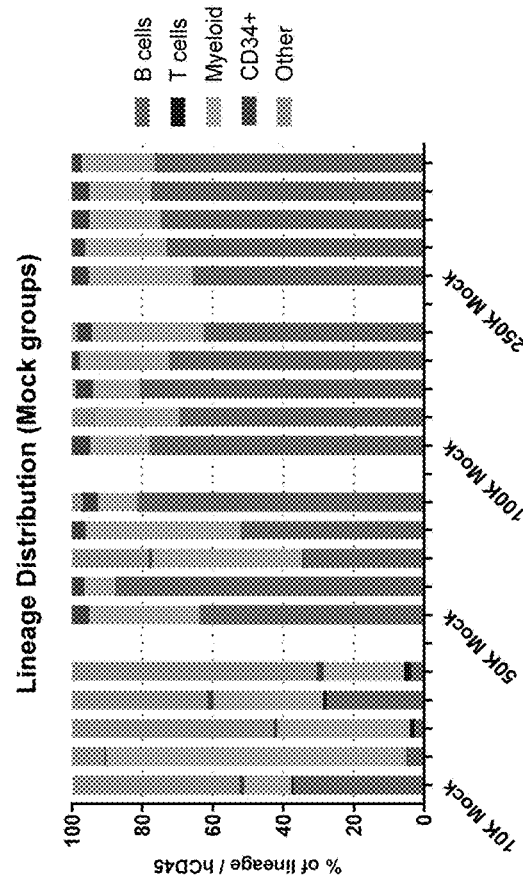
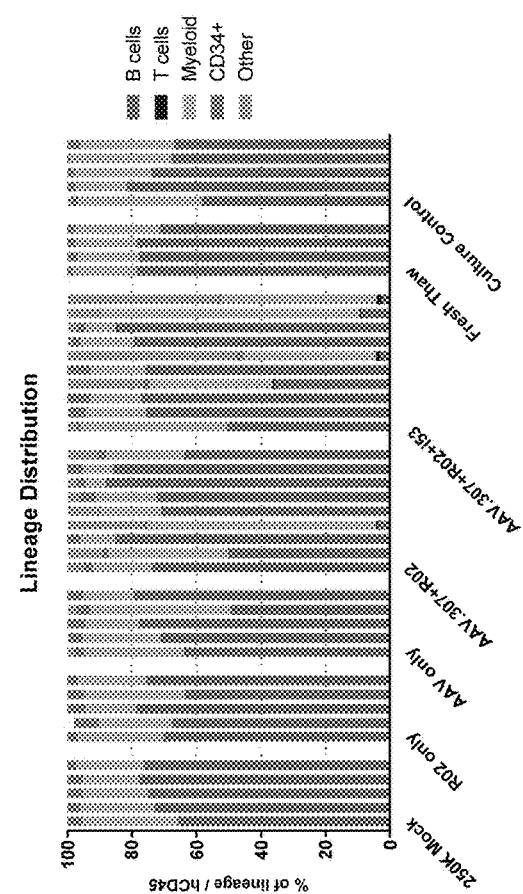
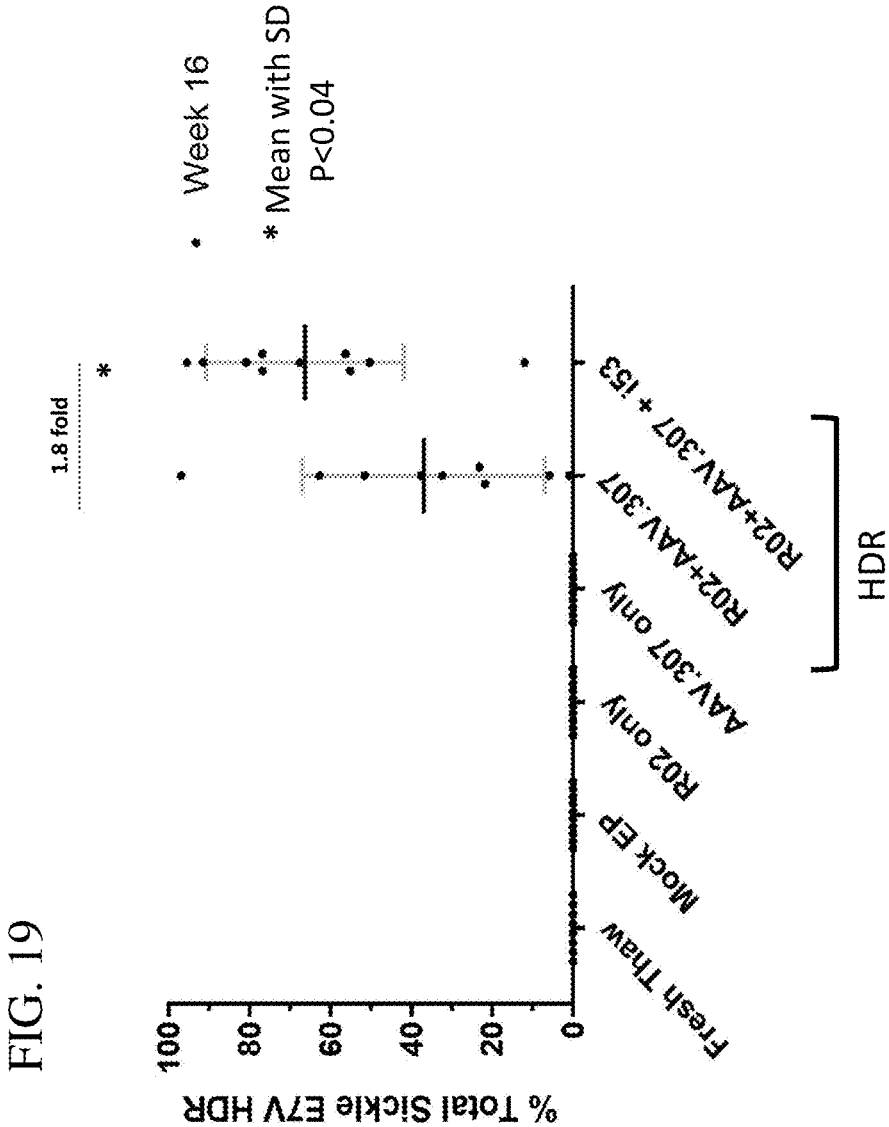


FIG. 18B





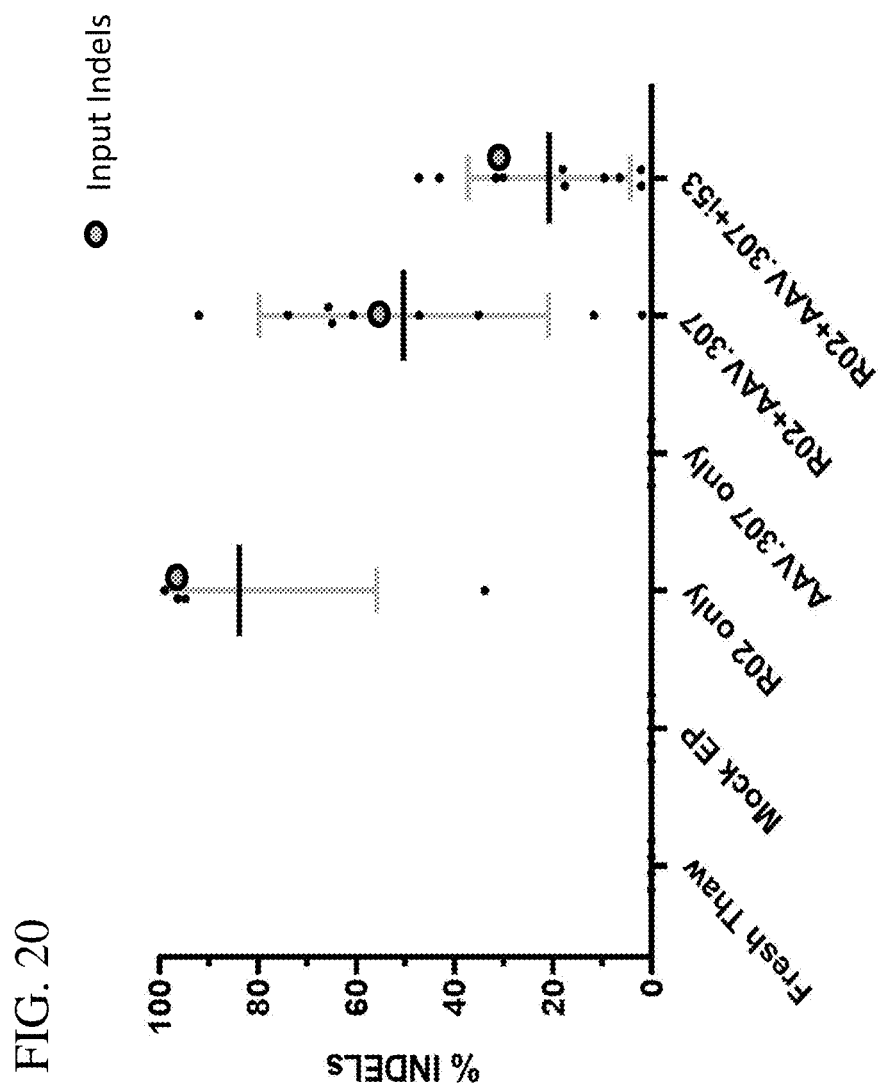


FIG. 21

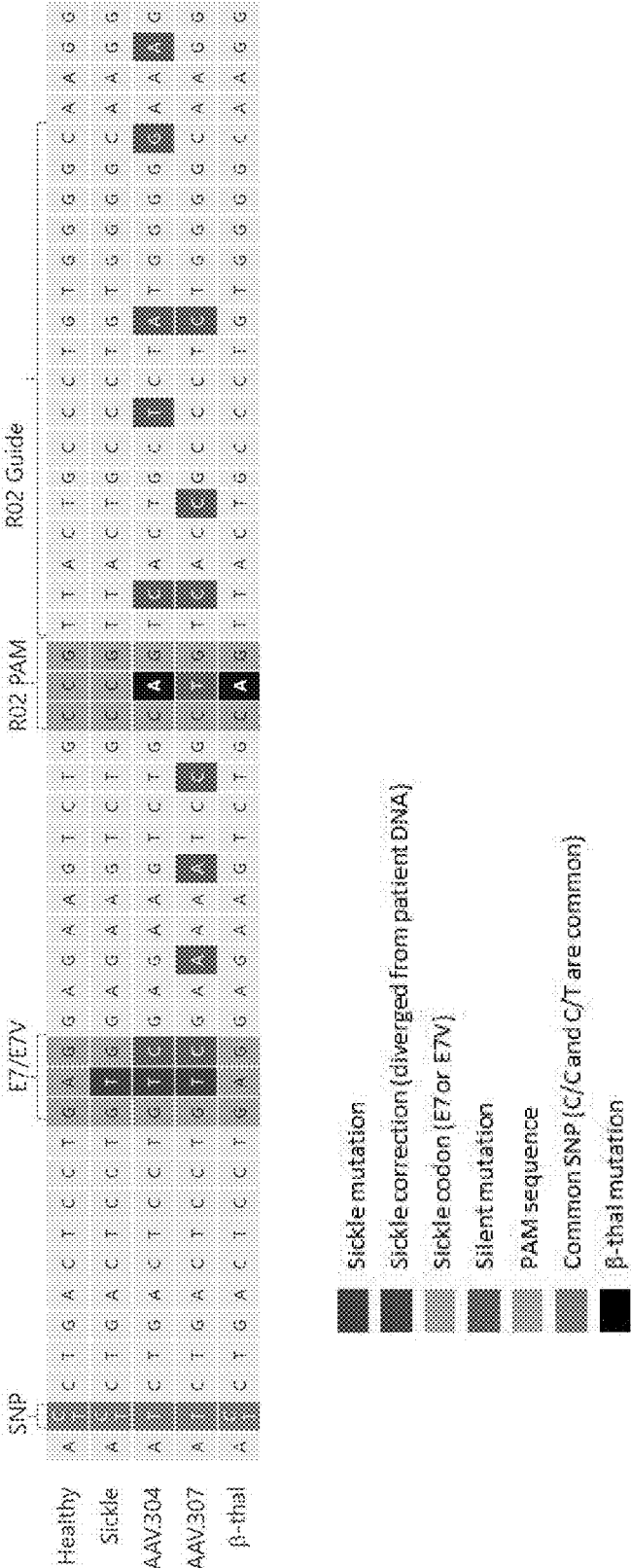


FIG. 22A

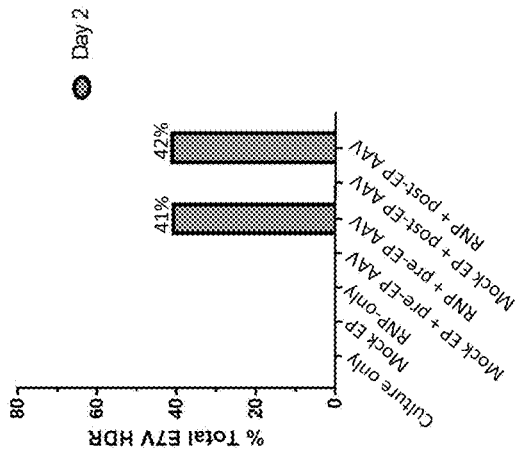
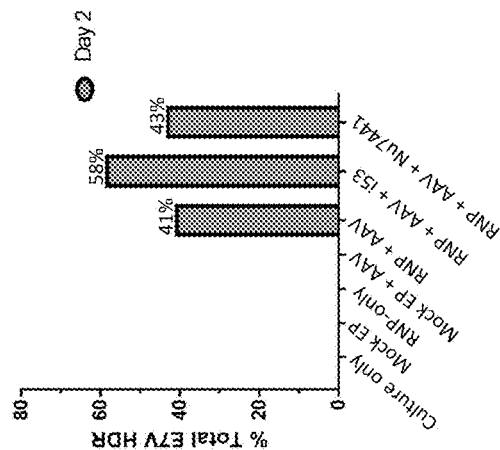


FIG. 22B



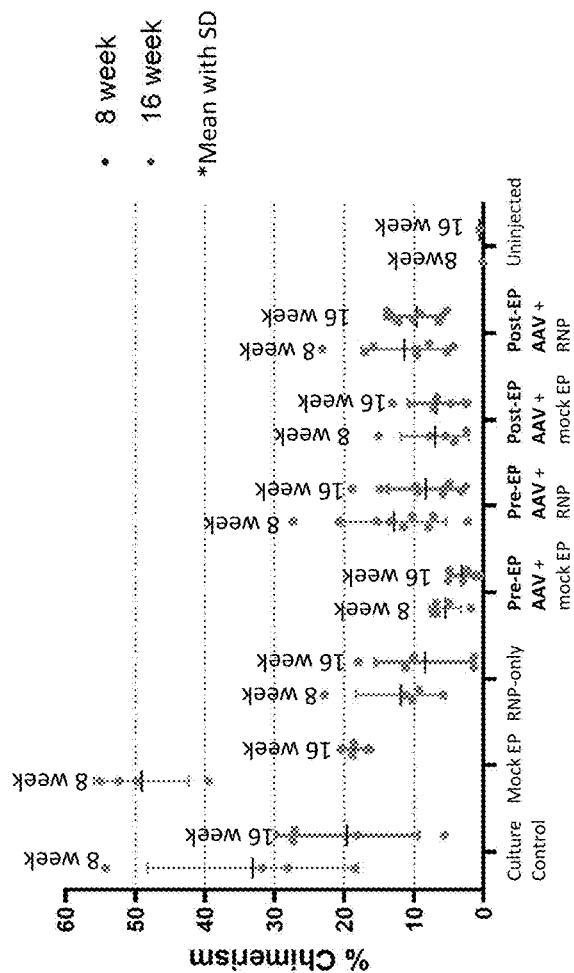


FIG. 23A

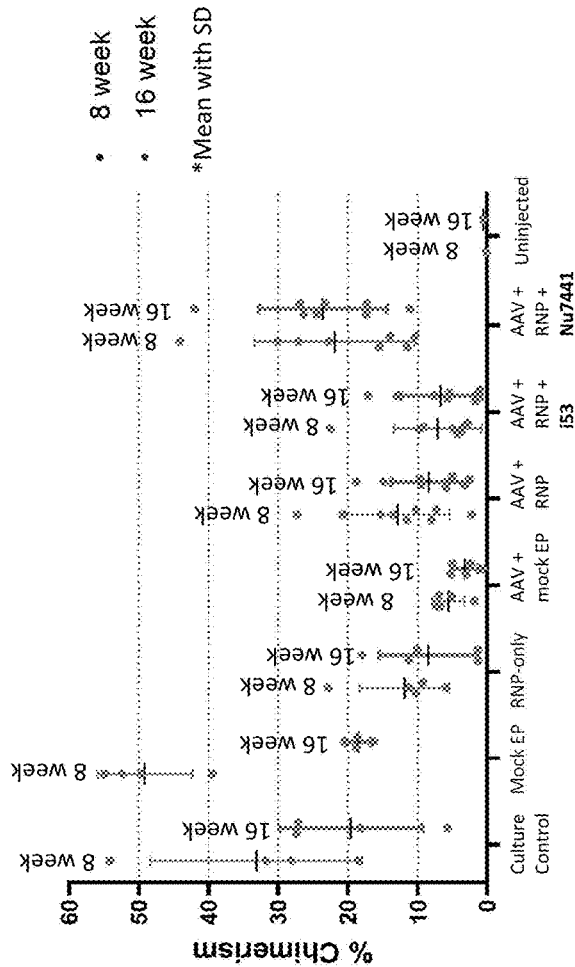


FIG. 23B

FIG. 24A

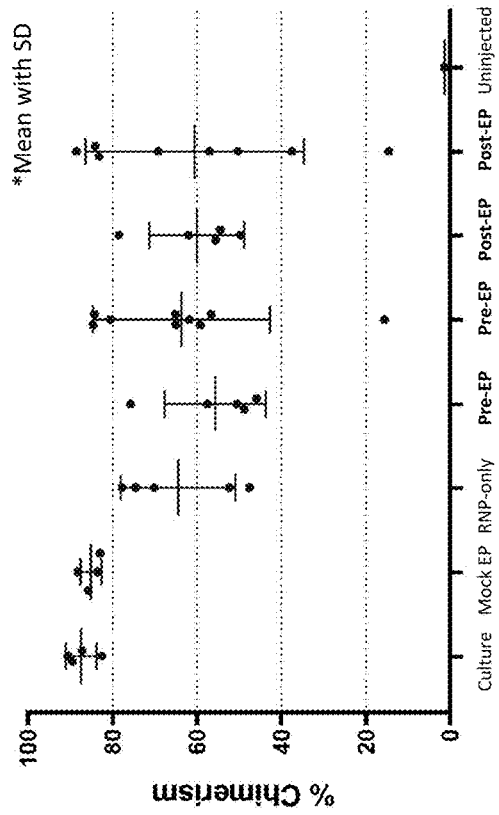


FIG. 24B

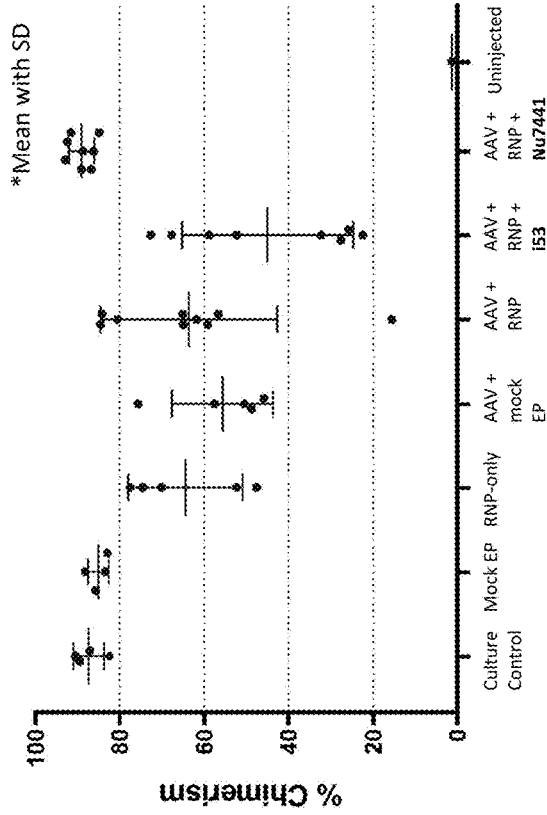
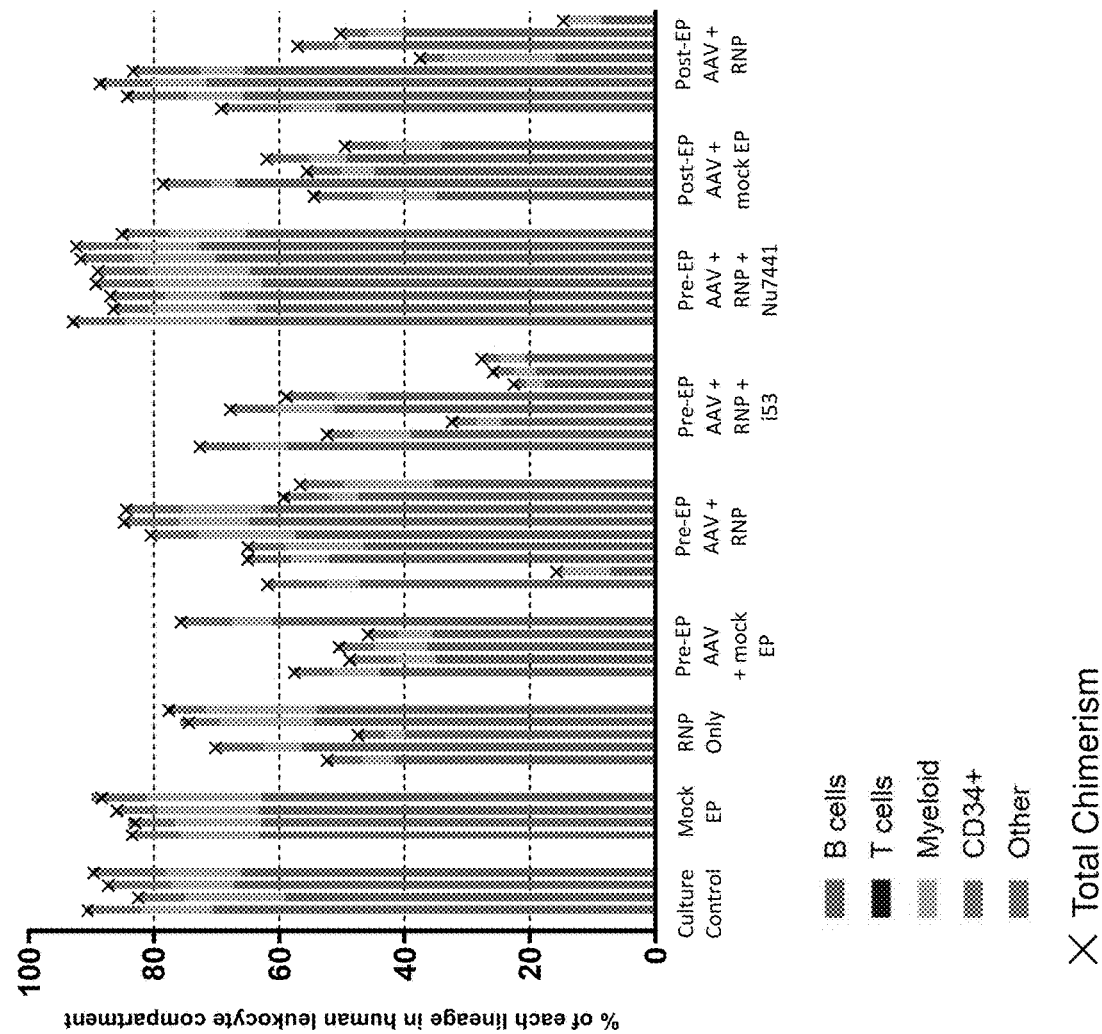
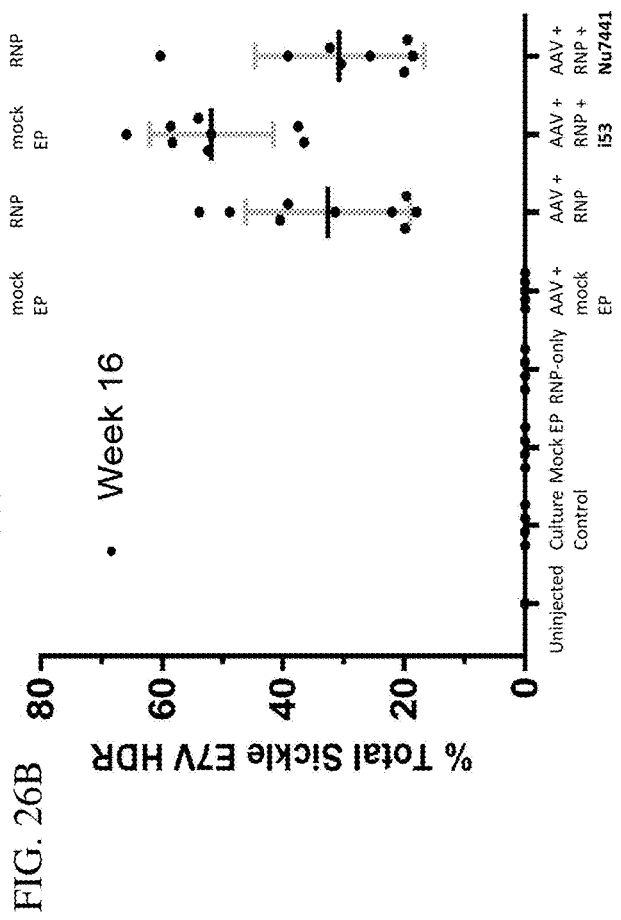
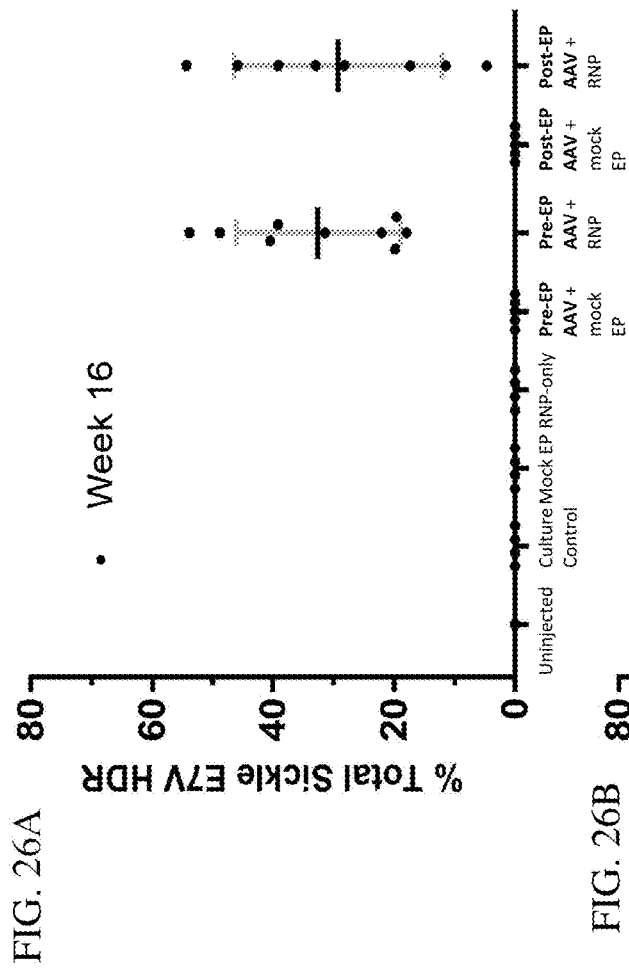


FIG. 25





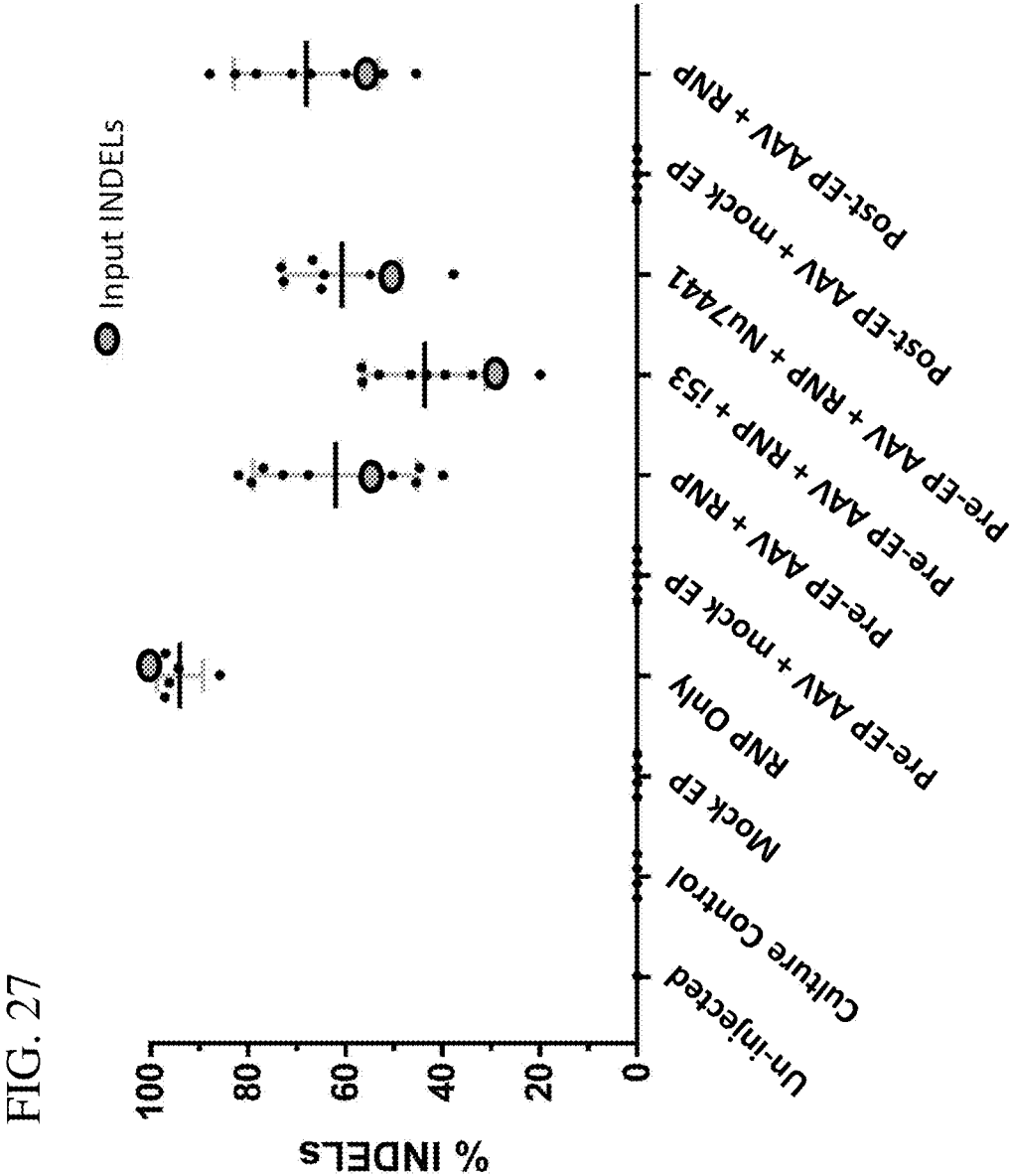


FIG. 28

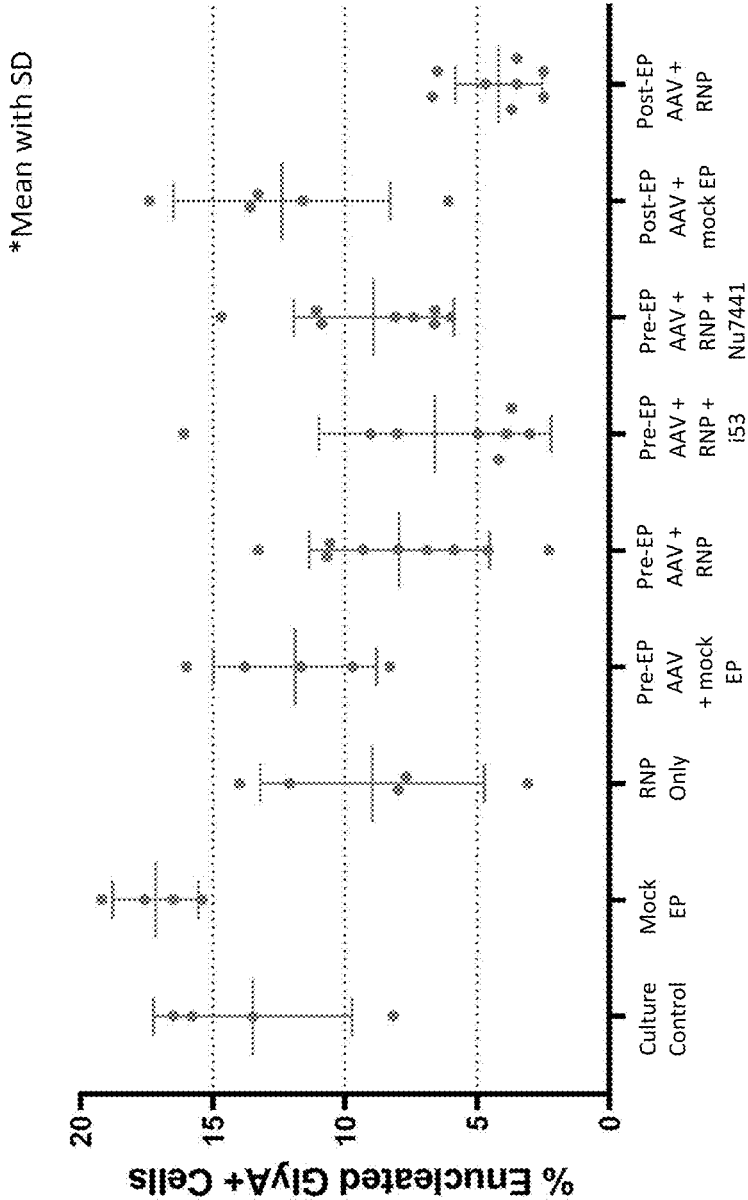


FIG. 29

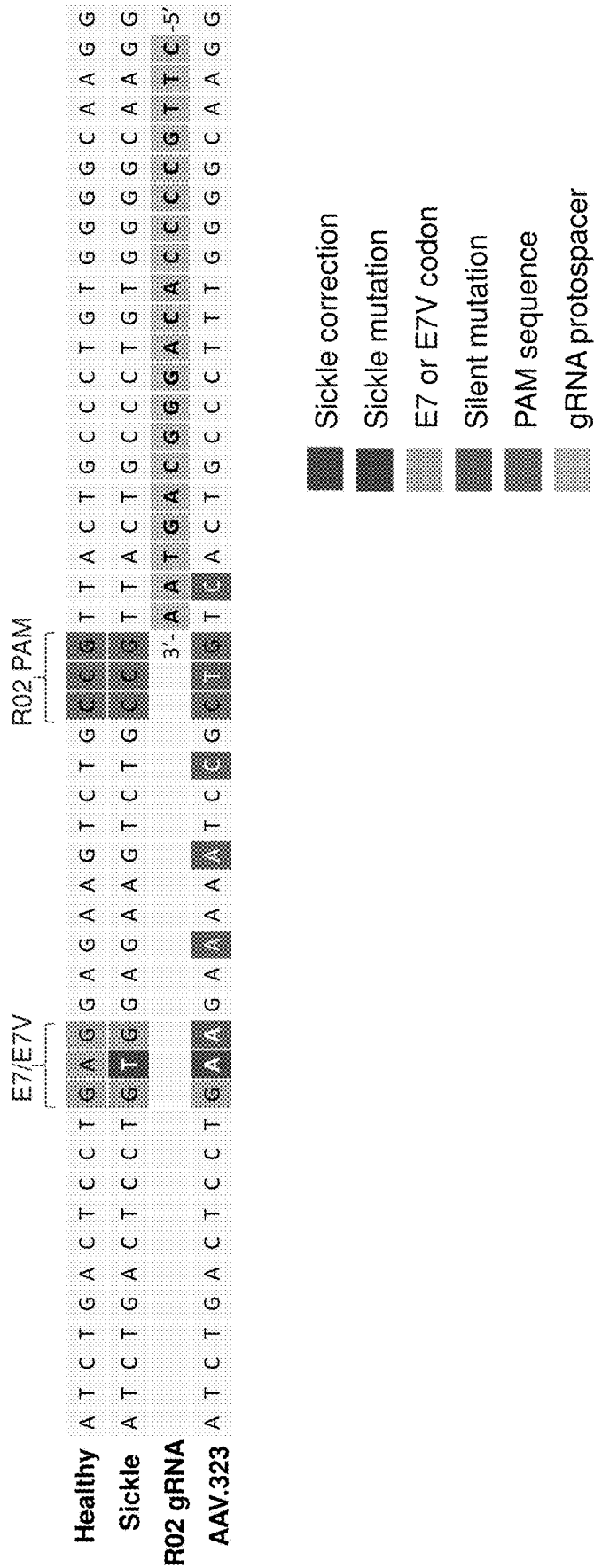


FIG. 30A

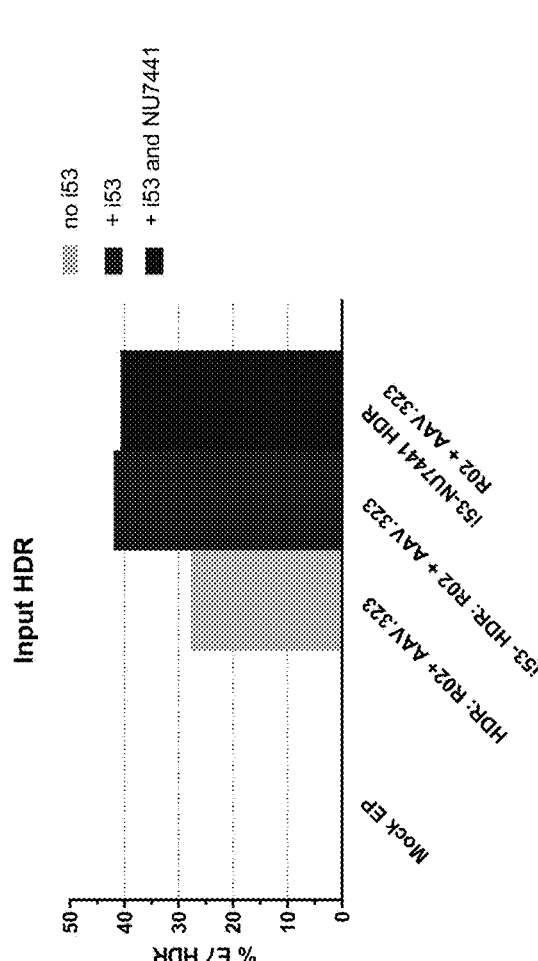
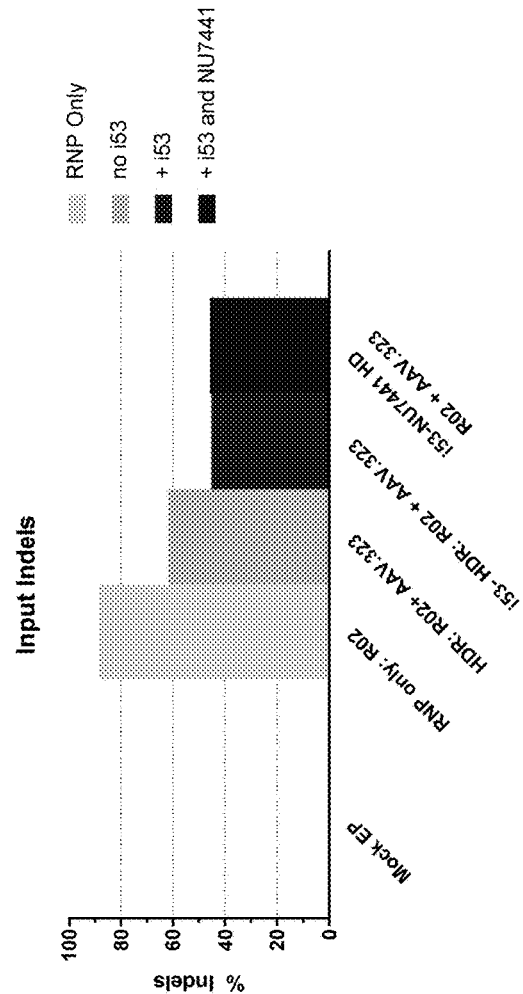


FIG. 30B



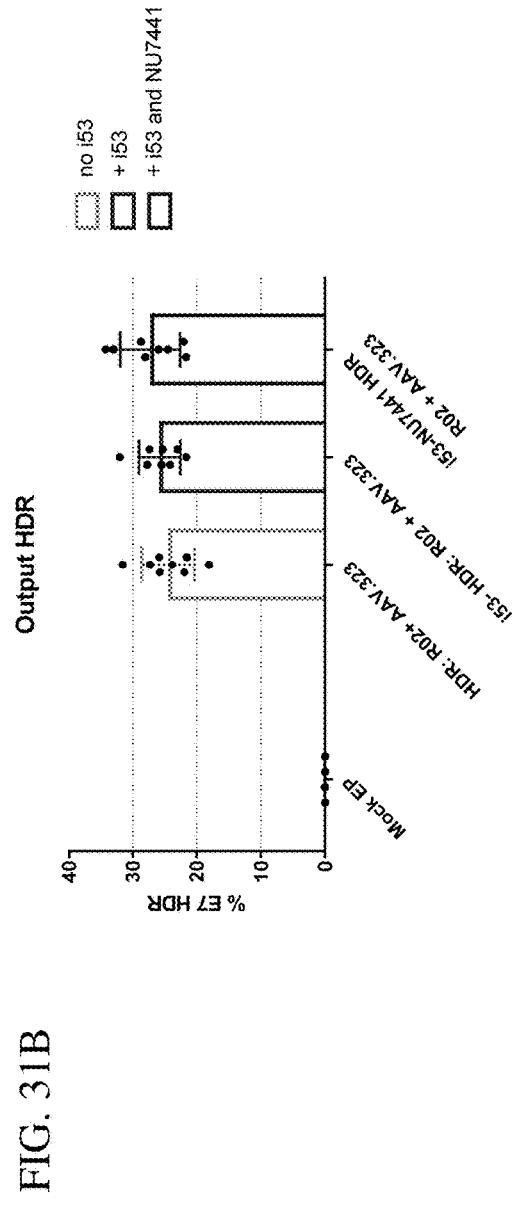
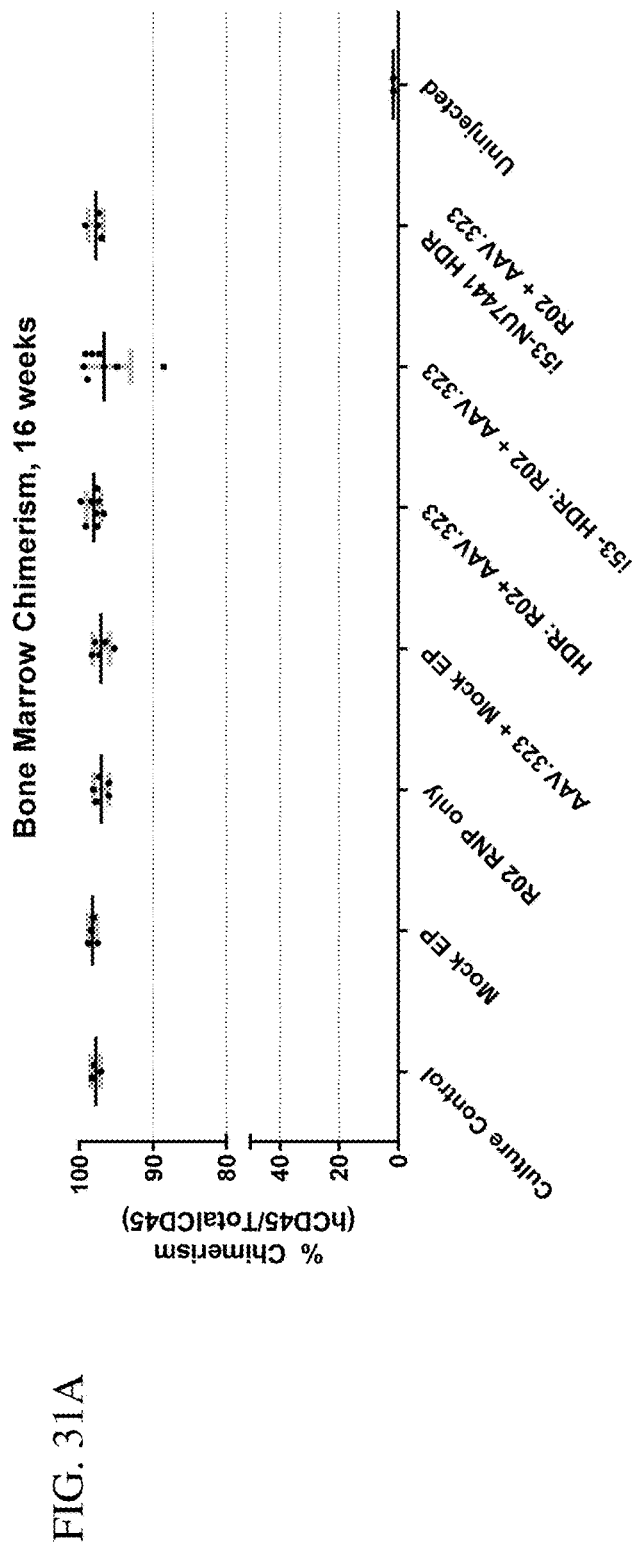


FIG. 32A

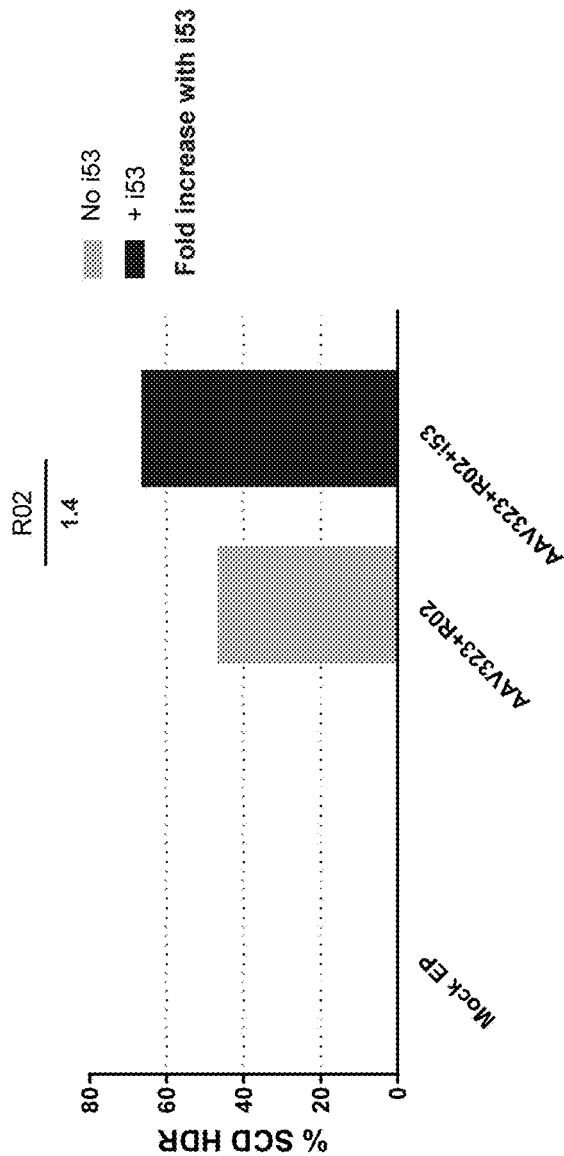
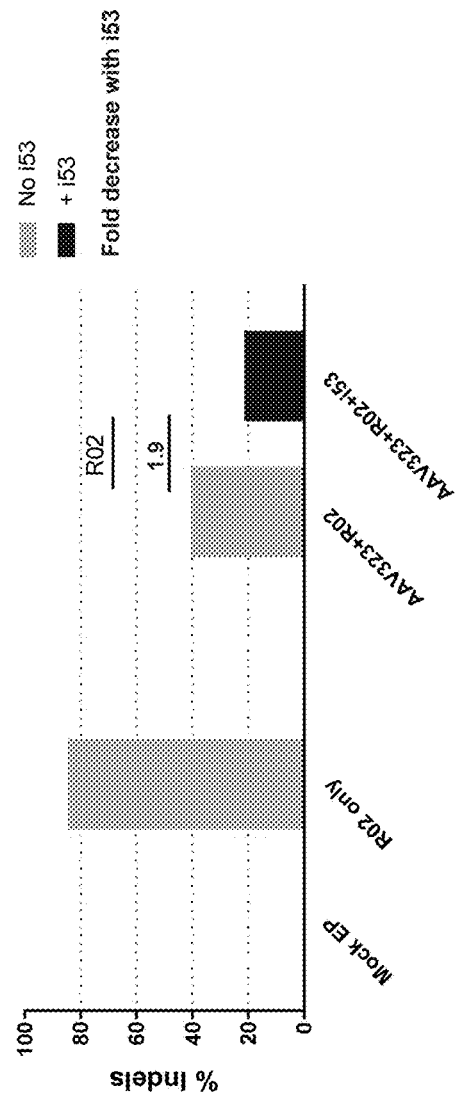


FIG. 32B



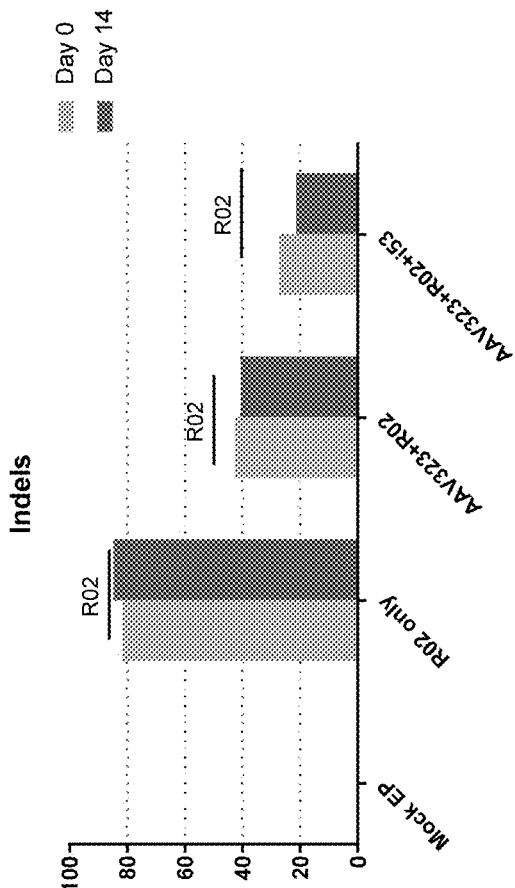
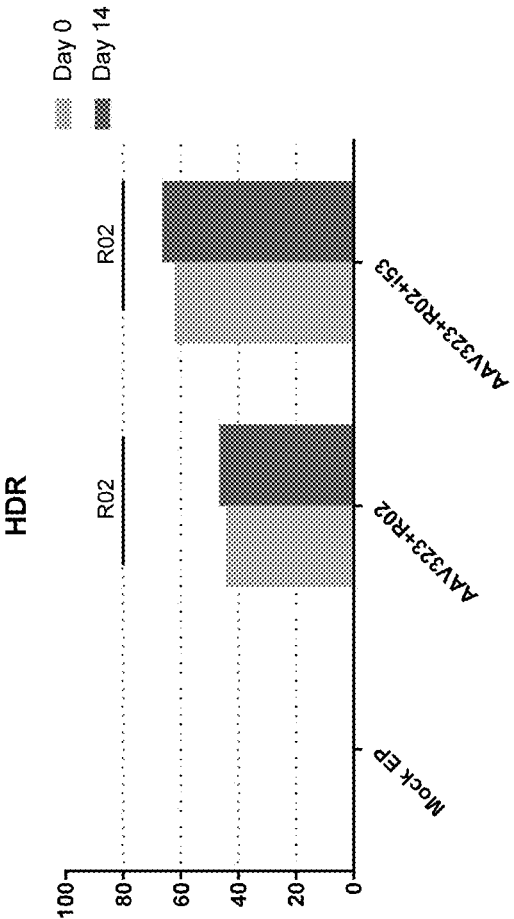


FIG. 34A

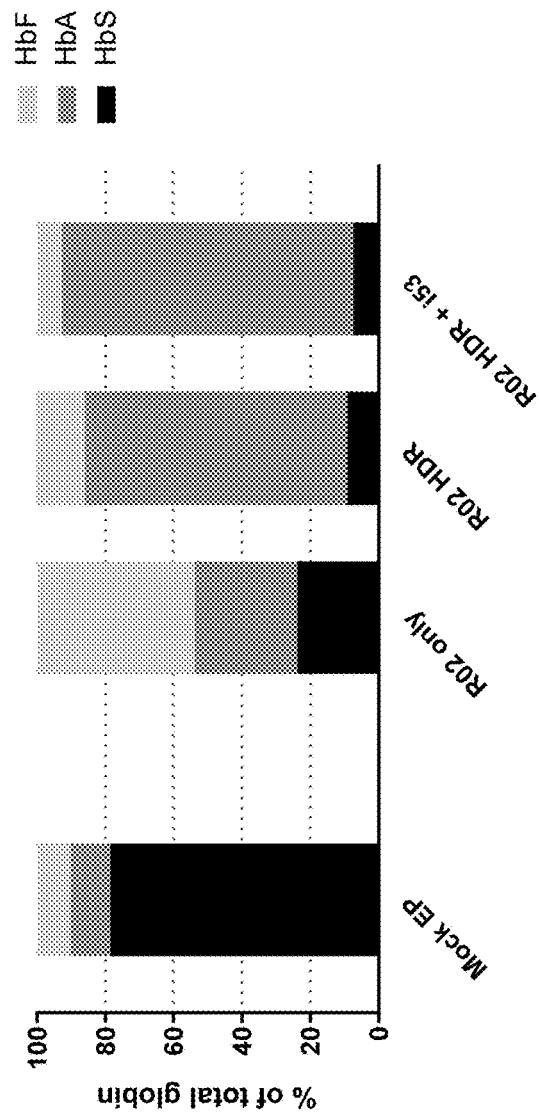


FIG. 34B

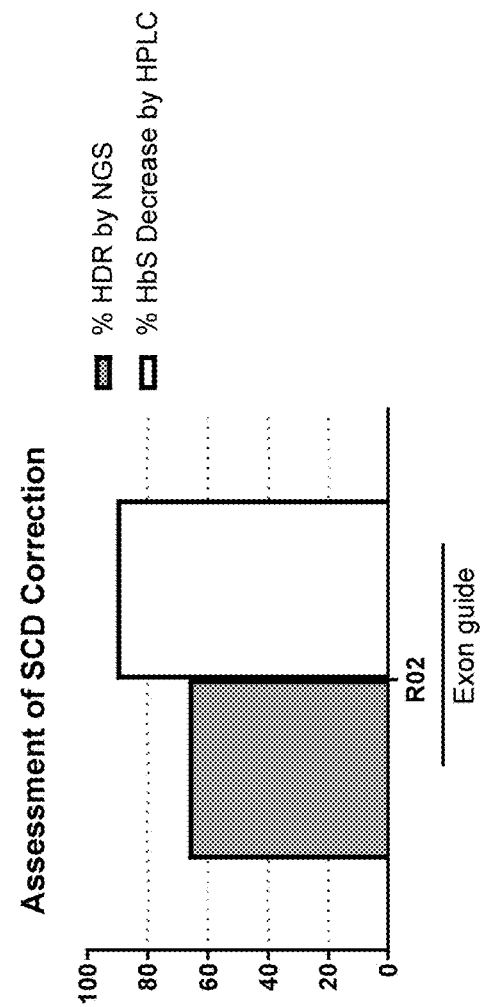


FIG. 35A

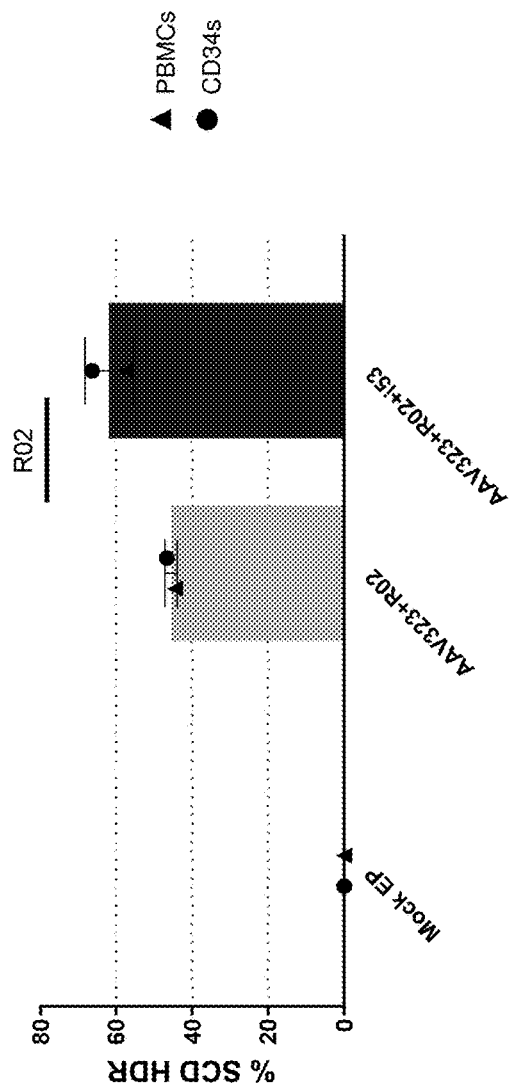


FIG. 35B

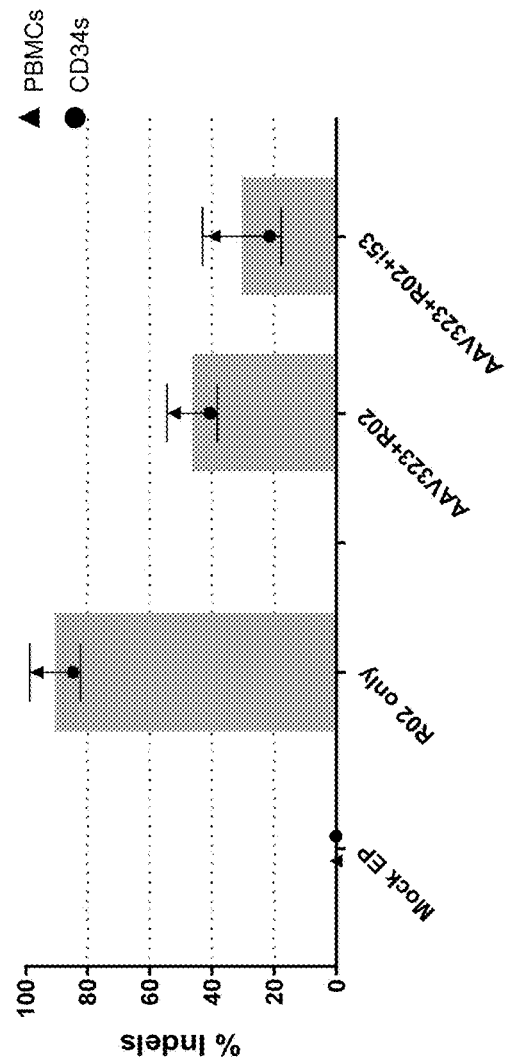
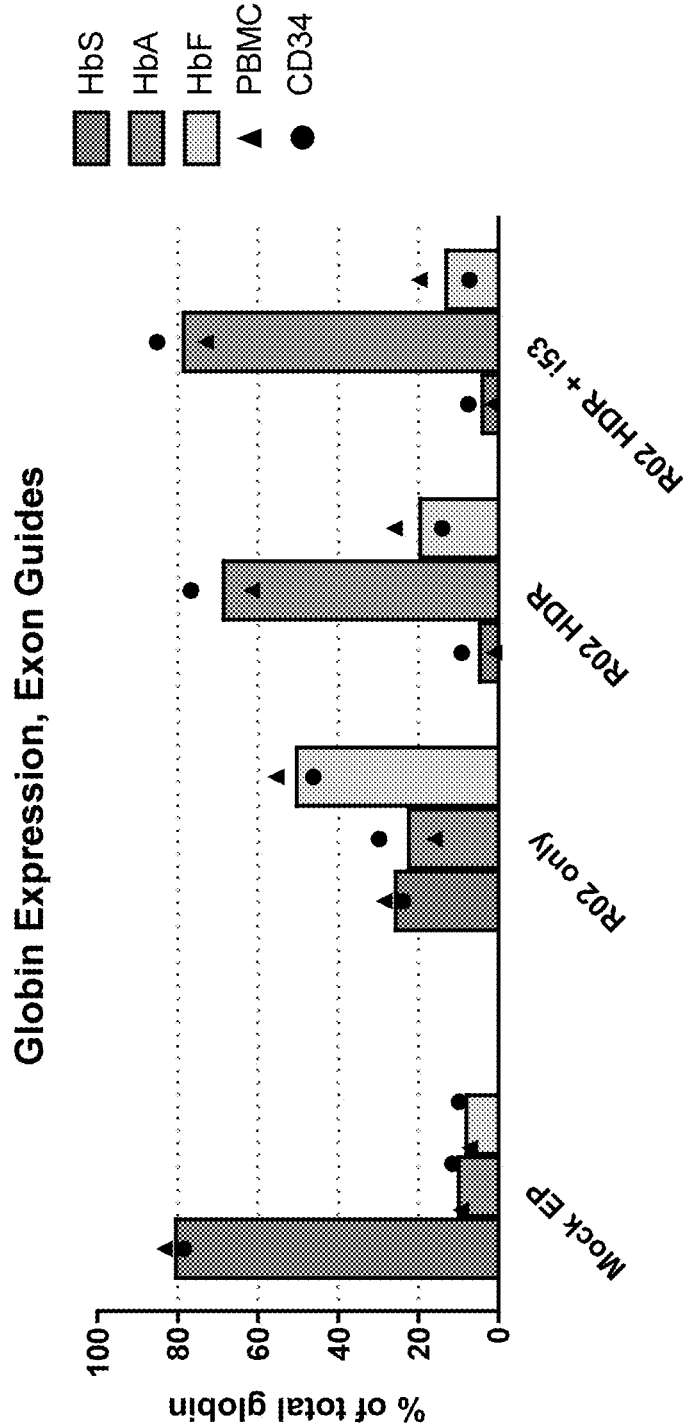


FIG. 36



METHODS AND COMPOSITIONS FOR IMPROVED HOMOLOGY DIRECTED REPAIR

RELATED APPLICATION

[0001] This application claims the benefit of U.S. Provisional Application Ser. No. 62/862,539, filed on Jun. 17, 2019, which is hereby incorporated by reference in its entirety.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format via EFS-Web, and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Oct. 2, 2020, is named CRTN_115A_Sequence_Listing.txt and is 114123 bytes in size.

BACKGROUND

[0003] The dominant pathways for repair of a DNA double-strand break (DSB) are the non-homologous end joining (NHEJ) pathway and the homology-directed repair (HDR) pathway (e.g., also known as homologous recombination or HR). The NHEJ pathway results in direct ligation of broken DNA ends and is error-prone, often introducing small insertions or deletions (e.g., indels) at the site of a DNA DSB. By contrast, the HDR pathway is a high-fidelity repair pathway that uses a homologous DNA sequence, such as that present on an exogenous donor polynucleotide, to repair a DNA DSB, thus enabling precise reconstruction of a broken DNA duplex according to a sequence provided by a homology donor DNA.

[0004] Selection of repair pathway is largely determined by mutual antagonism between p53-binding protein 1 (53BP1), a factor that favors the NHEJ pathway, and BRCA1, a factor that favors the HDR pathways and is associated with genomic instability in certain cancers. 53BP1 is recruited to DSBs by recognition of damaged chromatin where it functions to antagonize the formation of 3' single-stranded DNA tails, the rate-limiting step in the initiation of HDR repair. Thus, 53BP1 favors the NHEJ pathway by inhibiting DNA end-resection and HDR.

[0005] By using a site-directed nuclease (e.g., a CRISPR/Cas9 endonuclease) to induce a site-specific DSB in a target gene, a precise gene edit can be introduced by HDR repair using a homology donor polynucleotide as a template. However, HDR repair is often inefficient, largely due to competition with the NHEJ pathway, and achieving high levels of genome editing can be difficult. This is particularly true in non-dividing cells (e.g., G0 or G1 phase cells) wherein the NHEJ pathway is dominant. Thus, methods are needed to improve the efficiency of HDR for targeted integration of a transgene or other polynucleotide sequence using a genome editing system.

SUMMARY OF THE DISCLOSURE

[0006] The present disclosure provides methods for increasing homology directed repair (HDR) of a double-strand break (DSB) in a target gene in a cell, the method comprising contacting a cell with a 53BP1 inhibitor, wherein the cell is a quiescent human cell that is induced to divide,

and wherein the DSB is mediated by a site-directed nuclease, thereby increasing HDR of the DSB in the target gene in cell.

[0007] In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a population of cells, the method comprising contacting the population of cells with a 53BP1 inhibitor, wherein the population of cells is a population of quiescent human cells that is induced to divide, and wherein DSB is mediated by a site-directed nuclease, thereby increasing HDR of the DSB in the target gene in the population of cells.

[0008] In some embodiments, the methods of the disclosure feature use of a 53BP1 inhibitor which comprises a 53BP1 binding polypeptide that inhibits 53BP1 recruitment to the DSB in the cell. In some embodiments, the methods of the disclosure feature use of a 53BP1 binding polypeptide comprising an amino acid sequence selected from a group consisting of: SEQ ID NO: 70, SEQ ID NO: 74, SEQ ID NO: 77, SEQ ID NO: 80, SEQ ID NO: 83, or SEQ ID NO: 86. In some embodiments, the 53BP1 binding polypeptide comprises the amino acid sequence of SEQ ID NO: 70.

[0009] In some embodiments, the methods of the disclosure feature use of a 53BP1 inhibitor which comprises a nucleic acid comprising a nucleotide sequence encoding a 53BP1 binding polypeptide that inhibits 53BP1 recruitment to the DSB site in the cell. In some embodiments, the nucleic acid comprising a nucleotide sequence encoding a 53BP1 binding polypeptide comprises a nucleotide sequence selected from a group consisting of: SEQ ID NO: 69, SEQ ID NO: 73, SEQ ID NO: 76, SEQ ID NO: 79, SEQ ID NO: 82, SEQ ID NO: 85, or SEQ ID NO: 88. In some embodiments, the nucleic acid comprising a nucleotide sequence encoding a 53BP1 binding polypeptide comprises the nucleotide sequence of SEQ ID NO: 69.

[0010] In some embodiments, the methods of the disclosure feature use of a 53BP1 inhibitor which comprises a vector comprising a nucleotide sequence encoding the 53BP1 binding polypeptide. In some embodiments, the vector comprises a nucleotide sequence selected from a group consisting of: SEQ ID NO: 68, SEQ ID NO: 72, SEQ ID NO: 75, SEQ ID NO: 78, SEQ ID NO: 81, SEQ ID NO: 84, or SEQ ID NO: 87. In some embodiments, the vector is an adeno associated viral vector (AAV).

[0011] In some embodiments, the methods of the disclosure feature use of a 53BP1 inhibitor which comprises a small interfering ribonucleic acid (siRNA) targeting 53BP1.

[0012] In some embodiments, the methods of the disclosure feature use of a 53BP1 inhibitor which comprises a genome-editing system comprising a site-directed nuclease that disrupts the gene encoding 53BP1.

[0013] In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, wherein the cell or population of cells is isolated from a tissue sample obtained from a human donor. In some embodiments, the tissue sample is a peripheral blood sample.

[0014] In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, wherein the cell or population of cells is a hematopoietic stem or progenitor cell (HSPC). In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, wherein the cell or population of cells is a long-term HSPC (LT-HSPC). In some embodiments, the

disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, wherein the cell or population of cells is a CD34 expressing cell. In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, wherein the cell or population of cells is a white blood cell. In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, wherein the cell or population of cells is in the G0 phase of the cell cycle.

[0015] In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, wherein the cell or population of cells is induced to divide by treatment with one or more extrinsic signals. In some embodiments, the treatment comprises contacting the cell with an extrinsic signal that stimulates a cell signaling receptor. In some embodiments, the extrinsic signal is a mitogen. In some embodiments, the extrinsic signal is an environmental factor.

[0016] In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, the method comprising contacting the cell or population of cells with a 53BP1 inhibitor, wherein the population of cells is a population of quiescent human cells that is induced to divide, and wherein DSB is mediated by a site-directed nuclease, wherein the site-directed nuclease is selected from a zinc-finger nuclease (ZFN), a TALEN, and a hybrid meganuclease-TALEN (megaTAL), thereby increasing HDR of the DSB in the target gene in the population of cells.

[0017] In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or population of cells, the method comprising contacting the cell or population of cells with a 53BP1 inhibitor, wherein the population of cells is a population of quiescent human cells that is induced to divide, and wherein DSB is mediated by a Cas9 DNA endonuclease and one or more guide RNAs (gRNA), thereby increasing HDR of the DSB in the target gene in the population of cells. In some embodiments, the DSB is mediated by a Cas9 DNA endonuclease and one or more guide RNAs (gRNA). In some embodiments, the Cas9 DNA endonuclease is *Streptococcus pyogenes* Cas9 (SpCas9).

[0018] In some embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a cell or a population of cells, the method comprising (i) contacting the cell or population of cells with a 53BP1 inhibitor, wherein the population of cells is a population of quiescent human cells that is induced to divide, and wherein DSB is mediated by a site-directed nuclease, and (ii) contacting the cell or population of cells with a donor polynucleotide comprising a nucleotide sequence that corrects or induces a mutation when incorporated at the site of the DSB, thereby increasing HDR of the DSB in the target gene in the population of cells. In some embodiments, the donor polynucleotide comprises a double-stranded DNA (dsDNA). In some embodiments, the donor polynucleotide comprises a single-stranded oligonucleotide (ssODN). In some embodiments, the donor polynucleotide comprises a mutation in a protospacer adjacent motif (PAM) sequence to prevent re-cutting of the target gene by a site-directed nuclease. In some embodiments, the donor polynucleotide comprises a vector comprising the donor polynucleotide. In some embodiments, the vector comprises a recombinant single-stranded

adeno-associated viral vector (rAAV). In some embodiments, the rAAV is serotype 6 (AAV6).

[0019] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the cell or population of cells with an inhibitor of DNA-PK, thereby increasing HDR of the DSB in the target gene in the population of cells. In some embodiments, the inhibitor of DNA-PK is Nu7441.

[0020] In any of the foregoing or related methods, the disclosure provides a method wherein the cell or population of cells is induced to divide prior to contact with the 53BP1 inhibitor. In some embodiments, the cell or population of cells is contacted with the 53BP1 inhibitor prior to being induced to divide. In some embodiments, the cell or population of cells is contacted with the 53BP1 inhibitor concurrent with being induced to divide.

[0021] In any of the foregoing or related methods, the disclosure provides a method wherein the cell or population of cells is contacted ex vivo and the method further comprises administering the cell or population of cells to a subject in need thereof.

[0022] In any of the foregoing or related methods, the disclosure provides a method wherein the cell or population of cells is contacted with the 53BP1 inhibitor in vivo.

[0023] In some embodiments, the disclosure provides a method for increasing HDR in a target gene in a population of cells, the method comprising contacting the population of cells with a 53BP1 inhibitor, wherein the inhibitor is a 53BP1 binding polypeptide or a nucleic acid comprising a nucleotide sequence encoding a 53BP1 binding polypeptide, wherein the 53BP1 binding polypeptide comprises an amino acid sequence selected from a group consisting of: SEQ ID NO: 70, SEQ ID NO: 74, SEQ ID NO: 77, SEQ ID NO: 80, SEQ ID NO: 83, or SEQ ID NO: 86, wherein the population of cells are quiescent human cells that are induced to divide, and wherein the HDR is mediated by a site-directed nuclease, thereby increasing HDR in the target gene in the population of cells. In some embodiments, the nucleotide sequence encoding the 53BP1 binding polypeptide is selected from a group consisting of: SEQ ID NO: 69, SEQ ID NO: 73, SEQ ID NO: 76, SEQ ID NO: 79, SEQ ID NO: 82, SEQ ID NO: 85, or SEQ ID NO: 88. In some embodiments, the nucleic acid comprises a vector comprising a nucleotide sequence encoding the 53BP1 binding polypeptide. In some embodiments, the vector comprises a nucleotide sequence selected from a group consisting of: SEQ ID NO: 68, SEQ ID NO: 72, SEQ ID NO: 75, SEQ ID NO: 78, SEQ ID NO: 81, SEQ ID NO: 84, or SEQ ID NO: 87. In some embodiments, the vector is an adeno associated viral vector (AAV).

[0024] In some embodiments, the disclosure provides a method for increasing HDR in a target gene in a population of cells, the method comprising (i) contacting the population of cells with a 53BP1 inhibitor, wherein the inhibitor is a 53BP1 binding polypeptide or a nucleic acid comprising a nucleotide sequence encoding a 53BP1 binding polypeptide, wherein the 53BP1 binding polypeptide comprises an amino acid sequence selected from a group consisting of: SEQ ID NO: 70, SEQ ID NO: 74, SEQ ID NO: 77, SEQ ID NO: 80, SEQ ID NO: 83, or SEQ ID NO: 86, wherein the population of cells are quiescent human cells that are induced to divide, and (ii) contacting the population of cells with a site-directed nuclease that mediates a DSB in a target gene in the cell, thereby increasing HDR in the target gene in the population

of cells. In some embodiments, the site-directed nuclease is a Cas9 DNA endonuclease. In some embodiments, the method further comprises contacting the population of cells with one or more gRNAs to effect a DSB within a target gene in the cell. In some embodiments, the one or more gRNAs is a crRNA and a tracrRNA, a single-molecule gRNA (sgRNA), or a combination thereof. In some embodiments, the Cas9 DNA endonuclease is a polypeptide. In some embodiments, the Cas9 DNA endonuclease is pre-complexed with one or more gRNAs. In some embodiments, the Cas9 and one or more gRNAs are electroporated into the cell.

[0025] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the population of cells with a donor polynucleotide comprising a nucleotide sequence that corrects or induces a mutation when incorporated at the site of the DSB. In some embodiments, the donor polynucleotide comprises a double-stranded DNA (dsDNA). In some embodiments, the donor polynucleotide comprises a single-stranded oligonucleotide (ssODN). In some embodiments, the donor polynucleotide comprises a mutation in a protospacer adjacent motif (PAM) sequence to prevent re-cutting of the target gene by a site-directed nuclease. In some embodiments, the donor polynucleotide comprises a vector comprising the donor polynucleotide. In some embodiments, the vector comprises a rAAV. In some embodiments, the rAAV is AAV6.

[0026] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the population of cells with an inhibitor of DNA-PK. In some embodiments, the inhibitor of DNA-PK is Nu7441.

[0027] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is induced to divide prior to contact with the 53BP1 inhibitor.

[0028] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted with the 53BP1 inhibitor prior to being induced to divide.

[0029] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted with the 53BP1 inhibitor concurrent with being induced to divide.

[0030] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted ex vivo and the method further comprises administering the cell to a subject in need thereof.

[0031] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted with the 53BP1 inhibitor in vivo.

[0032] In other embodiments, the disclosure provides methods for increasing HDR of a DSB in a target gene in a population of human CD34+ cells, the method comprising contacting the population of human CD34+ cells with a 53BP1 inhibitor, wherein the inhibitor is a 53BP1 binding polypeptide that inhibits 53BP1 recruitment to the DSB site in the cells, and wherein the DSB is mediated by a site-directed nuclease, thereby increasing HDR of the DSB in the target gene in the population of human CD34+ cells. In some embodiments, the 53BP1 binding polypeptide comprises an amino acid sequence selected from a group consisting of: SEQ ID NO: 70, SEQ ID NO: 74, SEQ ID NO: 77, SEQ ID NO: 80, SEQ ID NO: 83, or SEQ ID NO: 86. In some embodiments, the 53BP1 binding polypeptide comprises the

amino acid sequence of SEQ ID NO: 70. In some embodiments, the 53BP1 inhibitor comprises a nucleic acid comprising a nucleotide sequence encoding the 53BP1 binding polypeptide. In some embodiments, the nucleic acid comprises a nucleotide sequence selected from a group consisting of: SEQ ID NO: 69, SEQ ID NO: 73, SEQ ID NO: 76, SEQ ID NO: 79, SEQ ID NO: 82, SEQ ID NO: 85, or SEQ ID NO: 88. In some embodiments, the nucleic acid comprises the nucleotide sequence of SEQ ID NO: 69. In some embodiments, the nucleic acid comprises a vector comprising a nucleotide sequence encoding the 53BP1 binding polypeptide. In some embodiments, the nucleic acid comprises a nucleotide sequence selected from a group consisting of: SEQ ID NO: 68, SEQ ID NO: 72, SEQ ID NO: 75, SEQ ID NO: 78, SEQ ID NO: 81, SEQ ID NO: 84, or SEQ ID NO: 412. In some embodiments, the nucleic acid comprises the nucleotide sequence of SEQ ID NO: 68. In some embodiments, the vector is a rAAV. In some embodiments, the rAAV is AAV6.

[0033] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells comprises HSPCs. In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells comprises LT-HSPCs. In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is isolated from a tissue sample obtained from a human donor. In some embodiments, the tissue sample is a peripheral blood sample. In some embodiments, the tissue sample is obtained from a human donor administered an HSPC mobilizing agent prior to obtaining the tissue sample. In some embodiments, the HSPC mobilizing agent is Plurixifor. In some embodiments, the tissue sample is obtained from a human donor administered a combination of HSPC mobilizing agents prior to obtaining the tissue sample. In some embodiments, the combination of HSPC mobilizing agents comprises Plerixafor and granulocyte colony stimulating factor (G-CSF).

[0034] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is induced to divide by treatment with one or more extrinsic signals selected from a group consisting of: cytokines, mitogens and environmental factors. In some embodiments, the extrinsic signal is a cytokine or one or more cytokines. In some embodiments, the cytokine is interleukin-3 (IL-3). In some embodiments, the cytokine is stem cell factor (SCF) or IL-3 and a SCF. In some embodiments, the cytokine is Fms-like tyrosine kinase 3 (Flt3) ligand or IL-3, SCF and Flt3. In some embodiments, the cytokine is thrombopoietin or IL-3, SCF, Flt3 and thrombopoietin. In some embodiments, the extrinsic signal is a mitogen. In some embodiments, the extrinsic signal is an environmental factor. In some embodiments, the environmental factor is hypoxia.

[0035] In any of the foregoing or related methods, the disclosure provides a method wherein the DSB is mediated by a Cas9 DNA endonuclease and one or more gRNAs. In some embodiments, the Cas9 DNA endonuclease is SpCas9.

[0036] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the population of cells with a donor polynucleotide comprising a nucleotide sequence that corrects or induces a mutation when incorporated at the site of the DSB. In some embodiments, the donor polynucleotide comprises a nucleotide sequence that corrects a disease-causing single nucleotide polymorphism (SNP). In some embodiments, the donor

polynucleotide comprises a dsDNA. In some embodiments, the donor polynucleotide comprises a ssODN. In some embodiments, the donor polynucleotide comprises a mutation in a PAM sequence to prevent re-cutting of the target gene by a site-directed nuclease. In some embodiments, the donor polynucleotide comprises a vector comprising the donor polynucleotide. In some embodiments, the vector comprises a rAAV. In some embodiments, the vector comprises AAV6.

[0037] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the population of cells with an inhibitor of DNA-PK. In some embodiments, the inhibitor of DNA-PK is Nu7441.

[0038] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is induced to divide prior to contact with the 53BP1 inhibitor.

[0039] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted with the 53BP1 inhibitor prior to being induced to divide.

[0040] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted with the 53BP1 inhibitor concurrent with being induced to divide.

[0041] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted ex vivo and the method further comprises administering the cell to a subject in need thereof.

[0042] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted with the 53BP1 inhibitor in vivo.

[0043] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in an increase in HDR frequency in the cell population to 50% or more. In some embodiments, contacting the population of cells with the 53BP1 inhibitor results in an increase in HDR frequency in the cell population of 65% or more.

[0044] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in a 1-2 fold increase in HDR frequency in the cell population. In some embodiments, contacting the population of cells with the 53BP1 inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population.

[0045] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in a 1-2 fold decrease in indel frequency in the cell population. In some embodiments, contacting the population of cells with the 53BP1 inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold decrease in indel frequency in the cell population. In some embodiments, contacting the population of cells with the 53BP1 inhibitor results in a 2-3 fold, 2-4 fold, 2-5 fold, 2-6 fold, 2-7 fold, 2-8 fold, 2-9 fold, or 2-10 fold decrease in indel frequency in the cell population.

[0046] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in an increase in engraftment in vivo following administration.

[0047] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in an increase in chimerism in vivo following administration.

[0048] In other embodiments, the disclosure provides a method for increasing HDR of a DSB in a target gene in a population of human CD34+ cells, the method comprising contacting the population of human CD34+ cells with (i) an inhibitor of 53BP1, wherein the inhibitor is nucleic acid comprising a nucleotide sequence encoding a 53BP1 binding polypeptide comprising the amino acid sequence of SEQ ID NO: 70, and (ii) a Cas9 DNA endonuclease and one or more gRNAs to effect a DSB within the target gene in the population of cells, thereby increasing HDR of a DSB in a target gene in a population of human CD34+ cells. In some embodiments, the nucleic acid comprising a nucleotide sequence encoding the 53BP1 binding polypeptide is an RNA. In some embodiments, the nucleic acid is a vector. In some embodiments, the vector is an AAV. In some embodiments, the AAV is an AAV of the DJ serotype (AAV-DJ). In some embodiments, the one or more gRNAs is a crisprRNA and a tracrRNA, a single-molecule gRNA (sgRNA), or a combination thereof. In some embodiments, the Cas9 DNA endonuclease is a polypeptide. In some embodiments, the Cas9 DNA endonuclease is pre-complexed with one or more gRNAs. In some embodiments, the Cas9 DNA endonuclease and one or more gRNAs are electroporated into the cells. In some embodiments, the cell population is contacted with the nucleic acid encoding the 53BP1 polypeptide prior to electroporation. In some embodiments, the cell population is contacted with the nucleic acid encoding the 53BP1 polypeptide during electroporation. In some embodiments, the cell population is contacted with the nucleic acid encoding the 53BP1 polypeptide subsequent to electroporation.

[0049] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the population of cells with a donor polynucleotide comprising a nucleotide sequence that corrects or induces a mutation when incorporated at the site of the DSB. In some embodiments, the donor polynucleotide comprises a nucleotide sequence that corrects a disease-causing single nucleotide polymorphism (SNP). In some embodiments, the donor polynucleotide comprises a dsDNA. In some embodiments, the donor polynucleotide comprises a ssODN. In some embodiments, the donor polynucleotide comprises a mutation in a PAM sequence to prevent re-cutting of the target gene by a site-directed nuclease. In some embodiments, the donor polynucleotide comprises a vector comprising the donor polynucleotide. In some embodiments, the vector comprises rAAV. In some embodiments, the rAAV is AAV6.

[0050] In any of the foregoing or related methods, the disclosure provides a method wherein the donor polynucleotide is administered to the cells prior to electroporation. In any of the foregoing or related methods, the disclosure provides a method wherein the donor polynucleotide is administered to cells following electroporation.

[0051] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the population of cells with an inhibitor of DNA-PK. In some embodiments, the inhibitor of DNA-PK is Nu7441.

[0052] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is induced to divide by treatment with one or more extrinsic signals selected from a group consisting of: cytokines,

mitogens and environmental factors. In some embodiments, the extrinsic signal is a cytokine or one or more cytokines. In some embodiments, the cytokine is interleukin-3 (IL-3). In some embodiments, the cytokine is stem cell factor (SCF) or IL-3 and a SCF. In some embodiments, the cytokine is Fms-like tyrosine kinase 3 (Flt3) ligand or IL-3, SCF and Flt3. In some embodiments, the cytokine is thrombopoietin or IL-3, SCF, Flt3 and thrombopoietin. In some embodiments, the cells are cultured with a combination of cytokines comprising thrombopoietin, Flt3 ligand, SCF, and IL-3. In some embodiments, the extrinsic signal is a mitogen. In some embodiments, the extrinsic signal is an environmental factor. In some embodiments, the environmental factor is hypoxia. In some embodiments, the cells are cultured under hypoxic conditions comprising atmospheric oxygen content less than 5%.

[0053] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted ex vivo and the method further comprises administering the cell to a subject in need thereof.

[0054] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is contacted with the 53BP1 inhibitor in vivo.

[0055] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in an increase in HDR frequency in the cell population to 50% or more. In some embodiments, contacting the population of cells with the 53BP1 inhibitor results in an increase in HDR frequency in the cell population of 65% or more.

[0056] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in a 1-2 fold increase in HDR frequency in the cell population. In some embodiments, contacting the population of cells with the 53BP1 inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population. In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in a 1-2 fold decrease in indel frequency in the cell population. In some embodiments, contacting the population of cells with the 53BP1 inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold decrease in indel frequency in the cell population. In some embodiments, contacting the population of cells with the 53BP1 inhibitor results in a 2-3 fold, 2-4 fold, 2-5 fold, 2-6 fold, 2-7 fold, 2-8 fold, 2-9 fold, or 2-10 fold decrease in indel frequency in the cell population.

[0057] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in an increase in engraftment in vivo following administration.

[0058] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the 53BP1 inhibitor results in an increase in chimerism in vivo following administration.

[0059] In other embodiments, the disclosure provides a method for increasing HDR of a DSB in a target gene in a population of human CD34+ cells, the method comprising contacting the population of human CD34+ cells with an effective amount of an inhibitor of the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) DNA-PKcs

inhibitor, and wherein the DSB is mediated by a site-directed nuclease, thereby increasing HDR of the DSB in the target gene in the population of human CD34+ cells. In some embodiments, the DNA-PKcs inhibitor inhibits the kinase activity of DNA-PKcs. In some embodiments, the DNA-PKcs inhibitor inhibits repair of the DSB in a target gene by the NHEJ repair pathway. In some embodiments, the DNA-PKcs inhibitor increases repair of the DSB in a target gene by the HDR repair pathway. In some embodiments, the DNA-PKcs inhibitor is a small molecule. In some embodiments, the DNA-PKcs inhibitor is Nu7441.

[0060] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells comprises LT-HSCs. In some embodiments, the population of cells is isolated from a tissue sample obtained from a human donor. In some embodiments, the tissue sample is a peripheral blood sample. In some embodiments, the tissue sample is obtained from a human donor administered an HSPC mobilizing agent prior to obtaining the tissue sample. In some embodiments, the HSPC mobilizing agent is Plurixifor. In some embodiments, the tissue sample is obtained from a human donor administered a combination of HSPC mobilizing agents prior to obtaining the tissue sample. In some embodiments, the combination of HSPC mobilizing agents comprises Plerixafor and granulocyte colony stimulating factor (G-CSF).

[0061] In any of the foregoing or related methods, the disclosure provides a method wherein the population of cells is induced to divide by treatment with one or more extrinsic signals selected from a group consisting of: cytokines, mitogens and environmental factors. In some embodiments, the extrinsic signal is a cytokine or one or more cytokines. In some embodiments, the cytokine is interleukin-3 (IL-3). In some embodiments, the cytokine is stem cell factor (SCF) or IL-3 and a SCF. In some embodiments, the cytokine is Fms-like tyrosine kinase 3 (Flt3) ligand or IL-3, SCF and Flt3. In some embodiments, the cytokine is thrombopoietin or IL-3, SCF, Flt3 and thrombopoietin. In some embodiments, the cells are cultured with a combination of cytokines comprising thrombopoietin, Flt3 ligand, SCF, and IL-3. In some embodiments, the extrinsic signal is a mitogen. In some embodiments, the extrinsic signal is an environmental factor. In some embodiments, the environmental factor is hypoxia. In some embodiments, the cells are cultured under hypoxic conditions comprising atmospheric oxygen content less than 5%.

[0062] In other embodiments, the disclosure provides a method for increasing HDR of a DSB in a target gene in a population of human CD34+ cells, the method comprising contacting the population of human CD34+ cells with (i) a DNA-PKcs inhibitor, and (ii) a Cas9 DNA endonuclease and one or more gRNAs to effect a DSB within the target gene in the population of cells, thereby increasing HDR of a DSB in a target gene in a population of human CD34+ cells.

[0063] In some embodiments, the DNA-PKcs inhibitor inhibits the kinase activity of DNA-PKcs. In some embodiments, the DNA-PKcs inhibitor inhibits repair of the DSB in a target gene by the NHEJ repair pathway. In some embodiments, the DNA-PKcs inhibitor increases repair of the DSB in a target gene by the HDR repair pathway. In some embodiments, the DNA-PKcs inhibitor is a small molecule. In some embodiments, the DNA-PKcs inhibitor is Nu7441. In some embodiments, the one or more gRNAs is a crRNA and a tracrRNA, a single-molecule gRNA (sgRNA),

or a combination thereof. In some embodiments, the Cas9 DNA endonuclease is a polypeptide. In some embodiments, the Cas9 DNA endonuclease is pre-complexed with one or more gRNAs. In some embodiments, the Cas9 DNA endonuclease and one or more gRNAs are electroporated into the cells. In some embodiments, the cell population is contacted with the DNA-PKcs inhibitor prior to electroporation. In some embodiments, the cell population is contacted with the DNA-PKcs inhibitor during electroporation. In some embodiments, the cell population is contacted with the DNA-PKcs inhibitor subsequent to electroporation.

[0064] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the population of cells with a donor polynucleotide comprising a nucleotide sequence that corrects or induces a mutation when incorporated at the site of the DSB. In some embodiments, the donor polynucleotide comprises a nucleotide sequence that corrects a disease-causing single nucleotide polymorphism (SNP). In some embodiments, the donor polynucleotide comprises a dsDNA. In some embodiments, the donor polynucleotide comprises a ssODN. In some embodiments, the donor polynucleotide comprises a mutation in a PAM sequence to prevent re-cutting of the target gene by a site-directed nuclease. In some embodiments, the donor polynucleotide comprises a vector comprising the donor polynucleotide. In some embodiments, the vector comprises rAAV. In some embodiments, the rAAV is AAV6.

[0065] In any of the foregoing or related methods, the disclosure provides a method wherein the donor polynucleotide is administered to the cells prior to electroporation. In any of the foregoing or related methods, the disclosure provides a method wherein the donor polynucleotide is administered to cells following electroporation.

[0066] In any of the foregoing or related methods, the disclosure provides a method wherein the population of human CD34+ cells is induced to divide prior to contact with the DNA-PKcs inhibitor. In any of the foregoing or related methods, the disclosure provides a method wherein the population of human CD34+ cells is contacted with the DNA-PKcs inhibitor prior to being induced to divide. In any of the foregoing or related methods, the disclosure provides a method wherein the population of human CD34+ cells is contacted with the DNA-PKcs inhibitor concurrent with being induced to divide.

[0067] In any of the foregoing or related methods, the disclosure provides a method wherein the population of human CD34+ cells is contacted ex vivo and the method further comprises administering the cells to a subject in need thereof.

[0068] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in an increase in HDR frequency in the cell population to 50% or more.

[0069] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in a 1-2 fold increase in HDR frequency in the cell population. In some embodiments, contacting the population of cells with the DNA-PKcs inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population.

[0070] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in a 2-10 fold decrease in indel frequency in the cell population.

[0071] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in an increase in engraftment in vivo following administration.

[0072] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in an increase in chimerism in vivo following administration.

[0073] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in at least 20% of bone marrow cells in a patient comprising ex vivo edited cells at one week or longer following administration.

[0074] In other embodiments, the disclosure provides a method for increasing HDR of a DSB in a target gene in a population of human CD34+ cells, the method comprising contacting the population of human CD34+ cells with (i) an inhibitor of DNA-PKcs, wherein the inhibitor is a small molecule inhibitor that inhibits the kinase activity of DNA-PKcs; and (ii) a Cas9 DNA endonuclease and one or more gRNAs to effect a DSB within the target gene in the population of cells, thereby increasing HDR of a DSB in a target gene in a population of human CD34+ cells. In some embodiments the DNA-PKcs inhibitor is Nu7441. In some embodiments, the one or more gRNAs is a crisprRNA and a tracrRNA, a single-molecule gRNA (sgRNA), or a combination thereof. In some embodiments, the Cas9 DNA endonuclease is a polypeptide. In some embodiments, the Cas9 DNA endonuclease is pre-complexed with one or more gRNAs. In some embodiments, the Cas9 DNA endonuclease and one or more gRNAs are electroporated into the cells. In some embodiments, wherein the cell population is contacted with small molecule inhibitor of DNA-PKcs prior to electroporation. In some embodiments, the cell population is contacted with the small molecule inhibitor of DNA-PKcs during electroporation. In some embodiments, the cell population is contacted with the small molecule inhibitor of DNA-PKcs subsequent to electroporation.

[0075] In any of the foregoing or related methods, the disclosure provides a method further comprising contacting the population of cells with a donor polynucleotide comprising a nucleotide sequence that corrects or induces a mutation when incorporated at the site of the DSB. In some embodiments, the donor polynucleotide comprises a nucleotide sequence that corrects a disease-causing single nucleotide polymorphism (SNP). In some embodiments, the donor polynucleotide comprises a dsDNA. In some embodiments, the donor polynucleotide comprises a ssODN. In some embodiments, the donor polynucleotide comprises a mutation in a PAM sequence to prevent re-cutting of the target gene by a site-directed nuclease. In some embodiments, the donor polynucleotide comprises a vector comprising the donor polynucleotide. In some embodiments, the vector comprises rAAV. In some embodiments, the rAAV is AAV6.

[0076] In any of the foregoing or related methods, the disclosure provides a method wherein the donor polynucleotide is administered to the cells prior to electroporation. In any of the foregoing or related methods, the disclosure provides a method wherein the donor polynucleotide is administered to cells following electroporation.

[0077] In any of the foregoing or related methods, the disclosure provides a method wherein the population of human CD34+ cells is induced to divide prior to contact with the DNA-PKcs inhibitor. In any of the foregoing or related methods, the disclosure provides a method wherein the population of human CD34+ cells is contacted with the DNA-PKcs inhibitor prior to being induced to divide. In any of the foregoing or related methods, the disclosure provides a method wherein the population of human CD34+ cells is contacted with the DNA-PKcs inhibitor concurrent with being induced to divide.

[0078] In any of the foregoing or related methods, the disclosure provides a method wherein the population of human CD34+ cells is contacted ex vivo and the method further comprises administering the cells to a subject in need thereof.

[0079] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in an increase in HDR frequency in the cell population to 50% or more.

[0080] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in a 1-2 fold increase in HDR frequency in the cell population. In some embodiments, contacting the population of cells with the DNA-PKcs inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population.

[0081] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in a 2-10 fold decrease in indel frequency in the cell population.

[0082] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in an increase in engraftment in vivo following administration.

[0083] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in an increase in chimerism in vivo following administration.

[0084] In any of the foregoing or related methods, the disclosure provides a method wherein contacting the population of cells with the DNA-PKcs inhibitor results in at least 20% of bone marrow cells in a patient comprising ex vivo edited cells at one week or longer following administration.

[0085] In other embodiments, the disclosure provides a population of cells produced by a method of the disclosure, e.g., a population of human CD34+ cells. In other embodiments, the disclosure provides a population of cells comprising LT-HSCs produced by a method of the disclosure.

[0086] In other embodiments, the disclosure provides a population of human CD34+ cells characterized by increased HDR repair of a DSB in a target gene mediated by a site-directed nuclease, wherein the population of CD34+ cells have increased HDR frequency of 50% or more as a result of contact with a 53BP1 inhibitor or a DNA-PKcs inhibitor, or a combination of a 53BP1 inhibitor and a DNA-PKcs inhibitor. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in a 1-2 fold increase in HDR frequency in the cell population. In some embodiments, contact with the 53BP1

inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in a 1-2 fold increase in HDR frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor results in a 1-2 fold decrease in indel frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold decrease in indel frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor results in a 2-3 fold, 2-4 fold, 2-5 fold, 2-6 fold, 2-7 fold, 2-8 fold, 2-9 fold, or 2-10 fold decrease in indel frequency in the cell population. In some embodiments, contact with the DNA-PKcs inhibitor results in a 2-10 fold decrease in indel frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in an increase in engraftment in vivo following administration. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in an increase in chimerism in vivo following administration.

[0087] In other embodiments, the disclosure provides a population of human LT-HSCs characterized by increased HDR repair of a DSB in a target gene mediated by a site-directed nuclease, wherein the population of human LT-HSCs have increased HDR frequency of 50% or more as a result of contact with a 53BP1 inhibitor or a DNA-PKcs inhibitor, or a combination of a 53BP1 inhibitor and a DNA-PKcs inhibitor. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in a 1-2 fold increase in HDR frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in a 1-2 fold increase in HDR frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor results in a 1-2 fold decrease in indel frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold decrease in indel frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor results in a 2-3 fold, 2-4 fold, 2-5 fold, 2-6 fold, 2-7 fold, 2-8 fold,

2-9 fold, or 2-10 fold decrease in indel frequency in the cell population. In some embodiments, contact with the DNA-PKcs inhibitor results in a 2-10 fold decrease in indel frequency in the cell population. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in an increase in engraftment in vivo following administration. In some embodiments, contact with the 53BP1 inhibitor or the DNA-PKcs inhibitor, or a combination of the 53BP1 inhibitor and the DNA-PKcs inhibitor results in an increase in chimerism in vivo following administration.

[0088] In other embodiments, the disclosure provides use of a population of cells of the disclosure, and an optional pharmaceutically acceptable carrier, for treating or delaying progression of a disease caused by a SNP in a subject in need thereof.

[0089] In other embodiments, the disclosure provides a kit comprising a container comprising a population of cells of the disclosure, and an optional pharmaceutically acceptable carrier, and a package insert comprising instructions for administration of the population of cells for treating or delaying progression of a disease caused by a SNP in an individual.

BRIEF DESCRIPTION OF THE DRAWINGS

[0090] FIGS. 1A-1D include bar graphs showing efficiency of HDR repair in HEK293 T cells using single-stranded oligodeoxynucleotide (ssODN) donor DNA that converts a gene in the AAVS1 locus encoding a blue fluorescent protein (BFP) to a gene encoding green fluorescent protein (GFP). FIG. 1A shows HDR efficiency in the presence of Nu7441 (e.g., an inhibitor of DNA-PKcs), SCR7 (e.g., an inhibitor of DNA Ligase IV), and RS1 (e.g., an agonist of Rad51). FIG. 1B shows HDR efficiency in the presence of Nu7441 or Veliparib (e.g., an inhibitor of PARP) with varied doses of inhibitor. FIG. 1C shows HDR efficiency in the presence of Nu7441 or L755,507 (e.g., an inhibitor of 33-adrenergic receptor) using two different ssODN templates with varied doses of inhibitor. FIG. 1D shows HDR efficiency in the presence of the i53 polypeptide inhibitor of 53BP1 at varied doses using two different ssODN donors.

[0091] FIGS. 2A-2D include bar graphs showing editing in HEK293 T cells following electroporation with Cas9/sgRNA RNP using single-stranded oligodeoxynucleotide (ssODN) donor DNA that converts a gene in the AAVS1 locus encoding GFP to a gene encoding BFP in the presence of Nu7441, SCR7 or RS1. FIGS. 2A-2B show the efficiency of HDR repair to convert GFP to BFP. FIGS. 2C-2D show indel formation in the AAVS1 locus.

[0092] FIG. 3 includes a bar graph showing efficiency of gene insertion into the GSD1a locus in HEK293 T cells using either ssODNs as homology donors that facilitate HDR or dsDNA donors that facilitate NHEJ repair. Repair efficiency was evaluated in the presence of Nu7441, SCR7, or RS-1 using two different ssODN donor templates and two different dsDNA donor templates.

[0093] FIGS. 4A-4B include bar graphs showing mutations at the site of a DSB induced by Cas9/gRNA in the CFTR locus in HEK293 T cells resulting from DSB repair in the presence of a donor ssODN only (FIG. 4A) or donor ssODN and the DNA-PK inhibitor Nu7441 (FIG. 4B).

[0094] FIG. 5 include a bar graphs showing mutations at the site of DSB induced by Cas9/gRNA in the CFTR locus in HEK293 T cells resulting from DSB repair in the presence of donor ssODN H3-95-30 (SEQ ID NO: 41) or donor ssODN N1-95-30 (SEQ ID NO: 42) with treatment of Nu7441. Control cells are electroporated in the absence of gene-editing components or Nu7441 ("mock+DMSO").

[0095] FIGS. 6A-6C include bar graphs showing HDR editing efficiency for insertion of donor DNA encoding GFP into the hemoglobin subunit beta (HBB) locus of CD34-expressing long-term repopulating hematopoietic stem cells (LT-HSPCs) using AAV-mediated delivery of donor DNA encoding GFP. FIG. 6A shows HDR efficiency in the presence of different doses of mRNA encoding i53 (e.g., inhibitor of 53BP1) relative to negative controls that include mock electroporation (EP), AAV donor DNA alone or RNP-only (i.e., no AAV donor DNA). FIG. 6B shows HDR efficiency in the presence of different doses of mRNA encoding i53, Cyren1 (e.g., inhibitor of Ku70/80), or Cyren2 (e.g., inhibitor of Ku70/80) relative to negative controls that include AAV alone (i.e., no RNP). FIG. 6C shows HDR efficiency in the presence of varied doses of Nu7441 relative to a DMSO-only control, mRNA encoding i53, or a control mRNA (DM) (i.e., absence of a modulator of DNA repair).

[0096] FIG. 7 includes a dot plot showing HDR editing efficiency with treatment of i53 for insertion of donor DNA encoding GFP delivered by AAV into the AAVS1 locus of hTERT RPE-1 cells.

[0097] FIG. 8 includes a schematic showing editing of the HBB locus using a homology DNA donor to introduce a sickle cell correction mutation by HDR repair of a DNA DSB formed by Cas9/gRNA complex.

[0098] FIG. 9 (SEQ ID NOS: 111, 112, 113, and 114) includes a schematic showing the sequence near the site of Cas9/gRNA gene editing within the HBB locus. Included is the sequence for a wild type gene and for a sickle cell mutant gene. The sequence targeted by the gRNA is highlighted, as well as the sequence of the donor DNA that includes the sickle cell mutation. Silent mutations encoded by the donor DNA are annotated.

[0099] FIG. 10 includes a bar graph showing HDR editing efficiency for insertion of donor DNA encoding a sickle cell mutation into the HBB locus of CD34-expressing HSPCs using AAV-mediated delivery of donor DNA. Shown is a comparison of HDR efficiency in the presence of i53 relative to RNP+AAV-only, AAV-only, or RNP-only.

[0100] FIG. 11 includes a bar graph showing NHEJ editing efficiency within the HBB locus in CD34-expressing LT-HSPCs following electroporation with gRNA/Cas9 RNP and transfection with a homology DNA donor delivered by AAV. Treatment with i53 is compared to RNP+AAV-only, AAV-only, RNP-only, and mock electroporation (e.g., no RNP or AAV).

[0101] FIGS. 12A-12B include bar graphs showing HDR editing efficiency for insertion of donor DNA encoding a sickle cell mutation into the HBB locus of CD34-expressing LT-HSPCs using AAV-mediated delivery of donor DNA. FIG. 12A includes a bar graph showing HDR editing efficiency in CD34-expressing LT-HSPCs isolated following mobilization with a combination of Mozobil and GCSF. FIG. 12B includes a bar graph showing HDR editing efficiency in LT-HSPCs isolated following mobilization with Mozobil alone.

[0102] FIG. 13 includes a bar graph showing growth of CD34-expressing LT-HSPCs during ex vivo culture following gene editing with Cas9/gRNA RNP and AAV, either with or without treatment of i53.

[0103] FIG. 14 includes a schematic showing a schedule for administration of gene-edited CD34-expressing LT-HSPCs into irradiated mice and subsequent analysis of mouse tissues for engraftment and HDR editing efficiency.

[0104] FIG. 15 includes scatter plots showing a flow cytometry gating strategy for quantification and lineage analysis of mouse tissue samples for cells derived from engrafted human LT-HSPCs.

[0105] FIG. 16 includes a bar graph showing % chimerism of human cells derived from engrafted LT-HSPCs in mouse bone marrow samples isolated at 16 weeks post-engraftment. CD34-expressing LT-HSPCs were administered to mice according to FIG. 14

[0106] FIG. 17 includes a bar graph showing % chimerism of human cells derived from engrafted LT-HSPCs in mouse blood samples isolated at 8 and 16 weeks post-engraftment. CD34-expressing LT-HSPCs were administered to mice according to FIG. 14

[0107] FIG. 18A-18B include bar graphs showing lineage distribution in mouse bone marrow isolated at 16 weeks post-engraftment of LT-HSPCs. Shown is the percentage of human CD45-expressing cells that are B cells, T cells, myeloid cells, or CD34-expressing hematopoietic stem/progenitor cells (HSPCs). CD34-expressing LT-HSPCs were administered to mice according to FIG. 14

[0108] FIG. 19 includes a dot plot showing HDR editing efficiency for insertion of donor DNA encoding a sickle cell mutation into the HBB locus, measured in mouse bone marrow isolated at 16 weeks post-engraftment of LT-HSPCs.

[0109] FIG. 20 includes a dot plot showing indel frequency in mouse bone marrow isolated at 16 weeks post-engraftment of LT-HSPCs relative to the indel frequency of LT-HSPCs prior to engraftment (e.g., input indels).

[0110] FIG. 21 (SEQ ID NOS: 111, 112, 115, 114, 116, respectively) includes a schematic showing the sequence near the site of Cas9/gRNA gene editing within the HBB locus. Included is the sequence for a wild type gene and for a sickle cell mutant gene. The sequence of different homology donor DNA templates that include either a sickle mutation or a sickle cell correction are shown. The donor DNA template with a sickle cell correction includes a β -thalassemia mutation.

[0111] FIGS. 22A-22B include bar graphs showing HDR editing efficiency for insertion of donor DNA encoding a sickle cell mutation into the HBB locus of CD34-expressing LT-HSPCs using AAV-mediated delivery of donor DNA. FIG. 22A shows a comparison of HDR efficiency for AAV given pre-EP or post-EP in combination with gRNA/Cas9 RNP. FIG. 22B shows a comparison of HDR efficiency in the presence of i53 or Nu7441 relative to RNP+AAV-only.

[0112] FIGS. 23A-23B include dot plots showing % chimerism of human cells derived from engrafted LT-HSPCs in mouse blood samples isolated at 8 weeks and 16 weeks post-engraftment. FIG. 23A shows % chimerism for LT-HSPCs edited with gRNA/Cas9 RNP and AAV given either pre-EP or post-EP. FIG. 23B shows % chimerism for LT-HSPCs edited with AAV and gRNA/Cas9 RNP in the presence of i53 or Nu7441 compared to RNP+AAV-only.

[0113] FIGS. 24A-24B include dot plots showing % chimerism of human cells derived from engrafted LT-HSPCs in mouse bone marrow samples isolated at 16 weeks post-engraftment. FIG. 24A shows % chimerism for LT-HSPCs edited with gRNA/Cas9 RNP and AAV given either pre-EP or post-EP. FIG. 24B shows % chimerism for LT-HSPCs edited with AAV and gRNA/Cas9 RNP in the presence of i53 or Nu7441 compared to RNP+AAV-only.

[0114] FIG. 25 includes a bar graph showing lineage distribution in mouse bone marrow isolated at 16 weeks post-engraftment of LT-HSPCs. Shown is total chimerism and percentage of human CD45-expressing cells that are B cells, T cells, myeloid cells, or CD34-expressing hematopoietic stem/progenitor cells (HSPCs). Lineage distribution is shown for LT-HSPCs edited with gRNA/Cas9 RNP and AAV given either pre-EP or post-EP. Also shown is lineage distribution for LT-HSPCs edited with AAV and gRNA/Cas9 RNP in the presence of i53 or Nu7441 compared to RNP+AAV-only.

[0115] FIGS. 26A-26B include dot plots showing HDR editing efficiency for insertion of donor DNA encoding a sickle cell mutation into the HBB locus, measured in mouse bone marrow isolated at 16 weeks post-engraftment of LT-HSPCs. FIG. 26A shows HDR editing efficiency for LT-HSPCs edited with gRNA/Cas9 RNP and AAV given either pre-EP or post-EP. FIG. 26B shows HDR editing efficiency for LT-HSPCs edited with AAV and gRNA/Cas9 RNP in the presence of i53 or Nu7441 compared to RNP+AAV-only.

[0116] FIG. 27 includes a dot plot showing indel frequency in mouse bone marrow isolated at 16 weeks post-engraftment of LT-HSPCs relative to the indel frequency of LT-HSPCs prior to engraftment (e.g., input indels). Shown is indel frequency for LT-HSPCs edited with gRNA/Cas9 RNP and AAV given either pre-EP or post-EP. Also shown is indel frequency for LT-HSPCs edited with AAV and gRNA/Cas9 RNP in the presence of i53 or Nu7441 compared to RNP+AAV-only.

[0117] FIG. 28 includes a dot plot showing erythroid cell enucleation in mouse bone marrow isolated at 16 weeks post-engraftment of LT-HSPCs. Shown is enucleation for LT-HSPCs edited with gRNA/Cas9 RNP and AAV given either pre-EP or post-EP. Also shown is enucleation for LT-HSPCs edited with AAV and gRNA/Cas9 RNP in the presence of i53 or Nu7441 compared to RNP+AAV-only.

[0118] FIG. 29 includes a schematic showing the sequence near the site of Cas9/gRNA gene editing within the HBB locus. Included is the sequence for wild type HBB (healthy), for HBB encoding an E6V mutation (sickle), spacer sequence of R02 gRNA, and sequence of homology donor DNA encoded by AAV.323 to provide correction of the E6V sickle cell disease (SCD) mutation.

[0119] FIGS. 30A-30B provide bar graphs quantifying the frequency of an E6 gene edit in HBB by HDR repair (FIG. 30A) and frequency of INDELs in the HBB gene (FIG. 30B) in CD34-expressing LT-HSPCs electroporated with R02 gRNA/Cas9 RNP+AAV.323 in the presence of i53 alone or in combination with Nu7441. Control cells were electroporated with R02 gRNA/Cas9 RNP+AAV.323, R02 gRNA/Cas9 RNP only, or without AAV or RNP (mock EP).

[0120] FIG. 31A provides a graph quantifying the proportion of human CD45+ cells among total CD45+ cells (percent chimerism) in mouse bone marrow samples that were isolated at 16 weeks following administration of edited

human CD34-expressing LT-HSPCs. Edited cells were electroporated with R02 gRNA/Cas9 RNP+AAV.323 in the presence of i53 alone or in combination with Nu7441. Control cells were electroporated with R02 gRNA/Cas9 RNP+AAV.323, R02 gRNA/Cas9 RNP only, or without RNP or AAV (mock EP). FIG. 31B provides a graph quantifying the frequency of an E6 gene edit in HBB by HDR as measured by an NGS assay in bone marrow isolated from mice administered edited cells as in FIG. 31A.

[0121] FIGS. 32A-32B provide bar graphs quantifying the frequency of a SCD gene correction (E6V to E6) in HBB by HDR repair (FIG. 32A) and frequency of INDELs in the HBB gene (FIG. 32B) in CD34-expressing LT-HSPCs derived from a patient donor with SCD mutation that were subsequently edited by electroporation with R02 gRNA/Cas9 RNP+AAV.323 in the presence of i53. Control cells were edited by electroporation with R02 gRNA/Cas9 RNP+AAV.323 only, R02 gRNA/Cas9 RNP only, or without AAV or RNP (mock EP).

[0122] FIGS. 33A-33B provide bar graphs quantifying the frequency of SCD gene correction by HDR repair (FIG. 33A) and frequency of INDELs in HBB (FIG. 33B) measured either the same day as gene-editing (Day 0) or at 14 days following gene-editing and maintenance by in vitro culture (Day 14) for cells edited as in FIGS. 32A-32B.

[0123] FIG. 34A provides a bar graph quantifying the proportion of total hemoglobin expressed by patient-derived CD34-expressing LT-HSPCs edited as in FIGS. 32A-32B that was HbF, HbA, or HbS as measured by HPLC analysis. FIG. 34B provides an assessment of SCD correction for patient-derived CD34-expressing LT-HSPCs edited with R02 gRNA/Cas9 RNP+AAV.323+i53 that is comparison of the frequency of SCD gene correction by HDR repair (“% HDR by NGS”) and percent decrease in HbS expression relative to mock EP (no RNP or AAV) control cells (“% HbS decrease by HPLC”).

[0124] FIGS. 35A-35B provide bar graphs quantifying the frequency of SCD gene correction by HDR repair (FIG. 35A) and frequency of INDELs in HBB (FIG. 35B) measured in PBMCs or CD34-expressing LT-HSPCs isolated from a patient donor with SCD mutation that were subsequently edited by electroporation with R02 gRNA/Cas9 RNP+AAV.323 in the presence of i53. Control cells were edited by electroporation with R02 gRNA/Cas9 RNP+AAV.323 only, R02 gRNA/Cas9 RNP only, or without AAV or RNP (mock EP).

[0125] FIG. 36 provides a bar graph quantifying the proportion of total hemoglobin expressed by patient-derived PBMCs or CD34-expressing LT-HSPCs edited as in FIGS. 34A-34B that was HbF, HbA, or HbS as measured by HPLC analysis.

DETAILED DESCRIPTION

[0126] Numerous diseases have the potential to be treated by gene-edited cells. Of the approximately 25,000 annotated genes in the human genome, mutations in over 3,000 genes have been linked to disease phenotypes (e.g., see www.omim.org/statistics/geneMap). The ability to correct mutations within disease-affected cells and tissues by genome-editing has the potential to treat and cure numerous monogenic diseases, including severe-combined immunodeficiency (SCID), haemophilia, sickle cell disease (SCD), cystic fibrosis, and enzyme-deficiency diseases. The present disclosure relates to methods and compositions for inducing

or correcting a mutation in a target gene using a site-directed nuclease that creates a DNA DSB. Once a DNA DSB is induced, a gene-edit can be made at the DSB site by introducing an exogenous donor polynucleotide encoding a desired gene-edit. Repair of the DSB using the donor polynucleotide as a template is done through the HDR pathway. Repair by HDR using an exogenous donor polynucleotide enables a precise gene-edit that comprises a correction of a gene mutation, a replacement of a mutated gene altogether, a replacement of a target gene with a mutant or variant version of the target gene, or the introduction of a therapeutic transgene within a target gene. In some embodiments, cells can be removed from a patient, corrected ex vivo, and transplanted back into the host, whereupon engraftment of the edited cells can replace diseased cells and tissues. While in some embodiments, cells can be corrected in vivo by administration of gene editing components.

[0127] However, one challenge with precise gene-editing within a target gene is that HDR repair is often out-competed by the NHEJ repair pathway. Several aspects of NHEJ repair make this a more efficient pathway than HDR repair. Namely NHEJ factors bind to broken dsDNA with high affinity and block processing of the dsDNA to form the single-stranded DNA tails that are necessary for initiation of HDR (Lieber, M. et al. (2010) *Annu Rev Biochem* 79:181-211; Symington, L. et al. (2011) *Annu Review Genetics* 45:247-271). Secondly, the pro-NHEJ factor 53BP1 is actively recruited to sites of damaged chromatin present at a DNA DSB where it functions to suppress the formation of 3' ssDNA tails and antagonize the action of BRCA1, a key factor in HDR (Escribano-Diaz, C. (2013) *Molecular cell* 49:872-883; Feng, L. et al. (2013) *J. Biol Chem.* 288:11135-11143).

[0128] Another challenge with precise gene-editing within a target gene is that HDR repair is particularly inefficient in non-dividing cells (e.g., cells in the G0 or G1 phase of the cell cycle). HDR primarily occurs during the S and G2 phases of the cell cycle, and is thus restricted to cells that are actively dividing (Cox et al (2015) *Nat Med* 21:121-131; Chapman et al (2012) *Mol Cell* 47:497-510; Rothkamm et al (2003) *Mol Cell Biol*; Sharma et al (2007) *Brain Res Bull* 73:48-54; Ciccio et al (2010) *Mol Cell* 40:179-204). In contrast, the NHEJ pathway is the dominant form of DNA DSB repair outside of the S and G2 cell cycle phases and remains a competitive DSB repair pathway even during the S and G2 phases (Beucher et al (2009) *Embo J* 28:3413-3427). The restriction of HDR to certain phases of the cell cycle presents a challenge for editing cells that are non-dividing or slowly dividing, a categorization that encompasses most adult somatic cells.

[0129] Finally, a challenge with precise gene editing is the risk of inducing a mutation within a target gene when NHEJ repair out-competes HDR repair of a DSB. Although NHEJ-mediated DSB repair can accurately rejoin two ends of a broken DNA strand, repeated cutting and repair at the same DSB by NHEJ machinery eventually results in the formation of indels at the break site. The formation of indels into the coding sequence often introduces a loss of function mutation to the target gene, resulting in permanent inactivation of the protein product encoded by the target gene (i.e., a gene knock-out). Thus, the challenges of precise genome-editing are two-fold: poor efficiency of HDR results in low frequencies of cells with the desired gene-edit, particularly in non-dividing cells; and high efficiency of NHEJ results in

the risk of knock-out of the target gene and cells with an undesirable or detrimental phenotype.

[0130] Accordingly, the disclosure provides methods for increasing the efficiency of HDR repair of a DSB introduced within a target gene by a site-directed nuclease using a donor polynucleotide, while also providing a decreased risk of indel formation. For example, the disclosure provides methods for increasing the efficiency of HDR repair of a DSB introduced within a target gene in a quiescent cell that has been induced to divide or a population of quiescent cells that has been induced to divide. Methods of improving efficiency of HDR of the disclosure use inhibition of 53BP1, a key regulator of repair pathway selection. As described herein, inhibition of 53BP1 using a polypeptide inhibitor that inhibits recruitment of 53BP1 to DSB sites resulted in increased HDR efficiency when evaluated in a variety of cell types and gene loci. As described in Examples 1-4, these included cancer cell lines (e.g., HEK293 T cells and human epithelial cells immortalized with hTERT e.g., hTERT RPE-1 cells), as well as human-derived CD34+ hematopoietic stem cells (HSCs). In the latter cell type, contacting the cells ex vivo with 53BP1 polypeptide inhibitor (SEQ ID NO: 70) resulted in a 1.4 fold increase in HDR repair of a DSB induced by a Cas9/gRNA gene-editing system. Significantly, the risk of indel formation within the target gene locus was reduced to less than 30% upon contacting the cells with a 53BP1 polypeptide inhibitor.

[0131] In some embodiments, the disclosure provides methods for increasing efficiency of HDR repair in quiescent cells that are CD34-expressing long-term HSPCs (LT-HSPCs), wherein the method comprises (i) contacting the cells with a 53BP1 inhibitor described herein, and (ii) inducing the cells to divide. Without being bound by theory, it is thought that one or more parameters used to culture LT-HSPCs as described herein (e.g., oxygen concentration, growth medium) induce a cell growth state that enables or promotes increased HDR repair using a method described herein (e.g., a method comprising contacting the cells with a 53BP1 inhibitor of the disclosure).

[0132] For a cell therapy to correct a disease, transplanted cells must survive following transplantation, differentiate into progenitor cell types, and retain the gene-edit through subsequent cycles of cell division and differentiation. While in no way being bound by theory, cells gene-edited in the presence of a 53BP1 inhibitor retain an ability to engraft and differentiate within a host. Experiments described in Example 5 and 7 demonstrated that CD34+ HSCs gene-edited ex vivo in the presence of a 53BP1 polypeptide inhibitor persist following administration to lympho-depleted mice, even up to 16 weeks following administration. Additionally, CD34+ HSCs gene-edited ex vivo in the presence of a 53BP1 polypeptide inhibitor retain the ability to differentiate, generating a similar distribution of progenitor cells when compared to CD34+ HSCs that were either not gene-edited or edited without treatment of a 53BP1 inhibitor. Thus, without being bound by theory, contacting cells with a 53BP1 inhibitor can provide improved HDR efficiency without hampering the fitness of the cells when transplanted into a host.

[0133] Moreover, cells gene-edited in the presence of a 53BP1 inhibitor and transplanted into a host results in a higher proportion of the engrafted tissue encoding the desired gene edit. Experiments described in Example 5 demonstrated that following transplantation of gene-edited

CD34+ HSCs, a higher proportion of cells in the host bone marrow had the desired gene-edit if the animal received CD34+ HSCs edited in the presence of 53BP1 inhibitor compared to without. Indeed, administration of CD34+ HSCs edited in the presence of a 53BP1 inhibitor produced bone marrow with 65% incorporation of the desired gene-edit, an increase in HDR frequency of 1.8-fold relative to cells edited in the absence of the 53BP1 inhibitor. Thus, without being bound by theory, contacting cells with a 53BP1 inhibitor can result in a transplant wherein gene-edited cells dominate the tissue compartment. Such an outcome is essential for transplanted cells to dominate a target tissue and correct a disease state.

[0134] Additionally, methods of improving efficiency of HDR of the disclosure use inhibition of DNA-PKcs, a key factor in mediating repair by the NHEJ pathway. As described herein, inhibition of DNA-PKcs, using a small molecule inhibitor that inhibits kinase function of DNA-PKcs, resulted in increased HDR efficiency when evaluated in a variety of cell types and gene loci. As described in Examples 1-2, these included cancer cell lines (e.g., HEK293 T cells) and CD34+ HSCs. In the latter cell type, contacting the cells ex vivo with a DNA-PKcs small molecule inhibitor resulted in a 1.3 fold increase in HDR repair of a DSB induced by a Cas9/gRNA gene-editing system. Significantly, the risk of indel formation within the target gene locus was reduced to less than 10% upon contacting the cells with a DNA-PKcs small molecule inhibitor.

[0135] While in no way bound by theory, contacting cells comprising a DSB in a target gene with a 53BP1 inhibitor and a DNA-PKcs inhibitor is believed to have beneficial combinatorial effects for gene-editing. Given that the function of 53BP1 is to antagonize the HDR repair pathway, inhibiting 53BP1 is thought to decrease its antagonism of HDR repair, thus improving the efficiency of repair by HDR. However, inhibition of 53BP1 is not expected to prevent repair of a DNA DSB by the NHEJ pathway. In contrast, inhibition of DNA-PKcs is thought to directly inhibit the NHEJ pathway by eliminating the function of a key factor necessary for NHEJ repair. Thus inhibition of 53BP1 and DNA-PKcs results in improved HDR efficiency by non-overlapping mechanisms—inhibition of 53BP1 removes an antagonism that blocks HDR and inhibition of DNA-PKcs directly eliminates the ability of NHEJ to repair a DSB. Thus, while in no way bound by theory, a combination of 53BP1 inhibition and DNA-PKcs inhibition is thought to have a more beneficial gene-editing effect that treatment of either inhibitor alone, wherein the efficiency of gene-editing by HDR repair is improved and the undesirable formation of indels by NHEJ repair is reduced.

Methods of Increasing Homology Directed Repair

Overview of Pathways of DSB Repair

[0136] The repair of DNA breaks (e.g., DSBs) in cells is accomplished primarily through two DNA repair pathways, namely the non-homologous end joining (NHEJ) repair pathway and homology-directed repair (HDR) pathway.

[0137] During NHEJ, the Ku70/80 heterodimers bind to DNA ends and recruit the DNA protein kinase (DNA-PK) (Cannan & Pederson (2015) J Cell Physiol 231:3-14). Once bound, DNA-PK activates its own catalytic subunit (DNA-PKcs) and further enlists the endonuclease Artemis (also known as SNM1c). At a subset of DSBs, Artemis removes

excess single-strand DNA (ssDNA) and generates a substrate that will be ligated by DNA ligase IV. DNA repair by NHEJ involves blunt-end ligation mechanism independent of sequence homology via the canonical DNA-PKcs/Ku70/80 complex.

[0138] During DNA repair by HDR, DSB ends are resected to expose 3' ssDNA tails, primarily by the MRE11-RAD50-NBS1 (MRN) complex (Heyer et al., (2010) *Annu Rev Genet* 44: 113-139). Under physiological conditions, the adjacent sister chromatid will be used as a repair template, providing a homologous sequence, and the ssDNA will invade the template mediated by the recombinase Rad51, displacing an intact strand to form a D-loop. D-loop extension is followed by branch migration to produce double-Holliday junctions, the resolution of which completes the repair cycle. HDR often requires error-prone polymerases yet is typically viewed as error-free (Li and Xu (2016) *Acta Biochim Biophys Sin* 48(7):641-646).

[0139] The NHEJ pathway limits HDR first by being a fast-acting repair pathway that seals the broken DNA ends through a DNA ligase IV-dependent mechanism. Secondly, in NHEJ the Ku70/Ku80 heterodimer binds to the DNA ends with high affinity to block their processing by the nucleases that generate the single-stranded DNA tails that are necessary for initiation of HDR (Lieber, M. et al. (2010) *Annu Rev Biochem* 79:181-211; Symington, L. et al. (2011) *Annu Review Genetics* 45:247-271). Thirdly, 53BP1 is actively recruited to sites of damaged chromatin present at a DNA DSB where it functions to suppress the formation of 3' ssDNA tails and antagonize the action of BRCA1, a factor involved in HDR (Escribano-Diaz, C. (2013) *Molecular cell* 49:872-883; Feng, L. et al. (2013) *J. Biol Chem.* 288:11135-11143).

[0140] During the cell cycle, NHEJ occurs predominantly during G0/G1 and G2 (Chiruvella et al., (2013) *Cold Spring Harb Perspect Biol* 5:a012757). Current studies have shown that NHEJ is the only DSB repair pathway active during G0 and G1, while HDR functions primarily during the S and G2 phases, playing a major role in the repair of replication-associated DSBs (Karanam et al., (2012) *Mol Cell* 47:320-329; Li and Xu (2016) *Acta Biochim Biophys Sin* 48(7): 641-646). NHEJ, unlike HDR, is active in both dividing and non-dividing cells, not just dividing cells, which enables the development of therapies based on genome editing for non-dividing adult cells, such as, for example, cells of the eye, brain, pancreas, or heart.

[0141] A third repair mechanism is microhomology-mediated end joining (MMEJ), also referred to as "Alternative NHEJ", in which the genetic outcome is similar to NHEJ in that small deletions and insertions can occur at the cleavage site. MMEJ makes use of homologous sequences of a few nucleotides flanking the DNA break site to drive a more favored DNA end joining repair outcome, and recent reports have further elucidated the molecular mechanism of this process (Cho and Greenberg, (2015) *Nature* 518:174-176; Mateos-Gomez et al., (2015) *Nature* 518, 254-257; Ceccaldi et al., (2015) *Nature* 528, 258-262). The key mechanistic steps are resection of DSB ends, annealing of microhomologous regions, removal of heterologous flaps, fill-in synthesis and ligation. PARP1 plays a key role in binding to DNA blunt ends and initiating the MMEJ pathway by recruiting DNA polymerase theta (Polθ). Polθ enables the formation of resected DNA ends, as well as enabling the fill-in synthesis (Wang, H. et al. (2017) *Cell Biosci* 7:6).

Inhibition of 53BP1

[0142] In some embodiments, the disclosure provides methods for increasing HDR of a DSB mediated by a site-directed nuclease in a target gene in a cell or population of cells, such as a quiescent cell that has been induced to divide or a population of quiescent cells that has been induced to divide, e.g., CD34+ HSCs, by inhibition of 53BP1.

[0143] The p53-binding protein 1 (53BP1) is a key regulator of cellular response to DNA damage. The choice of repair pathway for repair of a DNA DSB is largely controlled by an antagonism between 53BP1, a pro-NHEJ factor, and BRCA1, a pro-HDR factor (Chapman, J. et al. (2012) *Molecular cell* 47:497-510). 53BP1 promotes NHEJ repair over HDR repair by suppressing formation of 3' single-stranded DNA tails, which is the rate-limiting step in the initiation of the HDR pathway, and by inhibiting BRCA1 recruitment to DSB sites (Escribano-Diaz, C. et al. (2013) *Mol Cell.* 49:872-883; Feng, L. et al (2013) *J Biol Chem* 288:11135-11143). Loss of 53BP1 has been shown to increase HDR efficiency, (Canny, M. et al. (2018) *Nat Biotechnol.* 36(1):95-102). Thus, inhibition of 53BP1 is expected to reduce DSB repair by the NHEJ pathway and favor repair by the HDR pathway.

[0144] Distinct protein domains in the 53BP1 structure are required to enable its function as a pro-NHEJ factor (Zimmermann et al (2014) *Trends Cell Biol* 24:108-117). Human 53BP1 is a large (e.g., 200 kDa, 1972 amino acids) multi-domain protein that enables recruitment to DSB sites and binding of protein factors involved in DNA repair. The 53BP1 N-terminus is comprised of a large subunit that is heavily phosphorylated following DNA damage and facilitates binding interactions with DNA repair machinery. The central portion of 53BP1 comprises a focus-forming region that is essential for binding to damaged chromatin, which allows recruitment to DSB sites. It comprises a nuclear localization signal (NLS), a tandem Tudor domain that binds to di-methylated histone H4 lysine 20 (e.g., H4K20^{Me2}), and a ubiquitin-dependent recruitment (UDR) motif that recognizes histone H2A/H2AX ubiquitinated on lysine 15 (e.g., H2A(X)K15^{Ub}) (Botuyan, M. (2006) *Cell* 127:1361-1373; Fradet-Turcotte et al (2013) *Nature* 499:50-54). The focus-forming region extends from amino acids 1220-1711 of human 53BP1, with the tandem Tudor domain extending from amino acids 1484-1603 and the UDR extending from amino acids 1604-1631. The 53BP1 C-terminus is comprised of repeating BRCA1 C-terminus (BRCT) domains that are important for DNA repair in heterochromatin (Noon et al (2010) *Nat Cell Biol* 12:177-184) and mediate interactions with the tumor suppressor p53 that guides cellular response to DNA damage (Iwabuchi, et al (1994) *PNAS* 91:6098-6102).

[0145] The functionality of 53BP1 for promoting the NHEJ pathway requires recruitment to damaged chromatin through its tandem Tudor and UDR domains and binding to repair machinery through phosphorylation of the 53BP1 N-terminus.

[0146] Accordingly, the present disclosure provides 53BP1 inhibitors that inhibit NHEJ and promote HDR repair of a DSB in a target gene. In some embodiments, a 53BP1 inhibitor of the disclosure inhibits 53BP1 recruitment to DSB sites. In some embodiments, a 53BP1 inhibitor of the disclosure inhibits 53BP1 recruitment by inhibiting, reducing, disrupting or blocking an interaction of 53BP1 with

damaged chromatin. In some embodiments, a 53BP1 inhibitor of the disclosure inhibits, reduces, disrupts or blocks an interaction of the 53BP1 focus forming region (amino acids 1220-1711) with DSB sites. In some embodiments, a 53BP1 inhibitor of the disclosure inhibits, reduces, disrupts or blocks an interaction of the 53BP1 focus forming region (amino acids 1220-1711) with damaged chromatin. In some embodiments, a 53BP1 inhibitor of the disclosure inhibits, reduces, disrupts or blocks an interaction of the 53BP1 tandem Tudor domain with damaged chromatin (e.g., with methylated histone, H4K20^{Me2}). In some embodiments, a 53BP1 inhibitor of the disclosure inhibits, reduces, disrupts or blocks the interaction of the 53BP1 UDR motif with damaged chromatin (e.g., with ubiquitinated histone, H2A (X)K15^{Ub}).

[0147] In some embodiments, a 53BP1 inhibitor of the disclosure inhibits, reduces, disrupts or blocks protein-protein interactions with the 53BP1 BRCT domain. In some embodiments, a 53BP1 inhibitor of the disclosure inhibits, reduces, disrupts or blocks the interactions of the 53BP1 BRCT domain with the tumor suppressor p53.

[0148] In some embodiments, a 53BP1 inhibitor of the disclosure inhibits, reduces, disrupts or blocks the ability of 53BP1 to bind to DNA repair factors. In some embodiments, a 53BP1 inhibitor of the disclosure inhibits, reduces, disrupts or blocks phosphorylation of the 53BP1 N-terminus, thus inhibiting, reducing or preventing binding of DNA repair factors. In some embodiments, a 53BP1 inhibitor of the disclosure binds to phosphorylated sites on the 53BP1 N-terminus, thus inhibiting, reducing or preventing DNA repair factors from recognizing and binding to phosphorylated sites on the 53BP1 N-terminus. In some embodiments, a 53BP1 inhibitor of the disclosure reduces, eliminates or removes phosphorylated sites on the 53BP1 N-terminus (e.g., by promoting or catalyzing a dephosphorylation mechanism), thus reducing, eliminating or removing sites required for binding of DNA repair factors. In some embodiments, a 53BP1 inhibitor that binds to phosphorylated sites on 53BP1 and facilitates HDR is suppressor of cancer cell invasion (SCAI) or a fragment thereof. In some embodiments, binding of SCAI or a fragment thereof prevents binding of the DNA repair factor RAP1-interacting factor homolog (RIF1). In some embodiments, blocking RIF1 binding to 53BP1 results in increased HDR repair of a DNA DSB.

[0149] In some embodiments, the 53BP1 inhibitor of the disclosure inhibits, disrupts or blocks 53BP1 recruitment to DSB sites in the cell. In some embodiments, the 53BP1 inhibitor of the disclosure inhibits, disrupts or blocks an interaction of 53BP1 with damaged chromatin in the cell. In some embodiments, the 53BP1 inhibitor of the disclosure inhibits, disrupts or blocks binding of DNA repair factors to sites of phosphorylation on the 53BP1 N-terminus. In some embodiments, the 53BP1 inhibitor of the disclosure is a small molecule. In some embodiments, the 53BP1 inhibitor of the disclosure is a polypeptide. In some embodiments, the 53BP1 inhibitor of the disclosure is a nucleic acid.

[0150] In some embodiments, recruitment of 53BP1 to a DSB site occurs via recognition of damaged chromatin. In some embodiments, recruitment of 53BP1 to damaged chromatin occurs through recognition of H4K20me2 through the 53BP1 UDR motif. In some embodiments, recognition of damaged chromatin by 53BP1 is dependent upon ubiquitina-

tion of histones. In some embodiments, inhibition of histone ubiquitination results in inhibition of 53BP1 recruitment to DSB sites.

[0151] Acetylation of 53BP1 has been shown to inhibit 53BP1 binding to damaged chromatin (Guo et al (2018) *Nucleic Acids Res* 46:689-703). In some embodiments, an inhibitor of 53BP1 promotes post-translational modification of 53BP1. In some embodiments, an inhibitor of 53BP1 promotes post-translation modification of 53BP1 that prevents 53BP1 binding to damaged chromatin. In some embodiments, an inhibitor of 53BP1 promotes acetylation of 53BP1. In some embodiments, an inhibitor of 53BP1 promotes acetylation of the 53BP1 UDR motif. In some embodiments, acetylation of 53BP1 prevents 53BP1 recruitment to DSB sites.

[0152] In some embodiments, a 53BP1 inhibitor is identified by binding affinity for the 53BP1 polypeptide. Methods of measuring binding affinity of an inhibitor to a protein are known in the art. Non-limiting examples include measuring inhibitor affinity by enzyme-linked immunosorbent assay (e.g., ELISA), immunoblot, immunoprecipitation-based assay, fluorescence polarization assay, fluorescence resonance energy transfer assay, fluorescence anisotropy assay, yeast surface display (Gal (2007) *Curr Opin Struct Biol* 17:467-473), kinetic exclusion assay, surface plasmon resonance, or isothermal titration calorimetry. In some embodiments, a method of measuring binding affinity is an ELISA wherein an inhibitor is measured for affinity to the 53BP1 polypeptide. In some embodiments, binding affinity is evaluated by a competition-based ELISA wherein binding of an inhibitor to the 53BP1 polypeptide is measured in the presence of increasing concentrations of a known 53BP1 binding partner (e.g., a histone methyl-lysine peptide with affinity for 53BP1).

[0153] In some embodiments, a 53BP1 inhibitor is identified by binding affinity for a fragment of the 53BP1 polypeptide. In some embodiments, a fragment is a domain of the 53BP1 polypeptide. In some embodiments, the domain is the Tudor domain. In some embodiments, the domain is the UDR motif. In some embodiments, the domain comprises the N-terminus of the 53BP1 polypeptide.

[0154] In some embodiments, a 53BP1 inhibitor of the disclosure binds to the 53BP1 polypeptide. Methods of determining the structural interactions that enable binding of the inhibitor with the 53BP1 polypeptide are known in the art. Non-limiting examples include X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, electron microscopy, small-angle X-ray scattering (SAXS), and small-angle neutron scattering (SANS). In some embodiments, the structural interactions are determined by a mutagenesis experiment wherein residues of the 53BP1 polypeptide are mutated and the effect on inhibitor binding are evaluated. Such methods enable identification of key residues that contribute to binding.

[0155] In some embodiments, the 53BP1 inhibitor of the disclosure is a 53BP1 binding polypeptide that inhibits 53BP1 recruitment to the DSB in the cell. In some embodiments, a 53BP1 binding polypeptide of the disclosure inhibits, disrupts or blocks binding of 53BP1 to damaged chromatin in the cell. In some embodiments, a 53BP1 binding polypeptide of the disclosure inhibits, disrupts or blocks the 53BP1 tandem Tudor domain from binding to damaged chromatin in the cell. In some embodiments, a 53BP1

binding polypeptide of the disclosure inhibits, disrupts or blocks the 53BP1 UDR motif from binding to damaged chromatin in the cell.

[0156] In some embodiments, an inhibitor of 53BP1 is a polypeptide identified from a phage-display library or a variant thereof as described by US 2019/0010196A, which is incorporated by reference herein. In some embodiments, a polypeptide inhibitor of 53BP1 has binding affinity for the 53BP1 Tudor domain. The 53BP1 Tudor domain is involved in recognition of methylated residues on the histone core that facilitates recruitment of 53BP1 to a DNA DSB site. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure inhibits, reduces or prevents recruitment of 53BP1 to a DNA DSB by binding to the 53BP1 Tudor domain.

[0157] In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure is modified, by, for example, substitution of one or more amino acid residues, insertion of one or more amino acid residues, or deletion of one or more amino acid residues. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure is modified by chemical modifications. Techniques for modification of one or more amino acid residues are known to one skilled in the art. In some embodiments, a modification is substitution of one or more amino acid residues. In one embodiment, a modification increases binding affinity of the 53BP1 polypeptide inhibitor for the 53BP1 polypeptide or a fragment thereof.

[0158] In some embodiments, a modified polypeptide inhibitor of 53BP1 is identified by affinity for the 53BP1 Tudor domain. Affinity for the 53BP1 Tudor domain may be assessed by suitable assays known to one skilled in the art. In some embodiments, affinity is measured by a competitive immunoprecipitation assay against an endogenous polypeptide that binds 53BP1, for example, dimethylated histone H4 Lys20. In some embodiments, affinity is measured by isothermal calorimetry using recombinant 53BP1. In some embodiments, affinity is determined by assessing 53BP1 recruitment to DSB sites. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure has a quantifiable binding affinity for the 53BP1 Tudor domain of approximately 0.5 to 15×10^{-9} M, 0.5 to 25×10^{-9} M, 0.5 to 50×10^{-9} M, 0.5 to 100×10^{-9} M, 0.5 to 200×10^{-9} M, 1 to 200×10^{-9} M, 1 to 300×10^{-9} M, 1 to 400×10^{-9} M, 1 to 500×10^{-9} M, 100 to 250×10^{-9} M, 100 to 500×10^{-9} M, or 200 to 500×10^{-9} M. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure has a quantifiable binding affinity for the 53BP1 Tudor domain of approximately 200 to 500×10^{-9} M. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure has a quantifiable binding affinity for the 53BP1 Tudor domain of approximately 250×10^{-9} M.

[0159] In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure comprises a polypeptide sequence of SEQ ID NO: 70. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure comprises a polypeptide sequence that is at least about 50%, 60%, 70% or 80% identical to the polypeptide sequence of SEQ ID NO: 70. In some embodiments, a 53BP1 polypeptide inhibitor comprises a polypeptide sequence that is at least about 90%, 95%, 96%, 97%, 98% or 99% identical to the polypeptide sequence of SEQ ID NO: 70. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure comprises a polypeptide sequence that is at least about 95% identical to the polypeptide sequence of SEQ ID NO: 70. In some embodiments, a 53BP1 polypeptide inhibitor of the disclo-

sure comprises a polypeptide sequence that is at least about 96% identical to the polypeptide sequence of SEQ ID NO: 70. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure comprises a polypeptide sequence that is at least about 97% identical to the polypeptide sequence of SEQ ID NO: 70. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure comprises a polypeptide sequence that is at least about 98% identical to the polypeptide sequence of SEQ ID NO: 70. In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure comprises a polypeptide sequence that is at least about 99% identical to the polypeptide sequence of SEQ ID NO: 70. In some embodiments, percent identity is made by a comparison that is performed by a BLAST algorithm wherein the parameters of the algorithm are selected to encompass the largest match between the respective polypeptide sequences over the entire length of the polypeptide sequence as set forth by SEQ ID NO: 70. BLAST algorithms are often used for sequence analysis and are well known by one skilled in the art (Altschul, S., et al. (1990) *J. Mol. Biol.* 215:403-410; Gish, W. et al. (1993) *Nat. Genet.* 3:266-272; Madden, T. et al. (1996) *Meth. Enzymol.* 266:131-141; Altschul, S. et al. (1997) *Nucleic Acids Res.* 25:3389-3402; Zhang, J. et al. (1997) *Genome Res.* 7:649-656; Wootton, J. et al., (1993) *Comput. Chem.* 17:149-163; Hancock, J. et al. (1994) *Comput. Appl. Biosci.* 10:67-70).

[0160] In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure comprises a fragment of a polypeptide comprising the polypeptide sequence of SEQ ID NO: 70 that retains binding to the 53BP1 Tudor domain. In some embodiments, a fragment has at least 1-5, at least 1-10, at least 5-15, at least 10-20, at least 15-30, at least 15-40 fewer amino acid residues than a polypeptide comprising a polypeptide sequence as set forth by SEQ ID NO: 70.

[0161] In some embodiments, a 53BP1 polypeptide inhibitor of the disclosure comprises a fusion polypeptide comprising a polypeptide comprising the polypeptide sequence of SEQ ID NO: 70 that retains binding to the 53BP1 Tudor domain. In some embodiments, a fusion polypeptide is obtained by addition of amino acids or peptides or by substitutions of individual amino acids or peptides that enable by chemical coupling with suitable reagents to a fusion partner. In some embodiments, a fusion is prepared by preparation and expression of a vector comprising a gene encoding a polypeptide described herein and a gene encoding a fusion partner. In some embodiments, a fusion partner is a polypeptide, non-limiting examples include an enzyme, a fluorescent tag, a purification tag, a toxin, an antibody fragment, or an albumin fragment. In some embodiments, a fusion partner is a chemical label, non-limiting examples include a fluorescent dye, biotin, a radioactive label, a saccharide, or a phosphate.

[0162] In some embodiments, a 53BP1 polypeptide inhibitor as described herein is encoded by a polynucleotide. In some embodiments, a 53BP1 polypeptide inhibitor as described herein is provided as a nucleic acid comprising a nucleotide sequence encoding the 53BP1 polypeptide inhibitor. In some embodiments, the nucleic acid is a DNA molecule. In some embodiments, the nucleic acid is an RNA molecule. In some embodiments, the nucleic acid is a messenger RNA (mRNA). Methods of preparing mRNA or high expression of an encoded polypeptide are known in the art. In some embodiments, an mRNA comprises an open-reading frame (ORF) encoding an inhibitor of 53BP1. In

some embodiments, the nucleic acid encoding a 53BP1 polypeptide inhibitor comprises an mRNA comprising an ORF encoding the amino acid sequence of SEQ ID NO: 70.

[0163] In some embodiments, a nucleic acid comprising a nucleotide sequence encoding a 53BP1 polypeptide inhibitor is delivered to a cell by a vector. Methods of delivering nucleic acids to a cell using a vector are known in the art and are described herein.

[0164] In some embodiments, a 53BP1 inhibitor of the disclosure comprises a gene-editing system for disrupting a gene encoding 53BP1. In some embodiments, the 53BP1 inhibitor comprises a CRISPR/Cas9 gene editing system. Methods of using CRISPR-Cas gene editing technology to create a genomic deletion in a cell (e.g., a knock-out in a gene of a cell) are known (e.g., Bauer (2015) *Vis Exp* 95:e52118). In some embodiments, a knock-out of a gene encoding 53BP1 using CRISPR-Cas gene editing comprises contacting a cell with Cas9 polypeptide and a gRNA targeting the 53BP1 gene locus. In some embodiments, gRNA sequence targeting the 53BP1 gene locus is designed using the 53BP1 gene sequence using methods known in the art (see e.g., Briner (2014) *Molecular Cell* 56:333-339). In some embodiments, gRNAs targeting the 53BP1 gene locus create indels in the region of the 53BP1 gene that disrupt expression of 53BP1 in the cell. In some embodiments, 50-100%, 50-90%, 50-80%, 50-70%, 50-60%, 60-100%, 60-90%, 60-80%, 60-70%, 70-100%, 70-90%, 70-80%, 80-100%, 80-90%, or 90-100% of cells in the edited population lack detectable expression of 53BP1.

[0165] In some embodiments, a 53BP1 inhibitor of the disclosure comprises a small interfering RNA (siRNA) which silences 53BP1 expression. Methods of silencing 53BP1 expression using siRNA are taught by US 2019/0010196 which is incorporated by reference herein. Methods of delivering siRNA can be performed using non-viral or viral delivery methods as described in the art (e.g., Gao (2009) *Mol Pharm* 6:651-658; Oliveira (2006) *J Biomed Biotechnol* 2006:63675; Tatiparti (2017) *Nanomaterials* 7:77). In some embodiments, a cell is transfected with siRNA targeting 53BP1 mRNAs. In some embodiments, expression of 53BP1 is decreased by about 50%, by about 60%, by about 70%, by about 80%, by about 90%, or by about 100% following transfection with siRNA targeting 53BP1 mRNA.

[0166] Inhibition of DNA-PKcs

[0167] In some embodiments, the disclosure provides methods for increasing HDR of a DSB mediated by a site-directed nuclease in a target gene in a cell or population of cells, such as a quiescent cell that has been induced to divide or a population of quiescent cells that has been induced to divide, e.g., CD34+ HSCs, by inhibition of DNA-PKcs. In some embodiments, the disclosure provides methods for increasing HDR of a DSB mediated by a site-directed nuclease in a target gene in a cell or population of cells, such as a quiescent cell that has been induced to divide or a population of quiescent cells that has been induced to divide, e.g., CD34+ HSCs, by inhibition of 53BP1 and DNA-PKcs.

[0168] The DNA-PKcs is a member of the phosphatidylinositol-3 (PI-3) kinase-like kinase family (PIKK) and is a key kinase involved in NHEJ repair. DNA-PKcs is directed to DSB sites by binding to the Ku70/80 heterodimer that has high-affinity for broken dsDNA ends and is first recruited to DSB sites. The complex formed at the DSB comprising

DNA, Ku70/80 and DNA-PKcs is referred to as "DNA-PK" (Gottlieb (1993) *Cell* 72:131-142). The large DNA-PK complex is responsible for holding the two ends of a broken DNA molecule together. Additionally, binding of DNA-PKcs to the DNA-Ku70/80 complex results in activation of DNA-PKcs kinase activity (Yoo et al (1999) *Nucleic Acids Res* 27:4679-4686; Calsou (1999) *J Biol Chem* 274:7848-7856). DNA-PKcs phosphorylates numerous NHEJ repair factors, thus enabling their function in NHEJ repair.

[0169] Accordingly, the present disclosure provides DNA-PKcs inhibitors that inhibit NHEJ and promote HDR repair of a DSB in a target gene. In some embodiments, a DNA-PKcs inhibitor of the disclosure inhibits, reduces, disrupts, or blocks the ability of DNA-PKcs to a DSB site. In some embodiments, a DNA-PKcs inhibitor of the disclosure inhibits, reduces, disrupts, or blocks the ability of DNA-PKcs to bind to Ku70/80 to form a DNA-PK complex. In some embodiments, a DNA-PKcs inhibitor of the disclosure inhibits, reduces, disrupts, or blocks the function of the DNA-PKcs kinase domain. In some embodiments, a DNA-PKcs inhibitor of the disclosure inhibits, reduces, disrupts, or blocks phosphorylation of NHEJ factors by the DNA-PKcs kinase domain. In some embodiments, a DNA-PKcs inhibitor of the disclosure is a polypeptide. In some embodiments, a DNA-PKcs inhibitor is a nucleic acid. In some embodiments, a DNA-PKcs inhibitor is a small molecule. In some embodiments, a DNA-PKcs inhibitor of the disclosure is a small molecule that inhibits, disrupts or blocks the DNA-PKcs kinase domain.

[0170] In some embodiments, a DNA-PKcs inhibitor of the disclosure is identified by binding affinity for DNA-PKcs or a fragment thereof (e.g., a functional domain of DNA-PKcs). Methods of measuring binding affinity of an inhibitor for a protein domain are known in the art. Non-limiting examples include measuring inhibitor affinity by enzyme-linked immunosorbent assay (e.g., ELISA), immunoblot, immunoprecipitation-based assay, fluorescence polarization assay, fluorescence resonance energy transfer assay, fluorescence anisotropy assay, yeast surface display (Gai (2007) *Curr Opin Struct Biol* 17:467-473), kinetic exclusion assay, surface plasmon resonance, or isothermal titration calorimetry.

[0171] In some embodiments, a DNA-PKcs inhibitor of the disclosure binds to the DNA-PKcs polypeptide. Methods of determining the structural interactions that enable binding of the inhibitor with the DNA-PKcs polypeptide are known in the art. Non-limiting examples include X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, electron microscopy, small-angle X-ray scattering (SAXS), and small-angle neutron scattering (SANS). In some embodiments, the structural interactions are determined by a mutagenesis experiment wherein residues of the DNA-PKcs polypeptide are mutated and the effect on inhibitor binding are evaluated. Such methods enable identification of key residues that contribute to binding.

[0172] In some embodiments, a method of inhibition of DNA-PKcs function in a cell comprises contacting the cell with a small molecule inhibitor of DNA-PKcs. In some embodiments, the DNA-PKcs inhibitor of the disclosure is a small molecule inhibitor Nu7441 (e.g., Leahy (2004) *Bioorg Med Chem Lett* 14:6083-6087). In some embodiments, the DNA-PKcs inhibitor of the disclosure is a PI 3-kinase inhibitor LY294002, which has been found to inhibit DNA-PKcs function in vitro (Izzard (1999) *Cancer*

Res 59:2581-2586). In some embodiments, the DNA-PKcs inhibitor of the disclosure is a small molecule inhibitor capable of selectively inhibiting the activity of DNA-PKcs compared to PI 3-kinase. Non-limiting examples include 2-amino-chromen-4-ones that are described by WO 03/024949, which is incorporated by reference herein. In some embodiments, the DNA-PKcs inhibitor of the disclosure is a small molecule inhibitor of DNA-PKcs function, including 1 (2-hydroxy-4-morpholin-4-yl-phenyl)-ethanone (e.g., Kashishian (2003) *Mol Cancer Ther* 2:1257-1264). In some embodiments, the DNA-PKcs inhibitor of the disclosure is a small molecule inhibitor of DNA-PKcs function SU11752 (e.g., Ismail (2004) *Oncogene* 23:873-882). In some embodiments, the DNA-PKcs inhibitor of the disclosure is a small molecule inhibitor of DNA-PKcs function described in U.S. Pat. No. 9,592,232, incorporated herein by reference. In some embodiments, the DNA-PKcs inhibitor of the disclosure is a small molecule inhibitor of DNA-PKcs function described in U.S. Pat. No. 7,402,607, incorporated herein by reference. In some embodiments, the DNA-PKcs inhibitor of the disclosure is a small molecule inhibitor of DNA-PKcs function described in U.S. Pat. No. 6,893,821, incorporated herein by reference. In some embodiments, the DNA-PKcs inhibitor of the disclosure is a small molecule inhibitor of DNA-PKcs function described in US 2018/0194782.

Inhibition of Other Targets

[0173] In some embodiments, the disclosure provides methods for increasing HDR of a DSB mediated by a site-directed nuclease in a target gene in a cell or population of cells, such as a quiescent cell that has been induced to divide or a population of quiescent cells that has been induced to divide, e.g., CD34+ HSCs, by inhibition of the NHEJ pathway, alone or in combination with inhibition of 53BP1 and/or DNA-PKcs. In some embodiments, the disclosure provides a method of inhibiting the NHEJ pathway by inhibition of key NHEJ enzymes. For example, in some embodiments, the disclosure provides a method of inhibiting the NHEJ pathway by inhibition of Ku70/80. In some embodiments, the disclosure provides inhibitors of Ku70/80 including CYREN (e.g., Arnoult (2017) *Nature* 549:548-552). In some embodiments, the disclosure provides a method of inhibiting the NHEJ pathway by inhibition of DNA Ligase IV. In some embodiments, the disclosure provides inhibitors of DNA Ligase IV, including Scr7 (Maruyama (2015) *Nat Biotechnol* 33:538-542).

[0174] In some embodiments, the disclosure provides methods of increasing or improving repair of a DNA DSB by HDR by inhibition of the MMEJ pathway (e.g., methods of MMEJ inhibition reviewed in Sfeir (2015) 40:701-714). In some embodiments, the disclosure provides methods of inhibition of the MMEJ pathway by inhibition of DNA polymerase theta (Pol θ). In some embodiments, the disclosure provides method of inhibition of the MMEJ pathway by inhibition of PARP. In some embodiments, the disclosure provides PARP inhibitors, including molecules developed for the treatment of cancer, including Veliparib and Olaparib. In some embodiments, inhibition of the MMEJ pathway comprises inhibition of MRE11. In some embodiments, the disclosure provides MRE11 inhibitors, including Mirin and derivatives (e.g., Shibata (2014) *Molec Cell* 53:7-18).

[0175] In some embodiments, the disclosure provides methods for increasing HDR of a DSB mediated by a

site-directed nuclease in a target gene in a cell or population of cells, such as a quiescent cell that has been induced to divide or a population of quiescent cells that has been induced to divide, e.g., CD34+ HSCs, by treatment of a cell or population of cells with a compound that stimulates HDR efficiency. In some embodiments, the disclosure provides a stimulator of HDR, wherein the stimulator of HDR is an agonist that promotes the function of a factor in the HDR pathway. In some embodiments, the disclosure provides a stimulator of an HDR factor, wherein the HDR factor is RAD51. In some embodiments, the disclosure provides agonists of RAD51, including RS-1 (e.g., Jayatilaka (2008) *PNAS* 105:15848-15853).

Combination of Inhibitors

[0176] In some embodiments, the disclosure provides methods for increasing HDR of a DSB mediated by a site-directed nuclease in a target gene in a cell or population of cells, such as a quiescent cell that has been induced to divide or a population of quiescent cells that has been induced to divide, e.g., CD34+ HSCs, by treatment with an inhibitor of 53BP1 in combination with an inhibitor of the NHEJ pathway. In some embodiments, a method of increasing HDR is treatment with an inhibitor of 53BP1 in combination with an inhibitor of DNA-PKcs. In some embodiments, a method of increasing HDR is treatment with a polypeptide inhibitor of 53BP1 in combination with an inhibitor of DNA-PKcs. In some embodiments, a method of increasing HDR is treatment with a polypeptide inhibitor of 53BP1 comprising the amino acid sequence identified by SEQ ID NO: 70 in combination with a small molecule inhibitor of DNA-PKcs.

[0177] In some embodiments, a method of increasing HDR is treatment with an inhibitor of 53BP1 in combination with an inhibitor of Ku70/80. In some embodiments, a method of increasing HDR is treatment with a polypeptide inhibitor of 53BP1 comprising the amino acid sequence identified by SEQ ID NO: 70 in combination with an inhibitor of Ku70/80. In some embodiments, a method of increasing HDR is treatment with an inhibitor of 53BP1 in combination with an inhibitor of DNA Ligase IV. In some embodiments, a method of increasing HDR is treatment with a polypeptide inhibitor of 53BP1 comprising the amino acid sequence identified by SEQ ID NO: 70 in combination with an inhibitor of DNA Ligase IV.

[0178] In some embodiments, a method of increasing HDR is treatment with an inhibitor of 53BP1 in combination an inhibitor of the MMEJ pathway. In some embodiments, a method of increasing HDR is treatment with a polypeptide inhibitor of 53BP1 comprising the amino acid sequence identified by SEQ ID NO: 70 in combination with an inhibitor of the MMEJ pathway. In some embodiments, a method of increasing HDR is treatment with a polypeptide inhibitor of 53BP1 comprising the amino acid sequence identified by SEQ ID NO: 70 in combination with an inhibitor of PARP. In some embodiments, a method of increasing HDR is treatment with a polypeptide inhibitor of 53BP1 comprising the amino acid sequence identified by SEQ ID NO: 70 in combination with an inhibitor of DNA polymerase theta.

Engineered Human Cells

[0179] Provided herein are methods of gene-editing within a target gene by repair of a DNA DSB in the target gene by

the HDR pathway using a donor polynucleotide. In some embodiments, the target gene is edited to correct a mutation. In some embodiments, the target gene is edited to introduce a mutation. In some embodiments, the target gene is edited by replacement with a different polynucleotide sequence, such as a polynucleotide sequence encoding a different gene (e.g., a transgene) or a variant version of the target gene. In some embodiments, the target gene is edited by deletion and insertion of a different gene (e.g., a transgene). In some embodiments, the target gene is edited by insertion of a transgene comprising one or more exons and one or more introns. In some embodiments, the target gene is edited by insertion of a transgene comprising only exons.

[0180] In some embodiments, a target gene is edited using methods herein to correct a genetic mutation that results in a monogenic disease. A monogenic disease is characterized by a mutation in a single gene. Non-limiting examples of gene mutations that result in monogenic disease include mutation of the beta-globin (e.g., hemoglobin beta) gene that results in hemoglobinopathies, mutation of the cystic fibrosis conductance transmembrane regulator (CFTR) gene that results in cystic fibrosis, mutation of the huntingtin (HTT) gene that results in Huntington's disease, mutation of the dystrophin myotonic-protein kinase (DMPK) gene that results in Myotonic dystrophy type 1, mutation of the low-density lipoprotein receptor (LDLR) or apolipoprotein B (APOB) gene that results in hypercholesterolemia, mutation of coagulation factor VIII that results in hemophilia A, mutation of recombination activating gene 1 (RAG1) that results in severe combined immunodeficiency (SCID), or mutation in the dystrophin gene that results in Duchenne type muscular dystrophy. Further non-limiting examples of disorders associated with particular target genes that are edited using methods described herein are detailed in Table 1.

TABLE 1

Disorders Associated with Mutations in a Target Gene	
Monogenic Disorder	Target Gene
Sickle Cell Disease	Hemoglobin subunit beta (HBB)
Alpha-Thalassemia	Hemoglobin subunit alpha 1 (HBA1) and hemoglobin subunit alpha 2 (HBA2)
Beta-Thalassemia	Hemoglobin subunit beta (HBB)
Hemophilia A	Coagulation factor VIII
Cystic Fibrosis	Cystic fibrosis conductance transmembrane regulator (CFTR)
X-linked severe Combined Immunodeficiency (SCID)	Interleukin 2 receptor subunit gamma (IL2RG)
RAG1-deficient SCID (Omenn Syndrome)	Recombination activating 1 (RAG1)
JAK3-deficient SCID	Janus kinase 3 (JAK3)
ZAP70-related SCID	Zeta chain of T cell receptor associated protein kinase 70 (ZAP70)
Adenosine deaminase deficiency	Adenosine deaminase (ADA)
Familial Hypercholesterolemia	Low-density lipoprotein receptor (LDLR) or Apolipoprotein B (APOB)
Epidermolysis bullosa	Collagen type VII alpha 1 chain (COL7A1)
Muscular Dystrophy type 1	Dystrophin myotonic-protein kinase (DMPK)
Duchenne type muscular dystrophy	Dystrophin
Huntington's Disease	Huntingtin (HTT)
NGLY1 Deficiency	N-glycanase 1 (NGLY1)

[0181] In some embodiments, a monogenic disease is treated by administering gene-edited human cells to a patient. In some embodiments, human cells are taken from the patient and edited to correct a genetic mutation prior to being reintroduced to the patient for treatment of a monogenic disorder. In some embodiments, cells from a patient are somatic cells that are reprogrammed to generate induced pluripotent stem cells (iPSCs). In some embodiments, iPSCs are gene-edited to correct a mutation and then differentiated prior to administration to a patient. In some embodiments, cells from a patient are hematopoietic stem cells (HSCs) or hematopoietic progenitor cells (HPCs). In some embodiments, HSCs and HPCs are gene-edited and introduced to a patient for treatment of a monogenic disease.

[0182] In some embodiments, HSCs are engineered (e.g., gene-edited) for treatment of a hemoglobinopathy. Hemoglobinopathies encompass a number of anemias that are associated with changes in the genetically determined structure or expression of hemoglobin. These include changes to the molecule structure of the hemoglobin chain, such as occurs with sickle cell anemia, as well as changes in which synthesis of one or more chains is reduced or absent, such as occurs with various thalassemias.

[0183] Disorders specifically associated with the β -globin protein are referred to generally as β -hemoglobinopathies. For example, β -thalassemias result from a partial or complete defect in the expression of the β -globin gene, leading to deficient or absent hemoglobin A (HbA). HbA is the most common human hemoglobin tetramer and consists of two α -chains and two β -chains ($\alpha_2\beta_2$). β -thalassemias are due to mutations on the adult β -globin gene (HBB) on chromosome 11, and are inherited in an autosomal, recessive fashion.

[0184] Sickle cell disease (SCD) includes SCA, sickle hemoglobin C disease, sickle beta-plus-thalassemia, and sickle beta-zero-thalassemia. All forms of SCD are caused by mutations within the HBB gene. SCA is caused by a single missense mutation in the sixth codon (e.g., seventh codon when including the start codon) of the HBB gene (e.g., A to T), resulting in a substitution of glutamic acid by valine (e.g., Glu to Val). The mutant protein, when incorporated into hemoglobin, results in unstable hemoglobin HbS ($\alpha_2\beta_2^s$) in contrast to normal adult hemoglobin HbA ($\alpha_2\beta_2^A$). When HbS is the predominant form of hemoglobin, it results in red blood cells (RBCs) with distorted sickle shape. Sickled RBCs are less flexible than normal RBCs, and tend to get stuck in small blood vessels, resulting in vaso-occlusive events. These events are associated with tissue ischemia leading to acute and chronic pain.

[0185] In some embodiments, a patient is treated with gene-edited human cells to ameliorate a hemoglobinopathy (e.g., de Montalembert (2008) *BMJ*, 337:a1397; Sheth, et al. (2013) *British J. Haematology* 162:455-464). Methods towards treatment of hemoglobinopathies by production of genome-edited stem cells, including hematopoietic stem cells (HSCs), are taught by US 2018/0030438 and US 2018/0200387 which are incorporated by reference herein. In some embodiments, a method of treating a patient with hemoglobinopathy comprises administering gene-edited stem cells to the patient that give rise to a population of circulating RBCs that will be effective in ameliorating one or more clinical conditions associated with the patient's disease. In some embodiments, a gene-edited stem cell is an HSC, long-term repopulating hematopoietic cell or an LT-HSPC. In some embodiments, a gene-edited HSC or HPC

administered for treatment of a hemoglobinopathy comprises a gene-edit within the HBB locus for correction of a mutation.

[0186] Also provided herein are methods of gene-editing for inducing a mutation within a target gene in a cell or population of cells, such as a quiescent cell or population of quiescent cells induced to divide, by repair of a DNA DSB in the target gene by the HDR pathway using a donor polynucleotide. In some embodiments, a mutation in a target gene is an insertion of a trans-gene. In some embodiments a trans-gene is a chimeric antigen receptor (CAR). In some embodiments, T cells are engineered (e.g., gene-edited) using methods of gene-editing described herein to express a CAR. In some embodiments, a T cell is engineered by introducing a DSB within a target gene that is repaired by HDR with a donor polynucleotide encoding a chimeric antigen receptor (CAR). In some embodiments, a CAR is selected for a cancer of interest wherein the cancer expresses an antigen recognized by the CAR. Non-limiting examples of antigen recognized by CARs include CD19, CD33, CD70, BCMA, CD22, CD20, CD138, CD123, Lewis Y antigen, and inactive tyrosine protein kinase transmembrane receptor ROR1.

[0187] In some embodiments, the methods comprise delivering engineered T cells to a patient comprising a chimeric antigen receptor (CAR). In some embodiments, a population of CAR T cells is administered to a patient with cancer wherein the CAR is specific to an antigen expressed by the cancer. Non-limiting examples of cancers that may be treated as provided herein include multiple myeloma, leukemia (e.g., T cell leukemia, B-cell acute lymphoblastic leukemia (B-ALL), and/or chronic lymphocytic leukemia (C-CLL)), lymphoma (e.g., B-cell non-Hodgkin's lymphoma (B-NHL), Hodgkin's lymphoma, and/or T cell lymphoma), and/or clear cell renal cell carcinoma (ccRCC). Other non-limiting examples of cancers (e.g., solid tumors) that may be treated as provided herein include pancreatic cancer, gastric cancer, ovarian cancer, cervical cancer, breast cancer, renal cancer, thyroid cancer, nasopharyngeal cancer, non-small cell lung (NSCLC), glioblastoma, and/or melanoma.

Engineered Hematopoietic Stem Cells

[0188] In some embodiments, stem cells are engineered (e.g., gene-edited) using methods of the disclosure. In some embodiments, stem cells are engineered to correct a gene mutation and/or replace a target gene. In some embodiments, engineered stem cells are administered to a patient for treatment of a monogenic disease. In some embodiments, a stem cell (e.g., a stem cell) comprises an HSC. HSCs are defined by their pluripotency (e.g., capacity of a single HSC to generate any type of blood cell) and ability to self-renew. HSCs are comprised of two populations: short-term HSCs and long-term HSCs. Short term HSCs are capable of self-renewal for a short period of time, while LT-HSPCs are capable of indefinite self-renewal. LT-HSPCs are largely in a quiescent state, dividing only once every 145 days (Wilson, A. et al. (2008) *Cell* 135:1118-1129). In some embodiments, an HSC divides asymmetrically wherein one daughter cell remains in a stem state and one daughter cell expresses a distinct function or phenotype. In some embodiments, an HSC divides symmetrically wherein both daughter cells retain a stem state.

[0189] Early descendants of an HSC are termed hematopoietic progenitor cells. Hematopoietic progenitor cells (HPCs) retain the ability to differentiate into other cell types, but are not capable of self-renewal. In some embodiments, progenitor cells of an HSC are differentiated cells. In some embodiments, progenitor cells of an HSC comprise the same differentiation state. In some embodiments, progenitor cells of an HSC comprise different differentiation states. In some embodiments, progenitor cells of an HSC are lineage restricted precursor cells (e.g., a common myeloid progenitor cell, a common lymphoid progenitor cell). In some embodiments, lineage restricted precursor cells further differentiate. In some embodiments, an HSC differentiates into a common lymphoid progenitor cell that further differentiates into cell types comprising B cells, natural killer (NK) cells, and T cells. In some embodiments, an HSC differentiates into a common myeloid progenitor cell that further differentiates into cell types comprising dendritic cells (DCs), monocytes, myeloblasts, monocyte-derived DCs, macrophages, neutrophils, eosinophils, basophils, megakaryocyte-erythroid progenitor cells, erythrocytes, megakaryocytes, and platelets.

[0190] In some embodiments, an HSC of the disclosure has positive expression for the cell surface marker CD34. In some embodiments, an HSC of the disclosure has positive expression for cell surface markers comprising CD38, CD45RA, CD90, c-Kit tyrosine kinase receptor, stem cell antigen-1 (Sca-1), CD133 and CD49f. In some embodiments, an HSC of the disclosure has negative or low expression for cell surface markers comprising CD38, CD45RA, CD90, Thy-1.1 cell surface antigen and CD49f. In some embodiments, an HSC of the disclosure has negative or low expression of lineage cell surface markers comprising CD2, CD3, CD11b, CD11c, CD14, CD16, CD19, CD24, CD56, CD66b, CD235. In some embodiments, an HSC of the disclosure is an LT-HSC. In some embodiments, an LT-HSC has negative or low expression of lineage cell surface markers comprising CD2, CD3, CD11b, CD11c, CD14, CD16, CD19, CD24, CD56, CD66b, CD235. In some embodiments, an LT-HSC has negative or low expression of cell surface markers comprising CD45RA and CD38. In some embodiments, an LT-HSC has positive expression for cell surface markers comprising CD34 and CD90.

[0191] Methods for isolation of HSCs are known in the art as taught by U.S. Pat. Nos. 5,643,741, 5,087,570, 5,677,136, 7,790,458, 10,006,004, 10,086,045, 7,939,057, 10,058,573 that are incorporated by reference herein. In some embodiments, a population of cells comprising HSCs is derived from the patient (e.g., an autologous HSC). In some embodiments, a population of cells comprising HSCs is derived from a healthy donor (e.g., an allogenic HSC). In some embodiments, a population of cells comprising HSCs is derived from human cord blood. In some embodiments, a population of cells comprising HSCs is derived from bone marrow. In some embodiments, a population of cells comprising HSCs is derived from human peripheral blood.

[0192] In some embodiments, a population of cells comprising HSCs is derived following treatment of a subject (e.g., a patient, a healthy donor) with a stem cell mobilizer. In some embodiments, a stem cell mobilizer comprises a CXCR4 antagonist. The chemokine stromal cell derived factor-1 (e.g., CXCL12) is a chemokine that binds to CXCR4 on HSCs and HPCs and signals for retention in the bone marrow. By blocking this interaction with a CXCR4

antagonist, HSCs and HPCs rapidly mobilize to the blood (Broxmeyer, et al. (2005) *J. Exp Med* 18:1307-1318; Devine, S. et al (2008) *Blood* 112:990-998). Non-limiting examples of a CXCR4 antagonist include TG-0054 (TaiGen Biotechnology, Co., Ltd. (Taipei, Taiwan)), AMD3465, AMD3100 (e.g., wherein AMD or AMD3100 is used interchangeably with plerixafor, rINN, USAN, JM3100, and its trade name, Mozobil™, see U.S. Pat. Nos. 6,835,731 and 6,825,351), and NIBR1816 (Novartis, Basel, Switzerland). In some embodiments, a stem-cell mobilizer is plerixafor.

[0193] In some embodiments, a stem cell mobilizer comprises a colony stimulating factor. Non-limiting examples of a colony stimulating factor include, but are not limited to, granulocyte colony stimulating factor (G-CSF), granulocyte-macrophage colony stimulating factor (GM-CSF), macrophage colony stimulating factor (M-CSF), stem cell factor (SCF), FLT-3 ligand, or a combination thereof. Use of G-CSF as a stem cell mobilizing factor has demonstrated increased yield of stem cells from peripheral blood (Morton, et al (2001) *Blood* 98:3186; Smith, T. et al. (1997) *J. Clin. Oncol.* 15:5-10). In some embodiments, a stem cell mobilizer is a combination of a CXCR4 antagonist and a colony stimulating factor. In some embodiments, a stem cell mobilizer is a combination of Plerixafor and G-CSF.

[0194] In some embodiments, CD34+ HSCs are enriched following isolation from a subject (e.g., a patient, a healthy donor). In some embodiments, CD34+ HSCs are enriched from human blood, bone marrow, or cord blood. Methods of enriching CD34+ HSCs are known in the art. In some embodiments, CD34+ HSCs are enriched using a magnetic cell separator. In some embodiments, CD34+ HSCs are enriched by fluorescent activated cell sorting (FACS). In some embodiments, CD34+ HSCs are enriched by magnetic bead sorting for cells expressing CD34.

[0195] In some embodiments, an enriched population of CD34+ cells has a purity of at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 100%. In some embodiments, an enriched population of CD34+ cells has a purity of at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 90%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 91%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 92%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 93%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 94%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 95%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 96%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 97%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 98%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 99%. In some embodiments, an enriched population of CD34+ cells has a purity of at least about 100%.

[0196] In some embodiments, an enriched population of CD34+ cells comprises LT-HSPCs. In some embodiments, the proportion of the CD34+ population that is LT-HSPCs is 0.01-0.05%, 0.01-0.1%, 0.05-0.1%, 0.05-1%, 0.1-0.5%, 0.1-

0.7%, 0.1-1.0%, 0.1-1.5%, 0.1-2.0%, 0.5-1.5%, 0.5-2.0%, or 1-2%. In some embodiments, the proportion of the CD34+ population that is LT-HSPCs is 0.05-1%. In some embodiments, the proportion of the CD34+ population that is LT-HSPCs is 0.1-1%. In some embodiments, the proportion of the CD34+ population that is LT-HSPCs is 0.1-2%. In some embodiments, the proportion of the CD34+ population that is LT-HSPCs is at least about 0.01%, at least about 0.05%, at least about 0.1%, at least about 0.2%, at least about 0.3%, at least about 0.4%, at least about 0.5%, at least about 0.6%, at least about 0.7%, at least about 0.8%, at least about 0.9%, or at least about 1.0% of the population.

[0197] In some embodiments, gene-editing of HSCs is performed prior to enrichment of CD34+ HSCs. In some embodiments, gene-editing of HSCs is performed following enrichment of CD34+ HSCs. In some embodiments, following gene-editing, a method is used to selected for gene-edited HSCs from a population comprising CD34+ HSCs. In some embodiments, a method of isolating gene-edited HSCs comprises enrichment of HSCs expressing truncated nerve growth factor (tNGFR) as described in the art (Dever et al (2016) *Nature* 539:384-389).

[0198] For ex vivo therapy, transplantation requires clearance of bone-marrow niches for donor HSCs to engraft. Methods are known in the art for depletion of the bone-marrow niche, including methods of treating with radiation, chemotherapy or a combination thereof.

Engineered Induced Pluripotent Stem Cells

[0199] In some embodiments, genetically engineered human cells of the disclosure are derived from induced pluripotent stem cells (iPSCs). iPSCs are reprogrammed from somatic cells to a pluripotent state wherein they can differentiate into all three germ layers. An advantage of using iPSCs is that the cell can be derived from the same subject to which the progenitor cells are to be administered. That is, a somatic cell can be obtained from a subject, reprogrammed to an iPSC, and then re-differentiated into a progenitor cell to be administered to the subject for treatment of a disorder (e.g., an autologous progenitor). Since the progenitors are derived from an autologous source, the risk of engraftment rejection or allergic responses is reduced compared to the use of cells from another subject or group of subjects. Thus, an iPSC can be gene-edited and reintroduced into a patient for correction of a disease resulting from a somatic genetic mutation.

[0200] Briefly, human iPSCs can be obtained by transducing somatic cells with stem cell associated transcription factors that include OCT4, SOX2, and NANOG (Budnizky et al. (2014) *Stem Cells Transl Med* 3:448-457; Barret et al. *Stem Cells Trans Med* (2014) 3:1-6; Focosi et al. (2014) *Blood Cancer Journal* 4:e211). Exemplary methods for reprogramming somatic cells to generate iPSCs are known in the art as described by US 2019/0038771 which is incorporated by reference herein.

Engineered T Cells

[0201] In some embodiments, engineered (gene edited) CAR T cells are autologous ("self"). In some embodiments, engineered CAR T cells are non-autologous ("non-self," e.g., allogeneic, syngeneic or xenogeneic). "Autologous" refers to cells from the same subject. "Allogeneic" refers to cells of the same species as a subject, but that differ

genetically to the cells in the subject. In some embodiments, the T cells are obtained from a mammal. In some embodiments, the T cells are obtained from a human.

[0202] T cells can be obtained from a number of sources including, but not limited to, peripheral blood mononuclear cells, bone marrow, lymph nodes tissue, cord blood, thymus issue, tissue from a site of infection, ascites, pleural effusion, spleen tissue, and tumors. In certain embodiments, T cells can be obtained from a unit of blood collected from a subject using any number of techniques known to the skilled person, such as sedimentation, e.g., FICOLL™ separation.

[0203] In some embodiments, an isolated population of T cells is used. In some embodiments, after isolation of peripheral blood mononuclear cells (PBMC), both cytotoxic and helper T lymphocytes can be sorted into naive, memory, and effector T cell subpopulations either before or after activation, expansion, and/or genetic modification.

[0204] A specific subpopulation of T cells, expressing one or more of the following cell surface markers: TCRab, CD3, CD4, CD8, CD27, CD28, CD38, CD45RA, CD45RO, CD62L, CD127, CD122, CD95, CD197, CCR7, KLRG1, MCH-I proteins and/or MCH-II proteins, can be further isolated by positive or negative selection techniques. In some embodiments, a specific subpopulation of T cells, expressing one or more of the markers selected from the group consisting of TCRab, CD4 and/or CD8, is further isolated by positive or negative selection techniques. In some embodiments, the engineered T cell populations do not express or do not substantially express one or more of the following markers: CD70, CD57, CD244, CD160, PD-1, CTLA4, HM3, and LAG3. In some embodiments, subpopulations of T cells may be isolated by positive or negative selection prior to genetic engineering and/or post genetic engineering.

[0205] In some embodiments, an isolated population of T cells expresses one or more of the markers including, but not limited to a CD3+, CD4+, CD8+, or a combination thereof. In some embodiments, the T cells are isolated from a donor, or subject, and first activated and stimulated to proliferate in vitro prior to undergoing gene editing.

[0206] To achieve sufficient therapeutic doses of T cell compositions, T cells are often subjected to one or more rounds of stimulation, activation and/or expansion. T cells can be activated and expanded generally using methods as described, for example, in U.S. Pat. Nos. 6,352,694; 6,534,055; 6,905,680; 6,692,964; 5,858,358; 6,887,466; 6,905,681; 7,144,575; 7,067,318; 7,172,869; 7,232,566; 7,175,843; 5,883,223; 6,905,874; 6,797,514; and 6,867,041. In some embodiments, T cells are activated and expanded for about 1 day to about 4 days, about 1 day to about 3 days, about 1 day to about 2 days, about 2 days to about 3 days, about 2 days to about 4 days, about 3 days to about 4 days, or about 1 day, about 2 days, about 3 days, or about 4 days prior to introduction of the genome editing compositions into the T cells.

[0207] In some embodiments, T cells are activated and expanded for about 4 hours, about 6 hours, about 12 hours, about 18 hours, about 24 hours, about 36 hours, about 48 hours, about 60 hours, or about 72 hours prior to introduction of the gene editing compositions into the T cells.

[0208] In some embodiments, T cells are activated at the same time that genome editing compositions are introduced into the T cells. T cell populations or isolated T cells

generated by any of the gene editing methods described herein are also within the scope of the present disclosure.

Compositions of Cells and Methods of Inducing Cell Expansion

[0209] The methods of the disclosure enable introduction of a gene-edit to a cell or a population of cells, such as a quiescent cell that has been induced to divide or a population of quiescent cells that has been induced to divide, by repair of a DNA DSB in a target gene by the HDR pathway. In some embodiments, a gene-edit is introduced to a population comprising stem cell. In some embodiments, a gene-edit is introduced to a population comprising stem cells derived from a human tissue. Methods for deriving cultures of adult stem cells have been described, examples including stem cells derived from tissues such as the nervous system (McKay (1997) *Science* 276:66-71; Shihabuddin (1999) *Mol. Med Today* 5:474-480), bone marrow (Pittenger, et al. (1999) *Science* 284:143-147; Pittenger (2001) In: *Mesenchymal stem cells of human adult bone marrow*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 349-374), adipose tissue (Gronthos (2001) *J Cell Physiol* 189: 54-63), dermis (Toma (2001) *Nature Cell Biol* 3:778-784); pancreas and liver (Deutsch (2001) *Development* 128:871-881).

[0210] In some embodiments, culture of stem cells induces the stem cells to divide. In some embodiments, culture of stem cells maintains the stem cells in a pluripotent state. In some embodiments, culture of stem cell induces differentiation.

Hematopoietic Stem Cells

[0211] Methods of cell culture that are used for ex vivo expansion of HSC and other stem cells are described by U.S. Pat. Nos. 5,670,351; 5,851,984; 6,030,836; 7,790,458; 5,817,773; 8,895,299, 10,058,573, 9,943,545 which are incorporated by reference herein.

[0212] In some embodiments, a method of inducing ex vivo expansion (e.g., cell division) of a population of HSCs comprises stimulating the Wnt signaling pathway. The Wnt signaling cascade is important for promoting self-renewal of adult stem cells, and has been described in promoting self-renewal of HSCs and proliferation of progenitors (Staal (2016) *Expt Hematol* 44:451-457). Methods of inducing Wnt signaling are known in the art. Wnt signaling can be induced by treatment with a wnt5a protein (Nemeth, et al (2007) *Proc Natl Acad Sci* 104:15436-15441), by treatment with a wnt3a protein (Willert (2003) *Nature* 423:448), by inducing constitutive expression of β -catenin (Reya (2003) *Nature* 423:409), by treatment with prostaglandin E2 (Goessling (2009) *Cell* 136:1136-1147; Goessling (2011) 8:445-458), and by treatment with a glycogen synthase kinase 3 inhibitor in combination with a rapamycin inhibitor (Huang (2012) *Nat Med* 18:1778). In some embodiments, a method of stimulating Wnt signaling is used to promote ex vivo expansion of HSCs. In some embodiments, a method of stimulating Wnt signaling is titrated to prevent exhausting of HSCs in culture. In some embodiments, a method of stimulating Wnt signaling is titrated to promote reconstitution of HSCs in an irradiated patient.

[0213] In some embodiments, a method of inducing ex vivo expansion (e.g., cell division) of a population of HSCs comprises stimulating the Notch signaling pathway. The

Notch signaling cascade is important for controlling cell fate decisions, development, and hematopoiesis (Andersson (2011) *Development* 138:3593-3612). The Notch signaling cascade has also been demonstrated to play a role in HSC self-renewal (Kunisato (2003) *Blood* 101:1777-1783; Varnum-Finney (2003) *Blood* 101:1784-1789). Notch signaling is induced by a family of ligands comprising a Delta, Serrate, and Lag2 domain (e.g., DSL ligands). In some embodiments, a soluble DSL ligand is used for stimulation of Notch signaling.

[0214] In some embodiments, a method of inducing ex vivo expansion (e.g., cell division) of a population of HSCs is to culture in growth media comprising a cytokine. In some embodiments, a cytokine is selected that will bind to an HSC. In some embodiments, a cytokine is selected that will regulate HSC function, including quiescence, self-renewal, differentiation, apoptosis, and mobility. In some embodiments, a cytokine is stem cell factor (SCF). SCF is a cytokine expressed by a number of cell types. SCF binds to c-Kit, a tyrosine kinase receptor expressed on all HSCs, and prevents apoptosis of HSCs. SCF has been shown to potentiate the ability of HSCs to undergo symmetric self-renewal wherein both daughter cells of a dividing HSC retain stem-like properties (Bowe, M. (2007) *Blood* 109:5043-5048). In some embodiments, a cytokine is thrombopoietin (TPO). TPO is a cytokine that binds to the Mpl receptor expressed on HSCs. TPO promotes the survival of repopulating HSCs in vitro (Matsunaga (1998) *Blood*, 92:452-461). In some embodiments, a cytokine is interleukin-3 (IL-3). In some embodiments, a cytokine is interleukin-6 (IL-6).

[0215] In some embodiments, a method of inducing ex vivo expansion (e.g., cell division) of a population of HSCs is to culture in growth media comprising a growth factor. In some embodiments, a growth factor is Fms-like tyrosine kinase 3 (Flt3) ligand. Flt3 ligand is a growth factor that promotes HSC proliferation, differentiation, and survival (Hannum (1994) *Nature* 368:643-648). In some embodiments, a growth factor is a fibroblast growth factor (FGF). Both FGF-1 and FGF-2 support HSC expansion in vitro (de Haan (2003) *Dev Cell* 4:241-251; Yeoh (2006) *Stem Cells* 24:1564-1572). In some embodiments, a growth factor is a member of the angiopoietin (Ang) family comprising Ang1, Ang2, and Ang 4 in humans. In some embodiments, a growth factor is an Ang-like protein (Angptl). Non-limiting examples include Angptl7, Angptl2, Angptl3, Angptl5, and Mfap4. In some embodiments, a growth factor is insulin-like growth factor 2.

[0216] In some embodiments, a method of inducing ex vivo expansion (e.g., cell division) of a population of HSCs is to culture in growth media comprising a molecule identified through a high-throughput screen for HSC proliferation such as those described in the art (Boitano, *Science* (2010) 329:1345-1348; Wagner (2016) *Cell Stem Cell* 18:144-155; Fares (2014) *Science* 345:1509-1512). In some embodiments, a molecule is prostaglandin E2 (PGE2). In some embodiments, a molecule is Stemregenin 1 (SR1), an aryl hydrocarbon receptor antagonist. In some embodiments, a molecule is UM171.

[0217] In some embodiments, a method of inducing ex vivo expansion (e.g., cell division) of a population of HSCs is to culture in growth media comprising a regulator of epigenetic changes. Epigenetic regulation of DNA methylation and post-translational modifications is important for HSC cell fate decisions (Hodges (2011) *Mol Cell* 44:17-28;

Kulis (2015) *Nat Genet* 47:746). In some embodiments, a small molecule inhibitor of histone deacetylase (HDAC) is used to support ex vivo expansion of HSCs. In some embodiments, a HDAC inhibitor is valproic acid. In some embodiments, a HDAC inhibitor is trichostatin A. In some embodiments, a small molecule inhibitor of DNA methyltransferase is used to support ex vivo expansion of HSCs. In some embodiments, a DNA methyltransferase inhibitor is decitabine.

[0218] In some embodiments, a method of inducing ex vivo expansion (e.g., cell division) of a population of HSCs is to culture in growth media comprising a combination of factors. In some embodiments, a combination comprises one or more cytokines (e.g., IL-3, SCF, TBP) and one or more growth factors (e.g., Flt3 ligand, Ang protein, Insulin-like growth factor). In some embodiments, a combination comprises a combination of IL-3 and SCF. In some embodiments, a culture of HSCs comprises a combination of IL-3, SCF, and TBP. In some embodiments, a culture of HSCs comprises a combination of IL-3, SCF, TBP, and Flt3 ligand.

[0219] In some embodiments, a factor for inducing HSC ex vivo expansion (e.g., cell division) is combined with a factor that promotes differentiation. Non-limiting examples of factors that promote differentiation to myeloid progenitor cells include GM-CSF, G-CSF, and M-CSF. Non-limiting examples of factors that promote differentiation to lymphoid progenitor cells, including dendritic cell, B cell, T cell, and NK cell progenitors include IL-3, IL-4, IL-7, IL-11, IL-12, IL-15, GM-CSF, and TNF α .

[0220] In some embodiments, ex vivo expansion (e.g., cell division) of HSCs is performed by co-culture with a supportive niche comprising a stromal cell lines established from fetal or adult hematopoietic organs. Methods for co-culture with a stromal cell line have been described in the art (Moore (1997) *Blood* 89:4337-4347; Weisel (2006) *Exp Hematol* 34:1505-1516; Yoder (1995) *Blood* 86:1322-1330).

[0221] In some embodiments, a method of inducing ex vivo expansion (e.g., cell division) of a population of HSCs is modification of environmental factors used for incubation during culture. In some embodiments, an environmental factor is temperature. In some embodiments, a culture of HSCs is maintained at about 32° C., at about 33° C., at about 34° C., at about 35° C., at about 36° C., at about 37° C., at about 38° C., at about 39° C., or at about 40° C. In some embodiments, a culture of HSCs is maintained at about 37° C. In some embodiments, an environmental factor is oxygen level. In some embodiments, a culture of HSCs is maintained under oxygen levels comparable to physiological oxygen levels (e.g., normoxic culture). In some embodiments, a culture of HSCs is maintained under oxygen levels lower than physiological oxygen levels (e.g., hypoxic culture). Methods of identifying normoxic or hypoxic conditions are known in the art (Wenger (2015) *Hypoxia* 3:35-43).

[0222] In some embodiments, a population comprising HSCs is induced to expand ex vivo and gene-edited using a method of the disclosure. In some embodiments, a cell is cultured under conditions that induce cell-division at least 1-3 days, 1-5 days, 2-7 days, 5-15 days, or 5-20 days prior to gene-editing. In some embodiments, a cell is cultured under conditions that induce cell-division at least 1 day, at least 2 days, at least 3 days, at least 4 days, at least 5 days, at least 6 days, or at least 7 days prior to gene-editing. In some embodiments, a cell is cultured under conditions that induce cell-division at least 2 days prior to gene-editing. In

some embodiments, a cell is cultured under conditions that induce cell-division at least 3 days prior to gene-editing. In some embodiments, a cell is cultured under conditions that induce cell-division at least 4 days prior to gene-editing. In some embodiments, a gene-edited cell is cultured under conditions that induce cell division at least at least 1 day, at least 2 days, at least 3 days, at least 4 days, at least 5 days, at least 6 days, or at least 7 days prior to transplantation. In some embodiments, a gene-edited cell is cultured under conditions that induce cell division at least 2 days prior to transplantation. In some embodiments, a gene-edited cell is cultured under conditions that induce cell division at least 3 days prior to transplantation. In some embodiments, a gene-edited cell is cultured under conditions that induce cell division at least 4 days prior to transplantation.

Methods of Measuring Cell Quiescence

[0223] In some embodiments, methods of the disclosure are used for gene-editing of a quiescent cell or a population comprising quiescent cells. Methods of defining the phase of the cell cycle for a given cell in a population are described in the art (e.g., Nakamura-Ishizu (2014) Development 141:4656-4666) and are described briefly herein.

[0224] As used herein, the term quiescent state refers to the reversible state of a cell in which it does not divide, but retains the ability to re-enter the cycle of cell proliferation at some later time. A quiescent state is distinct from a state of terminal differentiation termed senescence wherein a cell has irreversibly exited the cycle of cell proliferation. A quiescent state occurs when a cell is in the G0 phase of the cell cycle. The G0 phase can be considered as an extended G1 phase. The G1 phase is the period between the end of a mitotic phase and the beginning of S phase, wherein DNA replication occurs. A quiescent state is defined by growth cessation. A cell may enter a quiescent state due to conditions of nutrient deprivation or high population density (Cheung, T. et al. (2013) *Nat Rev Mol Cell Biol* 14:329-340). A cell may also enter a quiescent state to preserve cell homeostasis and the ability to regenerate. A cell can change from a quiescent state by responding to an external stimuli and re-entering the cell cycle, undergoing cell differentiation, or entering a senescent state.

[0225] Several markers can be used to determine if a cell is in a quiescent state. One marker of a quiescent state is low DNA or RNA content (Huttmann, A. et al. (2001) *Exp Hematol.* 29:1109-1116 and Fukada, S. et al. (2007) *Stem Cells* 25:2448-2459). Another marker of a quiescent state is low abundance of cell proliferation markers (Geres, J. et al. (1983) *Int J Cancer* 31:13-20). Another marker of a quiescent state is retention of an exogenously administered marker that is incorporated into the cell, for example, 5'-bromo-2'-deoxyuridine (BrdU), tritiated thymidine, H2B-GFP, H2B-YFP. Retention of a cell marker is indicative of low turnover.

[0226] In some embodiments, a measure of DNA content in a cell is used to determine a cell's position in the cell cycle. The phase of the cell cycle is determined by measuring DNA content of a cell using a DNA-binding dye and flow cytometry as described in the art (e.g., Darzynkiewicz (2011) *Curr Protoc Cytom Unit* 7.2; Darzynkiewicz (2004) *Cytometry* 58A:21-32). Non-limiting examples of a DNA-binding dye include Hoechst 33342, Hoechst 33258, 4',6-diamidino-2-phenylindole (DAPI), propidium iodide, 4',6-diamidino-2-phenylindole, Nuclear Green, Nuclear Red,

and 7-aminoactinomycin D (7-AAD), propidium iodide, Vybrant DyeCycle stain FxCycle stains, SYTOX Green. Such assays are based upon the understanding that DNA content is different for pre-replicative phase cells (e.g., G0 or G1 phase), cells that are replicating (e.g., S phase), and cells that are post-replication (e.g., G2 and M phase). DNA content is defined using a "DNA-index". While cells that are in G0 or G1 phase have a DNA-index of 1.0, cells in the G2 and M phase have a higher DNA-index of 2.0. Cells that are in the S-phase have an intermediate DNA-index between 1.0 and 2.0. The result of a cellular DNA content measurement is often presented in the form of a frequency histogram. Discrimination of cells in particular phases of the cell cycle and their quantification, is based on difference in DNA content (e.g., deconvolution of the histogram). Thus, the proportion of a population that is in a given phase of the cell cycle can be determined.

[0227] In some embodiments, DNA content is measured relative to a proliferation-associated protein. Non-limiting examples of a proliferation-associated protein include cyclin D, cyclin E, cyclin, Ki-67, and cyclin B7. Measuring the ratio DNA content relative to a proliferation-associated protein can be used to determine the proportion of cells in the G0 phase relative to the G1 phase of the cell cycle (e.g., Kim, et al (2015) *Curr Protoc Mol Biol* 111:28.6.1-28.6.11). In some embodiments, DNA content is measured relative to RNA content. Measuring the ratio of DNA to RNA content can also be used to quantify cells during the G0 phase of the cell cycle relative to cells in the G1 phase of the cell cycle (e.g., Gothot (1997) *Blood* 90:4384-4393). In some embodiments, DNA content is measured relative to RNA content, wherein RNA content is measured using a dye that labels RNA in a cell. Non-limiting examples of a RNA-binding dye include SYTORNA Select, fluorescent styryl dyes, and Pyronin Y.

[0228] In some embodiments, a label incorporation-based assay is used to determine a cell's position in the cell cycle. A label incorporation-based assay comprises measuring cell proliferation using a DNA label. Non-limiting examples of labels incorporated into DNA are bromodeoxyuridine (BrdU), 5-ethynyl-2'-deoxyuridine (EdU) and tritiated thymidine (³H-thymidine or ³H-TdR). Cells that are actively proliferating will incorporate a DNA label into newly synthesized DNA strands. By combining a measure of cell proliferation using a label incorporated into DNA with a dye to measure DNA content, the cell cycle phase can be resolved (Cecchini (2012) *J Vis Exp* 59:3491). In some embodiments, a method of measuring the percentage of cells in S-phase comprises labeling cells with EdU as described by Pereira, et al. (2017) *Oncotarget* 8:40514-40532.

[0229] In some embodiments, a dye dilution assay is used to measure cell proliferation. A dye dilution assay comprises labeling cells with a fluorescent covalent dye wherein the dye is covalently attached to cellular proteins following labeling. During mitosis, the dye is split evenly between daughter cells and cellular fluorescence is reduced by half with each cell cycle. Thus, the level of dye fluorescence measured by a method of fluorescence detection is used to determine the number of cell divisions that occur following labeling. In some embodiments, a covalent dye is carboxyfluorescein diacetate (CFSE). In some embodiments, CFSE dilution is measured to determine number of cell cycles.

[0230] In some embodiments, a population of the disclosure comprises non-dividing cells, wherein non-dividing cells are measured as cells in the G0 or G1 phase of the cell cycle. In some embodiments, a population comprises about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90% or about 100% cells that are in the G0 or G1 phase of the cell cycle. In some embodiments, a population of the disclosure comprises quiescent cells, wherein quiescent cells are measured as cells in the G0 phase of the cell cycle. In some embodiments, a population comprises about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90% or about 100% cells that are in the G0 phase of the cell cycle.

[0231] In some embodiments, a population of the disclosure comprises non-dividing cells (e.g., G0 or G1 phase), wherein non-dividing cells of the population are induced to divide. In some embodiments, an extrinsic signal is provided to induce the cells to divide. In some embodiments, upon addition of an extrinsic signal, the proportion of cells that are non-dividing decreases by about 5%, by about 10%, by about 20%, by about 30%, by about 40%, by about 50%, by about 60%, by about 70%, by about 80%, by about 90%, or by about 100%. In some embodiments, a population of the disclosure comprises quiescent cells (e.g., G0 phase), wherein quiescent cells of the population are induced to divide. In some embodiments, an extrinsic signal is provided to induce the cells to divide. In some embodiments, upon addition of an extrinsic signal, the proportion of cells that are quiescent decreases by about 5%, by about 10%, by about 20%, by about 30%, by about 40%, by about 50%, by about 60%, by about 70%, by about 80%, by about 90%, or by about 100%.

Genome Editing

[0232] Genome editing generally refers to the process of editing or changing the nucleotide sequence of a genome, preferably in a precise, desirable and/or pre-determined manner. Examples of compositions, systems, and methods of genome editing described herein use of site-directed nucleases to cut or cleave DNA at precise target locations in the genome, thereby creating a double-strand break (DSB) in the DNA. Such breaks can be repaired by endogenous DNA repair pathways, such as homology directed repair (HDR) and/or non-homologous end-joining (NHEJ) repair (see e.g., Cox et al., (2015) *Nature Medicine* 21 (2):121-31). One of the major obstacles to efficient genome editing in non-dividing cells is lack of homology directed repair (HDR). Without HDR, non-dividing cells rely on non-homologous end joining (NHEJ) to repair double-strand breaks (DSB) that occur in the genome. The results of NHEJ-mediated DNA repair of DSBs can include correct repair of the DSB, or deletion or insertion of one or more nucleotides or polynucleotides.

Donor Polynucleotides

[0233] The disclosure provides donor polynucleotides that, upon insertion into a DSB, correct or induce a mutation in a target nucleic acid (e.g., a genomic DNA). In some embodiments, the donor polynucleotides provided by the disclosure are recognized and used by the HDR machinery of a cell to repair a double strand break (DSB) introduced into a target nucleic acid by a site-directed nuclease, wherein

repair of the DSB results in the insertion of the donor polynucleotide into the target nucleic acid. In some embodiments, the donor polynucleotides provided by the disclosure are recognized and used by the HDR machinery of a cell to repair a double strand break (DSB) introduced into a target nucleic acid (e.g., HBB gene) by a site-directed nuclease, wherein the region proximal to the DSB is exchanged for the corresponding region provided by the donor polynucleotide. Alternatively, a donor polynucleotide may have no regions of homology to the targeted location in the DNA and may be integrated by NHEJ-dependent end joining following cleavage at the target site.

[0234] A donor template can be DNA or RNA, single-stranded and/or double-stranded, and can be introduced into a cell in linear or circular form. 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.

[0235] A donor template 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. Moreover, a donor template 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 and integrase defective lentivirus (IDLV)).

[0236] A donor template, in some embodiments, is inserted so that its expression is driven by the endogenous promoter at the integration site, namely the promoter that drives expression of the endogenous gene into which the donor is inserted. In some embodiments, a donor template is integrated so that its expression is driven by the endogenous promoter at the integration site, namely the promoter that drives expression of the endogenous gene into which the donor is exchanged. However, in some embodiments, the donor template comprises an exogenous promoter and/or enhancer, for example a constitutive promoter, an inducible promoter, or tissue-specific promoter. In some embodiments, the exogenous promoter is an EF1a promoter comprising a sequence of SEQ ID NO: 59. Other promoters may be used.

[0237] Furthermore, exogenous sequences may also include transcriptional or translational regulatory sequences, for example, promoters, enhancers, insulators, internal ribosome entry sites, sequences encoding 2A peptides and/or polyadenylation signals.

[0238] In some embodiments, the donor polynucleotides comprise a nucleotide sequence which corrects or induces a mutation in a genomic DNA (gDNA) molecule in a cell, wherein when the donor polynucleotide is introduced into the cell in combination with a site-directed nuclease, a HDR DNA repair pathway inserts the donor polynucleotide into a double-stranded DNA break (DSB) introduced into the

gDNA by the site-directed nuclease at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the donor polynucleotides comprise a nucleotide sequence which corrects or induces a mutation in a genomic DNA (gDNA) molecule in a cell, wherein when the donor polynucleotide is introduced into the cell in combination with a site-directed nuclease, a HDR DNA repair pathway exchanges a region proximal to a double-stranded DNA break (DSB) for the corresponding region provided by the donor polynucleotide, by the site-directed nuclease at a location proximal to the mutation, thereby correcting the mutation.

[0239] In some embodiments, the donor polynucleotide comprises a nucleotide sequence which corrects or induces a mutation, wherein the nucleotide sequence that corrects or induces a mutation comprises a single nucleotide. In some embodiments, the nucleotide sequence which corrects or induces a mutation comprises two or more nucleotides. In some embodiments, the nucleotide sequence which corrects or induces a mutation comprises a codon. In some embodiments, the nucleotide sequence which corrects or induces a mutation is comprises one or more codons. In some embodiments, the nucleotide sequence which corrects or induces a mutation comprises an exonic sequence. In some embodiments, the donor polynucleotide comprises a nucleotide sequence which corrects or induces a mutation, wherein the nucleotide sequence which corrects or induces a mutation comprises an intronic sequence. In some embodiments, the nucleotide sequence which corrects or induces a mutation comprises all or a portion of an exonic sequence. In some embodiments, the nucleotide sequence which corrects or induces a mutation comprises all or a portion of an intronic sequence. In some embodiments, the nucleotide sequence which corrects or induces a mutation comprises all or a portion of an exonic sequence and all or a portion of an intronic sequence.

[0240] In some embodiments, the donor polynucleotide sequence is identical to or substantially identical to (having at least one nucleotide difference) an endogenous sequence of a target nucleic acid. In some embodiments, the endogenous sequence comprises a genomic sequence of the cell. In some embodiments, the endogenous sequence comprises a chromosomal or extrachromosomal sequence. In some embodiments, the donor polynucleotide sequence comprises a sequence that is substantially identical (comprises at least one nucleotide difference/change) to a portion of the endogenous sequence in a cell at or near the DSB. In some embodiments, repair of the target nucleic acid molecule with the donor polynucleotide results in an insertion, deletion, or substitution of one or more nucleotides of the target nucleic acid molecule. In some embodiments, the insertion, deletion, or substitution of one or more nucleotides results in one or more amino acid changes in a protein expressed from a gene comprising the target sequence. In some embodiments, the insertion, deletion, or substitution of one or more nucleotides results in one or more nucleotide changes in an RNA expressed from the target gene. In some embodiments, the insertion, deletion, or substitution of one or more nucleotides alters the expression level of the target gene. In some embodiments, the insertion, deletion, or substitution of one or more nucleotides results in increased or decreased expression of the target gene. In some embodiments, the insertion, deletion, or substitution of one or more nucleotides results in gene knockdown. In some embodiments, the insertion, dele-

tion, or substitution of one or more nucleotides results in gene knockout. In some embodiments, the repair of the target nucleic acid molecule with the donor polynucleotide results in replacement of an exon sequence, an intron sequence, a transcriptional control sequence, a translational control sequence, a sequence comprising a splicing signal, or a non-coding sequence of the target gene.

[0241] The donor polynucleotide is of a suitable length to correct or induce a mutation in a gDNA. In some embodiments, the donor polynucleotide comprises 10, 15, 20, 25, 50, 75, 100 or more nucleotides in length. In some embodiments (for example those described herein where a donor polynucleotide is incorporated into the cleaved nucleic acid as an insertion mediated by non-homologous end joining) the donor polynucleotide has no homology arms. In some embodiments, to facilitate HDR repair of a DSB, the donor polynucleotide has flanking homology arms (for example those described herein where a donor polynucleotide is incorporated into the cleaved nucleic acid as an insertion mediated by HDR repair). In some embodiments, the donor polynucleotide is about 10-100, about 20-80, about 30-70, or about 40-60 nucleotides in length. In some embodiments, the donor polynucleotide is about 10-100 nucleotides in length. In some embodiments, the donor polynucleotide is about 20-80 nucleotides in length. In some embodiments, the donor polynucleotide is about 30-70 nucleotides in length. In some embodiments, the donor polynucleotide is about 40-60 nucleotides in length. In some embodiments, the donor polynucleotide is 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 or 60 nucleotides in length. In some embodiments, the donor polynucleotide is 40 nucleotides in length. In some embodiments, the donor polynucleotide is 41 nucleotides in length. In some embodiments, the donor polynucleotide is 42 nucleotides in length. In some embodiments, the donor polynucleotide is 43 nucleotides in length. In some embodiments, the donor polynucleotide is 44 nucleotides in length. In some embodiments, the donor polynucleotide is 45 nucleotides in length. In some embodiments, the donor polynucleotide is 46 nucleotides in length. In some embodiments, the donor polynucleotide is 47 nucleotides in length. In some embodiments, the donor polynucleotide is 48 nucleotides in length. In some embodiments, the donor polynucleotide is 49 nucleotides in length. In some embodiments, the donor polynucleotide is 50 nucleotides in length. In some embodiments, the donor polynucleotide is 51 nucleotides in length. In some embodiments, the donor polynucleotide is 52 nucleotides in length. In some embodiments, the donor polynucleotide is 53 nucleotides in length. In some embodiments, the donor polynucleotide is 54 nucleotides in length. In some embodiments, the donor polynucleotide is 55 nucleotides in length. In some embodiments, the donor polynucleotide is 56 nucleotides in length. In some embodiments, the donor polynucleotide is 57 nucleotides in length. In some embodiments, the donor polynucleotide is 58 nucleotides in length. In some embodiments, the donor polynucleotide is 59 nucleotides in length. In some embodiments, the donor polynucleotide is 60 nucleotides in length.

[0242] In some embodiments, a donor polynucleotide comprising exogenous genetic material is flanked by homology arms to allow integration of the exogenous genetic material by HDR repair of a DSB in a target gene. The homology arms are designed to anneal to regions of gDNA that flank a DSB in a target gene. Methods of designing

homology arms that allow HDR repair of a DSB site in a target gene are taught in the art. See for example US 20110281361 which is incorporated by reference herein.

[0243] In some embodiments, for HDR repair of a DSB, a donor polynucleotide comprises a left and right flanking homology arms that allow annealing to gDNA. In some embodiments, the homology arms flank the mutation or correction being introduced at the site of a DSB. In some embodiments, the homology arms are at least 30-100, at least 50-200, at least 100-300, at least 100-500, at least 250-1000, at least 500-1500 nucleotides in length. In some embodiments, the homology arms are at least 100 nucleotides in length. In some embodiments, the homology arms are at least 200 nucleotides in length. In some embodiments, the homology arms are at least 300 nucleotides in length. In some embodiments, the homology arms are at least 400 nucleotides in length. In some embodiments, the homology arms are at least 500 nucleotides in length. In some embodiments, the homology arms are at least 600 nucleotides in length. In some embodiments, the homology arms are at least 700 nucleotides in length. In some embodiments, the homology arms are at least 800 nucleotides in length. In some embodiments, the homology arms are at least 900 nucleotides in length. In some embodiments, the homology arms are at least 1000 nucleotides in length. In some embodiments, the homology arms are at least 1500 nucleotides in length.

[0244] The rate of HDR is a function of the distance between the mutation at the DSB cut site. Thus, in some embodiments, the homology arms are designed to anneal to gDNA directly adjacent to the site of a DSB. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 1-10 nucleotides, 5-15, 10-30, 15-40, or 15-50 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 1 nucleotide from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 2 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 3 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 4 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 5 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 6 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 7 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 8 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 9 nucleotides from the DSB site in a target gene. In some embodiments, a left or right homology arm is designed to anneal to gDNA no more than 10 nucleotides from the DSB site in a target gene.

[0245] In some embodiments, the homology arms of a donor polynucleotide are fully complementary to gDNA flanking a DSB site in a target gene. In some embodiments,

the homology arms of a donor polynucleotide have sufficient complimentary to gDNA flanking a DSB site in a target gene to allow HDR repair.

[0246] In some embodiments, a donor polynucleotide provided by the disclosure comprises an intronic sequence. In some embodiments, the donor polynucleotide comprises an intronic sequence which corrects or induces a mutation in a gDNA. In some embodiments, the donor polynucleotide comprises an exonic sequence. In some embodiments, the donor polynucleotide comprises an exonic sequence which corrects or induces a mutation in a gDNA.

Methods of Making and Testing Donor Polynucleotides

[0247] The donor polynucleotides provided by the disclosure are produced by suitable DNA synthesis method or means known in the art. DNA synthesis is the natural or artificial creation of deoxyribonucleic acid (DNA) molecules. The term DNA synthesis refers to DNA replication, DNA biosynthesis (e.g., in vivo DNA amplification), enzymatic DNA synthesis (e.g., polymerase chain reaction (PCR); in vitro DNA amplification) or chemical DNA synthesis.

[0248] In some embodiments, each strand of the donor polynucleotide is produced by oligonucleotide synthesis. Oligonucleotide synthesis is the chemical synthesis of relatively short fragments or strands of single-stranded nucleic acids with a defined chemical structure (sequence). Methods of oligonucleotide synthesis are known in the art (see e.g., Reese (2005) *Organic & Biomolecular Chemistry* 3(21): 3851). The two strands can then be annealed together or duplexed to form a donor polynucleotide.

[0249] In some aspects, the insertion of a donor polynucleotide into a DSB is determined by a suitable method known in the art. For example, after the insertional event, the nucleotide sequence of PCR amplicons generated using PCR primer that flank the DSB site is analyzed for the presence of the nucleotide sequence comprising the donor polynucleotide. Next-generation sequencing (NGS) techniques are used to determine the extent of donor polynucleotide insertion into a DSB analyzing PCR amplicons for the presence or absence of the donor polynucleotide sequence. Further, since each donor polynucleotide is a linear, dsDNA molecule, which can insert in either of two orientations, NGS analysis can be used to determine the extent of insertion of the donor polynucleotide in either direction.

[0250] In some aspects, the insertion of the donor polynucleotide and its ability to correct a mutation is determined by nucleotide sequence analysis of mRNA transcribed from the gDNA into which the donor polynucleotide is inserted. An mRNA transcribed from gDNA containing an inserted donor polynucleotide is analyzed by a suitable method known in the art. For example, conversion of mRNA extracted from cells treated or contacted with a donor polynucleotide or system provided by the disclosure is enzymatically converted into cDNA, which is further by analyzed by NGS analysis to determine the extent of mRNA molecule comprising the corrected mutation.

[0251] In other aspects, the insertion of a donor polynucleotide and its ability to correct a mutation is determined by protein sequence analysis of a polypeptide translated from an mRNA transcribed from the gDNA into which the donor polynucleotide is inserted. In some embodiments, a donor polynucleotide corrects or induces a mutation by the incorporation of a codon into an exon that makes an amino acid

change in a gene comprising a gDNA molecule, wherein translation of an mRNA from the gene containing the inserted donor polynucleotide generates a polypeptide comprising the amino acid change. The amino acid change in the polypeptide is determined by protein sequence analysis using techniques including, but not limited to, Sanger sequencing, mass spectrometry, functional assays that measure an enzymatic activity of the polypeptide, or immunoblotting using an antibody reactive to the amino acid change.

Use of Donor Polynucleotides to Correct or Induce a Mutation

[0252] In some embodiments, a donor polynucleotide provided by the disclosure is used to correct or induce a mutation in a gDNA in a cell by insertion of the donor polynucleotide into a target nucleic acid (e.g., gDNA) at a cleavage site (e.g., a DSB) induced by a site-directed nuclease, such as those described herein. In some embodiments, a donor polynucleotide provided by the disclosure is used to correct or induce a mutation in a gDNA in a cell by exchanging a region proximal to a cleavage site (e.g., a DSB) for the corresponding region provided by the donor polynucleotide in a target nucleic acid (e.g., gDNA), induced by a site-directed nuclease, such as those described herein. In some embodiments, HDR DNA repair mechanisms of the cell repair the DSB using the donor polynucleotide, thereby inserting the donor polynucleotide into the DSB and adding the nucleotide sequence of the donor polynucleotide to the gDNA. In some embodiments, HDR DNA repair mechanisms of the cell repair the DSB using the donor polynucleotide, thereby exchanging a region proximal to a cleavage site (e.g., a DSB) for the corresponding region provided by the donor polynucleotide and adding the nucleotide sequence of the donor polynucleotide to the gDNA. In some embodiments, the donor polynucleotide comprises a nucleotide sequence which corrects a disease-causing mutation in a gDNA in a cell. In some embodiments, the donor polynucleotide is inserted at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the donor polynucleotide is exchanged at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the mutation is a substitution, missense, nonsense, insertion, deletion or frameshift mutation. In some embodiments the mutation is in an exon. In some embodiments, the mutation is a substitution, insertion or deletion and is located in an intron. In some embodiments, the mutation is proximal to a cleavage site in a gDNA. In some embodiments, the mutation is a protein-coding mutation. In some embodiments, the mutation is associated with or causes a disease.

[0253] In some embodiments, the donor polynucleotide is inserted into the DSB by HDR DNA repair. In some embodiments, the donor polynucleotide is exchanged a location proximal to the DSB by HDR DNA repair. In some embodiments, the donor polynucleotide, a portion of the donor polynucleotide is inserted into the target nucleic acid cleavage site by HDR DNA repair. In some embodiments, the donor polynucleotide, a portion of the donor polynucleotide is exchanged proximal to a target nucleic acid cleavage site by HDR DNA repair. In certain aspects, insertion of a donor polynucleotide into the target nucleic acid via HDR repair can result in, for example, mutations, deletions, alterations, integrations, gene correction, gene replacement, gene tagging, transgene insertion, nucleotide deletion, gene disruption,

translocations and/or gene mutation of the endogenous gene sequence. In certain aspects, exchange of a donor polynucleotide into the target nucleic acid via HDR repair can result in, for example, mutations, deletions, alterations, integrations, gene correction, gene replacement, gene tagging, transgene insertion, nucleotide deletion, gene disruption, translocations and/or gene mutation of the endogenous gene sequence.

[0254] In some embodiments, the disease-causing mutation in the HBB gene results in an E7V amino acid substitution in the human beta-globin protein. In some embodiments, the disclosure provides donor polynucleotides used to repair a DSB introduced into a target nucleic acid molecule (e.g., gDNA) by a site-directed nuclease (e.g., Cas9) in a cell. In some embodiments, the donor polynucleotide is used by the HDR repair pathway of the cell to repair the DSB in the target nucleic acid molecule. In some embodiments, the site-directed nuclease is a Cas nuclease. In some embodiments, the Cas nuclease is Cas9. The site-directed nucleases described herein can introduce DSB in target nucleic acids (e.g., genomic DNA) in a cell. The introduction of a DSB in the genomic DNA of a cell, induced by a site-directed nuclease, will stimulate the endogenous DNA repair pathways, such as those described herein. The HDR pathway can be used to insert a polynucleotide (e.g., a donor polynucleotide) into the DSB during repair.

[0255] Accordingly, in some embodiments, a single donor polynucleotide or multiple copies of the same donor polynucleotide are provided. In other embodiments, two or more donor polynucleotides are provided such that repair may occur at two or more target sites. For example, different donor polynucleotides are provided to repair a single gene in a cell, or two different genes in a cell. In some embodiments, the different donor polynucleotides are provided in independent copy numbers.

[0256] In some embodiments, the donor polynucleotide is incorporated into the target nucleic acid as an insertion mediated by HDR. In some embodiments, the donor polynucleotide sequence has no similarity to the nucleic acid sequence near the cleavage site. In some embodiments, a single donor polynucleotide or multiple copies of the same donor polynucleotide are provided. In other embodiments, two or more donor polynucleotides having different sequences are inserted at two or more sites by non-homologous end joining. In some embodiments, the different donor polynucleotides are provided in independent copy numbers.

Systems for Genome Editing

[0257] In some aspects, the disclosure provide systems for correcting a mutation in a genomic DNA molecule. In some embodiments, the system comprises a site-directed nuclease, such as a CRISPR/Cas system and optionally a gRNA, and a donor polynucleotide, such as those described herein. In some embodiments of the present disclosure, the system comprises an engineered nuclease. In some embodiments, the system comprises a site-directed nuclease. In some embodiments, the site-directed nuclease comprises a CRISPR/Cas nuclease system. In some embodiments, the Cas nuclease is Cas9. In some embodiments, the guide RNA comprising the CRISPR/Cas system is an sgRNA.

CRISPR/Cas Nuclease Systems

[0258] Naturally-occurring CRISPR/Cas systems are genetic defense systems that provides a form of acquired

immunity in prokaryotes. CRISPR is an abbreviation for Clustered Regularly Interspaced Short Palindromic Repeats, a family of DNA sequences found in the genomes of bacteria and archaea that contain fragments of DNA (spacer DNA) with similarity to foreign DNA previously exposed to the cell, for example, by viruses that have infected or attacked the prokaryote. These fragments of DNA are used by the prokaryote to detect and destroy similar foreign DNA upon re-introduction, for example, from similar viruses during subsequent attacks. Transcription of the CRISPR locus results in the formation of an RNA molecule comprising the spacer sequence, which associates with and targets Cas (CRISPR-associated) proteins able to recognize and cut the foreign, exogenous DNA. Numerous types and classes of CRISPR/Cas systems have been described (see e.g., Koonin et al., (2017) *Curr Opin Microbiol* 37:67-78).

[0259] Engineered versions of CRISPR/Cas systems has been developed in numerous formats to mutate or edit genomic DNA of cells from other species. The general approach of using the CRISPR/Cas system involves the heterologous expression or introduction of a site-directed nuclease (e.g.: Cas nuclease) in combination with a guide RNA (gRNA) into a cell, resulting in a DNA cleavage event (e.g., the formation a single-strand or double-strand break (SSB or DSB)) in the backbone of the cell's genomic DNA at a precise, targetable location. The manner in which the DNA cleavage event is repaired by the cell provides the opportunity to edit the genome by the addition, removal, or modification (substitution) of DNA nucleotide(s) or sequences (e.g., genes).

Cas Nuclease

[0260] In some embodiments, the disclosure provides compositions and systems (e.g., an engineered CRISPR/Cas system) comprising a site-directed nuclease, wherein the site-directed nuclease is a Cas nuclease. The Cas nuclease may comprise at least one domain that interacts with a guide RNA (gRNA). Additionally, the Cas nuclease are directed to a target sequence by a guide RNA. The guide RNA interacts with the Cas nuclease as well as the target sequence such that, once directed to the target sequence, the Cas nuclease is capable of cleaving the target sequence. In some embodiments, the guide RNA provides the specificity for the cleavage of the target sequence, and the Cas nuclease are universal and paired with different guide RNAs to cleave different target sequences.

[0261] In some embodiments, the CRISPR/Cas system comprise components derived from a Type-I, Type-II, or Type-III system. Updated classification schemes for CRISPR/Cas loci define Class 1 and Class 2 CRISPR/Cas systems, having Types I to V or VI (Makarova et al., (2015) *Nat Rev Microbiol*, 13(11):722-36; Shmakov et al., (2015) *Mol Cell*, 60:385-397). Class 2 CRISPR/Cas systems have single protein effectors. Cas proteins of Types II, V, and VI are single-protein, RNA-guided endonucleases, herein called "Class 2 Cas nucleases." Class 2 Cas nucleases include, for example, Cas9, Cpf1, C2c1, C2c2, and C2c3 proteins. The Cpf1 nuclease (Zetsche et al., (2015) *Cell* 163:1-13) is homologous to Cas9, and contains a RuvC-like nuclease domain.

[0262] In some embodiments, the Cas nuclease are from a Type-II CRISPR/Cas system (e.g., a Cas9 protein from a CRISPR/Cas9 system). In some embodiments, the Cas nuclease are from a Class 2 CRISPR/Cas system (a single-

protein Cas nuclease such as a Cas9 protein or a Cpf1 protein). The Cas9 and Cpf1 family of proteins are enzymes with DNA endonuclease activity, and they can be directed to cleave a desired nucleic acid target by designing an appropriate guide RNA, as described further herein.

[0263] A Type-II CRISPR/Cas system component are from a Type-IIA, Type-IIB, or Type-IIC system. Cas9 and its orthologs are encompassed. Non-limiting exemplary species that the Cas9 nuclease or other components are from include *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus* sp., *Staphylococcus aureus*, *Listeria innocua*, *Lactobacillus gasseri*, *Francisella novicida*, *Wolinella succinogenes*, *Sutterella wadsworthensis*, *Gamma proteobacterium*, *Neisseria meningitidis*, *Campylobacter jejuni*, *Pasteurella multocida*, *Fibrobacter succinogene*, *Rhodospirillum rubrum*, *Nocardiopsis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycoides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Lactobacillus buchneri*, *Treponema denticola*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas* sp., *Crocospaera watsonii*, *Cyanothece* sp., *Microcystis aeruginosa*, *Synechococcus* sp., *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor beccsii*, *Candidatus Desulfurudis*, *Clostridium botulinum*, *Clostridium difficile*, *Finegoldia magna*, *Natranaerobius thermophilus*, *Pelotomaculum rmopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochrochromatium vinosum*, *Marinobacter* sp., *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc* sp., *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira* sp., *Lyngbya* sp., *Microcoleus chthonoplastes*, *Oscillatoria* sp., *Petrotoga mobilis*, *Thermosipho africanus*, *Streptococcus pasteurianus*, *Neisseria cinerea*, *Campylobacter lari*, *Parvibaculum lavamentivorans*, *Corynebacterium diphtheria*, or *Acaryochloris marina*. In some embodiments, the Cas9 protein are from *Streptococcus pyogenes* (SpCas9). In some embodiments, the Cas9 protein are from *Streptococcus thermophilus* (StCas9). In some embodiments, the Cas9 protein are from *Neisseria meningitidis* (NmCas9). In some embodiments, the Cas9 protein are from *Staphylococcus aureus* (SaCas9). In some embodiments, the Cas9 protein are from *Campylobacter jejuni* (CjCas9).

[0264] In some embodiments, a Cas nuclease may comprise more than one nuclease domain. For example, a Cas9 nuclease may comprise at least one RuvC-like nuclease domain (e.g., Cpf1) and at least one HNH-like nuclease domain (e.g., Cas9). In some embodiments, the Cas9 nuclease introduces a DSB in the target sequence. In some embodiments, the Cas9 nuclease is modified to contain only one functional nuclease domain. For example, the Cas9 nuclease is modified such that one of the nuclease domains is mutated or fully or partially deleted to reduce its nucleic acid cleavage activity. In some embodiments, the Cas9 nuclease is modified to contain no functional RuvC-like nuclease domain. In other embodiments, the Cas9 nuclease is modified to contain no functional HNH-like nuclease domain. In some embodiments in which only one of the nuclease domains is functional, the Cas9 nuclease is a

nickase that is capable of introducing a single-stranded break (a “nick”) into the target sequence. In some embodiments, a conserved amino acid within a Cas9 nuclease domain is substituted to reduce or alter a nuclease activity. In some embodiments, the Cas nuclease nickase comprises an amino acid substitution in the RuvC-like nuclease domain. Exemplary amino acid substitutions in the RuvC-like nuclease domain include D10A (based on the *S. pyogenes* Cas9 nuclease). In some embodiments, the nickase comprises an amino acid substitution in the HNH-like nuclease domain. Exemplary amino acid substitutions in the HNH-like nuclease domain include E762A, H840A, N863A, H983A, and D986A (based on the *S. pyogenes* Cas9 nuclease). In some embodiments, the nuclease system described herein comprises a nickase and a pair of guide RNAs that are complementary to the sense and antisense strands of the target sequence, respectively. The guide RNAs directs the nickase to target and introduce a DSB by generating a nick on opposite strands of the target sequence (i.e., double nicking). Chimeric Cas9 nucleases are used, where one domain or region of the protein is replaced by a portion of a different protein. For example, a Cas9 nuclease domain is replaced with a domain from a different nuclease such as FokI. A Cas9 nuclease is a modified nuclease.

[0265] In alternative embodiments, the Cas nuclease is from a Type-I CRISPR/Cas system. In some embodiments, the Cas nuclease is a component of the Cascade complex of a Type-I CRISPR/Cas system. For example, the Cas nuclease is a Cas3 nuclease. In some embodiments, the Cas nuclease is derived from a Type-III CRISPR/Cas system. In some embodiments, the Cas nuclease is derived from Type-IV CRISPR/Cas system. In some embodiments, the Cas nuclease is derived from a Type-V CRISPR/Cas system. In some embodiments, the Cas nuclease is derived from a Type-VI CRISPR/Cas system.

Guide RNAs (gRNAs)

[0266] Engineered CRISPR/Cas systems comprise at least two components: 1) a guide RNA (gRNA) molecule and 2) a Cas nuclease, which interact to form a gRNA/Cas nuclease complex. A gRNA comprises at least a user-defined targeting domain termed a “spacer” comprising a nucleotide sequence and a CRISPR repeat sequence. In engineered CRISPR/Cas systems, a gRNA/Cas nuclease complex is targeted to a specific target sequence of interest within a target nucleic acid (e.g., a genomic DNA molecule) by generating a gRNA comprising a spacer with a nucleotide sequence that is able to bind to the specific target sequence in a complementary fashion (See Jinek et al., *Science*, 337, 816-821 (2012) and Deltcheva et al., *Nature*, 471, 602-607 (2011)). Thus, the spacer provides the targeting function of the gRNA/Cas nuclease complex.

[0267] In naturally-occurring type II-CRISPR/Cas systems, the “gRNA” is comprised of two RNA strands: 1) a CRISPR RNA (crRNA) comprising the spacer and CRISPR repeat sequence, and 2) a trans-activating CRISPR RNA (tracrRNA). In Type II-CRISPR/Cas systems, the portion of the crRNA comprising the CRISPR repeat sequence and a portion of the tracrRNA hybridize to form a crRNA: tracrRNA duplex, which interacts with a Cas nuclease (e.g., Cas9). As used herein, the terms “split gRNA” or “modular gRNA” refer to a gRNA molecule comprising two RNA strands, wherein the first RNA strand incorporates the crRNA function(s) and/or structure and the second RNA

strand incorporates the tracrRNA function(s) and/or structure, and wherein the first and second RNA strands partially hybridize.

[0268] Accordingly, in some embodiments, a gRNA provided by the disclosure comprises two RNA molecules. In some embodiments, the gRNA comprises a CRISPR RNA (crRNA) and a trans-activating CRISPR RNA (tracrRNA). In some embodiments, the gRNA is a split gRNA. In some embodiments, the gRNA is a modular gRNA. In some embodiments, the split gRNA comprises a first strand comprising, from 5' to 3', a spacer, and a first region of complementarity; and a second strand comprising, from 5' to 3', a second region of complementarity; and optionally a tail domain.

[0269] In some embodiments, the crRNA comprises a spacer comprising a nucleotide sequence that is complementary to and hybridizes with a sequence that is complementary to the target sequence on a target nucleic acid (e.g., a genomic DNA molecule). In some embodiments, the crRNA comprises a region that is complementary to and hybridizes with a portion of the tracrRNA.

[0270] In some embodiments, the tracrRNA may comprise all or a portion of a wild-type tracrRNA sequence from a naturally-occurring CRISPR/Cas system. In some embodiments, the tracrRNA may comprise a truncated or modified variant of the wild-type tracrRNA. The length of the tracrRNA may depend on the CRISPR/Cas system used. In some embodiments, the tracrRNA may comprise 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 length. In certain embodiments, the tracrRNA is at least 26 nucleotides in length. In additional embodiments, the tracrRNA is at least 40 nucleotides in length. In some embodiments, the tracrRNA may comprise certain secondary structures, such as, e.g., one or more hairpins or stem-loop structures, or one or more bulge structures.

Single Guide RNA (sgRNA)

[0271] Engineered CRISPR/Cas nuclease systems often combine a crRNA and a tracrRNA into a single RNA molecule, referred to herein as a “single guide RNA” (sgRNA), by adding a linker between these components. Without being bound by theory, similar to a duplexed crRNA and tracrRNA, an sgRNA will form a complex with a Cas nuclease (e.g., Cas9), guide the Cas nuclease to a target sequence and activate the Cas nuclease for cleavage the target nucleic acid (e.g., genomic DNA). Accordingly, in some embodiments, the gRNA may comprise a crRNA and a tracrRNA that are operably linked. In some embodiments, the sgRNA may comprise a crRNA covalently linked to a tracrRNA. In some embodiments, the crRNA and the tracrRNA is covalently linked via a linker. In some embodiments, the sgRNA may comprise a stem-loop structure via base pairing between the crRNA and the tracrRNA. In some embodiments, a sgRNA comprises, from 5' to 3', a spacer, a first region of complementarity, a linking domain, a second region of complementarity, and, optionally, a tail domain.

[0272] The sgRNA can comprise a 20 nucleotide spacer sequence at the 5' end of the sgRNA sequence. The sgRNA can comprise a less than 20 nucleotide spacer sequence at the 5' end of the sgRNA sequence. The sgRNA can comprise a more than 20 nucleotide spacer sequence at the 5' end of the sgRNA sequence. The sgRNA can comprise a variable length spacer sequence with 17-30 nucleotides at the 5' end of the sgRNA sequence as set forth by SEQ ID NO: 1.

[0273] The sgRNA can comprise no uracil at the 3' end of the sgRNA sequence. The sgRNA can comprise one or more uracil at the 3' end of the sgRNA sequence. For example, the sgRNA can comprise 1 uracil (U) at the 3' end of the sgRNA sequence. The sgRNA can comprise 2 uracil (UU) at the 3' end of the sgRNA sequence. The sgRNA can comprise 3 uracil (UUU) at the 3' end of the sgRNA sequence. The sgRNA can comprise 4 uracil (UUUU) at the 3' end of the sgRNA sequence. The sgRNA can comprise 5 uracil (UUUUU) at the 3' end of the sgRNA sequence. The sgRNA can comprise 6 uracil (UUUUUU) at the 3' end of the sgRNA sequence. The sgRNA can comprise 7 uracil (UUUUUUU) at the 3' end of the sgRNA sequence. The sgRNA can comprise 8 uracil (UUUUUUUU) at the 3' end of the sgRNA sequence.

[0274] The sgRNA can be unmodified or modified. For example, modified sgRNAs can comprise one or more 2'-O-methyl phosphorothioate nucleotides.

[0275] By way of illustration, guide RNAs used in the CRISPR/Cas system, or other smaller RNAs can be readily synthesized by chemical means, as illustrated herein and described in the art. While chemical synthetic procedures are continually expanding, purifications of such RNAs by procedures such as high performance liquid chromatography (HPLC, which avoids the use of gels such as PAGE) tends to become more challenging as polynucleotide lengths increase significantly beyond a hundred or so nucleotides. One approach used for generating RNAs of greater length is to produce two or more molecules that are ligated together. Much longer RNAs, such as those encoding a Cas9 endonuclease, are more readily generated enzymatically. Various types of RNA modifications can be introduced during or after chemical synthesis and/or enzymatic generation of RNAs, e.g., modifications that enhance stability, reduce the likelihood or degree of innate immune response, and/or enhance other attributes, as described in the art.

Spacers

[0276] In some embodiments, the gRNAs provided by the disclosure comprise a spacer sequence. A spacer sequence is a sequence that defines the target site of a target nucleic acid (e.g.: DNA). The target nucleic acid is a double-stranded molecule: one strand comprises the target sequence adjacent to a PAM sequence and is referred to as the "PAM strand," and the second strand is referred to as the "non-PAM strand" and is complementary to the PAM strand and target sequence. Both gRNA spacer and the target sequence are complementary to the non-PAM strand of the target nucleic acid. The gRNA spacer sequence hybridizes to the complementary strand (e.g.: the non-PAM strand of the target nucleic acid/target site). In some embodiments, the spacer is sufficiently complementary to the complementary strand of the target sequence (e.g.: non-PAM strand), as to target a Cas nuclease to the target nucleic acid. In some embodiments, the spacer is at least 80%, 85%, 90% or 95% complementary to the non-PAM strand of the target nucleic acid. In some embodiments, the spacer is 100% complementary to the non-PAM strand of the target nucleic acid. In some embodiments, the spacer comprises 1, 2, 3, 4, 5, 6 or more nucleotides that are not complementary with the non-PAM strand of the target nucleic acid. In some

embodiments, the spacer comprises 2 nucleotides that are not complementary with the non-PAM strand of the target nucleic acid.

[0277] In some embodiments, the 5' most nucleotide of gRNA comprises the 5' most nucleotide of the spacer. In some embodiments, the spacer is located at the 5' end of the crRNA. In some embodiments, the spacer is located at the 5' end of the sgRNA. In some embodiments, the spacer is about 15-50, about 20-45, about 25-40 or about 30-35 nucleotides in length. In some embodiments, the spacer is about 19-22 nucleotides in length. In some embodiments the spacer is about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 nucleotides in length. In some embodiments the spacer is 19 nucleotides in length. In some embodiments, the spacer is 20 nucleotides in length, in some embodiments, the spacer is 21 nucleotides in length.

[0278] In some embodiments, the nucleotide sequence of the target sequence and the PAM comprises the formula 5' N₁₉₋₂₁-N-R-G-3' (SEQ ID NO: 63), wherein N is any nucleotide, and wherein R is a nucleotide comprising the nucleobase adenine (A) or guanine (G), and wherein the three 3' terminal nucleic acids, N-R-G represent the *S. pyogenes* PAM (SEQ ID NO: 64). In some embodiments, the nucleotide sequence of the spacer is designed or chosen using a computer program. The computer program can use variables, such as predicted melting temperature, secondary structure formation, predicted annealing temperature, sequence identity, genomic context, chromatin accessibility, % GC, frequency of genomic occurrence (e.g., of sequences that are identical or are similar but vary in one or more spots as a result of mismatch, insertion or deletion), methylation status, and/or presence of SNPs.

[0279] In some embodiments, the spacer comprise at least one or more modified nucleotide(s) such as those described herein. The disclosure provides gRNA molecules comprising a spacer which may comprise the nucleobase uracil (U), while any DNA encoding a gRNA comprising a spacer comprising the nucleobase uracil (U) will comprise the nucleobase thymine (T) in the corresponding position(s).

Methods of Making gRNAs

[0280] The gRNAs of the present disclosure is produced by a suitable means available in the art, including but not limited to in vitro transcription (IVT), synthetic and/or chemical synthesis methods, or a combination thereof. Enzymatic (IVT), solid-phase, liquid-phase, combined synthetic methods, small region synthesis, and ligation methods are utilized. In one embodiment, the gRNAs are made using IVT enzymatic synthesis methods. Methods of making polynucleotides by IVT are known in the art and are described in International Application PCT/US2013/30062. Accordingly, the present disclosure also includes polynucleotides, e.g., DNA, constructs and vectors are used to in vitro transcribe a gRNA described herein.

[0281] In some aspects, non-natural modified nucleobases are introduced into polynucleotides, e.g., gRNA, during synthesis or post-synthesis. In certain embodiments, modifications are on internucleoside linkages, purine or pyrimidine bases, or sugar. In particular embodiments, the modification is introduced at the terminal of a polynucleotide; with chemical synthesis or with a polymerase enzyme. Examples of modified nucleic acids and their synthesis are disclosed in PCT application No. PCT/US2012/058519.

Synthesis of modified polynucleotides is also described in Verma and Eckstein, *Annual Review of Biochemistry*, vol. 76, 99-134 (1998).

[0282] In some aspects, enzymatic or chemical ligation methods are used to conjugate polynucleotides or their regions with different functional moieties, such as targeting or delivery agents, fluorescent labels, liquids, nanoparticles, etc. Conjugates of polynucleotides and modified polynucleotides are reviewed in Goodchild, *Bioconjugate Chemistry*, vol. 1(3), 165-187 (1990).

[0283] Certain embodiments of the invention also provide nucleic acids, e.g., vectors, encoding gRNAs described herein. In some embodiments, the nucleic acid is a DNA molecule. In other embodiments, the nucleic acid is an RNA molecule. In some embodiments, the nucleic acid comprises a nucleotide sequence encoding a crRNA. In some embodiments, the nucleotide sequence encoding the crRNA comprises a spacer flanked by all or a portion of a repeat sequence from a naturally-occurring CRISPR/Cas system. In some embodiments, the nucleic acid comprises a nucleotide sequence encoding a tracrRNA. In some embodiments, the crRNA and the tracrRNA is encoded by two separate nucleic acids. In other embodiments, the crRNA and the tracrRNA is encoded by a single nucleic acid. In some embodiments, the crRNA and the tracrRNA is encoded by opposite strands of a single nucleic acid. In other embodiments, the crRNA and the tracrRNA is encoded by the same strand of a single nucleic acid.

[0284] In some embodiments, the gRNAs provided by the disclosure are chemically synthesized by any means described in the art (see e.g., WO/2005/01248). While chemical synthetic procedures are continually expanding, purifications of such RNAs by procedures such as high performance liquid chromatography (HPLC, which avoids the use of gels such as PAGE) tends to become more challenging as polynucleotide lengths increase significantly beyond a hundred or so nucleotides. One approach used for generating RNAs of greater length is to produce two or more molecules that are ligated together.

[0285] In some embodiments, the gRNAs provided by the disclosure are synthesized by enzymatic methods (e.g., in vitro transcription, IVT).

[0286] Various types of RNA modifications can be introduced during or after chemical synthesis and/or enzymatic generation of RNAs, e.g., modifications that enhance stability, reduce the likelihood or degree of innate immune response, and/or enhance other attributes, as described in the art.

[0287] In certain embodiments, more than one guide RNA can be used with a CRISPR/Cas nuclease system. Each guide RNA may contain a different targeting sequence, such that the CRISPR/Cas system cleaves more than one target nucleic acid. In some embodiments, one or more guide RNAs may have the same or differing properties such as activity or stability within the Cas9 RNP complex. Where more than one guide RNA is used, each guide RNA can be encoded on the same or on different vectors. The promoters used to drive expression of the more than one guide RNA is the same or different.

[0288] The guide RNA may target any sequence of interest via the targeting sequence (e.g. spacer sequence) of the crRNA. In some embodiments, the degree of complementarity between the targeting sequence of the guide RNA and the target sequence on the target nucleic acid molecule is

about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%. In some embodiments, the targeting sequence of the guide RNA and the target sequence on the target nucleic acid molecule is 100% complementary. In other embodiments, the targeting sequence of the guide RNA and the target sequence on the target nucleic acid molecule may contain at least one mismatch. For example, the targeting sequence of the guide RNA and the target sequence on the target nucleic acid molecule may contain 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mismatches. In some embodiments, the targeting sequence of the guide RNA and the target sequence on the target nucleic acid molecule may contain 1-6 mismatches. In some embodiments, the targeting sequence of the guide RNA and the target sequence on the target nucleic acid molecule may contain 5 or 6 mismatches.

[0289] The length of the targeting sequence may depend on the CRISPR/Cas9 system and components used. For example, different Cas9 proteins from different bacterial species have varying optimal targeting sequence lengths. Accordingly, the targeting sequence may comprise 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, 35, 40, 45, 50, or more than 50 nucleotides in length. In some embodiments, the targeting sequence may comprise 18-24 nucleotides in length. In some embodiments, the targeting sequence may comprise 19-21 nucleotides in length. In some embodiments, the targeting sequence may comprise 20 nucleotides in length.

[0290] In some embodiments of the present disclosure, a CRISPR/Cas nuclease system includes at least one guide RNA. In some embodiments, the guide RNA and the Cas protein may form a ribonucleoprotein (RNP), e.g., a CRISPR/Cas complex. The guide RNA may guide the Cas protein to a target sequence on a target nucleic acid molecule (e.g., a genomic DNA molecule), where the Cas protein cleaves the target nucleic acid. In some embodiments, the CRISPR/Cas complex is a Cpf1/guide RNA complex. In some embodiments, the CRISPR complex is a Type-II CRISPR/Cas9 complex. In some embodiments, the Cas protein is a Cas9 protein. In some embodiments, the CRISPR/Cas9 complex is a Cas9/guide RNA complex.

Engineered Nucleases

[0291] In additional embodiments, the donor polynucleotides provided by the disclosure are used in combination with a site-directed nuclease, wherein the site-directed nuclease is an engineered nuclease. Exemplary engineered nucleases are meganuclease (e.g., homing endonucleases), ZFN, TALEN, and megaTAL.

[0292] Naturally-occurring meganucleases may recognize and cleave double-stranded DNA sequences of about 12 to 40 base pairs and are commonly grouped into five families. In some embodiments, the meganuclease are chosen from the LAGLIDADG family, the GIY-YIG family, the HNH family, the His-Cys box family, and the PD-(D/E)XK family. In some embodiments, the DNA binding domain of the meganuclease are engineered to recognize and bind to a sequence other than its cognate target sequence. In some embodiments, the DNA binding domain of the meganuclease are fused to a heterologous nuclease domain. In some embodiments, the meganuclease, such as a homing endonuclease, are fused to TAL modules to create a hybrid protein, such as a "megaTAL" protein. The megaTAL protein have improved DNA targeting specificity by recognizing

ing the target sequences of both the DNA binding domain of the meganuclease and the TAL modules.

[0293] ZFNs are fusion proteins comprising a zinc-finger DNA binding domain (“zinc fingers” or “ZFs”) and a nuclease domain. Each naturally-occurring ZF may bind to three consecutive base pairs (a DNA triplet), and ZF repeats are combined to recognize a DNA target sequence and provide sufficient affinity. Thus, engineered ZF repeats are combined to recognize longer DNA sequences, such as, e.g., 9-, 12-, 15-, or 18-bp, etc. In some embodiments, the ZFN comprise ZFs fused to a nuclease domain from a restriction endonuclease. For example, the restriction endonuclease is FokI. In some embodiments, the nuclease domain comprises a dimerization domain, such as when the nuclease dimerizes to be active, and a pair of ZFNs comprising the ZF repeats and the nuclease domain is designed for targeting a target sequence, which comprises two half target sequences recognized by each ZF repeats on opposite strands of the DNA molecule, with an interconnecting sequence in between (which is sometimes called a spacer in the literature). For example, the interconnecting sequence is 5 to 7 bp in length. When both ZFNs of the pair bind, the nuclease domain may dimerize and introduce a DSB within the interconnecting sequence. In some embodiments, the dimerization domain of the nuclease domain comprises a knob-into-hole motif to promote dimerization. For example, the ZFN comprises a knob-into-hole motif in the dimerization domain of FokI.

[0294] The DNA binding domain of TALENs usually comprises a variable number of 34 or 35 amino acid repeats (“modules” or “TAL modules”), with each module binding to a single DNA base pair, A, T, G, or C. Adjacent residues at positions 12 and 13 (the “repeat-variable di-residue” or RVD) of each module specify the single DNA base pair that the module binds to. Though modules used to recognize G may also have affinity for A, TALENs benefit from a simple code of recognition—one module for each of the 4 bases—which greatly simplifies the customization of a DNA-binding domain recognizing a specific target sequence. In some embodiments, the TALEN may comprise a nuclease domain from a restriction endonuclease. For example, the restriction endonuclease is FokI. In some embodiments, the nuclease domain may dimerize to be active, and a pair of TALENs is designed for targeting a target sequence, which comprises two half target sequences recognized by each DNA binding domain on opposite strands of the DNA molecule, with an interconnecting sequence in between. For example, each half target sequence is in the range of 10 to 20 bp, and the interconnecting sequence is 12 to 19 bp in length. When both TALENs of the pair bind, the nuclease domain may dimerize and introduce a DSB within the interconnecting sequence. In some embodiments, the dimerization domain of the nuclease domain may comprise a knob-into-hole motif to promote dimerization. For example, the TALEN may comprise a knob-into-hole motif in the dimerization domain of FokI.

Modified Nucleases

[0295] In certain embodiments, the nuclease is optionally modified from its wild-type counterpart. In some embodiments, the nuclease is fused with at least one heterologous protein domain. At least one protein domain is located at the N-terminus, the C-terminus, or in an internal location of the nuclease. In some embodiments, two or more heterologous protein domains are at one or more locations on the nuclease.

[0296] In some embodiments, the protein domain may facilitate transport of the nuclease into the nucleus of a cell. For example, the protein domain is a nuclear localization signal (NLS). In some embodiments, the nuclease is fused with 1-10 NLS(s). In some embodiments, the nuclease is fused with 1-5 NLS(s). In some embodiments, the nuclease is fused with one NLS. In other embodiments, the nuclease is fused with more than one NLS. In some embodiments, the nuclease is fused with 2, 3, 4, or 5 NLSs. In some embodiments, the nuclease is fused with 2 NLSs. In some embodiments, the nuclease is fused with 3 NLSs. In some embodiments, the nuclease is 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: 65) or PKKKRRV (SEQ ID NO: 66). In some embodiments, the NLS is a bipartite sequence, such as, e.g., the NLS of nucleoplasmin, KRPAATKKAGQAKKKK (SEQ ID NO: 67). In some embodiments, the NLS is genetically modified from its wild-type counterpart.

[0297] In some embodiments, the protein domain is capable of modifying the intracellular half-life of the nuclease. In some embodiments, the half-life of the nuclease may be increased. In some embodiments, the half-life of the nuclease is reduced. In some embodiments, the entity is capable of increasing the stability of the nuclease. In some embodiments, the entity is capable of reducing the stability of the nuclease. In some embodiments, the protein domain act as a signal peptide for protein degradation. In some embodiments, the protein degradation is mediated by proteolytic enzymes, such as, e.g., proteasomes, lysosomal proteases, or calpain proteases. In some embodiments, the protein domain comprises a PEST sequence. In some embodiments, the nuclease is modified by addition of ubiquitin or a polyubiquitin chain. In some embodiments, the ubiquitin is a ubiquitin-like protein (UBL). Non-limiting examples of ubiquitin-like proteins include small ubiquitin-like modifier (SUMO), ubiquitin cross-reactive protein (UCRP, also known as interferon-stimulated gene-15 (ISG15)), ubiquitin-related modifier-1 (URM1), neuronal-precursor-cell-expressed developmentally downregulated protein-8 (NEDD8, also called Rub 1 in *S. cerevisiae*), human leukocyte antigen F-associated (FAT10), autophagy-8 (ATG8) and -12 (ATG12), Fau ubiquitin-like protein (FUB1), membrane-anchored UBL (MUB), ubiquitin fold-modifier-1 (UFM1), and ubiquitin-like protein-5 (UBL5).

[0298] In some embodiments, the protein domain is a marker domain. Non-limiting examples of marker domains include fluorescent proteins, purification tags, epitope tags, and reporter gene sequences. In some embodiments, the marker domain is a fluorescent protein. Non-limiting examples of suitable fluorescent proteins include green fluorescent proteins (e.g., GFP, GFP-2, tagGFP, turboGFP, sfGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), yellow fluorescent proteins (e.g., YFP, EYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), blue fluorescent proteins (e.g., EBFP, EBFP2, Azurite, mKalamal, GFPuv, Sapphire, T-sapphire.), cyan fluorescent proteins (e.g., ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), red fluorescent proteins (e.g., mKate, mKate2, mPlum, DsRed monomer, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), and orange fluorescent proteins (mOrange, mKO,

Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato) or any other suitable fluorescent protein. In other embodiments, the marker domain is a purification tag and/or an epitope tag. Non-limiting exemplary tags include glutathione-S-transferase (GST), chitin binding protein (CBP), maltose binding protein (MBP), thioredoxin (TRX), poly(NANP), tandem affinity purification (TAP) tag, myc, AcV5, AU1, AUS, E, ECS, E2, FLAG (SEQ ID NO: 95), HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis (SEQ ID NO: 94), biotin carboxyl carrier protein (BCCP), and calmodulin. Non-limiting exemplary reporter genes include glutathione-S-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT), beta-galactosidase, beta-glucuronidase, luciferase, or fluorescent proteins.

[0299] In additional embodiments, the protein domain may target the nuclease to a specific organelle, cell type, tissue, or organ.

[0300] In further embodiments, the protein domain is an effector domain. When the nuclease is directed to its target nucleic acid, e.g., when a Cas9 protein is directed to a target nucleic acid by a guide RNA, the effector domain may modify or affect the target nucleic acid. In some embodiments, the effector domain is chosen from a nucleic acid binding domain, a nuclease domain, an epigenetic modification domain, a transcriptional activation domain, or a transcriptional repressor domain.

[0301] Certain embodiments of the invention also provide nucleic acids encoding the nucleases (e.g., a Cas9 protein) described herein provided on a vector. In some embodiments, the nucleic acid is a DNA molecule. In other embodiments, the nucleic acid is an RNA molecule. In some embodiments, the nucleic acid encoding the nuclease is an mRNA molecule. In certain embodiments, the nucleic acid is an mRNA encoding a Cas9 protein.

[0302] In some embodiments, the nucleic acid encoding the nuclease is codon optimized for efficient expression in one or more eukaryotic cell types. In some embodiments, the nucleic acid encoding the nuclease is codon optimized for efficient expression in one or more mammalian cells. In some embodiments, the nucleic acid encoding the nuclease is codon optimized for efficient expression in human cells. Methods of codon optimization including codon usage tables and codon optimization algorithms are available in the art.

Target Sites

[0303] In some embodiments, the site-directed nucleases described herein are directed to and cleave (e.g., introduce a DSB) a target nucleic acid molecule. In some embodiments, a Cas nuclease is directed by a guide RNA to a target site of a target nucleic acid molecule (gDNA), where the guide RNA hybridizes with the complementary strand of the target sequence and the Cas nuclease cleaves the target nucleic acid at the target site. In some embodiments, the complementary strand of the target sequence is complementary to the targeting sequence (e.g.: spacer sequence) of the guide RNA. In some embodiments, the degree of complementarity between a targeting sequence of a guide RNA and its corresponding complementary strand of the target sequence is about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%. In some embodiments, the complementary strand of the target sequence and the targeting sequence of the guide RNA is

100% complementary. In other embodiments, the complementary strand of the target sequence and the targeting sequence of the guide RNA contains at least one mismatch. For example, the complementary strand of the target sequence and the targeting sequence of the guide RNA contain 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mismatches. In some embodiments, the complementary strand of the target sequence and the targeting sequence of the guide RNA contain 1-6 mismatches. In some embodiments, the complementary strand of the target sequence and the targeting sequence of the guide RNA contain 5 or 6 mismatches.

[0304] The length of the target sequence may depend on the nuclease system used. For example, the target sequence for a CRISPR/Cas system comprise 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, 35, 40, 45, 50, or more than 50 nucleotides in length. In some embodiments, the target sequence comprises 18-24 nucleotides in length. In some embodiments, the target sequence comprises 19-21 nucleotides in length. In some embodiments, the target sequence comprises 20 nucleotides in length. When nickases are used, the target sequence comprises a pair of target sequences recognized by a pair of nickases on opposite strands of the DNA molecule.

[0305] In some embodiments, the target sequence for a meganuclease comprises 12-40 or more nucleotides in length. When ZFNs are used, the target sequence comprises two half target sequences recognized by a pair of ZFNs on opposite strands of the DNA molecule, with an interconnecting sequence in between. In some embodiments, each half target sequence for ZFNs independently comprise 9, 12, 15, 18, or more nucleotides in length. In some embodiments, the interconnecting sequence for ZFNs comprise 4-20 nucleotides in length. In some embodiments, the interconnecting sequence for ZFNs comprise 5-7 nucleotides in length.

[0306] When TALENs are used, the target sequence may similarly comprise two half target sequences recognized by a pair of TALENs on opposite strands of the DNA molecule, with an interconnecting sequence in between. In some embodiments, each half target sequence for TALENs may independently comprise 10-20 or more nucleotides in length. In some embodiments, the interconnecting sequence for TALENs may comprise 4-20 nucleotides in length. In some embodiments, the interconnecting sequence for TALENs may comprise 12-19 nucleotides in length.

[0307] The target nucleic acid molecule is any DNA molecule that is endogenous or exogenous to a cell. As used herein, the term "endogenous sequence" refers to a sequence that is native to the cell. In some embodiments, the target nucleic acid molecule is a genomic DNA (gDNA) molecule or a chromosome from a cell or in the cell. In some embodiments, the target sequence of the target nucleic acid molecule is a genomic sequence from a cell or in the cell. In other embodiments, the cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is 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. In further embodiments, the target sequence may be a viral sequence. In yet other embodiments, the target sequence may be a synthesized sequence. In some embodiments, the target sequence may be on a eukaryotic chromosome, such as a human chromosome.

[0308] In some embodiments, the target sequence may be located in a coding sequence of a gene, an intron sequence

of a gene, a transcriptional control sequence of a gene, a translational control sequence of a gene, or a non-coding sequence between genes. In some embodiments, the gene may be a protein coding gene. In other embodiments, the gene may be a non-coding RNA gene. In some embodiments, the target sequence may comprise all or a portion of a disease-associated gene.

[0309] In some embodiments, the target sequence may be located in a non-genic functional site in the genome that controls aspects of chromatin organization, such as a scaffold site or locus control region. In some embodiments, the target sequence may be a genetic safe harbor site, i.e., a locus that facilitates safe genetic modification.

[0310] In some embodiments, the target sequence may be adjacent to a protospacer adjacent motif (PAM), a short sequence recognized by a CRISPR/Cas9 complex. In some embodiments, the PAM may be adjacent to or within 1, 2, 3, or 4, nucleotides of the 3' end of the target sequence. The length and the sequence of the PAM may depend on the Cas9 protein used. For example, the PAM may be selected from a consensus or a particular PAM sequence for a specific Cas9 nuclease or Cas9 ortholog, including those disclosed in FIG. 1 of Ran et al., (2015) *Nature*, 520:186-191 (2015), which is incorporated herein by reference. In some embodiments, the PAM may comprise 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in length. Non-limiting exemplary PAM sequences include NGG (SpCas9 WT, SpCas9 nickase, dimeric dCas9-FokI, SpCas9-HF1, SpCas9 K855A, eSpCas9 (1.0), eSpCas9 (1.1)), NGAN or NGNG (SpCas9 VQR variant), NGAG (SpCas9 EQR variant), NGCG (SpCas9 VRER variant), NAAG (SpCas9 QQR1 variant), NNGRRRT or NNGRRN (SaCas9), NNNRRT (KKH SaCas9), NNNNRYAC (CjCas9), NNAGAAW (St1Cas9), NAAAAC (TdCas9), NGGNG (St3Cas9), NG (FnCas9), NAAAAN (TdCas9), NNAAAAW (StCas9), NNNNACA (CjCas9), GNNNCNNA (PmCas9), and NNNNGATT (NmCas9) (see e.g., Cong et al., (2013) *Science* 339:819-823; Kleinstiver et al., (2015) *Nat Biotechnol* 33:1293-1298; Kleinstiver et al., (2015) *Nature* 523:481-485; Kleinstiver et al., (2016) *Nature* 529:490-495; Tsai et al., (2014) *Nat Biotechnol* 32:569-576; Slaymaker et al., (2016) *Science* 351:84-88; Anders et al., (2016) *Mol Cell* 61:895-902; Kim et al., (2017) *Nat Comm* 8:14500; Fonfara et al., (2013) *Nucleic Acids Res* 42:2577-2590; Garneau et al., (2010) *Nature* 468:67-71; Magadan et al., (2012) *PLoS ONE* 7:e40913; Esvelt et al., (2013) *Nat Methods* 10(11):1116-1121 (wherein N is defined as any nucleotide, W is defined as either A or T, R is defined as a purine (A) or (G), and Y is defined as a pyrimidine (C) or (T)). In some embodiments, the PAM sequence is NGG. In some embodiments, the PAM sequence is NGAN. In some embodiments, the PAM sequence is NGNG. In some embodiments, the PAM is NNGRRRT. In some embodiments, the PAM sequence is NGGNG. In some embodiments, the PAM sequence may be NNAAAAW.

Modified Donor Polynucleotides

[0311] In some embodiments, donor polynucleotides are provided with chemistries suitable for delivery and stability within cells. Furthermore, in some embodiments, chemistries are provided that are useful for controlling the pharmacokinetics, biodistribution, bioavailability and/or efficacy of the donor polynucleotides described herein. Accordingly, in some embodiments, donor polynucleotides described herein may be modified, e.g., comprise a modified sugar

moiety, a modified internucleoside linkage, a modified nucleoside, a modified nucleotide and/or combinations thereof. In addition, the modified donor polynucleotides may exhibit one or more of the following properties: are not immune stimulatory; are nuclease resistant; have improved cell uptake compared to unmodified donor polynucleotides; and/or are not toxic to cells or mammals.

[0312] Nucleotide and nucleoside modifications have been shown to make a polynucleotide (e.g., a donor polynucleotide) into which they are incorporated more resistant to nuclease digestion than the native polynucleotide and these modified polynucleotides have been shown to survive intact for a longer time than unmodified polynucleotides. Specific examples of modified oligonucleotides include those comprising modified backbones (i.e. modified internucleoside linkage), for example, phosphorothioates, phosphotriesters, methyl phosphonates, short chain alkyl or cycloalkyl intersugar linkages or short chain heteroatomic or heterocyclic intersugar linkages. In some embodiments, oligonucleotides may have phosphorothioate backbones; heteroatom backbones, such as methylene(methylimino) or MMI backbones; amide backbones (see e.g., De Mesmaeker et al., *Acc. Chem. Res.* 1995, 28:366-374); morpholino backbones (see Summerton and Weller, U.S. Pat. No. 5,034,506); or peptide nucleic acid (PNA) backbones (wherein the phosphodiester backbone of the polynucleotide is replaced with a polyamide backbone, the nucleotides being bound directly or indirectly to the aza nitrogen atoms of the polyamide backbone, see Nielsen et al., *Science* 1991, 254, 1497). Phosphorus-containing modified linkages include, but are not limited to, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates comprising 3'alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates comprising 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'; see U.S. Pat. Nos. 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5, 177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321, 131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,031,272.1 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; and 5,625,050.

[0313] Morpholino-based oligomeric compounds are described in Dwaine A. Braasch and David R. Corey, *Biochemistry*, 2002, 41(14), 4503-4510; *Genesis*, volume 30, issue 3, 2001; Heasman, J., *Dev. Biol.*, 2002, 243, 209-214; Nasevicius et al., *Nat. Genet.*, 2000, 26, 216-220; Lacerra et al., *Proc. Natl. Acad. Sci.*, 2000, 97, 9591-9596; and U.S. Pat. No. 5,034,506, issued Jul. 23, 1991. In some embodiments, the morpholino-based oligomeric compound is a phosphorodiamidate morpholino oligomer (PMO) (e.g., as described in Iverson, *Curr. Opin. Mol. Ther.*, 3:235-238, 2001; and Wang et al., *J. Gene Med.*, 12:354-364, 2010).

[0314] Cyclohexenyl nucleic acid oligonucleotide mimetics are described in Wang et al., *J. Am. Chem. Soc.*, 2000, 122, 8595-8602.

[0315] Modified oligonucleotide backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl inter-

nucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These comprise those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH₂ component parts; see U.S. Pat. Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; and 5,677,439, each of which is herein incorporated by reference.

[0316] In some embodiments, the donor polynucleotides of the disclosure are stabilized against nucleolytic degradation such as by the incorporation of a modification (e.g., a nucleotide modification). In some embodiments, donor polynucleotides of the disclosure include a phosphorothioate at least the first, second, and/or third internucleotide linkage at the 5' and/or 3' end of the nucleotide sequence. In some embodiments, donor polynucleotides of the disclosure include one or more 2'-modified nucleotides, e.g., 2'-deoxy-2'-fluoro, 2'-O-methyl, 2'-O-methoxyethyl (2'-O-MOE), 2'-O-aminopropyl (2'-O-AP), 2'-O-dimethylaminoethyl (2'-O-DMAOE), 2'-O-dimethylaminopropyl (2'-O-DMAP), 2'-O-dimethylaminoethoxyethyl (2'-O-DMAEOE), or 2'-O-N-methylacetamido (2'-O-NMA). In some embodiments, donor polynucleotides of the disclosure include a phosphorothioate and a 2'-modified nucleotide as described herein.

[0317] Any of the modified chemistries described herein can be combined with each other, and that one, two, three, four, five, or more different types of modifications can be included within the same molecule. In some embodiments, the donor polynucleotide comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or modifications.

mRNA Components

[0318] In some embodiments, the systems provided by the disclosure comprise an engineered nuclease encoded by an mRNA. In some embodiments, the compositions provided by the disclosure comprise a nuclease system, wherein the nuclease comprising the nuclease system is encoded by an mRNA. In some embodiments, the compositions provided by the disclosure comprise a 53BP1 polypeptide inhibitor, wherein the 53BP1 polypeptide inhibitor is encoded by an mRNA. In some embodiments, the mRNA may be a naturally or non-naturally occurring mRNA. In some embodiments, the mRNA may include one or more modified nucleobases, nucleosides, or nucleotides, as described below, in which case it may be referred to as a "modified mRNA". In some embodiments, the mRNA may include a 5' untranslated region (5'-UTR), a 3' untranslated region (3'-UTR), and/or a coding region (e.g., an open reading frame). An mRNA may include any suitable number of base pairs, including tens (e.g., 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100), hundreds (e.g., 200, 300, 400, 500, 600, 700, 800, or 900) or thousands (e.g., 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10,000) of base pairs. Any number (e.g., all, some, or none) of nucleobases, nucleosides, or nucleotides may be an analog of a canonical species, substituted, modified,

or otherwise non-naturally occurring. In certain embodiments, all of a particular nucleobase type may be modified. In some embodiments, an mRNA as described herein may include a 5' cap structure, a chain terminating nucleotide, optionally a Kozak or Kozak-like sequence (also known as a Kozak consensus sequence), a stem-loop, a polyA sequence, and/or a polyadenylation signal.

[0319] A 5' cap structure or cap species is a compound including two nucleoside moieties joined by a linker and may be selected from a naturally occurring cap, a non-naturally occurring cap or cap analog, or an anti-reverse cap analog (ARCA). A cap species may include one or more modified nucleosides and/or linker moieties. For example, a natural mRNA cap may include a guanine nucleotide and a guanine (G) nucleotide methylated at the 7 position joined by a triphosphate linkage at their 5' positions, e.g., m⁷G(5')ppp(5')G, commonly written as m⁷GpppG. A cap species may also be an anti-reverse cap analog. A non-limiting list of possible cap species includes m⁷GpppG, m⁷Gpppm⁷G, m⁷3'dGpppG, m₂^{7,03'}GpppG, m₂^{7,03'}GppppG, m₂^{7,02'}GppppG, m⁷Gpppm⁷G, m⁷3'dGpppG, m₂^{7,03'}GpppG, m₂^{7,03'}GppppG, and m₂^{7,02'}GppppG.

[0320] An mRNA may instead or additionally include a chain terminating nucleoside. For example, a chain terminating nucleoside may include those nucleosides deoxygenated at the 2' and/or 3' positions of their sugar group. Such species may include 3'-deoxyadenosine (cordycepin), 3'-deoxyuridine, 3'-deoxycytosine, 3'-deoxyguanosine, 3'-deoxythymine, and 2',3'-dideoxynucleosides, such as 2',3'-dideoxyadenosine, 2',3'-dideoxyuridine, 2',3'-dideoxycytosine, 2',3'-dideoxyguanosine, and 2',3'-dideoxythymine. In some embodiments, incorporation of a chain terminating nucleotide into an mRNA, for example at the 3'-terminus, may result in stabilization of the mRNA, as described, for example, in International Patent Publication No. WO 2013/103659.

[0321] An mRNA may instead or additionally include a stem loop, such as a histone stem loop. A stem loop may include 2, 3, 4, 5, 6, 7, 8, or more nucleotide base pairs. For example, a stem loop may include 4, 5, 6, 7, or 8 nucleotide base pairs. A stem loop may be located in any region of an mRNA. For example, a stem loop may be located in, before, or after an untranslated region (a 5' untranslated region or a 3' untranslated region), a coding region, or a polyA sequence or tail. In some embodiments, a stem loop may affect one or more function(s) of an mRNA, such as initiation of translation, translation efficiency, and/or transcriptional termination.

[0322] An mRNA may instead or additionally include a polyA sequence and/or polyadenylation signal. A polyA sequence may be comprised entirely or mostly of adenine nucleotides or analogs or derivatives thereof. A polyA sequence may be a tail located adjacent to a 3' untranslated region of an mRNA. In some embodiments, a polyA sequence may affect the nuclear export, translation, and/or stability of an mRNA.

Modified RNA

[0323] In some embodiments, an RNA of the disclosure (e.g.: gRNA or mRNA) comprises one or more modified nucleobases, nucleosides, nucleotides or internucleoside linkages. In some embodiments, modified mRNAs and/or gRNAs may have useful properties, including enhanced stability, intracellular retention, enhanced translation, and/or

the lack of a substantial induction of the innate immune response of a cell into which the mRNA and/or gRNA is introduced, as compared to a reference unmodified mRNA and/or gRNA. Therefore, use of modified mRNAs and/or gRNAs may enhance the efficiency of protein production, intracellular retention of nucleic acids, as well as possess reduced immunogenicity.

[0324] In some embodiments, an mRNA and/or gRNA includes one or more (e.g., 1, 2, 3 or 4) different modified nucleobases, nucleosides, nucleotides or internucleoside linkages. In some embodiments, an mRNA and/or gRNA includes one or more (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, or more) different modified nucleobases, nucleosides, or nucleotides. In some embodiments, the modified gRNA may have reduced degradation in a cell into which the gRNA is introduced, relative to a corresponding unmodified gRNA. In some embodiments, the modified mRNA may have reduced degradation in a cell into which the mRNA is introduced, relative to a corresponding unmodified mRNA.

[0325] In some embodiments, the modified nucleobase is a modified uracil. Exemplary nucleobases and nucleosides having a modified uracil include pseudouridine (ψ), pyridin-4-one ribonucleoside, 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s^2U), 4-thio-uridine (s^4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho^5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), 3-methyl-uridine (m^3U), 5-methoxy-uridine (mo^5U), uridine 5-oxyacetic acid (cmo^5U), uridine 5-oxyacetic acid methyl ester ($mcmo^5U$), 5-carboxymethyl-uridine (cm^5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm^5U), 5-carboxyhydroxymethyl-uridine methyl ester ($mchm^5U$), 5-methoxycarbonylmethyl-uridine (mcm^5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm^5s^2U), 5-aminomethyl-2-thio-uridine (nm^5s^2U), 5-methylaminomethyl-uridine (mm^5U), 5-methylaminomethyl-2-thio-uridine (mm^5s^2U), 5-methylaminomethyl-2-seleno-uridine (mm^5se^2U), 5-carbamoylmethyl-uridine (ncm^5U), 5-carboxymethylaminomethyl-uridine ($cmnm^5U$), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm^5s^2U$), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine (tm^5U), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine (tm^5s^2U), 1-taurinomethyl-4-thio-pseudouridine, 5-methyl-uridine (m^5U , i.e., having the nucleobase deoxythymine), 1-methyl-pseudouridine ($m^1\psi$), 5-methyl-2-thio-uridine (m^5s^2U), 1-methyl-4-thio-pseudouridine ($m^1s^4\psi$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m^3\psi$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m^5D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp^3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine ($acp^3\psi$), 5-(isopentenylaminomethyl)uridine (inm^5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm^5s^2U), α -thio-uridine, 2'-O-methyl-uridine (Um), 5,2'-O-dimethyl-uridine (m^5Um), 2'-O-methyl-pseudouridine (ψm), 2-thio-2'-O-methyl-uridine (s^2Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm^5Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm^5Um), 5-carboxymethylaminomethyl-

2'-O-methyl-uridine ($cmnm^5Um$), 3,2'-O-dimethyl-uridine (m^3Um), and 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm^5Um), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-uridine, 2'-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, and 5-[3-(1-E-propenylamino)] uridine.

[0326] In some embodiments, the modified nucleobase is a modified cytosine. Exemplary nucleobases and nucleosides having a modified cytosine include 5-aza-cytidine, 6-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine (m^3C), N4-acetyl-cytidine (ac^4C), 5-formyl-cytidine (f^5C), N4-methyl-cytidine (m^4C), 5-methyl-cytidine (m^5C), 5-halo-cytidine (e.g., 5-iodo-cytidine), 5-hydroxymethyl-cytidine (hm^5C), 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine (s^2C), 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, 4-methoxy-1-methyl-pseudoisocytidine, lysidine (k^2C), α -thio-cytidine, 2'-O-methyl-cytidine (Cm), 5,2'-O-dimethyl-cytidine (m^5Cm), N4-acetyl-2'-O-methyl-cytidine (ac^4Cm), N4,2'-O-dimethyl-cytidine (m^4Cm), 5-formyl-2'-O-methyl-cytidine (f^5Cm), N4,N4,2'-O-trimethyl-cytidine (m^4_2Cm), 1-thio-cytidine, 2'-F-ara-cytidine, 2'-F-cytidine, and 2'-OH-ara-cytidine.

[0327] In some embodiments, the modified nucleobase is a modified adenine. Exemplary nucleobases and nucleosides having a modified adenine include a-thio-adenosine, 2-amino-purine, 2, 6-diaminopurine, 2-amino-6-halo-purine (e.g., 2-amino-6-chloro-purine), 6-halo-purine (e.g., 6-chloro-purine), 2-amino-6-methyl-purine, 8-azido-adenosine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-amino-purine, 7-deaza-8-aza-2-amino-purine, 7-deaza-2, 6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyl-adenosine (m^1A), 2-methyl-adenine (m^2A), N6-methyl-adenosine (m^6A), 2-methylthio-N6-methyl-adenosine (ms^2m^6A), N6-isopentenyl-adenosine (i^6A), 2-methylthio-N6-isopentenyl-adenosine (ms^2i^6A), N6-(cis-hydroxyisopentenyl)adenosine (io^6A), 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine (ms^2io^6A), N6-glycylcarbamoyl-adenosine (g^6A), N6-threonylcarbamoyl-adenosine (t^6A), N6-methyl-N6-threonylcarbamoyl-adenosine (m^6t^6A), 2-methylthio-N6-threonylcarbamoyl-adenosine (ms^2g^6A), N6,N6-dimethyl-adenosine (m^6_2A), N6-hydroxynorvalylcarbamoyl-adenosine (hn^6A), 2-methylthio-N6-hydroxynorvalylcarbamoyl-adenosine (ms^2hn^6A), N6-acetyl-adenosine (ac^6A), 7-methyl-adenine, 2-methylthio-adenine, 2-methoxy-adenine, α -thio-adenosine, 2'-O-methyl-adenosine (Am), N6,2'-O-dimethyl-adenosine (m^6Am), N6,N6,2'-O-trimethyl-adenosine (m^6_2Am), 1,2'-O-dimethyl-adenosine (m^1Am), 2'-O-ribosyladenosine (phosphate) (Ar(p)), 2-amino-N6-methyl-purine, 1-thio-adenosine, 8-azido-adenosine, 2'-F-ara-adenosine, 2'-F-adenosine, 2'-OH-ara-adenosine, and N6-(19-aminopentaaxanodecyl)-adenosine.

[0328] In some embodiments, the modified nucleobase is a modified guanine. Exemplary nucleobases and nucleosides having a modified guanine include a-thio-guanosine, inosine (I), 1-methyl-inosine (m^1I), wyosine (imG), methylwyosine (mimG), 4-demethyl-wyosine (imG-14), isowyosine (imG2), wybutosine (yW), peroxywybutosine (o_2yW),

hydroxywybutosine (OhyW), undermodified hydroxywybutosine (OhyW*), 7-deaza-guanosine, queuosine (Q), epoxy-queuosine (oQ), galactosyl-queuosine (galQ), mannosyl-queuosine (manQ), 7-cyano-7-deaza-guanosine (preQo), 7-aminomethyl-7-deaza-guanosine (preQi), archaeosine (G⁺), 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine (m⁷G), 6-thio-7-methyl-guanosine, 7-methyl-inosine, 6-methoxy-guanosine, 1-methyl-guanosine (m¹G), N2-methyl-guanosine (m²G), N2,N2-dimethyl-guanosine (m²₂G), N2,7-dimethyl-guanosine (m^{2,7}G), N2, N2,7-dimethyl-guanosine (m^{2,2,7}G), 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, N2,N2-dimethyl-6-thio-guanosine, α-thio-guanosine, 2'-O-methyl-guanosine (Gm), N2-methyl-2'-O-methyl-guanosine (m²Gm), N2,N2-dimethyl-2'-O-methyl-guanosine (m²₂Gm), 1-methyl-2'-O-methyl-guanosine (m¹Gm), N2,7-dimethyl-2'-O-methyl-guanosine (m^{2,7}Gm), 2'-O-methyl-inosine (Im), 1,2'-O-dimethyl-inosine (m¹Im), 2'-O-ribosylguanosine (phosphate) (Gr(p)), 1-thio-guanosine, O6-methyl-guanosine, 2'-F-ara-guanosine, and 2'-F-guanosine.

[0329] In some embodiments, an mRNA and/or gRNA of the disclosure includes a combination of one or more of the aforementioned modified nucleobases (e.g., a combination of 2, 3 or 4 of the aforementioned modified nucleobases.)

[0330] In some embodiments, the modified nucleobase is pseudouridine (w), N1-methylpseudouridine (m¹ψ), 2-thiouridine, 4'-thiouridine, 5-methylcytosine, 2-thio-1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-pseudouridine, 2-thio-5-aza-uridine, 2-thio-dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-pseudouridine, 4-methoxy-2-thio-pseudouridine, 4-methoxy-pseudouridine, 4-thio-1-methyl-pseudouridine, 4-thio-pseudouridine, 5-aza-uridine, dihydropseudouridine, 5-methoxyuridine, or 2'-O-methyl uridine. In some embodiments, an mRNA of the disclosure includes a combination of one or more of the aforementioned modified nucleobases (e.g., a combination of 2, 3 or 4 of the aforementioned modified nucleobases.) In one embodiment, the modified nucleobase is N1-methylpseudouridine (m¹ψ) and the mRNA of the disclosure is fully modified with N1-methylpseudouridine (m¹ψ). In some embodiments, N1-methylpseudouridine (m¹ψ) represents from 75-100% of the uracils in the mRNA. In some embodiments, N1-methylpseudouridine (m¹ψ) represents 100% of the uracils in the mRNA.

[0331] In some embodiments, the modified nucleobase is a modified cytosine. Exemplary nucleobases and nucleosides having a modified cytosine include N4-acetyl-cytidine (ac⁴C), 5-methyl-cytidine (m⁵C), 5-halo-cytidine (e.g., 5-iodo-cytidine), 5-hydroxymethyl-cytidine (hm⁵C), 1-methyl-pseudoisocytidine, 2-thio-cytidine (s²C), 2-thio-5-methyl-cytidine. In some embodiments, an mRNA of the disclosure includes a combination of one or more of the aforementioned modified nucleobases (e.g., a combination of 2, 3 or 4 of the aforementioned modified nucleobases.)

[0332] In some embodiments, the modified nucleobase is a modified adenine. Exemplary nucleobases and nucleosides having a modified adenine include 7-deaza-adenine, 1-methyl-adenosine (m¹A), 2-methyl-adenine (m²A), N6-methyl-adenosine (m⁶A). In some embodiments, an mRNA of the disclosure includes a combination of one or

more of the aforementioned modified nucleobases (e.g., a combination of 2, 3 or 4 of the aforementioned modified nucleobases.)

[0333] In some embodiments, the modified nucleobase is a modified guanine. Exemplary nucleobases and nucleosides having a modified guanine include inosine (I), 1-methyl-inosine (m¹I), wyosine (imG), methylwyosine (mimG), 7-deaza-guanosine, 7-cyano-7-deaza-guanosine (preQo), 7-aminomethyl-7-deaza-guanosine (preQi), 7-methyl-guanosine (m⁷G), 1-methyl-guanosine (m¹G), 8-oxo-guanosine, 7-methyl-8-oxo-guanosine. In some embodiments, an mRNA of the disclosure includes a combination of one or more of the aforementioned modified nucleobases (e.g., a combination of 2, 3 or 4 of the aforementioned modified nucleobases.)

[0334] In some embodiments, the modified nucleobase is 1-methyl-pseudouridine (m¹ψ), 5-methoxy-uridine (mo⁵U), 5-methyl-cytidine (m⁵C), pseudouridine (ψ), α-thio-guanosine, or α-thio-adenosine. In some embodiments, an mRNA of the disclosure includes a combination of one or more of the aforementioned modified nucleobases (e.g., a combination of 2, 3 or 4 of the aforementioned modified nucleobases.)

[0335] In certain embodiments, an mRNA and/or a gRNA of the disclosure is uniformly modified (i.e., fully modified, modified through-out the entire sequence) for a particular modification. For example, an mRNA can be uniformly modified with N1-methylpseudouridine (m¹ψ) or 5-methyl-cytidine (m⁵C), meaning that all uridines or all cytosine nucleosides in the mRNA sequence are replaced with N1-methylpseudouridine (m¹ψ) or 5-methyl-cytidine (m⁵C). Similarly, mRNAs of the disclosure can be uniformly modified for any type of nucleoside residue present in the sequence by replacement with a modified residue such as those set forth above.

[0336] In some embodiments, an mRNA of the disclosure may be modified in a coding region (e.g., an open reading frame encoding a polypeptide). In other embodiments, an mRNA may be modified in regions besides a coding region. For example, in some embodiments, a 5'-UTR and/or a 3'-UTR are provided, wherein either or both may independently contain one or more different nucleoside modifications. In such embodiments, nucleoside modifications may also be present in the coding region.

Ribonucleoproteins

[0337] In certain aspects, the site-directed polypeptide (e.g.: Cas nuclease) and genome-targeting nucleic acid (e.g.: gRNA or sgRNA) may each be administered separately to a cell or a subject. In certain aspects, the site-directed polypeptide may be pre-complexed with one or more guide RNAs, or one or more sgRNAs. Such pre-complexed material is known as a ribonucleoprotein particle (RNP). In some embodiments, the nuclease system comprises a ribonucleoprotein (RNP). In some embodiments, the nuclease system comprises a Cas9 RNP comprising a purified Cas9 protein in complex with a gRNA. Cas9 protein can be expressed and purified by any means known in the art. Ribonucleoproteins are assembled in vitro and can be delivered directly to cells using standard electroporation or transfection techniques known in the art.

Vectors

[0338] In some embodiments, one or more components of a genome editing system described herein are provided by

one or more vectors. In some embodiments, an inhibitor described herein is provided by a vector. In some embodiments, the site-directed nuclease (e.g., Cas nuclease), the donor polynucleotide, and/or a nucleic acid encoding an inhibitor described herein (e.g., a nucleic acid encoding a 53BP1 polypeptide inhibitor) are provided by one or more vectors. In some embodiments, the site-directed nuclease (e.g., Cas nuclease) and the donor polynucleotide may be provided by one or more vectors. In some embodiments, one or more gRNAs described herein are provided by a vector. In some embodiments, the vector may be a DNA vector. In some embodiments, the vector may be circular. In other embodiments, the vector may be linear. Non-limiting exemplary vectors include plasmids, phagemids, cosmids, artificial chromosomes, minichromosomes, transposons, viral vectors, and expression vectors.

[0339] 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 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 any helper virus. In some embodiments, the vector system described herein may also encode the viral components required for virus amplification and packaging.

[0340] Non-limiting exemplary viral vectors include adeno-associated virus (AAV) vector, lentivirus vectors, adenovirus vectors, 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 be a lentivirus vector. 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 (T) 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 30 kb-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 a Cas9 protein, while a second AAV vector may contain one or more guide sequences and one or more copies of donor polynucleotide.

[0341] A recombinant adeno-associated virus (AAV) vector can be used for delivery. Techniques to produce rAAV particles, in which an AAV genome to be packaged that includes the polynucleotide to be delivered, rep and cap genes, and helper virus functions are provided to a cell are standard in the art. Production of rAAV typically requires that the following components are present within a single cell (denoted herein as a packaging cell): a rAAV genome, AAV rep and cap genes separate from (i.e., not in) the rAAV genome, and helper virus functions. The AAV rep and cap genes may be from any AAV serotype for which recombinant virus can be derived, and may be from a different AAV serotype than the rAAV genome ITRs, including, but not limited to, AAV serotypes AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-6, AAV-7, AAV-8, AAV-9, AAV-10, AAV-11, AAV-12, AAV-13 and AAV rh.74. Production of pseudotyped rAAV is disclosed in, for example, international patent application publication number WO 01/83692. In some embodiments, the vector is AAV6.

[0342] A method of generating a packaging cell involves creating a cell line that stably expresses all of the necessary components for AAV particle production. For example, a plasmid (or multiple plasmids) comprising a rAAV genome lacking AAV rep and cap genes, AAV rep and cap genes separate from the rAAV genome, and a selectable marker, such as neomycin resistance gene, are integrated into the genome of a cell. AAV genomes have been introduced into bacterial plasmids by procedures such as GC tailing (Samulski et al., 1982, Proc. Natl. Acad. Sci. USA, 79:2077-2081), addition of synthetic linkers containing restriction endonuclease cleavage sites (Laughlin et al., 1983, Gene, 23:65-73) or by direct, blunt-end ligation (Senapathy & Carter, 1984, J. Biol. Chem., 259:4661-4666). The packaging cell line can then be infected with a helper virus, such as adenovirus. The advantages of this method are that the cells are selectable and are suitable for large-scale production of rAAV. Other examples of suitable methods employ adenovirus or baculovirus, rather than plasmids, to introduce rAAV genomes and/or rep and cap genes into packaging cells.

[0343] In certain embodiments, a viral vector may be modified to target a particular tissue or cell type. For example, viral surface proteins may be altered to decrease or eliminate viral protein binding to its natural cell surface

receptor(s). The surface proteins may also be engineered to interact with a receptor specific to a desired cell type. Viral vectors may have altered host tropism, including limited or redirected tropism. Certain engineered viral vectors are described, for example, in WO2011130749, WO2015009952, U.S. Pat. No. 5,817,491, WO2014135998, and WO2011125054. In some embodiments, the vector may be capable of driving expression of one or more coding sequences in a cell. 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.

[0344] In some embodiments, the vector may comprise a nucleotide sequence encoding the nuclease described herein. 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. In some embodiments, the nucleotide sequence encoding the nuclease may be operably linked to at least one transcriptional or translational control sequence.

[0345] 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- α (EF1 α) 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 EF1 α 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). In some embodiments, the promoter may be a tissue-specific promoter. In some embodiments, the tissue-specific promoter is exclusively or predominantly expressed in liver tissue. Non-limiting exemplary tissue-specific promoters include B29 promoter, CD14 promoter, CD43 promoter, CD45 promoter, CD68 promoter, desmin promoter, elastase-1 promoter, endoglin promoter,

fibronectin promoter, Flt-1 promoter, GFAP promoter, GPIIb promoter, ICAM-2 promoter, INF- β promoter, Mb promoter, Nphs1 promoter, OG-2 promoter, SP-B promoter, SYN1 promoter, and WASP promoter.

[0346] In some embodiments, the nuclease encoded by the vector may be a Cas protein, such as a Cas9 protein or Cpf1 protein. The vector system may further comprise a vector comprising a nucleotide sequence encoding the guide RNA described herein. In some embodiments, the vector system may comprise one copy of the guide RNA. In other embodiments, the vector system may comprise 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 have other different properties, such as activity or stability within the Cas9 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. In some embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to at least one promoter. In some embodiments, the promoter may be recognized by RNA polymerase III (Pol III). Non-limiting examples of Pol III promoters include U6, H1 and tRNA 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 some embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to a mouse or human tRNA 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 tracrRNA of the guide RNA may be provided on the same vector. In some embodiments, the nucleotide encoding the crRNA and the nucleotide encoding the tracrRNA may be driven by the same promoter. In some embodiments, the crRNA and tracrRNA may be transcribed into a single transcript. For example, the crRNA and tracrRNA may be processed from the single transcript to form a double-molecule guide RNA. Alternatively, the crRNA and tracrRNA may be transcribed into a single-molecule guide RNA. In other embodiments, the crRNA and the tracrRNA may be driven by their corresponding promoters on the same vector. In yet other embodiments, the crRNA and the tracrRNA may be encoded by different vectors.

[0347] In some embodiments, the nucleotide sequence encoding the guide RNA may be located on the same vector comprising the nucleotide sequence encoding a Cas9 protein. In some embodiments, expression of the guide RNA and of the Cas9 protein may be driven by different promoters. In some embodiments, expression of the guide RNA may be driven by the same promoter that drives expression of the Cas9 protein. In some embodiments, the guide RNA and the Cas9 protein transcript may be contained within a single transcript. For example, the guide RNA may be within an untranslated region (UTR) of the Cas9 protein transcript. In some embodiments, the guide RNA may be within the 5' UTR of the Cas9 protein transcript. In other embodiments, the guide RNA may be within the 3' UTR of the Cas9 protein transcript. In some embodiments, the intracellular half-life of the Cas9 protein 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 Cas9 protein 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 Cas9 protein and the guide RNA in close proximity on the same vector may facilitate more efficient formation of the CRISPR complex.

[0348] In some embodiments, the vector system may further comprise a vector comprising the donor polynucleotide described herein. In some embodiments, the vector system may comprise one copy of the donor polynucleotide. In other embodiments, the vector system may comprise more than one copy of the donor polynucleotide. In some embodiments, the vector system may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, or more copies of the donor polynucleotide. The multiple copies of the donor polynucleotide may be located on the same or different vectors. The multiple copies of the donor polynucleotide may also be adjacent to one another, or separated by other nucleotide sequences or vector elements.

[0349] In some embodiments, the vector system further comprises a vector comprising a nucleotide sequence encoding an inhibitor described herein (e.g., a 53BP1 polypeptide inhibitor). In some embodiments, the vector system comprises one or more copies (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more copies) of the nucleotide sequence. In some embodiments, the multiple copies of the nucleotide sequences are located on the same or different vectors. In some embodiments, the multiple copies of the nucleotide sequence are adjacent or separated by other nucleotide sequences or vector elements.

[0350] A vector system may comprise 1-3 vectors. 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 or donor polynucleotides are used for multiplexing, or when multiple copies of the guide RNA or the donor polynucleotide are used, the vector system may comprise more than three vectors. In some embodiments, the vector system comprises a vector comprising a nucleotide sequence encoding an inhibitor described herein (e.g., a 53BP1 polypeptide inhibitor).

[0351] In some embodiments, the nucleotide sequence encoding a Cas9 protein, a nucleotide sequence encoding the guide RNA, and a donor polynucleotide may be located on the same or separate vectors. In some embodiments, a nucleotide sequence encoding an inhibitor described herein (e.g., a 53BP1 polypeptide inhibitor) is further located on the same or separate vectors as the Cas9 protein, the nucleotide sequence encoding the guide RNA, or the donor polynucleotide. In some embodiments, all of the sequences may be located on the same vector. In some embodiments, two or more sequences may be located on the same vector. The sequences may be oriented in the same or different directions and in any order on the vector. In some embodiments, the nucleotide sequence encoding the Cas9 protein and the nucleotide sequence encoding the guide RNA may be located on the same vector. In some embodiments, the nucleotide sequence encoding the Cas9 protein and the donor polynucleotide may be located on the same vector. In some embodiments, the nucleotide sequence encoding the Cas9 protein, the nucleotide sequences encoding the guide RNA, and/or the donor polynucleotide are located on the

same vector as a nucleotide sequence encoding an inhibitor described herein (e.g., a 53BP1 polypeptide inhibitor). In some embodiments, the nucleotide sequence encoding the guide RNA and the donor polynucleotide may be located on the same vector. In some embodiments, the vector system may comprise a first vector comprising the nucleotide sequence encoding the Cas9 protein, and a second vector comprising the nucleotide sequence encoding the guide RNA and the donor polynucleotide. In some embodiments, the vector system comprises a third vector comprising the nucleotide sequence encoding an inhibitor (e.g., a 53BP1 polypeptide inhibitor) described herein.

Editing of a Target Gene

[0352] Provided herein are methods of gene-editing within a target gene by repair of a DNA DSB in the target gene by the HDR pathway using a donor polynucleotide. In some embodiments, a gene-edit comprises a correction of a mutation in a target gene. In some embodiments, a gene-edit comprises a replacement of a target gene with a different gene or a variant of the target gene. In some embodiments, a mutation in a target gene results in a monogenic disorder. Non-limiting examples of monogenic disorders caused by mutations in a target gene provide in Table 1.

[0353] In some embodiments, a gene-edit of the disclosure comprises a disruption to the HBB gene in an HSC. In some embodiments, a disruption of the HBB gene comprises editing of a mutation in the HBB gene. In some embodiments, editing of a mutation in the HBB gene comprises repair of a DSB in the HBB gene using a donor polynucleotide, wherein the donor polynucleotide encodes a correction of the HBB gene. In some embodiments, correction of the HBB gene in an HSC renders the engineered HSC suitable for treatment of a hemoglobinopathy. Methods of editing the HBB gene are described by US20180021413 which is incorporated by reference herein.

[0354] In some embodiments, a gene-edit of the disclosure comprises a disrupted T cell receptor alpha constant region (TRAC) gene in a T cell. This disruption leads to loss of function of the T cell receptor (TCR) and renders the engineered T cell non-alloreactive and suitable for allogeneic transplantation, minimizing the risk of graft versus host disease. In some embodiments, expression of the endogenous TRAC gene is eliminated to prevent a graft-versus-host response.

[0355] In some embodiments, a disruption in the TRAC gene expression is created by knocking a chimeric antigen receptor (CAR) into the TRAC gene (e.g., using an adeno-associated viral (AAV) vector and donor template). In some embodiments, a disruption in the TRAC gene expression is created with a nuclease and gRNAs targeting the TRAC genomic region. In some embodiments, a genomic deletion in the TRAC gene is created by HDR, wherein a chimeric antigen receptor (CAR) replaces a segment of the TRAC gene (e.g., using an adeno-associated viral (AAV) vector and donor template). In some embodiments, a disruption in the TRAC gene expression is created with a nuclease and gRNAs targeting the TRAC genomic region, and knocking a chimeric antigen receptor (CAR) into the TRAC gene.

[0356] In some embodiments, an engineered T cell comprises a disrupted beta-2-microglobulin ($\beta 2M$) gene. $\beta 2M$ is a common (invariant) component of MHC I complexes. Disrupting its expression by gene editing will prevent host versus therapeutic allogeneic T cells responses leading to

increased allogeneic T cell persistence. In some embodiments, expression of the endogenous $\beta 2M$ gene is eliminated to prevent a host-versus-graft response.

[0357] Methods of editing the TRAC gene locus and $\beta 2M$ gene locus are described by International Application No. PCT/US2018/032334, filed May 11, 2018, incorporated herein by reference.

Pharmaceutical Compositions

[0358] The present disclosure includes pharmaceutical compositions comprising a donor polynucleotide, a gRNA, and a Cas9 protein, in combination with one or more pharmaceutically acceptable excipient, carrier or diluent. In some embodiments, the pharmaceutical composition further comprises one or more inhibitors described herein (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor). In particular embodiments, the donor polynucleotides encapsulated in a nanoparticle, e.g., a lipid nanoparticle. In some embodiments, the gRNA is encapsulated in a nanoparticle. In some embodiments, a Cas nuclease (e.g., SpCas9) is encapsulated in a nanoparticle. In some embodiments, the one or more inhibitors is encapsulated in a nanoparticle. In particular embodiments, an mRNA encoding a Cas nuclease or nanoparticle encapsulating a Cas nuclease is present in a pharmaceutical composition. In some embodiments, an mRNA encoding an inhibitor described herein (e.g., a 53BP1 polypeptide inhibitor) is present in a pharmaceutical composition of the disclosure. In various embodiments, the one or more mRNA present in the pharmaceutical composition is encapsulated in a nanoparticle, e.g., a lipid nanoparticle. In particular embodiments, the molar ratio of the first mRNA to the second mRNA is about 1:50, about 1:25, about 1:10, about 1:5, about 1:4, about 1:3, about 1:2, about 1:1, about 2:1, about 3:1, about 4:1, or about 5:1, about 10:1, about 25:1 or about 50:1. In particular embodiments, the molar ratio of the first mRNA to the second mRNA is greater than one.

[0359] In some embodiments, the ratio between the lipid composition and the donor polynucleotide can be about 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 21:1, 22:1, 23:1, 24:1, 25:1, 26:1, 27:1, 28:1, 29:1, 30:1, 31:1, 32:1, 33:1, 34:1, 35:1, 36:1, 37:1, 38:1, 39:1, 40:1, 41:1, 42:1, 43:1, 44:1, 45:1, 46:1, 47:1, 48:1, 49:1, 50:1, 51:1, 52:1, 53:1, 54:1, 55:1, 56:1, 57:1, 58:1, 59:1 or 60:1 (wt/wt). In some embodiments, the wt/wt ratio of the lipid composition to the polynucleotide is about 20:1 or about 15:1.

[0360] In one embodiment, the lipid nanoparticles described herein can comprise polynucleotides (e.g., donor polynucleotide) in a lipid:polynucleotide weight ratio of 5:1, 10:1, 15:1, 20:1, 25:1, 30:1, 35:1, 40:1, 45:1, 50:1, 55:1, 60:1 or 70:1, or a range or any of these ratios such as, but not limited to, 5:1 to about 10:1, from about 5:1 to about 15:1, from about 5:1 to about 20:1, from about 5:1 to about 25:1, from about 5:1 to about 30:1, from about 5:1 to about 35:1, from about 5:1 to about 40:1, from about 5:1 to about 45:1, from about 5:1 to about 50:1, from about 5:1 to about 55:1, from about 5:1 to about 60:1, from about 5:1 to about 70:1, from about 10:1 to about 15:1, from about 10:1 to about 20:1, from about 10:1 to about 25:1, from about 10:1 to about 30:1, from about 10:1 to about 35:1, from about 10:1 to about 40:1, from about 10:1 to about 45:1, from about 10:1 to about 50:1, from about 10:1 to about 55:1, from about 10:1 to about 60:1, from about 10:1 to about

70:1, from about 15:1 to about 20:1, from about 15:1 to about 25:1, from about 15:1 to about 30:1, from about 15:1 to about 35:1, from about 15:1 to about 40:1, from about 15:1 to about 45:1, from about 15:1 to about 50:1, from about 15:1 to about 55:1, from about 15:1 to about 60:1 or from about 15:1 to about 70:1.

[0361] In one embodiment, the lipid nanoparticles described herein can comprise the polynucleotide in a concentration from approximately 0.1 mg/ml to 2 mg/ml such as, but not limited to, 0.1 mg/ml, 0.2 mg/ml, 0.3 mg/ml, 0.4 mg/ml, 0.5 mg/ml, 0.6 mg/ml, 0.7 mg/ml, 0.8 mg/ml, 0.9 mg/ml, 1.0 mg/ml, 1.1 mg/ml, 1.2 mg/ml, 1.3 mg/ml, 1.4 mg/ml, 1.5 mg/ml, 1.6 mg/ml, 1.7 mg/ml, 1.8 mg/ml, 1.9 mg/ml, 2.0 mg/ml or greater than 2.0 mg/ml.

Methods of Treatment

[0362] Provided herein are methods of treating a patient with a disease by gene-editing genomic DNA molecule, such as correcting a mutation in genomic DNA molecule. In some embodiments, the method may comprise introducing a donor polynucleotide, system, vector, or pharmaceutical composition described herein into a cell. In some embodiments, the method further comprises introducing one or more inhibitors described herein (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) to the cell. In some embodiments, the method may comprise administering a donor polynucleotide, system, vector, or pharmaceutical composition to a subject in need thereof (e.g., a patient having a disease caused by a mutation). In some embodiments, the method further comprises administering one or more inhibitors described herein (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) to the subject in need thereof.

[0363] Embodiments of the disclosure encompass methods for editing a target nucleic acid molecule (a genomic DNA) in a cell. In some embodiments, the method comprises introducing a donor polynucleotide described herein into a cell. In some embodiments, the method further comprises introducing one or more inhibitors described herein (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the one or more inhibitors (e.g., a vector, an mRNA) to the cell. In some embodiments, the method comprises contacting the cell with a pharmaceutical composition described herein. In some embodiments, the method comprises generating a stable cell line comprising a targeted edited nucleic acid molecule. In some embodiments, the cell is a eukaryotic cell. Non-limiting examples of eukaryotic cells include yeast cells, plant cells, insect cells, cells from an invertebrate animal, cells from a vertebrate animal, mammalian cells, rodent cells, mouse cells, rat cells, and human cells. 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. Similarly, the target sequence may be from any such cells or in any such cells.

[0364] The donor polynucleotide, system, vector, or pharmaceutical composition described herein may be introduced into the cell via any methods known in the art, such as, e.g., viral or bacteriophage infection, transfection, conjugation, protoplast fusion, lipofection, electroporation, calcium phosphate precipitation, polyethyleneimine (PEI)-mediated transfection, DEAE-dextran-mediated transfection, liposome-mediated transfection, particle gun technology, calcium phosphate precipitation, shear-driven cell permeation,

fusion to a cell-penetrating peptide followed by cell contact, microinjection, and nanoparticle-mediated delivery. In some embodiments, the vector system may be introduced into the cell via viral infection. In some embodiments, the vector system may be introduced into the cell via bacteriophage infection.

[0365] Embodiments of the invention also encompass treating a patient with donor polynucleotide, system, vector, or pharmaceutical composition described herein. In some embodiments, the method may comprise administering the donor polynucleotide, system, vector, or pharmaceutical composition described herein to the patient. In some embodiments, the method further comprises administering one or more inhibitors described herein (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) to the patient. The method may be used as a single therapy or in combination with other therapies available in the art. In some embodiments, the patient may have a mutation (such as, e.g., insertion, deletion, substitution, chromosome translocation) in a disease-associated gene. In some embodiments, administration of the donor polynucleotide, system, vector, or pharmaceutical composition may result in a mutation comprising an insertion, deletion, or substitution of one or more nucleotides of the disease-associated gene in the patient. Certain embodiments may include methods of repairing the patient's mutation in the disease-associated gene. In some embodiments, the mutation may result in one or more amino acid changes in a protein expressed from the disease-associated gene. In some embodiments, the mutation may result in one or more nucleotide changes in an RNA expressed from the disease-associated gene. In some embodiments, the mutation may alter the expression level of the disease-associated gene. In some embodiments, the mutation may result in increased or decreased expression of the gene. In some embodiments, the mutation may result in gene knockdown in the patient. In some embodiments, the administration of the donor polynucleotide, system, vector, or pharmaceutical composition may result in the correction of the patient's mutation in the disease-associated gene. In some embodiments, the administration of the donor polynucleotide, system, vector, or pharmaceutical composition may result in gene knockout in the patient. In some embodiments, the administration of the donor polynucleotide, system, vector, or pharmaceutical composition system may result in replacement of an exon sequence, an intron sequence, a transcriptional control sequence, a translational control sequence, or a non-coding sequence of the disease-associated gene.

[0366] In some embodiments, the administration of the donor polynucleotide, system, vector, or pharmaceutical composition may result in integration of an exogenous sequence (e.g., the donor polynucleotide sequence) into the patient's genomic DNA. In some embodiments, the exogenous sequence may comprise a protein or RNA coding sequence operably linked to an exogenous promoter sequence such that, upon integration of the exogenous sequence into the patient's genomic DNA, the patient is capable of expressing the protein or RNA encoded by the integrated sequence. The exogenous sequence may provide a supplemental or replacement protein coding or non-coding sequence. For example, the administration of the donor polynucleotide, system, vector, or pharmaceutical composition may result in the replacement of the mutant portion of the disease-associated gene in the patient. In some embodi-

ments, the mutant portion may include an exon of the disease-associated gene. In other embodiments, the integration of the exogenous sequence may result in the expression of the integrated sequence from an endogenous promoter sequence present on the patient's genomic DNA. For example, the administration of the donor polynucleotide, system, vector, or pharmaceutical composition may result in supply of a functional gene product of the disease-associated gene to rectify the patient's mutation. In yet other embodiments, the administration of the donor polynucleotide, system, vector, or pharmaceutical composition may result in integration of an exon sequence, an intron sequence, a transcriptional control sequence, a translational control sequence, or a non-coding sequence into the patient's genomic DNA.

[0367] Additional embodiments of the invention also encompass methods of treating the patient in a tissue-specific manner. In some embodiments, the method may comprise administering the donor polynucleotide, system, vector, or pharmaceutical composition comprising a tissue-specific promoter as described herein to the patient. Non-limiting examples of suitable tissues for treatment by the methods include the immune system, neuron, muscle, pancreas, blood, kidney, bone, lung, skin, liver, and breast tissues.

[0368] In some embodiments, the disclosure provides a method to correct a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising contacting the cell with a donor polynucleotide described herein, a system comprising a donor polynucleotide, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, wherein when the donor polynucleotide, system or composition contacts the cell, an HDR DNA repair pathway inserts the donor polynucleotide into a double-stranded DNA break introduced into the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises contacting the cell with an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway.

[0369] In some embodiments, the disclosure provides a method to correct a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising contacting the cell with a donor polynucleotide or recombinant vector described herein, a system comprising a donor polynucleotide, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, wherein when the donor polynucleotide or a recombinant vector, system or composition contacts the cell, an HDR DNA repair pathway exchanges the donor polynucleotide for a corresponding nucleic acid region in the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises contacting the cell with an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway.

[0370] In some embodiments, the disclosure provides a method to correct a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising contacting the cell with a donor polynucleotide or recombinant vector

described herein, a system comprising a donor polynucleotide, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, wherein when the donor polynucleotide or recombinant vector, system or composition contacts the cell, an HDR DNA repair pathway exchanges a region around a double-stranded DNA break introduced into the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises contacting the cell with an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway.

[0371] In some embodiments, the disclosure provides a method of treating a patient with a disease by correcting a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising isolating a cell from the patient, contacting the cell with a donor polynucleotide described herein, a system comprising a donor polynucleotide, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, wherein, when the donor polynucleotide, system or composition contacts the cell, an HDR DNA repair pathway inserts the donor polynucleotide into a double-stranded DNA break introduced into the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises contacting the cell with an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway.

[0372] In some embodiments, the disclosure provides a method of treating a patient with a disease by correcting a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising isolating a cell from the patient, contacting the cell with a donor polynucleotide or recombinant vector described herein, a system comprising a donor polynucleotide or recombinant vector, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, an HDR DNA repair pathway exchanges the donor polynucleotide or recombinant vector for a corresponding nucleic acid region in the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises contacting the cell with an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway.

[0373] In some embodiments, the disclosure provides a method of treating a patient with a disease by correcting a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising isolating a cell from the patient, contacting the cell with a donor polynucleotide or recombinant vector described herein, a system comprising a donor polynucleotide or recombinant vector, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, an HDR DNA repair pathway exchanges a region around a double-stranded DNA break introduced into the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises contacting the cell with an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK

inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway. In some embodiments, the disclosure provides a method of treating a patient with a disease by correcting a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising administering to the patient an effective amount of a donor polynucleotide described herein, a system comprising a donor polynucleotide, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, wherein, when the donor polynucleotide, system or composition is administered, an HDR DNA repair pathway inserts the donor polynucleotide into a double-stranded DNA break introduced into the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises administering to the patient an effective amount of an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway. In some embodiments, the disclosure provides a method of treating a patient with a disease by correcting a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising administering to the patient an effective amount of a donor polynucleotide or recombinant vector described herein, a system comprising a donor polynucleotide, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, wherein, when the donor polynucleotide, system or composition is administered, an HDR DNA repair pathway exchanges the donor polynucleotide or recombinant vector for a corresponding nucleic acid region in the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises administering to the patient an effective amount of an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway.

[0374] In some embodiments, the disclosure provides a method of treating a patient with a disease by correcting a mutation in a genomic DNA molecule (gDNA) in a cell, the method comprising administering to the patient an effective amount of a donor polynucleotide or recombinant vector described herein, a system comprising a donor polynucleotide, a gRNA, and a site-directed nuclease, according to the disclosure, or a pharmaceutical composition described herein, wherein, when the donor polynucleotide, system or composition is administered, an HDR DNA repair pathway exchanges a region around a double-stranded DNA break introduced into the gDNA at a location proximal to the mutation, thereby correcting the mutation. In some embodiments, the method further comprises administering to the patient an effective amount of an inhibitor (e.g., a 53BP1 polypeptide inhibitor, a DNA-PK inhibitor) or a nucleic acid encoding the inhibitor (e.g., a vector, e.g., an mRNA) described herein to increase the efficiency of the HDR DNA repair pathway.

[0375] In some embodiments, the cell is a hematopoietic stem cell. In some embodiments, the cell is a patient-specific induced pluripotent stem cell (iPSC). In some embodiments, the cell is a hepatocyte. In some embodiments, the method further comprises differentiating the iPSC comprising the corrected mutation into a differentiated cell; and implanting

the differentiated cell into a patient. In some embodiments, treatment results in the translation of an mRNA transcribed from the genomic DNA molecule (gDNA) comprising the inserted donor polynucleotide, wherein the translation results in the formation of a translation product (protein) that alleviates the disease or that does not cause or contribute to the disease.

Definitions

[0376] Terms used in the claims and specification are defined as set forth below unless otherwise specified. In the case of direct conflict with a term used in a parent provisional patent application, the term used in the instant application shall control.

[0377] It must be noted that, as used in the specification and the appended claims, the singular forms “a,” “an” and “the” include plural referents unless the context clearly dictates otherwise. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular.

[0378] About: As used herein, the term “about” (alternatively “approximately”) will be understood by persons of ordinary skill and will vary to some extent depending on the context in which it is used. If there are uses of the term which are not clear to persons of ordinary skill given the context in which it is used, “about” will mean up to plus or minus 10% of the particular value.

[0379] Amino acid: As used herein, the term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that function in a manner similar to a naturally occurring amino acid.

[0380] Amino acids can be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, can be referred to by their commonly accepted single-letter codes.

[0381] Amino acid substitution: As used herein, an “amino acid substitution” refers to the replacement of at least one existing amino acid residue in a predetermined amino acid sequence (an amino acid sequence of a starting polypeptide) with a second, different “replacement” amino acid residue. An “amino acid insertion” refers to the incorporation of at least one additional amino acid into a predetermined amino acid sequence. While the insertion will usually consist of the insertion of one or two amino acid residues, larger “peptide insertions,” can also be made, e.g., insertion of about three to about five or even up to about ten, fifteen, or twenty amino

acid residues. The inserted residue(s) may be naturally occurring or non-naturally occurring as disclosed above. An “amino acid deletion” refers to the removal of at least one amino acid residue from a predetermined amino acid sequence.

[0382] Base Composition: As used herein, the term “base composition” refers to the proportion of the total bases of a nucleic acid consisting of guanine+cytosine or thymine (or uracil)+adenine nucleobases.

[0383] Base Pair: As used herein, the term “base pair” refers to two nucleobases on opposite complementary polynucleotide strands, or regions of the same strand, that interact via the formation of specific hydrogen bonds. As used herein, the term “Watson-Crick base pairing”, used interchangeably with “complementary base pairing”, refers to a set of base pairing rules, wherein a purine always binds with a pyrimidine such that the nucleobase adenine (A) forms a complementary base pair with thymine (T) and guanine (G) forms a complementary base pair with cytosine (C) in DNA molecules. In RNA molecules, thymine is replaced by uracil (U), which, similar to thymine (T), forms a complementary base pair with adenine (A). The complementary base pairs are bound together by hydrogen bonds and the number of hydrogen bonds differs between base pairs. As is known in the art, guanine (G)-cytosine (C) base pairs are bound by three (3) hydrogen bonds and adenine (A)-thymine (T) or uracil (U) base pairs are bound by two (2) hydrogen bonds.

[0384] Base pairing interactions that do not follow these rules can occur in natural, non-natural, and synthetic nucleic acids and are referred to herein as “non-Watson-Crick base pairing” or alternatively “non-canonical base pairing”. A “wobble base pair” is a pairing between two nucleobases in RNA molecules that does not follow Watson-Crick base pair rules. For example, inosine is a nucleoside that is structurally similar to guanosine, but is missing the 2-amino group. Inosine is able to form two hydrogen bonds with each of the four natural nucleobases (Oda et al., (1991) *Nucleic Acids Res* 19:5263-5267) and it is often used by researchers as a “universal” base, meaning that it can base pair with all the naturally-occurring or canonical bases. The four main wobble base pairs are the guanine-uracil (G-U) base pair, the hypoxanthine-uracil (I-U) base pair, the hypoxanthine-adenine (I-A) base pair, and the hypoxanthine-cytosine (I-C) base pair. In order to maintain consistency of nucleic acid nomenclature, “I” is used for hypoxanthine because hypoxanthine is the nucleobase of inosine; nomenclature otherwise follows the names of nucleobases and their corresponding nucleosides (e.g., “G” for both guanine and guanosine—as well as for deoxyguanosine). The thermodynamic stability of a wobble base pair is comparable to that of a Watson-Crick base pair. Wobble base pairs play a role in the formation of secondary structure in RNA molecules.

[0385] Blunt-end: As used herein, the term “blunt-end” “blunt-ended” refers to the structure of an end of a duplexed or double-stranded nucleic acid (e.g., DNA), wherein both complementary strands comprising the duplex terminate, at least at one end, in a base pair. Hence, neither strand comprising the duplex extends further from the end than the other.

[0386] Chimeric antigen receptor (CAR): As used herein, the term “chimeric antigen receptor” or “CAR” refers to an artificial immune cell receptor that is engineered to recognize and bind to an antigen expressed by tumor cells.

Generally, a CAR is designed for a T cell and is a chimera of a signaling domain of the T-cell receptor (TCR) complex and an antigen-recognizing domain (e.g., a single chain fragment (scFv) of an antibody or other antibody fragment) (Enblad et al., *Human Gene Therapy*. 2015; 26(8):498-505). A T cell that expresses a CAR is referred to as a CAR T cell. CARs have the ability to redirect T-cell specificity and reactivity toward a selected target in a non-MHC-restricted manner. The non-MHC-restricted antigen recognition gives T-cells expressing CARs the ability to recognize an antigen independent of antigen processing, thus bypassing a major mechanism of tumor escape. Moreover, when expressed in T-cells, CARs advantageously do not dimerize with endogenous T-cell receptor (TCR) alpha and beta chains.

[0387] There are four generations of CARs, each of which contains different components. First generation CARs join an antibody-derived scFv to the CD3zeta (ξ or z) intracellular signaling domain of the T-cell receptor through hinge and transmembrane domains. Second generation CARs incorporate an additional domain, e.g., CD28, 4-1BB (41BB), or ICOS, to supply a costimulatory signal. Third-generation CARs contain two costimulatory domains fused with the TCR CD3 ξ chain. Third-generation costimulatory domains may include, e.g., a combination of CD3 ξ , CD27, CD28, 4-1BB, ICOS, or OX40. CARs, in some embodiments, contain an ectodomain (e.g., CD3 ξ), commonly derived from a single chain variable fragment (scFv), a hinge, a transmembrane domain, and an endodomain with one (first generation), two (second generation), or three (third generation) signaling domains derived from CD3Z and/or co-stimulatory molecules (Maude et al., *Blood*. 2015; 125(26):4017-4023; Kakarla and Gottschalk, *Cancer J*. 2014; 20(2):151-155).

[0388] CARs typically differ in their functional properties. The CD3t signaling domain of the T-cell receptor, when engaged, will activate and induce proliferation of T-cells but can lead to anergy (a lack of reaction by the body's defense mechanisms, resulting in direct induction of peripheral lymphocyte tolerance). Lymphocytes are considered anergic when they fail to respond to a specific antigen. The addition of a costimulatory domain in second-generation CARs improved replicative capacity and persistence of modified T-cells. Similar antitumor effects are observed in vitro with CD28 or 4-1BB CARs, but preclinical in vivo studies suggest that 4-1BB CARs may produce superior proliferation and/or persistence. Clinical trials suggest that both of these second-generation CARs are capable of inducing substantial T-cell proliferation in vivo, but CARs containing the 4-1BB costimulatory domain appear to persist longer. Third generation CARs combine multiple signaling domains (costimulatory) to augment potency.

[0389] Codon: As used herein, the term "codon" refers to a sequence of three nucleotides that together form a unit of genetic code in a DNA or RNA molecule. A codon is operationally defined by the initial nucleotide from which translation starts and sets the frame for a run of successive nucleotide triplets, which is known as an "open reading frame" (ORF). For example, the string GGGAAACCC, if read from the first position, contains the codons GGG, AAA, and CCC; if read from the second position, it contains the codons GGA and AAC; and if read from the third position, GAA and ACC. Thus, every nucleic sequence read in its 5'→3' direction comprises three reading frames, each producing a possibly distinct amino acid sequence (in the given

example, Gly-Lys-Pro, Gly-Asn, or Glu-Thr, respectively). DNA is double-stranded defining six possible reading frames, three in the forward orientation on one strand and three reverse on the opposite strand. Open reading frames encoding polypeptides are typically defined by a start codon, usually the first AUG codon in the sequence.

[0390] Corrects or induces a mutation: As used herein, the term "corrects or induces a mutation" refers to a function of a donor polynucleotide, such as those described herein, to incorporate a desired alteration into a nucleotide sequence comprising a genomic DNA (gDNA) molecule upon insertion of the donor polynucleotide into a double-strand break (DSB) induced in the gDNA molecule, thereby changing the nucleotide sequence of the gDNA.

[0391] The term "corrects a mutation" refers to an incorporation of a desired alteration by a donor polynucleotide that results in a change of one or more nucleotides in a gDNA that comprises a mutation (e.g., a deleterious or disease-causing mutation) such that the mutation is reverted or transmuted in a desired manner. The identification of a mutation to correct can be determined by comparison of the nucleotide sequence of a gDNA known, or suspected to, comprise the mutation to the nucleotide sequence of a wild-type gDNA.

[0392] The term "induces a mutation" refers to an incorporation of a desired alteration by a donor polynucleotide that results in a change of one or more nucleotides in a gDNA such that the gDNA is mutated in a desired manner. A mutation induced by a donor polynucleotide may be any type of mutation known in the art. In some embodiments, the induction of a mutation is for therapeutic purposes or results in a therapeutic effect.

[0393] Covalently linked: As used herein, the term "covalently linked" (alternatively "conjugated", "linked", "attached", "fused", or "tethered"), when used with respect to two or more moieties, means that the moieties are physically associated or connected with one another, by whatever means including chemical conjugation, recombinant techniques or enzymatic activity, either directly or via one or more additional moieties that serves as a linking agent, to form a structure that is sufficiently stable so that the moieties remain physically associated under the conditions in which the structure is used, e.g., physiological conditions.

[0394] Complementary: As used herein, the term "complementary" or "complementarity" refers to a relationship between the sequence of nucleotides comprising two polynucleotide strands, or regions of the same polynucleotide strand, and the formation of a duplex comprising the strands or regions, wherein the extent of consecutive base pairing between the two strands or regions is sufficient for the generation of a duplex structure. It is known that adenine (A) forms specific hydrogen bonds, or "base pairs", with thymine (T) or uracil (U). Similarly, it is known that a cytosine (C) base pairs with guanine (G). It is also known that non-canonical nucleobases (e.g., inosine) can hydrogen bond with natural bases. A sequence of nucleotides comprising a first strand of a polynucleotide, or a region, portion or fragment thereof, is said to be "sufficiently complementary" to a sequence of nucleotides comprising a second strand of the same or a different nucleic acid, or a region, portion, or fragment thereof, if, when the first and second strands are arranged in an antiparallel fashion, the extent of base pairing between the two strands maintains the duplex structure under the conditions in which the duplex structure

is used (e.g., physiological conditions in a cell). It should be understood that complementary strands or regions of polynucleotides can include some base pairs that are non-complementary. Complementarity may be “partial,” in which only some of the nucleobases comprising the polynucleotide are matched according to base pairing rules. Or, there may be “complete” or “total” complementarity between the nucleic acids. Although the degree of complementarity between polynucleotide strands or regions has significant effects on the efficiency and strength of hybridization between the strands or regions, it is not required for two complementary polynucleotides to base pair at every nucleotide position. In some embodiments, a first polynucleotide is 100% or “fully” complementary to a second polynucleotide and thus forms a base pair at every nucleotide position. In some embodiments, a first polynucleotide is not 100% complementary (e.g., is 90%, or 80% or 70% complementary) and contains mismatched nucleotides at one or more nucleotide positions. While perfect complementarity is often desired, some embodiments can include one or more but preferably 6, 5, 4, 3, 2, or 1 mismatches.

[0395] Contacting: As used herein, the term “contacting” means establishing a physical connection between two or more entities. For example, contacting a cell with an agent (e.g., an RNA, a lipid nanoparticle composition, or other pharmaceutical composition of the disclosure) means that the cell and the agent are made to share a physical connection. Methods of contacting cells with external entities both in vivo, in vitro, and ex vivo are well known in the biological arts. In exemplary embodiments of the disclosure, the step of contacting a mammalian cell with a composition (e.g., an isolated RNA, nanoparticle, or pharmaceutical composition of the disclosure) is performed in vivo. For example, contacting a lipid nanoparticle composition and a cell (for example, a mammalian cell) which may be disposed within an organism (e.g., a mammal) may be performed by any suitable administration route (e.g., parenteral administration to the organism, including intravenous, intramuscular, intradermal, and subcutaneous administration). For a cell present in vitro, a composition (e.g., a lipid nanoparticle or an isolated RNA) and a cell may be contacted, for example, by adding the composition to the culture medium of the cell and may involve or result in transfection. Moreover, more than one cell may be contacted by an agent.

[0396] Culture: As used herein, the term “culture” can be used interchangeably with the terms “culturing”, “grow”, “growing”, “maintain”, “maintaining”, “expand”, “expanding” when referring to a cell culture or the process of culturing. The term refers to a cell (e.g., a primary cell) that is maintained outside its normal environment (e.g., a tissue in a living organism) under controlled conditions. Cultured cells are treated in a manner that enables survival. Culturing conditions can be modified to alter cell growth, homeostasis, differentiation, division, or a combination thereof in a controlled and reproducible manner. The term does not imply that all cells in the culture survive, grow, or divide as some may die, enter a state of quiescence, or enter a state of senescence. Cells are typically cultured in media, which can be changed during the course of the culture. Components can be added to the media or environmental factors (e.g., temperature, humidity, atmospheric gas levels) to promote cell survival, growth, homeostasis, division, or a combination thereof.

[0397] Denaturation: As used herein, the term “denaturation” refers to the process by which the hydrogen bonding between base paired nucleotides in a nucleic acid is disrupted, resulting in the loss of secondary and/or tertiary nucleic acid structure (e.g., the separation of previously annealed strands). Denaturation can occur by the application of an external substance, energy, or biochemical process to a nucleic acid.

[0398] Double-strand break: As used herein the term, “double-strand break” (DSB) refers to a DNA lesion generated when the two complementary strands of a DNA molecule are broken or cleaved, resulting in two free DNA ends or termini. DSBs may occur via exposure to environmental insults (e.g., irradiation, chemical agents, or UV light) or generated deliberately (e.g., via a site-directed nuclease) and for a defined biological purpose (e.g., the insertion of a donor polynucleotide to correct a mutation).

[0399] Duplex: As used herein, the term “duplex” refers to a structure formed by complementary strands of a double-stranded polynucleotide, or complementary regions of a single-stranded polynucleotide that folds back on itself. The duplex structure of a nucleic acid arises as a consequence of complementary nucleotide sequences being bound together, or hybridizing, by base pairing interactions.

[0400] EC₅₀: As used herein, the term “EC₅₀” refers to the concentration of a composition which induces a response, either in an in vitro or an in vivo assay, which is 50% of the maximal response, i.e., halfway between the maximal response and the baseline.

[0401] Effective dose: As used herein, the term “effective dose” or “effective dosage” is defined as an amount sufficient to achieve or at least partially achieve the desired effect.

[0402] Engraftment: As used herein, the term “engraftment” is used interchangeably with the term “chimerism” and refers to the process wherein donor stem cells are administered to (e.g., transplanted) a host, traffic to a tissue compartment, and establish within that compartment by undergoing self-renewal and generating differentiated cells for reconstitution of the tissue compartment. Often the term engraftment refers to the success of a hematopoietic stem cell (HSC) transplant (e.g., a bone marrow transplant). The term “engraftment” in this context refers to the persistence of transplanted HSCs and their progenitors following administration. The term engraftment can also refer to the success of a T cell therapy wherein ex vivo manipulated T cells are administered to a host. The term “engraftment” in this context refers to the persistence of transplanted donor T cells and their progenitors following administration.

[0403] Genome editing: As used herein, the term genome editing generally refers to the process of editing or changing the nucleotide sequence of a genome, preferably in a precise or predetermined manner. Examples of methods of genome editing described herein include methods of using site-directed nucleases to cut genomic DNA at a precise target location or sequence within a genome, thereby creating a DNA break (e.g., a DSB) within the target sequence, and repairing the DNA break such that the nucleotide sequence of the repaired genome has been changed at or near the site of the DNA break.

[0404] Double-strand DNA breaks (DSBs) can be and regularly are repaired by natural, endogenous cellular processes such as homology-directed repair (HDR) and non-

homologous end-joining (NHEJ) (see e.g., Cox et al., (2015) *Nature Medicine* 21(2):121-131).

[0405] DNA repair by HDR utilizes a polynucleotide (often referred to as a “repair template” or “donor template”) with a nucleotide sequence that is homologous to the sequences flanking the DSB. DNA repair by HDR mechanisms involves homologous recombination between the repair template and the cut genomic DNA molecule. Repair templates may be designed such that they insert or delete nucleotides in the genomic DNA molecule or change the nucleotide sequence of the genomic DNA molecule.

[0406] NHEJ mechanisms can repair a DSB by directly joining or ligating together the DNA ends that result from the DSB. Repair of a DSB by NHEJ can involve the random insertion or deletion of one or more nucleotides (i.e. indels). This aspect of DNA repair by NHEJ is often leveraged in genome editing methods to disrupt gene expression. NHEJ can also repair a DSB by insertion of an exogenous polynucleotide into the cut site in a homology-independent manner.

[0407] A third repair mechanism is microhomology-mediated end joining (MMEJ), also referred to as “alternative NHEJ”, in which the genetic outcome is similar to NHEJ in that small deletions and insertions can occur at the cleavage site. MMEJ makes use of homologous sequences of a few basepairs flanking the DNA break site to drive a more favored DNA end joining repair outcome (see e.g., Cho and Greenberg, (2015) *Nature* 518, 174-176); Mateos-Gomez et al., *Nature* 518, 254-57 (2015); Ceccaldi et al., *Nature* 528, 258-62 (2015). In some instances it may be possible to predict likely repair outcomes based on analysis of potential microhomologies at the site of the DNA break. Each of the aforementioned DNA repair mechanisms can be used in genome editing methods to create desired genomic alterations. The first step in the genome editing process is to create typically one or two DNA breaks in a target sequence as close as possible to the site of intended mutation or alteration. This can be achieved via the use of a site-directed nuclease, as described and illustrated herein.

[0408] Hemoglobinopathy: As used herein, the term “hemoglobinopathy” refers to any defect in the structure, function, or expression of any hemoglobin of an individual, and includes defects in the primary, secondary, tertiary or quaternary structure of hemoglobin caused by any mutation, such as deletion mutations or substitution mutations in the coding regions of the β -globin gene, or mutations in, or deletions of, the promoters or enhancers of such genes that cause a reduction in the amount of hemoglobin produced as compared to a normal condition. The term further comprises any decrease in the amount or effectiveness of hemoglobin, whether normal or abnormal, caused by external factors such as disease, chemotherapy, toxins, poisons, or the like. B-hemoglobinopathies contemplated herein include, but are not limited to, sickle cell disease (SCD, also referred to as a sickle cell anemia or SCA), sickle cell trait, hemoglobin C disease, hemoglobin C trait, hemoglobin S/C disease, hemoglobin D disease, hemoglobin E disease, thalassemias, hemoglobins with increased oxygen affinity, hemoglobins with decreased oxygen affinity, unstable hemoglobin disease, and methemoglobinemia.

[0409] In need: As used herein, a subject “in need of prevention,” “in need of treatment,” or “in need thereof,” refers to one, who by the judgment of an appropriate medical practitioner (e.g., a doctor, a nurse, or a nurse practitioner in

the case of humans; a veterinarian in the case of non-human mammals), would reasonably benefit from a given treatment.

[0410] Insertion: As used herein, an “insertion” or an “addition” refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to a molecule as compared to a reference sequence, for example, the sequence found in a naturally-occurring molecule.

[0411] Intron: As used herein, the term “intron” refers to any nucleotide sequence within a gene that is removed by RNA splicing mechanisms during maturation of the final RNA product (e.g., an mRNA). An intron refers to both the DNA sequence within a gene and the corresponding sequence in a RNA transcript (e.g., a pre-mRNA). Sequences that are joined together in the final mature RNA after RNA splicing are “exons”. As used herein, the term “intronic sequence” refers to a nucleotide sequence comprising an intron or a portion of an intron. Introns are found in the genes of most eukaryotic organisms and can be located in a wide range of genes, including those that generate proteins, ribosomal RNA (rRNA), and transfer RNA (tRNA). When proteins are generated from intron-containing genes, RNA splicing takes place as part of the RNA processing pathway that follows transcription and precedes translation.

[0412] Lipid: As used herein, the term “lipid” refers to a small molecule that has hydrophobic or amphiphilic properties. Lipids may be naturally occurring or synthetic. Examples of classes of lipids include, but are not limited to, fats, waxes, sterol-containing metabolites, vitamins, fatty acids, glycerolipids, glycerophospholipids, sphingolipids, saccharolipids, and polyketides, and prenol lipids. In some instances, the amphiphilic properties of some lipids leads them to form liposomes, vesicles, or membranes in aqueous media.

[0413] Modified: As used herein “modified” or “modification” refers to a changed state or change in structure resulting from a modification of a polynucleotide, e.g., DNA. Polynucleotides may be modified in various ways including chemically, structurally, and/or functionally. For example, the DNA molecules of the present disclosure may be modified by the incorporation of a chemically-modified base that provides a biological activity. In one embodiment, the DNA is modified by the introduction of non-natural or chemically-modified bases, nucleosides and/or nucleotides, e.g., as it relates to the natural nucleobases adenine (A), guanine (G), cytosine (C), and thymine (T).

[0414] mRNA: As used herein, an “mRNA” refers to a messenger ribonucleic acid. An mRNA may be naturally or non-naturally occurring or synthetic. For example, an mRNA may include modified and/or non-naturally occurring components such as one or more nucleobases, nucleosides, nucleotides, or linkers. An mRNA may include a cap structure, a 5' transcript leader, a 5' untranslated region, an initiator codon, an open reading frame, a stop codon, a chain terminating nucleoside, a stem-loop, a hairpin, a polyA sequence, a polyadenylation signal, and/or one or more cis-regulatory elements. An mRNA may have a nucleotide sequence encoding a polypeptide. Translation of an mRNA, for example, in vivo translation of an mRNA inside a mammalian cell, may produce a polypeptide. Traditionally, the basic components of a natural mRNA molecule include

at least a coding region, a 5'-untranslated region (5'-UTR), a 3'UTR, a 5' cap and a polyA sequence.

[0415] Naturally occurring: As used herein, the term “naturally occurring” as applied to an object refers to the fact that an object can be found in nature. For example, a polypeptide or polynucleotide sequence (e.g., a splice site), or components thereof such as amino acids or nucleotides, that is present in an organism (including viruses) that can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally occurring.

[0416] Non-homologous end joining: As used herein, the term “non-homologous end joining” refers to a pathway that repairs double-strand breaks (DSBs) in DNA. NHEJ is referred to as “non-homologous” because the DNA termini are directly ligated without the need for a homologous template, in contrast to homology directed repair (HDR), which requires a homologous sequence to guide repair.

[0417] Non-replicative: As used herein, the term “non-replicative” refers to the characteristic of a DNA molecule as being unable to replicate within a cell or an organism. Certain DNA molecules (e.g., plasmids, viral genomes) contain sequence elements (e.g., origins of replications) that impart the DNA molecule with the ability to be copied, or replicated, by a cell or organism. The term “non-replicative” connotes those DNA molecules that do not contain such sequence elements.

[0418] Nucleic acid: As used herein, the term “nucleic acid” refers to deoxyribonucleotides or ribonucleotides and polymers or oligomers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Polymers of nucleotides are referred to as “polynucleotides”. Exemplary nucleic acids or polynucleotides of the disclosure include, but are not limited to, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), DNA-RNA hybrids, RNAi-inducing agents, RNAi agents, siRNAs, shRNAs, miRNAs, antisense RNAs, ribozymes, catalytic DNA, RNAs that induce triple helix formation, threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs, including LNA having a β -D-ribo configuration, α -LNA having an α -L-ribo configuration (a diastereomer of LNA), 2'-amino-LNA having a 2'-amino functionalization, and 2'-amino- α -LNA having a 2'-amino functionalization) or hybrids thereof.

[0419] Polynucleotides used herein can be composed of any polyribonucleotide or polydeoxyribonucleotide, which can be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that can be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide can also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. “Modified” bases include, for example, tritylated bases. “Modified nucleosides”

include, for example, as inosine and thymine, when the latter is found in or comprises RNA. A variety of modifications can be made to DNA and RNA; thus, “polynucleotide” embraces chemically, enzymatically, or metabolically modified forms.

[0420] Nucleobase: As used herein, the term “nucleobase” (alternatively “nucleotide base” or “nitrogenous base”) refers to a purine or pyrimidine heterocyclic compound found in nucleic acids, including any derivatives or analogs of the naturally occurring purines and pyrimidines that confer improved properties (e.g., binding affinity, nuclease resistance, chemical stability) to a nucleic acid or a portion or segment thereof. Adenine, cytosine, guanine, thymine, and uracil are the primary or canonical nucleobases predominately found in natural nucleic acids. Other natural, non-natural, non-canonical and/or synthetic nucleobases, can be incorporated into nucleic acids, such as those disclosed herein.

[0421] Nucleoside/Nucleotide: As used herein, the term “nucleoside” refers to a compound containing a sugar molecule (e.g., a ribose in RNA or a deoxyribose in DNA), or derivative or analog thereof, covalently linked to a nucleobase (e.g., a purine or pyrimidine), or a derivative or analog thereof. As used herein, the term “nucleotide” refers to a nucleoside covalently linked to a phosphate group. As used herein, the term “ribonucleoside” refers to a nucleoside that comprise a ribose and a nucleobase (e.g., adenosine (A), cytidine (C), guanosine (G), 5-methyluridine (m^5U), uridine (U), or inosine (I)).

[0422] Operably linked: As used herein, a nucleic acid, or fragment or portion thereof, such as a polynucleotide or oligonucleotide is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence, or fragment or portion thereof.

[0423] Polynucleotide/oligonucleotide: As used herein, the terms “polynucleotide” and “oligonucleotide” are used interchangeably and refer to a single-stranded or double-stranded polymer or oligomer of nucleotides or nucleoside monomers consisting of naturally-occurring bases, sugars and intersugar (backbone) linkages. The terms “polynucleotide” and “oligonucleotide” also includes polymers and oligomers comprising non-naturally occurring bases, sugars and intersugar (backbone) linkages, or portions thereof, which function similarly. Polynucleotides are not limited to any particular length of nucleotide sequence, as the term “polynucleotides” encompasses polymeric forms of nucleotides of any length. Short polynucleotides are typically referred to in the art as “oligonucleotides”. In the context of the present disclosure, such modified or substituted polynucleotides and oligonucleotides are often preferred over native forms because the modification increases one or more desirable or beneficial biological properties or activities including, but not limited to, enhanced cellular uptake and/or increased stability in the presence of nucleases. In some embodiments, the agonists of the disclosure comprise polynucleotides and oligonucleotides that contain at least one region of modified nucleotides that confers one or more beneficial properties or increases biological activity (e.g., increased nuclease resistance, increased uptake into cells, increased duplex stability, increased binding affinity to a target polypeptide).

[0424] Palindromic sequence: As used herein, the term “palindromic sequence” (alternatively “palindrome”) refers to a sequence of nucleotides that is self-complementary;

wherein the sequence of nucleotides in the 5' to 3' direction is the same as the sequence of nucleotides comprising the complementary strand, when read in the 5' to 3'. For example, the sequence 5'-ACCTAGGT-3' is a palindromic sequence because its complementary sequence, 3'-TG-GATCCA-5', when read in the 5' to 3' direction, is the same as the original sequence. In contrast, the sequence 5'-AGTGGCTG-3' is not a palindromic sequence because its complementary sequence, 3'-TCACCGAC-5', when read in the 5' to 3' direction, is not the same as the original sequence.

[0425] Parenteral administration: As used herein, “parenteral administration,” “administered parenterally,” and other grammatically equivalent phrases, refer to modes of administration other than enteral and topical administration, usually by injection, and include, without limitation, intravenous, intranasal, intraocular, intramuscular, intraarterial, intrathecal, intracapsular, intraorbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticular, subcapsular, subarachnoid, intraspinal, epidural, intracerebral, intracranial, intracarotid and intrasternal injection and infusion.

[0426] Percent identity: As used herein, the term “percent identity,” in the context of two or more nucleic acid or polypeptide sequences, refers to two or more sequences or subsequences that have a specified percentage of nucleotides or amino acid residues that are the same, when compared and aligned for maximum correspondence, as measured using one of the sequence comparison algorithms described below (e.g., BLASTP and BLASTN or other algorithms available to persons of skill) or by visual inspection. Depending on the application, the “percent identity” can exist over a region of the sequence being compared, e.g., over a functional domain, or, alternatively, exist over the full length of the two sequences to be compared. For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters. The percent identity between two sequences is a function of the number of identical positions shared by the sequences (i.e., $\% \text{ homology} = \# \text{ of identical positions} / \text{total} \# \text{ of positions} \times 100$), taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm, as described in the non-limiting examples below.

[0427] Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by visual inspection (see generally Ausubel et al., *infra*).

[0428] One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul et al., *J. Mol. Biol.* 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information website. The percent identity between two nucleotide sequences can be determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. The percent identity between two nucleotide or amino acid sequences can also be determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. In addition, the percent identity between two amino acid sequences can be determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6.

[0429] The nucleic acid and protein sequences of the present disclosure can further be used as a “query sequence” to perform a search against public databases to, for example, identify related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12 to obtain nucleotide sequences homologous to the nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to the protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

[0430] Pharmaceutically acceptable: As used herein, the term “pharmaceutically acceptable” refers to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues, organs, and/or bodily fluids of human beings and animals without excessive toxicity, irritation, allergic response, or other problems or complications commensurate with a reasonable benefit/risk ratio.

[0431] Pharmaceutically acceptable carrier: As used herein, the term “pharmaceutically acceptable carrier” refers to, and includes, any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like that are physiologically compatible. The compositions can include a pharmaceutically acceptable salt, e.g., an acid addition salt or a base addition salt (see, e.g., Berge et al. (1977) *J Pharm Sci* 66:1-19).

[0432] Polypeptide: As used herein, the terms “polypeptide,” “peptide,” and “protein” are used interchangeably to refer to a polymer of amino acid residues. The terms apply

to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

[0433] Preventing: As used herein, the term “preventing” or “prevent” when used in relation to a condition, refers to administration of a composition which reduces the frequency of, or delays the onset of, symptoms of a medical condition in a subject relative to a subject which does not receive the composition.

[0434] Purified: As used herein, the term “purified” or “isolated” as applied to any of the proteins (antibodies or fragments) described herein refers to a polypeptide that has been separated or purified from components (e.g., proteins or other naturally-occurring biological or organic molecules) which naturally accompany it, e.g., other proteins, lipids, and nucleic acid in a prokaryote expressing the proteins. Typically, a polypeptide is purified when it constitutes at least 60 (e.g., at least 65, 70, 75, 80, 85, 90, 92, 95, 97, or 99) %, by weight, of the total protein in a sample.

[0435] Reprogramming: As used herein, the term “reprogramming” refers to a process that alters or reverses the differentiation state of a differentiated cell (e.g., a somatic cell). Stated another way, reprogramming refers to a process of driving the differentiation of a cell backwards to a more undifferentiated or more primitive type of cell. It should be noted that placing many primary cells into culture can lead to some loss of fully differentiated characteristics. Thus, simply culturing such cells included in the term differentiated cells does not render these cells non-differentiated cells (e.g., undifferentiated) or pluripotent cells. The transition of a differentiated cell to pluripotency requires a reprogramming stimulus beyond the stimuli that lead to partial loss of differentiated character in culture. Reprogrammed cells also have the characteristic of the capacity of extending passaging without loss of growth potential, relative to primary cell parents, which generally have capacity for only a limited number of divisions in culture.

[0436] Sense strand: As used herein the term “sense strand” or “coding strand” refers to a segment within double-stranded DNA (e.g., genomic DNA) with a 5' to 3' directionality and has the same nucleotide sequence as an mRNA transcribed from the segment. The transcription product is pre-mRNA transcript, which contains a sequence of nucleotides that is identical to that of the sense strand, with the exception that uracil will be incorporated into the mRNA at those positions where thymine is located in the DNA. The sense strand is complementary to the antisense strand of DNA, or template strand, which runs from 3' to 5'.

[0437] Site-directed nuclease: As used herein, the term “site-directed nuclease” refers to one of several distinct classes of nucleases that can be programmed or engineered to recognize a specific target site (i.e., a target nucleotide sequence) in a DNA molecule (e.g., a genomic DNA molecule) and generate a DNA break (e.g., a DSB) within the DNA molecule at, near or within the specific site. Site-directed nucleases are useful in genome editing methods, such as those described herein. Site-directed nucleases include, but are not limited to, the zinc finger nucleases (ZFNs), transcription activator-like effector (TALE) nucleases, CRISPR/Cas nucleases (e.g., Cas9), homing endonucleases (also termed meganucleases), and other nucleases (see, e.g., Hafez and Hausner, *Genome* 55, 553-69 (2012);

Carroll, *Ann. Rev. Biochem.* 83, 409-39 (2014); Gupta and Musunuru, *J. Clin. Invest.* 124, 4154-61 (2014); and Cox et al., *supra*. These differ mainly in the way they bind DNA and create the targeted, site-specific DNA break. Site-directed nucleases known in the art may produce a single-strand break (SSB) or a DSB. For the purposes of the present invention, the disclosure's reference to a “site-directed nuclease” refers to those nucleases that produce a DSB. After creation of a DSB, essentially the same natural cellular DNA repair mechanisms of NHEJ or HDR are co-opted to achieve the desired genetic modification. Therefore, it is contemplated that genome editing technologies or systems using site-directed nucleases can be used to achieve genetic and therapeutic outcomes described herein.

[0438] Stem cell: As used herein, the term “stem cell” is used interchangeably with the term “hematopoietic stem cell” (HSC). Stem cells are distinguished from other cell types by two important characteristics. First, stem cells are unspecialized cells capable of renewing themselves through cell division, sometimes after periods of inactivity (e.g., quiescent state). Second, under certain physiologic or experimental conditions, stem cells can be induced to become tissue- or organ-specific cells with special functions. In some organs, such as the gut and bone marrow, stem cells regularly divide to repair and replace worn out or damaged tissues. In other tissues, such as the pancreas and heart, stem cells only differentiate under certain conditions.

[0439] The term “HSC” can refer to multipotent stem cell that is capable of differentiating into all blood cells including erythrocytes, leukocytes, and platelets. HSCs are contained not only in the bone marrow, but also in umbilical cord blood derived cells.

[0440] Stem cell mobilizer: As used herein, the term “stem cell mobilizer” is used interchangeably with the terms “mobilizer of hematopoietic stem or progenitor cells” or “mobilize” and refers to any agent, whether it is a small organic molecule, a polypeptide (e.g., a growth factor or colony stimulating factor or an active fragment or mimic thereof), a nucleic acid, a carbohydrate, an antibody, that acts to enhance the migration of stem cells from the bone marrow into the peripheral blood. A stem cell mobilizer may increase the number of HSCs or hematopoietic progenitor/precursor cells in the peripheral blood, thus allowing for a more accessible source of stem cells. In some embodiments, a stem cell mobilizer refers to any agent that mobilizes CD34+ stem cells. It is further understood that an agent may have stem cell mobilizing activity in addition to one or more other biological activities including, but not limited to, immunosuppression.

[0441] Subject: As used herein, the term “subject” includes any human or non-human animal. For example, the methods and compositions of the present invention can be used to treat a subject with a disorder (e.g., a genetic disorder). The term “non-human animal” includes all vertebrates, e.g., mammals and non-mammals, such as non-human primates, sheep, dog, cow, chickens, amphibians, reptiles, etc.

[0442] Therapeutic agent: As used herein, the term “therapeutic agent” refers to any agent that, when administered to a subject, has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect.

[0443] Therapeutically effective amount: As used herein, the terms “therapeutically effective amount” or “therapeu-

tically effective dose,” or similar terms used herein are intended to mean an amount of an agent that will elicit the desired biological or medical response, such as, for example, at least partially arresting the condition or disease and its complications in a patient already suffering from the disease (e.g., an improvement in one or more symptoms of a cancer). Amounts effective for this use will depend on the severity of the disorder being treated and the general state of the patient’s own immune system.

[0444] Treat: The terms “treat,” “treating,” and “treatment,” as used herein, refer to therapeutic measures described herein. The methods of “treatment” employ administration of a composition of the disclosure to a subject, in need of such treatment, in order to, cure, delay, reduce the severity of, or ameliorate one or more symptoms of the disorder or recurring disorder, or in order to prolong the survival of a subject beyond that expected in the absence of such treatment.

[0445] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure pertains. Preferred methods and materials are described below, although methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the presently disclosed methods and compositions.

EQUIVALENTS AND SCOPE

[0446] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments, described herein. The scope of the present disclosure is not intended to be limited to the above Description, but rather is as set forth in the appended claims.

[0447] In the claims articles such as “a,” “an,” and “the” may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include “or” between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The disclosure includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The disclosure includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process. Furthermore, it is to be understood that the disclosure encompasses all variations, combinations, and permutations in which one or more limitations, elements, clauses, descriptive terms, etc., from one or more of the listed claims is introduced into another claim. For example, any claim that is dependent on another claim can be modified to include one or more limitations found in any other claim that is dependent on the same base claim. Furthermore, where the claims recite a composition, it is to be understood that methods of using the composition for any of the purposes disclosed herein are included, and methods of making the composition according to any of the methods of making disclosed herein or other methods known in the art are included, unless otherwise indicated or unless it would be evident to one of ordinary skill in the art that a contradiction or inconsistency would arise.

[0448] Where elements are presented as lists, e.g., in Markush group format, it is to be understood that each subgroup of the elements is also disclosed, and any element (s) can be removed from the group. It should be understood that, in general, where the invention, or aspects of the invention, is/are referred to as comprising particular elements, features, etc., certain embodiments of the invention or aspects of the invention consist, or consist essentially of, such elements, features, etc. For purposes of simplicity those embodiments have not been specifically set forth in haec verba herein.

[0449] It is also noted that the term “comprising” is intended to be open and permits but does not require the inclusion of additional elements or steps. When the term “comprising” is used herein, the term “consisting of” is thus also encompassed and disclosed.

[0450] Where ranges are given, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different embodiments of the invention, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

[0451] In addition, it is to be understood that any particular embodiment of the present invention that falls within the prior art may be explicitly excluded from any one or more of the claims. Since such embodiments are deemed to be known to one of ordinary skill in the art, they may be excluded even if the exclusion is not set forth explicitly herein. Any particular embodiment of the compositions of the invention (e.g., any nucleic acid or protein encoded thereby; any method of production; any method of use; etc.) can be excluded from any one or more claims, for any reason, whether or not related to the existence of prior art.

[0452] All cited sources, for example, references, publications, databases, database entries, and art cited herein, are incorporated into this application by reference, even if not expressly stated in the citation. In case of conflicting statements of a cited source and the instant application, the statement in the instant application shall control.

EXAMPLES

Example 1. In Vitro Screen of DNA DSB Repair Modulators for Improved HDR in T Cells

[0453] Multiple pathways are used for repair of DNA double stranded breaks (DSBs). The homology directed repair (HDR) pathway uses homologous donor DNA (e.g., a sister chromatid or exogenous donor DNA) for high fidelity repair. The efficiency of HDR is generally low due to competition with other repair pathways, notably the non-homologous end-joining (NHEJ) pathway. HDR is predominantly active in the S/G2 phases of the cell cycle, whereas NHEJ repair is active in each phase of the cell cycle and is the predominant repair pathway in G1 cells. Thus, HDR efficiency is poor in non-dividing or slowly dividing cells, for example, long-term repopulating hematopoietic cells (LT-HSPCs), lung progenitor cells, or hepatic cells. Given that NHEJ repair is error-prone, frequently resulting in small nucleotide insertions or deletions (indels) that can cause a frameshift mutation, it is undesirable for generating precise modification of a gene (i.e., specific nucleotide changes or knock-in of a gene).

[0454] An in vitro assay was conducted to determine the ability of various inhibitors and enhancers to improve the efficiency of HDR repair in HEK293 T cells. Specifically, a fluorescent reporter-based screening approach was developed to identify molecules that enhance gene-editing by HDR. A reporter system was generated in HEK293 T cells comprising a genomic AAVS1 locus that encoded either a blue fluorescent protein (BFP) or a green fluorescent protein (GFP) within the AAVS1 locus. A gene encoding BFP can be converted to a gene encoding GFP by substitution of cytosine at position 199 to uridine (e.g., a C to U transition at position 199). By introducing a DSB within the BFP gene using a CRISPR/Cas9 gene-editing system, a homology donor DNA encoding the nucleotide substitution can be used to edit the BFP gene to a GFP gene by HDR repair. Thus, a gene edit that induces a C to U transition at position 199 results in a change in cellular fluorescence. A change in the fluorescence of the cell measured by flow cytometry can be used to quantify efficiency of HDR repair. Additionally, the reverse gene edit can be performed for an AAVS1 locus encoding a GFP gene, wherein a GFP can be converted to BFP by substitution of uridine at position 199 to cytosine.

[0455] To create the reporter system, an AAV-based homology donor DNA encoding either BFP or GFP was created. The AAV-BFP or AAV-GFP donor was flanked by homology arms that were 1000 base pairs in length and expression of BFP or GFP was driven by the MND promoter (e.g., a synthetic promoter that contains the U3 region of a modified Moloney murine leukemia virus long term terminal repeat with myeloproliferative sarcoma virus enhancer and deleted negative control region, SEQ ID NO: 58). The viral DNA was packaged into capsids of adeno-associated virus serotype 6 (AAV6) vector. Either the BFP or GFP donor was introduced using the AAV6 vector into the AAVS1 locus in HEK293 T cells by homologous recombination using CRISPR-Cas9 as follows: purified Cas9 protein was complexed with two different single gRNAs (sgRNAs) targeting the AAVS1 locus at SEQ ID NO: 3 (spacer sequence identified by SEQ ID NO: 4; gRNA obtained from Maxcyte) or SEQ ID NO: 5 (spacer sequence identified by SEQ ID NO: 6; gRNA obtained from Thermo Fisher). Cas9 protein used in the Exemplary section provided herein refers to Cas9 polypeptide derived from *S. pyogenes* (SpCas9), unless indicated otherwise.

[0456] The Cas9-sgRNA complex was electroporated into the HEK293 T cells using the Lonza Nucleofector program CM-130. Approximately, 2 hours after electroporation, the cells were infected with various viral doses of AAV6 encoding a BFP (SEQ ID NO: 44) or GFP homology donor (SEQ ID NO: 47). One week later, cells were analyzed by flow cytometry to verify integration and expression of BFP or GFP in the AAVS1 locus. HEK293 T cells expressing BFP or GFP were sorted into single cells in 96-well plates. After ~2 weeks of growth, DNA extracted from individual cells was analyzed for precise integration of the respective BFP or GFP gene into the AAVS1 locus by long-range PCR. This PCR analysis allowed for determination of whether the BFP/GFP gene was integrated in one or both alleles of the AAVS1 locus.

[0457] To introduce a gene edit that converts BFP to GFP, cells were electroporated with ribonucleoprotein (RNP) comprised of Cas9 and sgRNA that targets the BFP gene encoded in the AAVS1 locus. The sgRNA targeted a sequence in the BFP gene identified by SEQ ID NO: 7

(sgRNA spacer sequence identified by SEQ ID NO: 8). The sgRNA were prepared using a standard sgRNA cassette specific to SpCas9, as indicated by SEQ ID NO: 2 (wherein a, c, g, u represent 2' O-methyl phosphorothioate nucleotides; s represents phosphorothioate nucleotides; and A, C, G, U, N represent canonical RNA nucleotides). The gRNAs used in the Exemplary section provided herein were sgRNAs prepared with the SpCas9 sgRNA cassette unless indicated otherwise.

[0458] The cells were also transfected with single-stranded oligodeoxynucleotide (ssODN) that encoded the gene correction necessary to convert BFP to GFP and homology arms complimentary to the sequence upstream and downstream of the target gene cut site. The efficiency of HDR repair was determined by measuring the level of cell GFP fluorescence.

[0459] Molecules that manipulate targets in DSB repair pathways were evaluated and are listed in Table 2. These included molecules that inhibit targets that facilitate repair by the NHEJ pathway, including i53 (polypeptide inhibitor of 53BP1), Nu7441 (DNA-PKcs inhibitor), SCR7 (DNA Ligase IV inhibitor), CYREN1 (Ku70/80 inhibitor), and CYREN2 (Ku70/80 inhibitor). Also evaluated were molecules that enhance repair by the HDR pathway, including RS-1 (Rad51 agonist). Additionally, molecules were evaluated that inhibit repair by the alternative end joining (A-EJ) pathway, including siRNA targeting DNA polymerase θ and veliparib (PARP inhibitor). Also evaluated were molecules that affect cell cycle, including XL 413 (CDC7 inhibitor). Finally, also tested was the β 3 adrenergic receptor agonist L755,507 that was previously reported to improve HDR efficiency (Yu, C. et al. (2015) *Cell Stem Cell* 16:142-147).

TABLE 2

Pathway	Molecule	Target	Reference
NHEJ	i53	53BP1	SEQ ID NO: 70
NHEJ	Nu7441	DNA-PKcs	CAS 503468-95-9; Tocris 3712
NHEJ	SCR7	DNA Ligase IV	CAS 533426-72-0; Stemcell 74102
NHEJ	CYREN1 and 2	Ku70/80	Arnoult et al <i>Nature</i> (2017) 549: 548-552
HDR	RS-1	Rad51	CAS 312756-74-4 Sigma R9782
MMEJ	siRNA	Pol θ	
MMEJ	Veliparib	PARP	CAS 912444-00-9 Santa Cruz 394457
Unknown	L755,507	β 3-adrenergic receptor	CAS 159182-43-1 Tocris 2197
Cell Cycle	XL 413	CDC7	CAS 1169562-71-3 Tocris 5493

[0460] Shown in FIG. 1A is a comparison of HDR editing efficiency for cells treated with Nu7441, SCR7, or RS-1. The cells were transfected with four different ssODNs. 2×10^5 cells were used per reaction, and treated with 100 μ M of donor DNA. These included ssODN1 (SEQ ID NO: 21), ssODN2 (SEQ ID NO: 22) and ssODN4 (SEQ ID NO: 24) that are complimentary to the DNA strand containing the a PAM sequence and ssODN3 (SEQ ID NO: 23) that is complimentary to the DNA strand not containing the PAM sequence. A concentrated stock solution of each inhibitor was prepared in DMSO and diluted to a final concentration of 2.5 μ M Nu7441, 1.25 μ M SCR7, or 10 μ M RS-1. HDR efficiency was compared relative to treatment with DMSO alone or to treatment with RNP-only. The cells were treated with RNP containing 1 μ M Cas9 and 1.5 μ M sgRNA.

[0461] As shown, no improvement in HDR efficiency was seen with treatment of SCR7 or RS-1 for any of the ssODN constructs tested. However, treatment with Nu7441 resulted in approximately 3-fold higher HDR efficiency for each of the ssODN constructs tested. The improvement in HDR efficiency for treatment with Nu7441 was dose-dependent as shown in FIG. 1B. Cells treated with high concentration of Nu7441 (2.5 μ M) had an approximately 1.5-fold higher HDR efficiency than cells treated with low concentrations of Nu7441 (0.6 μ M). This improvement with higher dose was seen for cells transfected with either ssODN 91-36 (SEQ ID NO: 26) or ssODN 91-61 (SEQ ID NO: 27) (FIG. 1C). In each case, treatment with concentrations of Nu7441 higher than 2.5 μ M resulted in no further improvement in HDR efficiency.

[0462] A protein inhibitor of 53BP1 was also evaluated for improved HDR efficiency. The choice of repair pathway for repair of a DNA DSB is largely controlled by an antagonism between p53-binding protein 1 (53BP1), a pro-NHEJ factor, and BRCA1, a pro-HDR factor (Chapman, J. et al. (2012) *Molecular cell* 47:497-510). 53BP1 promotes NHEJ repair over HDR repair by suppressing formation of 3' single-stranded DNA tails, which is the rate-limiting step in the initiation of the HDR pathway. Loss of 53BP1 has been shown to increase HDR efficiency, (Canny, M. et al. (2018) *Nat Biotechnol.* 36(1):95-102). Thus, inhibition of 53BP1 is expected to reduce DSB repair by the NHEJ pathway and favor repair by the HDR pathway. An inhibitor of 53BP1 is the i53 polypeptide (SEQ ID NO: 70). Using the same in vitro assay assessing increased HDR by a BFP to GFP gene conversion, the effect of inhibition of the i53 polypeptide inhibitor (SEQ ID NO: 70) was evaluated. HEK293 T cells were transfected with two different ssODNs homologous to the AAVS1 locus: Hn-91-61 (SEQ ID NO: 25) and Ht-39-88 (SEQ ID NO: 28). 2×10^5 cells were edited with 1 μ M Cas9 and 1.5 μ M sgRNA, and with 100 μ M ssODN. Cells were treated with different doses of an mRNA encoding the i53 polypeptide mRNA open reading frame encoding i53 polypeptide identified by SEQ ID NO: 69). HDR efficiency increased with treatment of mRNA encoding i53 as shown in FIG. 1D.

[0463] Additional molecules were tested for improved HDR efficiency by assessing a BFP to GFP gene edit that had no effect. Cells were treated with different concentrations of veliparib using ssODN identified by SEQ ID NO: 25, but no improvement in HDR efficiency was seen as shown in FIG. 1B. Cells were treated with different concentrations of L755,507 using either 91-36 ssODN (SEQ ID NO: 26) or 91-61 ssODN (SEQ ID NO: 27), but no improvement in HDR efficiency was seen as shown in FIG. 1C. Cells were treated with siRNA targeting DNA polymerase θ (Pol θ). However, no improvement in HDR efficiency was seen for any siRNA dose tested.

Example 2. Increased HDR and Decreased Indel Formation with Treatment by DNA PKcs Inhibitor Nu7441

[0464] The effect of DNA PK inhibition by Nu7441 correction of a DSB by the NHEJ pathway (e.g., introduce an indel at the DSB site) or HDR pathway (e.g., introduce a gene mutation encoded by a homology donor at the DSB site) was assessed using the reporter system described in Example 1. In this case gene-editing was evaluated in HEK293 T cells expressing GFP in the AAVS1 locus. To introduce a gene edit that converts GFP to BFP, cells were electroporated with ribonucleoprotein (RNP) comprised of

Cas9 and gRNA1 (GFP target sequence identified by SEQ ID NO: 9; sgRNA spacer sequence identified by SEQ ID NO: 10) or gRNA2 (GFP target sequence identified by SEQ ID NO: 11; sgRNA spacer identified by SEQ ID NO: 12) that targets the GFP gene encoded in the AAVS1 locus. The cells were also transfected with ssODNs that encoded the gene correction necessary to convert GFP to BFP and homology arms complementary to the sequence upstream and downstream of the target gene cut site. The efficiency of HDR repair was determined by measuring the level of cell BFP fluorescence. The efficiency of cutting (indel information) was monitored by TIDE analysis.

[0465] Shown in FIGS. 2A-2B is a comparison of HDR editing efficiency for cells treated with Nu7441, SCR7, or RS-1. Shown in FIG. 2A are cells were electroporated with RNP comprising Cas9 and gRNA1 (GFP target sequence shown in SEQ ID NO: 9; sgRNA spacer sequence identified by SEQ ID NO: 10). 2×10^5 cells were edited with 1 μ M Cas9 and 1.5 μ M gRNA1. The cells were transfected with four different ssODNs. These included ssODN 1067 (SEQ ID NO: 29) and ssODN 1069 (SEQ ID NO: 31) that are complementary to the DNA strand containing the PAM sequence and ssODN 1068 (SEQ ID NO: 30) and ssODN 1070 (SEQ ID NO: 32) that are complementary to the DNA strand not containing the PAM sequence. Cells were edited with 100 μ M ssODN. A concentrated stock solution of each inhibitor was prepared in DMSO and diluted to a final concentration of 2.5 μ M Nu7441, 1.25 μ M SCR7, or 10 μ M RS-1. HDR efficiency was compared relative to treatment with DMSO alone or to treatment with RNP-only. As shown, no improvement in HDR efficiency was seen with treatment of SCR7 or RS-1 for any of the ssODN constructs tested. However, treatment with Nu7441 resulted in improved HDR efficiency for each of the ssODN constructs tested FIG. 2A.

[0466] Shown in FIG. 2B are cells were electroporated with RNP comprising Cas9 and gRNA2 (GFP target sequence shown in SEQ ID NO: 11; sgRNA spacer sequence shown in SEQ ID NO: 12). The cells were transfected with four different ssODNs. These included ssODN 1061 (SEQ ID NO: 33) and ssODN 1063 (SEQ ID NO: 35) that are complementary to the DNA strand containing the PAM sequence and ssODN 1062 (SEQ ID NO: 34) and ssODN 1064 (SEQ ID NO: 36) that are complementary to the DNA strand not containing the PAM sequence. Similar to above, treatment with 2.5 μ M Nu7441 resulted in improved HDR efficiency for each of the ssODN constructs tested FIG. 2A.

[0467] Shown in FIGS. 2C-2D is a measure of indel formation performed by TIDE analysis for the same edits that are described for FIGS. 2A-2B. Regardless of the ssODN or gRNA used, treatment with Nu7441 resulted in decreased indel formation at the DSB site, while treatment with SCR7 or RS-1 resulted in reduction compared to a DMSO control. Thus, the DNA PK inhibitor Nu7441 achieves reduced repair of a DSB by the NHEJ pathway when a homology donor is provided.

[0468] Having demonstrated improved HDR efficiency for a gene-edit in the AAV1 locus upon treatment with Nu7441, its effect on HDR for gene-editing at an additional gene locus was evaluated. HEK293 T cells were electroporated with RNP comprised of Cas9 and a gRNA targeting a sequence in the GSD1a locus shown in SEQ ID NO: 13 (spacer sequence identified by SEQ ID NO: 14). 2×10^5 cells were edited with 1 μ M Cas9 and 1.5 μ M gRNA. The cells were transfected with two different ssODNs homology donors: 93-50 (SEQ ID NO: 39) or 25-100 (SEQ ID NO: 40). These two ssODN donors contain homology arms spanning both sides of the double-stranded break induced by

the Cas9-guide and facilitate correction of a point mutation in the G6PC gene sequence by HDR. Cells were edited with 100 μ M ssODN. The cells were treated with 2.5 μ M Nu7441, 1.25 μ M SCR7, or 10 μ M RS-1 and the effect on HDR efficiency was evaluated. While treatment with Nu7441 resulted in an approximately 1.7-fold increase in HDR efficiency over DMSO alone, treatment with SCR7 or RS-1 had no effect (FIG. 3).

[0469] The effect of Nu7441 treatment on gene correction by the NHEJ pathway was also evaluated. To do so, the cells were transfected with two different dsDNA donors: 50-0 (SEQ ID NO: 37) or 150-0 (SEQ ID NO: 38). Cells were edited with 1.5 μ g dsDNA donor. These dsDNA donors, lacking homology arms, introduce a second 3' splice site into exon 2 at the GSD1a locus when inserted into the cut site induced by the Cas9-guide complex by NHEJ repair. The cells were treated with 2.5 μ M Nu7441, 1.25 μ M SCR7, or 10 μ M RS-1 and gene correction by NHEJ repair was evaluated. Treatment with Nu7441 resulted in a substantial decrease in gene correction for either dsDNA donor compared to treatment with DMSO alone, demonstrating that Nu7441 is inhibiting NHEJ repair following a Cas9/gRNA-mediated DNA DSB (FIG. 3). Treatment with SCR7 or RS-1 had no effect over DMSO alone.

[0470] The effect of Nu7441 treatment on HDR was also evaluated at the CFTR gene locus in HEK293 T cells. To do so cells were electroporated with RNP comprising Cas9 and a gRNA targeting the CFTR gene locus (SEQ ID NO: 18; sgRNA target sequence identified by SEQ ID NO: 19). 2×10^5 cells were edited with 1 μ M Cas9 and 1.5 μ M gRNA. The cells were transfected with a ssODN donor (SEQ ID NO: 42). Cells were edited with 100 μ M ssODN. The ssODN contains homology arms spanning both sides of the DSB induced by the Cas9-guide and is designed to include 3 additional base pairs (GCA) into the CFTR gene to aid detection of HDR. Cells were also treated with 5 μ M Nu7441. Gene correction was assessed by TIDE analysis. TIDE analysis uses a pair of PCR reactions and standard capillary sequencing runs to identify mutations induced at the site of a DSB (see e.g., Brinkman (2014) Nucleic Acids Res 42:e168). The type of mutation induced at the DSB was indicative of the pathway used to repair the DSB. Formation of an indel comprising either an insertion or a deletion of

bases was considered due to NHEJ repair; a deletion of 2-3 base pairs was considered due to MMEJ repair; while an insertion of 3 base pairs was considered due to HDR repair. Shown in FIGS. 4A-4B are mutations introduced at the DSB for cells treated with Nu7441 (FIG. 4B) compared to a DMSO negative control (FIG. 4A). While indel formation (e.g., +1, 0, or -1 base pair) was high in the negative control (FIG. 4A), indicating high levels of NHEJ repair, indel formation was dramatically reduced in cells treated with Nu7441 (FIG. 4B). Additionally, cells treated with Nu7441 had much higher levels of HDR repair (+3 base pair insertion) at the DSB. This reduction in NHEJ repair in the presence of Nu7441 was evaluated with an additional ssODN donor (SEQ ID NO: 41). As shown in FIG. 5, treatment with either ssODN donor in the presence of Nu7441 resulted in decreased levels of indel formation due to NHEJ repair with a concurrent increase in HDR repair. **[0471]** Together, these data demonstrate that treatment with Nu7441, a small molecule inhibitor of DNA PKcs, results in inhibition of NHEJ repair by CRISPR/Cas9 gene-editing and increased HDR editing efficiency at multiple gene loci.

Example 3. Efficient Gene Editing by HDR Using i53 at Multiple Gene Loci and in Multiple Cell Types

[0472] Having demonstrated improved HDR efficiency with 53BP1 inhibition by the i53 polypeptide, its effect on HDR efficiency at the hemoglobin subunit beta (e.g., β -globin) (HBB) locus in CD34-expressing LT-HSPCs was investigated.

[0473] Frozen CD34-expressing LT-HSPCs derived from plerixafor (i.e., Mozobil®)+G-CSF-dual mobilized peripheral blood obtained from healthy human donors were purchased from a commercial vendor. LT-HSPCs were maintained in culture media comprised of the reagents shown in Table 4 and were incubated at 37° C., 5% carbon dioxide, 4% oxygen. The cells were electroporated with RNP comprised of Cas9 and gRNA targeting the HBB locus (R02 gRNA, target sequence shown in SEQ ID NO: 15). 2×10^5 cells were edited with 3 μ g Cas9 and 3 μ g gRNA. The target gene sequence (including target sequence with PAM), R02 spacer sequence, and R02 sgRNA sequence are identified in Table 3.

TABLE 3

Sequences of R02 sgRNA		
Name/ Description	Sequence	SEQ ID NO
HBB Target Sequence	CTTGCCCCACAGGGCAGTAA	15
HBB Target Sequence with PAM	CTTGCCCCACAGGGCAGTAACGG	20
R02 Spacer Sequence	CUUGCCCCACAGGGCAGUAA	16
R02 sgRNA (spacer in bold)	csususGCCCCACAGGGCAGUAA GUUUUAGAGCUAGAAA UAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAA AAAGUGGCACCGAGUCGGUGCusususU	17

a, c, g, u: 2' O-methyl phosphorothioate nucleotides
s: phosphorothioate nucleotides
A, C, G, U, N: canonical RNA nucleotides

[0474] The cells were transfected with a dsDNA homology donor encoding GFP under a SFFV promoter (SEQ ID NO: 60) that was delivered by AAV (SEQ ID NO: 56). Cells were edited with an AAV dose of 5,000 MOI. The HDR efficiency was determined by measuring the level of GFP fluorescence in the cells following electroporation. To determine the effect of 53BP1 inhibition on HDR efficiency, the 53BP1 inhibitor i53 was introduced as an mRNA-encoded protein to the cells during electroporation with Cas9/gRNA RNP and AAV DNA that serves as a donor for homology directed repair. As shown in FIG. 6A, treatment with mRNA encoding i53 polypeptide (SEQ ID NO: 70) resulted in an approximately 1.3-fold increase in HDR efficiency over cells treated with RNP+AAV-only. Additionally, improved HDR efficiency was seen at each dose of i53 mRNA tested (0.5, 1, and 2 µg mRNA). No GFP expression over background was seen in cells treated with RNP-only or AAV-only.

TABLE 4

Media components used to culture LT-HSPCs	
Component	Concentration
Thrombopoietin	100 ng/mL
Fms-like tyrosine kinase 3 ligand	100 ng/mL
Stem cell factor	100 ng/mL
Interleukin-3	60 ng/mL

[0475] The effect of 53BP1 inhibition on HDR efficiency was compared to other mRNA-encoded proteins that inhibit the NHEJ pathway. These included the proteins CYREN1 and CYREN2 that inhibit Ku70/80, a heterodimer that binds DNA blunt ends to prevent processing of 3' single-stranded DNA tails necessary for HDR repair. Cells were electroporated with RNP, an AAV-GFP homology donor and mRNA encoding either i53 (mRNA ORF shown in SEQ ID NO: 69), CYREN1, or CYREN2. Treatment with i53 mRNA resulted in increased HDR efficiency when administered at 0.3 µg or 1 µg. Treatment with CYREN1 or CYREN2 mRNA resulted in no improvement in HDR efficiency over RNP+AAV-only (FIG. 6B).

[0476] An inhibitor of the cell division cycle7-related (CDCl₇) protein kinase was also evaluated for improved HDR editing of the HBB gene locus. CDCl₇ is an initiator of the G1/S transition. Inhibition of CDCl₇ using the small molecule XL413 has been shown to improve HDR efficiency by inducing an early S-phase cell cycle arrest (Wienert, B. et al. (2018) bioRxiv 500462). However, for gene-editing of the HBB locus in LT-HSPCs, no improvement in HDR efficiency was seen for any dose of XL413 tested (data not shown).

[0477] The effect of treatment with i53 and Nu7441 was compared for improving efficiency of HDR repair of the HBB locus. Cells were treated with different doses of Nu7441 or with mRNA encoding i53 (ORF shown by SEQ ID NO: 69). A comparable increase in HDR efficiency was seen upon treatment with 0.75 µg of mRNA encoding i53 and 5 µM Nu7441 (FIG. 6C). No improvement in HDR efficiency was seen with treatment of a mRNA encoding a non-functional mutant i53 (e.g, DM, mRNA ORF shown by SEQ ID NO: 71).

[0478] The effect of treatment of i53 on HDR efficiency was evaluated in additional cell types, including editing of the AAVS1 locus in human epithelial cells immortalized with hTERT (hTERT RPE-1, ATCC CRL-4000) cells. RPE1 cells were lipofected with RNP comprised of Cas9 protein

and gRNA targeting the AAVS1 locus (target sequence shown in SEQ ID NO: 3; spacer sequence shown in SEQ ID NO: 4). Cells were edited with 1 µg Cas9 and 1 µg gRNA. Cells were infected with AAV containing homology donor DNA encoding GFP (DJ serotype) as well as mRNA encoding i53 (mRNA ORF shown in SEQ ID NO: 69). Cells were edited with an AAV dose of 25,000 MOI. HDR efficiency was determined by measuring the level of GFP fluorescence in the cells following gene-editing. Treatment of cells with AAV and mRNA encoding i53 results in an increase in HDR efficiency over treatment with AAV alone (FIG. 7).

Example 4. Efficient Gene Editing of the Hemoglobin Beta Subunit (HBB) Locus in CD34-Expressing LT-HSPCs Using i53 In Vitro

[0479] Improved HDR efficiency for CRISPR/Cas9 gene-editing of the HBB locus with i53 treatment in CD34-expressing LT-HSPCs was evaluated with donor DNA encoding a sickle cell mutation. A sickle mutation is a change of a GAG codon encoding Glu at position 7 of the beta-globin protein to a GUG codon encoding Val (codon 7 of the HBB open reading frame; E7V mutation). As is well known in the art and used herein the term “E7V” refers to a single nucleotide polymorphism (SNP) in the HBB gene that occurs in the seventh codon downstream the transcription start site (i.e. the seventh codon of HBB if including the AUG start codon), wherein the SNP converts the wild-type codon encoding Glu to a codon encoding Val. Correspondingly, a beta-globin polypeptide with an “E7V” mutation refers to substitution of Glu to Val occurring in the seventh amino acid residue of the beta-globin polypeptide if including the initial methionine amino acid. As used herein, the term “E6V” refers to a SNP in the HBB gene that occurs in the sixth codon downstream the AUG start codon (i.e., the sixth codon of the HBB open reading frame downstream the start codon). wherein the SNP converts the wild-type encoding Glu to a codon encoding Val. Correspondingly, a beta-globin polypeptide with an “E6V” mutation refers to substitution of Glu to Val occurring at the sixth amino acid residue of the beta-globin polypeptide, not including the initial methionine amino acid. Accordingly, as readily understood by one of ordinary skill in the art, the terms “E7V” and “E6V” refer to the same mutation in the HBB gene, and are used interchangeably herein when used in reference to the sickle mutation. FIG. 8 shows editing of the HBB locus using a Cas9/gRNA complex to introduce a site-specific DSB into exon 1. The homology donor DNA provided introduces a gene correction into exon 1 when repair of the DSB occurs by the HDR pathway. For wild type cells, a homology donor DNA encoding a sickle cell mutation (E7V) can be provided to introduce the sickle mutation into the HBB gene. For cells with the sickle cell mutation, a homology donor DNA encoding a sickle cell correction (E7) can be provided to introduce a sickle correction to the HBB gene.

[0480] LT-HSPCs were maintained in culture and gene-editing was performed following two days of culture. Cells were electroporated with RNP comprised of Cas9 and gRNA targeting the HBB locus (R02 gRNA, target sequence shown by SEQ ID NO: 15). 1×10⁶ were edited per reaction using 20 µg Cas9 and 20 µg gRNA. The cells were electroporated with 1 µg mRNA encoding i53. Electroporation was performed using the Maxcyte HSC-3 program. AAV encoding homology donor DNA (AAV.307) was administered prior to electroporation (pre-EP). Cells were edited with an AAV dose of 10,000 MOI. The donor DNA comprised homology arms to the HBB locus and encoded a sickle cell mutation

(SEQ ID NO: 53). FIG. 9 shows the sequence of the HBB gene in the region of the gene edit as well as the sickle cell mutation that is introduced following gene editing. The downstream PAM recognition site on the HBB locus is indicated, as well as the sequence recognized by the R02 gRNA spacer. Additionally, a portion of the sequence of the AAV.307 homology donor DNA is shown, including gene changes that are incorporated into the HBB locus by HDR editing. The homology donor incorporates an edit to the PAM sequence to prevent re-cutting of the HBB locus by Cas9/gRNA following editing by HDR. Sequences of the AAV.307 donor are provided in Table 5.

TABLE 5

Sequence of AAV.307 Homology Donor	
Name/Description	SEQ ID NO
5' ITR	110
Left Homology Arm (LHA)	52
Gene-edit (E7 → E7V)	53
Right Homology Arm (RHA)	54
3' ITR	105
LHA to RHA	108
AAV.307 full AAV	109

[0481] HDR efficiency at the HBB locus was evaluated upon treatment with i53. For assessing frequency of E7V modification of the HBB allele in samples of edited cells, a next-generating sequencing (NGS) assay using three PCR reactions was performed

[0482] Cells were treated with AAV (e.g., AAV.307) and RNP (e.g., Cas9/R02 gRNA) and treated with 1 µg of mRNA encoding the i53 polypeptide (SEQ ID NO: 70). Treatment with i53 resulted in 68% incorporation of the E7V gene edit, an increase of 1.4-fold over RNP+AAV alone (FIG. 10).

[0483] Editing efficiency at the HBB locus was evaluated by measuring indel formation by TIDE analysis. Electroporation of CD34-expressing LT-HSPCs with RNP comprised of Cas9 and R02 gRNA resulted in 94% indel formation, demonstrating the Cas9/gRNA yields high cutting efficiency within the desired target gene (FIG. 11). Notably, the level of indel formation decreased for cells treated with RNP+AAV. The lowest level of indel formation was seen for cells treated with RNP+AAV in the presence of i53, indicating that as the repair pathway shifts towards HDR, indel formation by the NHEJ pathway decreases. This group had 1.7-fold reduction in INDEL frequency relative to cells treated with RNP+AAV.

[0484] The effect of i53 on HDR efficiency was evaluated for CD34-expressing LT-HSPCs isolated from human peripheral blood using different mobilization methods. HDR efficiency was evaluated for LT-HSPCs isolated from human donors following administration of either Mozobil+GCSF or Mozobil-alone and gene-edited with AAV+RNP with or without inclusion of mRNA encoding the i53 polypeptide (SEQ ID NO: 70). Editing was performed by electroporation with RNP containing 20 µg Cas9 and 20 µg R02 gRNA and homology donor with a E7V mutation encoded by AAV (AAV.307 or AAV.304 comprising a gene-edit identified by SEQ ID NO: 50) at a dose of 10,000 MOI. Treatment with i53 resulted in approximately 60% incorporation of the sickle cell gene-edit by HDR in cells isolated by Mozobil+GCSF, approximately 1.5-fold increase in HDR efficiency over treatment with RNP+AAV alone (FIG. 12A). HDR efficiency in LT-HSPCs isolated using Mozobil and treated with i53 was even higher, with approximately 70% of cells incorporating the sickle cell gene-edit (FIG. 12B). These

results demonstrate that by isolating LT-HSPCs from healthy donors using mobilization by Mozobil, along with editing in the presence of i53, high levels of HDR efficiency are achieved.

[0485] The growth of CD34+ cells from the time the cells were thawed to right before they were subjected to editing conditions (black bars) was evaluated and no changes in cell growth were observed. Manipulations of the cells by the addition of CRISPR reagents used for editing caused a decrease in the fold growth between when they were thawed and injected into mice (blue bars) (FIG. 13).

Example 5. Evaluation of LT-HSPCs Edited with i53 Following Administered In Vivo

[0486] As described above, HDR efficiency for editing the HBB locus in CD34-expressing LT-HSPCs is improved in the presence of i53. The effect of gene-editing with i53 was evaluated on the ability of the cells to engraft and retain the gene-edit following administration in vivo. Also evaluated was the effect of LT-HSPC dose on engraftment following administration in vivo.

[0487] Human LT-HSPCs were administered to mice following electroporation with Cas9/gRNA (R02 gRNA, target sequence identified in SEQ ID NO: 15) and AAV encoding a sickle cell mutation (E7V) as shown in FIG. 14. Briefly, LT-HSPCs were mobilized from healthy donors using plerixafor. LT-HSPCs were gene-edited following 2 days in culture under conditions described in Example 3. To perform the gene-edit, cells were electroporated with RNP comprised of Cas9 and gRNA targeting the HBB locus (e.g., R02 gRNA, target sequence SEQ ID NO: 15) and homology donor DNA encoding a sickle cell mutation delivered by AAV (e.g., AAV.307). 1×10⁶ cells were edited with 20 µg Cas9, 20 µg gRNA and an AAV dose of 10,000 MOI. The cells were incubated with AAV.307 for 1 hour prior to electroporation. Effect of gene-editing with i53 was determined by treating cells with mRNA encoding i53 polypeptide (SEQ ID NO: 70) during electroporation. Cells were edited with 1 µg mRNA.

[0488] The cells were administered by intravenous injection to cKit mice at 2 days following electroporation. Recipient mice were treated with sublethal irradiation (100 cGy) at 1 day prior to administration of LT-HSPCs to eliminate hematopoietic cells in the bone marrow and enable engraftment of the donor cells. Animals were evaluated for presence of human hematopoietic cells in peripheral blood at 8 and 16 weeks following LT-HSPC administration. The bone marrow was also evaluated for engraftment and maintenance of the HBB gene-edit at 16 weeks following LT-HSPC administration.

[0489] Presence of human hematopoietic cells was measured by flow cytometry in mouse blood or bone marrow samples. The antibodies used for labeling cell-surface markers are shown in Table 6. The gating strategy used to quantify cells by flow cytometry is shown in FIG. 15. Cells were gated on singlet, live cells. Mouse and human CD45-expressing hematopoietic cells were distinguished by antibodies targeting mouse or human CD45. Engraftment was measured as percent chimerism which was defined as the quantity of human CD45 positive cells divided by the total number of CD45 positive cells (e.g., human and mouse CD45 expressing cells combined). The lineage of human CD45 positive cells was determined using markers for CD19 (e.g., B cells), CD3 (e.g., T cells), CD33 (e.g., myeloid cells), and CD34 (hematopoietic stem/progenitor cells (HSPCs)).

TABLE 6

Antibody	Clone	Fluorophore	Catalog #
Anti-mouse CD45	30-F11	APC	103112
Anti-human CD45	HI30	BV786	563716
Anti-human CD19	HIB19	PE-Cy7	302216
Anti-human CD3	UCHT1	APC-Cy7	300426
Anti-human CD33	P67.6	PE	366608
Anti-human CD34	581	BV421	562577

[0490] The effect of titrating the dose of LT-HSPCs on percent chimerism was evaluated in mouse bone marrow samples collected at 16 weeks following administration of LT-HSPCs. Animals received a dose of 0.01×10^6 , 0.05×10^6 , 0.1×10^6 , or 0.25×10^6 LT-HSPCs that were treated with electroporation, but neither AAV or RNP. As shown in FIG. 16, increasing the dose of LT-HSPCs administered to the animals resulted in increased levels of chimerism. Administration of 0.25×10^6 LT-HSPCs resulted in approximately 80% human cells within all CD45-expressing cells in the bone marrow. The effect of treatment of i53 on percent chimerism was also evaluated in mouse bone marrow collected at 16 weeks following administration of LT-HSPCs. Animals received a dose of 0.25×10^6 LT-HSPCs that were gene-edited with RNP+AAV either with or without i53. Percent chimerism for cells treated with RNP+AAV was lower than for cells treated with RNP-only or AAV-only (FIG. 16). Inclusion of i53 resulted in no further decrease in chimerism compared to AAV+RNP alone.

[0491] The effect of titrating the dose of LT-HSPCs on percent chimerism was also evaluated in mouse blood samples collected at 8 and 16 weeks following administration of LT-HSPCs. As seen in the bone marrow, increasing the dose of LT-HSPCs resulted in higher proportion of human cells among CD45-expressing in the blood (FIG. 17). Additionally, the proportion of human CD45-expressing cells among total CD45-expressing cells in the blood was lower for LT-HSPCs gene-edited with RNP+AAV alone or in combination with i53, with approximately 2-3% human CD45-expressing cells present. These data indicate that cells edited by HDR repair exhibit lower levels of engraftment than unedited cells or cells edited with RNP only.

[0492] Shown in FIGS. 18A-18B is lineage analysis of engrafted CD45-expressing human cells in mouse bone marrow samples collected at 16 weeks post administration of LT-HSPCs. The lineage of CD45-expressing human leukocytes was evaluated for unedited LT-HSPCs administered at a dose of 0.01×10^6 , 0.05×10^6 , 0.1×10^6 , or 0.25×10^6 cells (FIG. 18A). The lineage of CD45-expressing human leukocytes was also evaluated for LT-HSPCs edited with AAV+RNP either in the presence or absence of i53 (FIG. 18B). In both cases no gross changes in lineage distribution were observed for engrafted cells that were edited LT-HSPCs compared to un-edited LT-HSPCs.

[0493] Maintenance of gene-editing was evaluated in mouse bone marrow collected at 16 weeks post-administration of LT-HSPCs. Incorporation of a sickle mutation (E7V) in the HBB locus was evaluated in DNA isolated from mouse bone marrow samples using the next generation sequencing assay described in Example 4. Shown in FIG. 19 is a comparison of HDR efficiency for LT-HSPCs edited either with or without i53. LT-HSPCs were electroporated with Cas9/gRNA and treated with AAV. Administration of LT-HSPCs edited with i53 produced a bone marrow compartment with levels of gene-editing in the HBB locus that were substantially higher than AAV+RNP alone, with 65% incorporation of the gene edit by HDR. LT-HSPCs edited in

the presence of i53 had 1.8-fold higher HDR frequency in the bone marrow compared to cells edited with AAV+RNP alone. These results demonstrate that LT-HSPCs edited with i53 have higher levels of HDR efficiency and the gene-edit is maintained in cells derived from these progenitor cells following administration in vivo.

[0494] NHEJ editing of the HBB locus was also evaluated in mouse bone marrow collected at 16 weeks post engraftment. Indel formation at the site of Cas9/gRNA cutting was evaluated by TIDE analysis in bone marrow samples and compared to indel formation of LT-HSPCs prior to administration. Regardless of the method used to edit the LT-HSPCs, indel formation was similar at 16 weeks post- engraftment to the level present prior to administration, demonstrating persistence of gene-editing following engraftment of LT-HSPCs (FIG. 20). Interestingly, the level of indel formation was lowest for LT-HSPCs gene-edited in the presence of i53, demonstrating that i53 is an effective inhibitor of the NHEJ pathway.

Example 6. Evaluation of Gene Editing of the Hemoglobin Beta Subunit (HBB) Locus in CD34-Expressing LT-HSPCs In Vitro in the Presence of i53 or Nu7441

[0495] A direct comparison was made of the effect of i53 and Nu7441 on HDR efficiency for CRISPR/Cas9 gene-editing at the HBB locus in CD34-expressing LT-HSPCs using a homology donor DNA encoding a sickle cell mutation. LT-HSPCs were maintained in culture as described in Example 3 and gene-editing was performed following two days in culture as described in Example 4. Cells were electroporated with RNP comprised of Cas9 and gRNA targeting the HBB locus (R02 gRNA, target sequence shown by SEQ ID NO: 15). AAV encoding homology donor DNA (SEQ ID NO: 50) was administered either prior to electroporation (pre-EP) or post electroporation (post-EP). Additionally, where indicated, cells were treated with 5 μ M Nu7441 or 1 μ g of mRNA encoding i53 polypeptide (SEQ ID NO: 70) during electroporation.

[0496] The donor DNA comprised homology arms to the HBB locus and encoded a sickle cell mutation. FIG. 21 shows the sequence of the HBB gene in the region of the gene edit as well as the sickle cell mutation that is introduced following gene editing. The downstream PAM recognition site on the HBB locus is indicated, as well as the sequence of the homology donor DNA (AAV.304), including gene changes that are incorporated into the HBB locus by HDR editing. The homology donor incorporates an edit to the PAM sequence to prevent re-cutting of the HBB locus by Cas9/gRNA following editing by HDR. Sequences of the AAV.304 donor are provided in Table 7.

TABLE 7

Sequence of AAV.304 Homology Donor	
Name/Description	SEQ ID NO
5' ITR	110
Left Homology Arm (LHA)	49
Gene-edit (E7 \rightarrow E7V)	50
Right Homology Arm (RHA)	51
3' ITR	105
LHA to RHA	106
AAV.304 full AAV	107

[0497] HDR efficiency was evaluated for AAV administered either pre-EP or post-EP following gene-editing of

LT-HSPCs in vitro. For pre-EP, cells were incubated with AAV for 1 hour prior to electroporation. For post-EP, cells were incubated with AAV for 1 hour immediately following electroporation. HDR efficiency was evaluated by NGS assay as described in Example 4. Treatment with RNP and AAV administered either before or after electroporation resulted in a comparable level of incorporation of the gene edit by HDR, approximately 40% (FIG. 22A). Treatment with AAV-only either pre- or post-EP resulted in no incorporation of the donor DNA in the HBB locus.

[0498] HDR efficiency at the HBB locus was evaluated upon treatment with either i53 or Nu7441 following gene-editing of LT-HSPCs in vitro. Cells were treated with pre-EP AAV and RNP and treated with either Nu7441 or mRNA encoding i53. Treatment with Nu7441 resulted in no improvement of HDR efficiency over RNP+AAV alone. However, treatment with i53 resulted in 58% incorporation of the E7V gene edit, an increase of 1.4-fold over RNP+AAV alone (FIG. 22B).

Example 7. Evaluation of LT-HSPCs Edited with i53 or Nu7441 Following Administered In Vivo

[0499] As described in Example 6, HDR efficiency for editing the HBB locus in CD34-expressing LT-HSPCs is improved in the presence of i53, but not in the presence of Nu7441. The effect of gene-editing with either i53 or Nu7441 was evaluated on the ability of the cells to engraft and retain the gene-edit following administration in vivo. Also evaluated was the effect of gene-editing with treatment of AAV prior to electroporation or following electroporation on the ability of the edited cells to engraft and maintain the gene edit following administration in vivo.

[0500] Human LT-HSPCs were administered to mice following electroporation with Cas9/gRNA and AAV encoding a sickle cell mutation (E7V) as shown in FIG. 21. LT-HSPCs were mobilized from healthy donors using plerixafor. LT-HSPCs were gene-edited following 2 days in culture. To perform the gene-edit, cells were electroporated with RNP comprised of Cas9 and gRNA targeting the HBB locus (e.g., R02 gRNA, target sequence shown by SEQ ID NO: 15) and AAV encoding a sickle cell mutation (AAV.304). 1×10^6 cells were edited with 20 μ g Cas9 and 20 μ g gRNA. The AAV was administered either prior to electroporation (pre-EP) or following electroporation (post-EP). Cells were edited with an AAV dose of 10,000 MOI. Effect of gene-editing with i53 or Nu7441 was determined by treating cells with 1 μ g of mRNA encoding i53 polypeptide (SEQ ID NO: 70) or 5 μ M Nu7441 during electroporation.

[0501] A dose of 0.5×10^6 cells was administered by intravenous injection to cKit mice at 2 days following electroporation. Recipient mice were treated with sublethal irradiation (100 cGy) at 1 day prior to administration of LT-HSPCs to eliminate hematopoietic cells in the bone marrow and enable engraftment of the donor cells. Animals were evaluated for presence of human hematopoietic cells in peripheral blood at 8 and 16 weeks following LT-HSPC administration. The bone marrow was also evaluated for engraftment and maintenance of the HBB gene-edit at 16 weeks following LT-HSPC administration.

[0502] Percent chimerism was evaluated as described in Example 5 in mouse blood samples collected at 8 and 16 weeks post-administration of LT-HSPCs and compared for LT-HSPCs edited under different conditions. Shown in FIG. 23A is a comparison of LT-HSPCs edited with AAV administered either pre-EP or post-EP. LT-HSPCs were electroporated with Cas9/gRNA and treated with AAV prior to electroporation or following electroporation. LT-HSPCs

electroporated with RNP demonstrated decreased chimerism relative to cells that were not electroporated. Treatment with AAV either pre-EP or post-EP resulted in no improvement in chimerism relative to treatment with RNP alone. Shown in FIG. 23B is a comparison of LT-HSPCs edited with pre-EP AAV and RNP in the presence of either i53 or Nu7441. Treatment with i53 resulted in levels of chimerism comparable to RNP+AAV alone. However, treatment with Nu7441 resulted in improved chimerism of approximately 25% at both 8 and 16 weeks, a 2.5-fold increase over RNP+AAV alone.

[0503] Additionally, percent chimerism was evaluated in mouse bone marrow samples collected at 16 weeks following administration of LT-HSPCs. Shown in FIG. 24A is a comparison of LT-HSPCs edited with AAV administered either pre-EP or post-EP. Similar to the chimerism seen in mouse blood samples as described above, LT-HSPCs electroporated with RNP demonstrated decreased chimerism relative to LT-HSPCs that were not electroporated. Furthermore, treatment with AAV either pre-EP or post-EP resulted in no improvement in chimerism relative to treatment with RNP alone. The chimerism in bone marrow at 16 weeks for LT-HSPCs edited in the presence of i53 or Nu7441 was also evaluated (FIG. 24B). Treatment with i53 resulted in levels of chimerism comparable to RNP+AAV alone. However, treatment with Nu7441 resulted in chimerism that was higher than RNP+AAV alone and comparable to culture LT-HSPCs that were not electroporated prior to engraftment. Combined, these results demonstrate an unexpected improvement in engraftment for LT-HSPCs gene-edited with treatment of Nu7441.

[0504] Shown in FIG. 25 is lineage analysis of engrafted CD45-expressing human cells in mouse bone marrow samples collected at 16 weeks post administration of LT-HSPCs. The lineage of CD45-expressing human leukocytes was compared for LT-HSPCs that were gene-edited in the presence of i53 or Nu7441. In addition to providing higher levels of engraftment, treatment with Nu7441 resulted in a greater proportion of CD34-expressing cells and myeloid cells among human leukocytes in the bone marrow compared to treatment with i53.

[0505] Maintenance of gene-editing was evaluated in mouse bone marrow collected at 16 weeks post-administration of LT-HSPCs. Incorporation of a sickle mutation (E7V) in the HBB locus was evaluated in DNA isolated from mouse bone marrow samples using the next generation sequencing assay described in Example 4. Shown in FIG. 26A is a comparison of HDR efficiency for LT-HSPCs edited with AAV administered either pre-EP or post-EP. LT-HSPCs were electroporated with Cas9/gRNA and treated with AAV prior to electroporation or following electroporation. Incorporation of the gene-edit in bone marrow samples was similar for LT-HSPCs edited with RNP and AAV given either pre-EP or post-EP. Additionally, incorporation of the gene edit in the HBB locus by HDR was compared for LT-HSPCs edited in the presence of i53 or Nu7441 (FIG. 26B). Administration of LT-HSPCs edited with Nu7441 produced a bone marrow compartment with levels of gene-editing in the HBB locus comparable to AAV+RNP alone. However, administration of LT-HSPCs edited with i53 produced levels of gene-editing in the HBB locus that were substantially higher than AAV+RNP alone. These results demonstrate that LT-HSPCs edited with i53 have higher levels of HDR efficiency and the gene-edit is maintained in cells derived from these progenitor cells following administration in vivo.

[0506] NHEJ editing of the HBB locus was also evaluated in mouse bone marrow collected at 16 weeks post engraftment. Indel formation at the site of Cas9/gRNA cutting was evaluated by TIDE analysis in bone marrow samples and compared to indel formation of LT-HSPCs prior to administration. Regardless of the method used to edit the LT-HSPCs, indel formation was similar at 16 weeks post-engraftment to the level present prior to administration (FIG. 27). Interestingly, the level of indel formation was lowest for LT-HSPCs gene-edited in the presence of i53, demonstrating that i53 is an effective inhibitor of the NHEJ pathway.

[0507] The functionality of the hematopoietic compartment in recipient mice was evaluated by measuring erythroid cell enucleation in bone marrow collected at 16 weeks post-engraftment. Mammalian erythrocytes extrude their nucleus prior to entering circulation. Human CD34-expressing LT-HSPCs are expected to differentiate into erythrocytes following engraftment, however the efficiency of enucleation can be low. Assessment of erythroid cell enucleation provides a measure of the ability of edited CD34 expressing cells to differentiate into erythroid cells compared to the unedited controls (i.e., ability to differentiate into functional cell types). Percent enucleation was compared for LT-HSPCs gene edited with Cas9/gRNA RNP and AAV given pre-EP or post-EP. Levels of enucleation were similar to cells treated with RNP-only or AAV-only (FIG. 28). Additionally, levels of enucleation were compared for LT-HSPCs gene edited with RNP+AAV in the presence of i53 or Nu7441. Levels of enucleation were also similar to cells treated with RNP-only or AAV-only.

Example 8: Evaluation of i53 and Nu7441
Combination for Editing HBB Using a Homology
Donor Encoding an HBB Gene Correction

[0508] A direct comparison was made of the effect of i53 and Nu7441 on HDR efficiency for CRISPR/Cas9 gene-editing at the HBB locus in wild-type CD34-expressing LT-HSPCs using a homology donor DNA encoding a correction to the sickle cell mutation in the HBB gene (i.e., E6V to E6) delivered by AAV. The AAV-encoded homology donor used for correction is referred to as “AAV.323” and is identified by sequence in Table 8. As shown in FIG. 29, AAV.323 encodes glutamate at position 6 of the HBB open reading frame (i.e., E6). However, the codon for E6 is “GAA” rather than wild-type “GAG”, allowing the correction encoded by the AAV.323 to be detected in wild-type cells or cells encoding the E6V mutation in the HBB gene.

TABLE 8

Sequence of AAV323 Homology Donor Encoding SCD Correction	
Name/Description	SEQ ID NO
5' ITR	104
Left Homology Arm (LHA)	99
Gene-edit (E6V → E6)	102
Right Homology Arm (RHA)	100
3' ITR	105
LHA to RHA	98
AAV.323 full AAV	103

[0509] Briefly, LT-HSPCs were maintained in culture and gene-editing was performed following two days in culture as described in Example 3. 1×10⁶ cells were electroporated with RNP containing 20 μg SpCas9 and 20 μg R02 sgRNA. Cells were incubated with AAV.323 at a dose of 10,000 MOI

for one hour prior to electroporation. Additionally, cells were treated with 5 μM Nu7441, 1 μg of i53 mRNA, or both during electroporation.

[0510] On day 2 post-editing, the efficiency of HDR for insertion of the AAV.323 gene-edit was evaluated by NGS assay as described in Example 4. The frequency of INDELs at the R02 cut site was also evaluated by NGS. Editing was compared for cells treated R02+AAV.323 and either i53 or i53+Nu7441 and compared to control cells edited with R02 only or R02+AAV.323 only.

[0511] As shown in FIG. 30A, cells edited in the presence of i53 inhibitor had an approximately 1.5-fold increase in HDR efficiency for incorporation of the AAV.323 gene-edit compared to cells edited with R02+AAV.323 only. However, editing performed with the combination of i53 and Nu7441 did not increase HDR efficiency further compared to editing performed with i53 only. Additionally, as shown in FIG. 30B, cells edited in the presence of i53 had decreased frequency of INDELs at the R02 cut site compared to cells edited with R02+AAV.323 only. Editing performed with the combination of i53 and Nu7441 did not further decrease INDEL frequency compared to editing performed with i53 only.

[0512] Edited cells were administered in vivo as described in Example 5, and the level of engraftment of human hematopoietic progenitor cells was evaluated in mouse bone marrow. Briefly, a dose of approximately 5×10⁶ per animal LT-HSPCs was administered by intravenous injection to cKit mice at two days post-editing. Recipient mice were treated with sublethal irradiation (100 cGy) at 1 day prior to administration of LT-HSPCs. At 16 weeks following administration, the presence of human hematopoietic cells in mouse bone marrow samples was evaluated by flow cytometry as described in Example 5. As shown in FIG. 31A, more than 90% of CD45-positive cells in mouse bone marrow were human CD45-positive cells for each treatment group evaluated, indicating high levels of engraftment regardless of the conditions used to edit the cells. These data suggest that edited LT-HSPCs can effectively establish and expand following transplantation. However, their ability to do so can depend on, for example, the donor from which the cells are obtained, method of isolation, culture conditions or cell cycle status.

[0513] The persistence of edits at the HBB gene locus in mouse bone marrow samples was also evaluated. Briefly, mouse bone marrow was isolated at 16 weeks post-administration of edited cells. Genomic DNA was isolated from the samples, and evaluated for incorporation of the edit encoded by AAV.323 at the HBB gene locus using the NGS assay as described in Example 4. Frequency of INDELs at the R02 cut site in the HBB gene locus was also assessed by NGS.

[0514] As shown in FIG. 31B, the frequency of the AAV.323 gene edit incorporated by HDR was maintained in bone marrow of mice administered LT-HSPCs edited with i53.

Example 9: Evaluation of i53 for In Vitro
Correction of Sickle Cell Mutation in Human
Patient-Derived Cells

[0515] The effect of i53 on HDR efficiency was evaluated in cells derived from patients with a sickle cell mutation in the HBB gene. Specifically, CD34-expressing LT-HSPCs derived from human patients with sickle cell disease were edited with SpCas9, R02 guide, and AAV.323 encoding a correction to the E6V mutation in the HBB gene.

[0516] Briefly, CD34-expressing LT-HSPCs were derived from plerixafor+GCSF-dual mobilized peripheral blood obtained from a human donor with sickle cell disease. The cells were seeded in Phase I media at a cell density of 2×10^5 cells/mL. Cells were cultured at 37° C. under normoxic conditions (i.e., oxygen 20%).

[0517] Editing of cells was performed following two days of in vitro culture. Briefly, 5×10^5 cells were electroporated with RNP containing 20 µg SpCas9 and 20 µg R02 sgRNA, AAV.323 at a dose of 10,000 MOI, and 1 µg mRNA encoding i53. The cells were incubated with AAV.323 for 1 hour prior to electroporation. The cells were edited by electroporation with R02+AAV.323+i53 mRNA and compared to control cells edited with R02 only, R02+AAV.323, or cells exposed to electroporation without RNP or AAV editing components (mock EP).

[0518] Following editing, the cells were differentiated to erythrocytes. Briefly, edited cells were plated in fresh Phase I media at a density of 2×10^5 cells/mL, and re-plated at similar density in fresh Phase I media on days 3 and 5 post-editing. On day 7 post-editing, the cells were incubated in Phase II media at a density of 2.5×10^5 cells/mL. On day 10 post-editing, the cells were incubated in Phase III media at a density of 1.2×10^6 cells/mL. Cell expansion during culture was monitored over time and cells electroporated with R02+AAV.323+i53 mRNA grew similarly to control cells (R02 only, R02+AAV.323, or mock EP cells) (data not shown). Additionally, cell viability was monitored at frequent time points beginning day 3 post-editing, and remained greater than 80% for each treatment group through approximately day 13 of culture.

Efficiency of Gene Edits

[0519] The efficiency of gene correction by HDR repair at the HBB gene locus was evaluated by NGS assay as described in Example 3. Frequency of INDELs at the R02 cut site was evaluated by NGS analysis. Treatment with i53 resulted in 66% incorporation of the E6V→E6 gene correction, an increase of 1.4-fold over RNP+AAV.323 alone (FIG. 32A). Additionally, frequency of INDELs at the R02 cut site was 1.9-fold lower for cells edited in the presence of i53 compared to cells edited with RNP+AAV.323 alone (FIG. 32B). Additionally, HDR repair and INDEL formation were compared at day 0 and day 14 post-editing. As shown in FIGS. 33A-33B, edit incorporated by HDR repair (FIG. 33A) and frequency of INDELs at the R02 cut site (FIG. 33B) was similar at day 0 and day 14 for each treatment group, indicating the edits were retained throughout in vitro differentiation to erythrocytes.

Hemoglobin Expression

[0520] Hemoglobin expressed by edited cells that were differentiated to erythrocytes was assessed. Hemoglobin A (HbA) is composed of 2 alpha-globin and 2 beta-globin units and is the dominant hemoglobin in adult humans. In human carriers of the HBB E6V mutation, a high proportion of total hemoglobin is hemoglobin S (HbS), which is composed of 2 alpha-globin units and 2 beta-globin units with E6V. Thus, proportion of HbS to HbA produced by edited cells was assessed using an HPLC-based quantification to determine if editing resulted in decreased levels of hemoglobin associated with sickle cell disease.

[0521] Briefly, on day 18 post-editing, 1×10^6 cells were harvested, centrifuged, and washed with PBS. The cells were prepared for HPLC analysis. Hemoglobin variants were quantified in cell samples using reverse-phase HPLC

chromatography and gradient elution. As shown in FIG. 34A, HbS levels were dramatically reduced and HbA levels increased for cells edited with R02+AAV.332 or R02+AAV.332+i53 compared to mock EP control cells. As shown in FIG. 34B, cells edited in the presence of i53 had 66% correction of HBB gene locus by HDR, but a 90% decrease in HbS levels relative to mock EP control cells. Thus, high levels of HDR achieved with i53 contribute to normalization of hemoglobin expression products.

Erythrocyte Functionality

[0522] The ability of edited cells to differentiate to functional erythrocytes was assessed by determining expression of erythrocyte-associated cell surface markers and enucleation using flow cytometry on day 18 post editing.

[0523] Briefly, 4×10^5 cells were obtained, and half were stained for erythrocyte cell-surface markers and half were used for detection of enucleation. For staining cell-surface markers, the cells were incubated in PBS containing 1% human serum albumin (PBS-A) and an antibody cocktail of anti-CD233 (BRIC6-Band3)-FITC, anti-CD71-PE, anti-CD235a(GlyA)-PE/Cy7, and anti-CD49d (a4)-VioBlue. For detection of enucleation, 2 drops of NucRed nuclear staining reagent was added to 1 mL PBS-A, and 100 µL was added to plated cells. Following incubation, both cell samples were labeled with Sytox Blue solution (1:1000 dilution in PBS-A) for live/dead analysis. Samples were then assessed by flow cytometry. Cells edited with R02 only, R02+AAV.332, or R02+AAV.332+i53 each demonstrated levels of enucleation comparable to mock EP control cells (>30% of cell population having enucleation). Additionally, the proportion of the population that was CD71⁺GlyA⁺ erythrocytes was similar for cells edited with R02+AAV.332 or R02+AAV.332+i53 compared to mock EP control cells (>30% of cell population CD71⁺GlyA⁺).

Editing Patient-Derived PBMCs

[0524] It was further evaluated if editing of patient-derived PBMCs in the presence of i53 would yield high levels of correction of the HBB gene. PBMCs were obtained from a human donor with sickle cell disease. The PBMCs were expanded in StemSpan SFEM II (1x)+StemSpan CC100 (1x)+Dexamethasone 1 µM+hEPO 2 IU/mL at 37° C. under normoxic conditions (20% O₂ concentration). The cells were edited following five days of in vitro culture. Patient-derived PBMCs were edited with R02, R02+AAV.332, or R02+AAV.332+i53 as described above. On day 8, the cells were transferred to Phase I media, and differentiation to erythrocytes was performed through day 18 as described above. Efficiency of HDR at the HBB gene locus was evaluated on day 12 using the NGS assay described in Example 4. Also evaluated was the frequency of INDELs at the R02 cut-site as measured by NGS.

[0525] As shown in FIG. 35A, the frequency of correction of the HBB gene by HDR repair in the presence of i53 was approximately 60% in patient-derived PBMCs. Additionally, the frequency of INDELs was reduced in patient-derived PBMCs edited in the presence of i53 compared to control cells (FIG. 35B). The level of HDR repair and the frequency of INDELs in HBB were comparable in PBMCs and CD34-expressing LT-HSPCs edited in the presence of i53.

[0526] Hemoglobin expression was measured by HPLC analysis for edited PBMCs as described above. As shown in FIG. 36, PBMCs edited in the presence of i53 had significant reduction in expression of HbS and increased expression of

HbA compared to mock EP control cells. The ratio of HbS to HbA was comparable for PBMCs edited in the presence of i53 to CD34-expressing LT-HSPCs edited with i53.

[0527] Additionally, the functionality of erythrocytes differentiated from edited PBMCs was evaluated by measuring cell-surface markers using flow cytometry as described above. PBMCs edited with either R02+AAV.323 or R02+AAV.232+i53 had similar levels of CD71⁺GlyA⁺ erythro-

cytes to control cells (R02 only or Mock EP cells), indicating edited cells undergoing HDR repair of the HBB locus properly differentiate to mature erythrocytes (data not shown).

Sequence Listing

[0528]

SEQ ID NO:	Identifier	Name/Description	Sequence
1	gRNA-related	sgRNA1	n ₍₁₇₋₃₀₎ GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA aaguggcaccgagucggugcu ₍₁₋₈₎
2	gRNA-related	SpCas9 sgRNA	csususN ₍₁₇₋₃₀₎ GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUU AUCAACUUGAAAAGUGGCACCGAGUCGGUGCusususU
3	gRNA-related	AAVS1 target sequence A	GGGGCCACTAGGGACAGGAT
4	gRNA-related	AAVS1 sgRNA spacer A	GGGGCCACUAGGGACAGGAU
5	gRNA-related	AAVS1 target sequence B	GCCAGTAGCCAGCCCCGTCC
6	gRNA-related	AAVS1 sgRNA spacer B	GCCAGUAGCCAGCCCCGUCC
7	gRNA-related	BFP target sequence	TGAAGCACTGCACGCCAT
8	gRNA-related	BFP sgRNA spacer	UGAAGCACUGCACGCCAU
9	gRNA-related	GFP target sequence A	GCTGAAGCACTGCACGCCGT
10	gRNA-related	GFP sgRNA spacer A	GCUGAAGCACUGCACGCCGU
11	gRNA-related	GFP target sequence B	CTCGTGACCACCCGTGACCTA
12	gRNA-related	GFP sgRNA spacer B	CUCGUGACCACCCUGACCUA
13	gRNA-related	GSD1a target sequence	TCTTTGGACAGCGTCCATAC
14	gRNA-related	GSD1a Ch32 gRNA spacer	UCUUUGGACAGCGUCCAUAUAC
15	gRNA-related	HBB target sequence	CTTGCCCCACAGGGCAGTAA
16	gRNA-related	HBB R02 sgRNA spacer	CUUGCCCCACAGGGCAGUAA
17	gRNA-related	HBB sgRNA	csususGCCCCACAGGGCAGUAGUUUUAGAGCUAGAAUAGCAAGU UAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAGUGGCACCGA GUCGGUGCusususU
18	gRNA-related	CFTR target sequence	TCTGTATCTATATTCATCAT
19	gRNA-related	CFTR sgRNA spacer	UCUGUAUCUAUAUUCAUCAU
20	gRNA-related	HBB Target Sequence (PAM in bold)	CTTGCCCCACAGGGCAGTA ACGG

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
21	Donor DNA	ssODN1 (Ht-CR282)	GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCA AGCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACGTACGG CGTGCAAGTGCTTCAGCCGCTACCCCGACCACATGA
22	Donor DNA	ssODN2 (Hn-CR283)	TCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTACGTCAG GGTGGTCACGAGGGTGGGCCAGGGCAGGGCAGCTTGCCGGTGGT GCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGC
23	Donor DNA	ssODN3 (Hn-39-88)	CGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGT GCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTA CGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGG
24	Donor DNA	ssODN4 (Ht-91-61)	GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCA AGCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACGTACGG CGTGCAAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGAC TTCTTCAAGTCCGC
25	Donor DNA	ssODN (Hn-91-61)	GCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGC TGAAGCACTGCACGCCGTACGTACGGGTGGTCACGAGGGTGGGCCA GGGCACGGGCAGCTTGCCGGTGGTCAGATGAACTTCAGGGTCAGC TTGCCGTAGGTGGC
26	Donor DNA	ssODN (Hn-91-36)	TCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTACGTCAG GGTGGTCACGAGGGTGGGCCAGGGCAGGGCAGCTTGCCGGTGGT GCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGC
27	Donor DNA	ssODN (Ht-91-61)	GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCA AGCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACGTACGG CGTGCAAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGAC TTCTTCAAGTCCGC
28	Donor DNA	ssODN (Ht-39-88)	CGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGT GCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGTA CGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGG
29	Donor DNA	ssODN 1067	GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCA AGCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGAGCCACGG GGTGCAAGTGCTTCAGCCGCTACCCCGACCACATGA
30	Donor DNA	ssODN 1068	TCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACCCCGTGGCTCAG GGTGGTCACGAGGGTGGGCCAGGGCAGGGCAGCTTGCCGGTGGT GCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGC
31	Donor DNA	ssODN 1069	TCATCTGCACCACCGGCAAGCTGCCCGTGCCTGGCCACCCCTCGT GACCACCCTGAGCCACGGGGTGAGTGCTTCAGCCGCTACCCCGAC CACATGAAGCAGCACGACTTCTTCAAGTCCGCCAT
32	Donor DNA	ssODN 1070	ATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGC GGCTGAAGCACTGCACCCCGTGGCTCAGGGTGGTCACGAGGGTGGG CCAGGGCACGGGAGCTTGCCGGTGGTCAGATGA
33	Donor DNA	ssODN 1061	CTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGC TGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACCCCGTGGC TCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGC
34	Donor DNA	ssODN 1062	GCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGAGCCACGGGGTG CAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCT TCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAG
35	Donor DNA	ssODN 1063	GACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGA AGCACTGCACCCCGTGGCTCAGGGTGGTCACGAGGGTGGGCCAGG GCACGGGCAGCTTGCCGGTGGTCAGATGAACTTCA
36	Donor DNA	ssODN 1064	TGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCTGGCCAC CCTCGTGACCACCCTGAGCCACGGGGTGAGTGCTTCAGCCGCTACC CCGACCACATGAAGCAGCACGACTTCTTCAAGTC
37	Donor DNA	50-0 dsDNA	CCCAGAAACTTGTCTGTTTTCCATAGGATTCCTTTGGACAGTGCC CCT

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
38	Donor DNA	150-0 dsDNA	AGGGCAGCTGTCCAAAGAGAATCCTATGGAAAAACAGAACAAAGTTTCTGGGGTTACTGAATGAATGCTTTTGCCCAAAGCCTACACCTTCAA GAAGAGTGTAGCTGAGAAGGATTTCACATGTTGCCTCTAGAAGGG AGAAGTGGGTGGC
39	Donor DNA	93-50 ssODN	TCTTGAAGGTGTAGGCTTTGGGCAAAGCATTTCATTAGTAACCCAGAACTGTCTGTCTTTTCCATAGGATTCTCTTTGGACAGTGCCCTTACTGGTGGGTCTGGTACTGACTACTACAGCAACACTTCCGTGCCCT
40	Donor DNA	25-100 ssODN	TAGGATTCTCTTTGGACAGTGCCCTTACTGGTGGGTCTGGTACTGACTACTACAGCAACACTTCCGTGCCCTGATAAAGCAGTTCCCTGT AACCTGTGAGACTGGACCAGGTAAGCGTCCCCA
41	Donor DNA	H3-95-30 ssODN	TAAGCACAGTGAAGAATTTTCATTCTGTTCTCAGTTTTCTGGATTA TGCCTGGCACCATTAAAGAAAATATCATAAGCTTTGGTGTGTGCTAT GATGAATATAGATACAGAAGCGTCATCAAAG
42	Donor DNA	N1-95-30 ssODN	AATTAAGCACAGTGAAGAATTTTCATTCTGTTCTCAGTTTTCTGGATTTATGCCTGGCACCATTAAAGAAAATATCATCTTTGGTGTGTGCTAG CATGATGAATATAGATACAGAAGCGTCATCA
43	Donor DNA	AAVS1 locus LHA (used for BFP donor)	CCCCAGCTCTTCTCTGTTTCAGCCCTAAGAATCCTGGCTCCAGCCCTCCTACTCTAGCCCCCAACCCCTAGCCACTAAGGCAATTGGGGTGC AGGAATGGGGGAGGGTACCAGCCTCACCAAGTGGTTGATAAACC CACGTGGGGTACCCCTAAGAACTTGGGAACAGCCACAGAGGGGGG CGATGCTTGGGGACCTGCCTGGAGAAGGATGCAGGACGAGAACA CAGCCCCAGGTGGAGAACTGGCCGGGAATCAAGAGTCACCCAGA GACAGTGACCAACCCTCCTGTTTTCCTAGGACTGAGGGTTTCAGT GCTAAAACTAGGCTGTCTTGGGCAACAGCATAAGCTGGTCACCCC ACACCCAGACCTGACCCAAACCCAGCTCCCCTGCTTCTTGGCCACG TAACCTGAGAAGGGGAATCCCTCCTCTCTGAACCCAGCCACCCCA ATGCTCCAGGCCTCCTGGGATACCCGAAGAGTGAGTTTGCCAAGC AGTCACCCACAGTTGGAGGAGAATCCACCCAAAAGGCAGCCTGGT AGACAGGGCTGGGGTGGCCTCTCGTGGGGTCCAGGCCAAGTAGGTG GCCTGGGGCTCTGGGGATGCAGGGGAAGGGGATGCAGGGGA CGGGGATGCAGGGGAACGGGGCTCAGTCTGAAGAGCAGAGCCAGG AACCCCTGTAGGGAAGGGGCAGGAGACCCAGGGGATGAGATGGT GGACGAGGAAGGGGACAGGGAAGCCTGAGCGCTCTCTGGGCT TGCCAAGGACTCAAACCCAGAAGCCAGAGCAGGGCCTTAGGGAA GCGGACCCCTGCTCTGGGCGGAGGAATATGTCCAGATAGCACTGG GGAATCTTAAAGGAAGAAGGATGGAGAAAGAGAAGGGAGTAGA GGCGGCCACGACCTGGTGAACACCTAGGACGACCATTTCTCAGAAA GGGAGTTTTCACACGGACACCCCTCTCTACCCAGCCCTG
44	Donor DNA	BFP locus donor	ATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGGTGGTGGCCATCC TGGTCGAGCTGGACGGCGACGTAACCGGCCACAAGTTTCAGCGTGT CGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAA GTTTCATCTGCACACCGCAAGCTGCCGTGCCCTGGCCACCCCTC GTGACCACTGACCCATGGCGTGAGTGCTTCAGCCGCTACCCCG ACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCTGGTGAAC CGCATCGAGCTGAAGGCGATCGACTTCAAGGAGGACGGCAACATC CTGGGGCACAAGCTGGAGTACAACACAGCCACAACGCTCTATA TCATGGCCGACAAAGCAGAAGAACCGCATCAAGGTGAACCTCAAGA TCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTA CCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGAC AACCCTACCTGAGCACCAGTCCAGCTGAGCAAGACCCCAACG AGAAGCGGATCAGATGGTCTGCTGGAGTTCTGTGACCGCCCGG GATCACTCTCGCATGGACGAGCTGTACAAG
45	Donor DNA	AAVS1 locus RHA (used for BFP donor)	ACTGTGGGTGGAGGGGACAGATAAAAGTACCCAGAACCAGAGCC ACATTAACCGCCCTGGGAATATAAGGTGGTCCAGCTCGGGGACA CAGGATCCCTGGAGGCAGCAACATGCTGTCTGAAGTGGACATAG GGGCCCGGTTGGAGGAAGAAGACTAGCTGAGCTCTCGGACCCCTG GAAGATGCCATGACAGGGGGCTGGAAGAGCTAGCACAGACTAGAG AGGTAAGGGGGGTAGGGGAGCTGCCCAATGAAAGGAGTGAGAGG TGACCCGAATCCACAGGAGAAGGGGTGTCAGGCAAGAAAGCA AGAGGATGGAGAGGTGGCTAAAGCCAGGGAGACGGGGTACTTTGG GGTTGTCCAGAAAACGGTGATGATGACGGCTACAAGAAGGGGA GGCGGGACGCAAGGAGACATCCGTGGAGAAGGCCATCTAAGA

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
			AACGAGAGATGGCACAGGCCCCAGAAGGAGAAGGAAAAGGGAAAC CCAGCGAGTGAAGACGGCATGGGGTTGGGTGAGGGAGGAGAGATG CCCGAGAGGAGCCAGACACGGGGAGGATCCGCTCAGAGGACATC ACGTGGTGCAGCGCCGAGAAGGAAGTGCTCCGGAAGAGCATCCT TGGGCAGCAACACAGCAGAGCAAGGGGAAGGGGAGTGGAGG AAGACGGAACCTGAAGGAGGCGGCAGGGAAGGATCTGGGCCAGCC GTAGAGGTGACCCAGGCCACAAGCTGCAGACAGAAAGCGGCACAG GCCCAGGGGAGAGAATGCAGGTGAGAGAAAGCAGGACCTGCCTGG GAAGGGGAAACAGTGGGCGAGAGGCGGCAGAAAGCCAGTAGAGC TCAAAGTGGTCCGGAAGCTCAGGAGAGAGACGGCAGCGTTAGAGGGC AGAGTTCGGCGGCACAGCAAGGGCACTCGGGGGCGAGAGGAGGG CAGCGCAAAGTGACAATGGCCAGGGCCAGGCAGATAGACCAGACT GAGCTATGG
46	Donor DNA	AAVS1 locus LHA (used for GFP donor)	CCCCAGCTCTTCTCTGTTTCAGCCCTAAGAATCCTGGCTCCAGCCCCCT CCTACTCTAGCCCCCAACCCCTAGCCACTAAGGCAATTGGGGTGC AGGAATGGGGGAGGGTACCAGCCTCACCAAGTGGTTGATAAAACC CACGTGGGGTACCCTAAGAACTTGGGAACAGCCACAGCAGGGGGG CGATGCTTGGGGACCTGCCTGGAGAAGGATGCAGGACGAGAACA CAGCCCCAGGTGGAGAACTGGCCGGGAATCAAGAGTCACCCAGA GACAGTGACCAACCATCCCTGTTTTCTTAGGACTGAGGGTTTCAGT GCTAAACTAGGCTGTCTTGGGCAACAGCATAGCTGGTCACCC ACACCCAGACCTGACCCAAACCCAGCTCCCTGCTTCTTGGCCACG TAACCTGAGAAAGGAATCCCTCTCTGAAACCCAGCCACCCCA ATGCTCCAGGCCTCTGGGATACCCGAAGAGTGAGTTTGCCAAGC AGTCACCCACAGTTGGAGGAGAATCCACCCAAAAGGCAGCCTGGT AGACAGGGCTGGGGTGGCCTCTCGTGGGGTCAGGCCAAGTAGGTG GCCTGGGGCTCTGGGGGATGCAGGGGAAGGGGATGCAGGGGA CGGGGATGCAGGGGAACGGGGCTCAGTCTGAAGAGCAGAGCCAGG AACCCCTGTAGGGAAGGGGACAGGAGCCAGGGGCATGAGATGGT GGACGAGGAAGGGGACAGGGAAGCCTGAGCGCTCTCTGGGCT TGCCAAGGACTCAAACCCAGAAGCCAGAGCAGGGCCTTAGGGAA GCGGGACCTGTCTGGGCGGAGGAATATGTCCAGATAGCACTGG GGACTCTTAAGGAAGAAGGATGGAGAAGAGAAAGGGAGTAGA GCGGGCCACGACCTGGTGAACACCTAGGACGCACCATTTCTCAGAAA GGGAGTTTTCACACGGACACCCCTCTCTACCCAGCCCTG
47	Donor DNA	GFP donor to AAVS1 locus	ATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGGTGGTGCCCATCC TGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTGAGCGTGT CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAA GTTCATCTGCACACCGCAAGCTGCCGTGCCCTGGCCACCCCTC GTGACCACTGACCTACGGCGTGAGTGCTTCAGCCGCTACCCCG ACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCGATGCCGGAAG CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAC AAGACCCGCGCGAGGTGAAGTTCGAGGGCGACACCTTGGTGAAC CGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATC CTGGGGCACAAGCTGGAGTACAATAACAAGCCACAACGTCTATA TCATGGCCGACAGCAGAAAGACGGCATCAAGGTGAATTCAGA TCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTA CCAGCAGAAACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGAC AACCCTACCTGAGCACCAGTCCGCCCCGAGCAAGACCCCAACG AGAAGCGCGATCACATGGTCTGCTGGAGTTCTGACCGCCGCCG GATCACTCTCGCATGGACGAGCTGTACAAGTAA
48	Donor DNA	AAVS1 locus RHA (used for GFP donor)	ACTGTGGGTGGAGGGGACAGATAAAAGTACCCAGAACCAGAGCC ACATTAAACCGGCCCTGGGAATATAAGGTGGTCCAGCTCGGGGACA CAGGATCCCTGGAGGCAGCAACATGCTGTCTGAAGTGAGCATAG GGGCCCGGTTGGAGGAAGAAGACTAGCTGAGCTCTCGGACCCCTG GAAGATGCCATGACAGGGGGCTGGAAGAGCTAGCACAGACTAGAG AGGTAAGGGGGTAGGGGAGTGCCCAATGAAGAGAGTGAGAGT TGACCCGAATCCACAGGAGAAGGGGTGTCAGGCAAGAAAGCA AGAGGATGGAGAGGTGGCTAAAGCCAGGGAGACGGGGTACTTTGG GGTTGTCCAGAAAAACGGTGATGATGCAGGCCCTACAAGAAGGGGA GGCGGGACGCAAGGGAGACATCCGTCGGAGAAGGCCATCTTAAGA AACGAGAGATGGCACAGGCCCCAGAAGGAGAAGGAAAAGGGAAAC CCAGCGAGTGAAGACGGCATGGGGTTGGGTGAGGGAGGAGAGATG CCCGAGAGGAGCCAGACACGGGGAGGATCCGCTCAGAGGACATC ACGTGGTGCAGCGCCGAGAAGGAAGTGCTCCGGAAGAGCATCCT TGGGCAGCAACACAGCAGAGAGCAAGGGGAAGGGGAGTGGAGG AAGACGGAACCTGAAGGAGGCGGCAGGGAAGGATCTGGGCCAGCC GTAGAGGTGACCCAGGCCACAAGCTGCAGACAGAAAGCGGCACAG GCCCAGGGGAGAGAATGCAGGTGAGAGAAAGCAGGACCTGCCTGG GAAGGGGAACAGTGGGCGAGGGCGGCAGAAAGCCAGTAGAGC

-continued

SEQ ID NO:	Identifi- fier	Name/ Description	Sequence
			TCAAAGTGGTCCGGACTCAGGAGAGAGACGGCAGCGTTAGAGGGC AGAGTTCCGGCGGCACAGCAAGGGCACTCGGGGCGAGAGGAGG CAGCGCAAAGTGACAATGGCCAGGGCCAGGCAGATAGACCAGACT GAGCTATGG
49	Donor DNA	HBB locus LHA (used for E7 to E7V AAV.304)	CTTGCTTTGACAATTTTGGTCTTTCAGAATACTATAAATATAACCTA TATTATAATTTTATAAAGTCTGTGCATTTTCTTTGACCCAGGATATT TGCAAAAAGACATATTCAAACTTCCGCAGAACACTTTATTTACATAT ACATGCCCTCTTATATCAGGGATGTGAAACAGGGTCTTGAAAACGT CTAAATCTAAAACAATGCTAATGCAGGTTTAAATTTAATAAATAA AATCCAAAATCTAACAGCCAAGTCAAATCTGCATGTTTAAACATTT AAAATATTTTAAAGACGTCTTTCCAGGATTCAACATGTGAAAATCT TTTCTCAGGGATACAGTGTGCCTAGATCCTCATTGCTTTAGTTTTTT ACAGAGGAATGAATATAAAAAGAAAATACTTAAATTTTATCCCTCT TACCTCTATAATCATACATAGGCATAATTTTAACTTAGGCTCCAG ATAGCCATAGAAGAACCACCACTTTCTGCGTGTGTGAGAATAATC AGAGTGAGATTTTTCACAAGTACCTGATGAGGGTTGAGACAGGTA GAAAAAGTGAGAGATCTCTATTATTAGCAATAATAGAGAAAGCA TTAAAGAGAATAAAGCAATGGAATAAGAAATTTGTAATTTCTCTT CTGATAACTAGAAATAGAGGATCCAGTTCTTTTGGTTAACCTAAAT TTTATTTTCAATTTATTTGTTTTATTTTATTTTATTTTATTTTGTG TAATCGTAGTTTTACAGAGTGTAGAGCTGAAAGGAAGAAGTAGGAG AAACATGCAAAGTAAAAGTATAACACTTTCTTACTAAACCGACAT GGGTTTCCAGGTAGGGGAGGATTGAGGATGACTGACAGGGCCCTT AGGGAACACTGAGACCTTACGCTGACCTCATAAATGCTTGTACCT TTGCTGTTTTAATTACATCTTTAATAGCAGGAAGCAGAATCTGCA CTTCAAAGTTTTTCTCTCACCTGAGGAGTTAATTTAGTACAAGGGG AAAAAAGTACAGGGGATGGGAGAAAGGCGATCACGTTGGGAAGCT ATAGAGAAAAGAAGTAAATTTAGTAAAGGAGGTTTAAACAAAC AAAATATAAAGAGAATAGGAACCTGAATCAAGGAAATGATTTTA AAACGCAGTATTCTTAGTGGACTAGAGGAAAAAATAATCTGAGCC AAGTAGAAGACCTTTTCCCTCCTACCCCTACTTTCTAAGTCACAGA GGCTTTTTGTTTCCCCCAGACACTCTTGCAGATTAGTCCAGGCAGAAA CAGTTAGATGTCCCAAGTTAACTCCTATTTGACACCACTGATTACC CCATTGATAGTCACACTTTGGGTTGTAAGTGACTTTTTATTTATTTGT ATTTTGTAGCTGATTAAGAGGCTCTAGTTTTTTATCTCTTGTTTCCC AAAACCTAATAAGTAACTAATGCACAGACACATTGATTGTATTATT ATTCTATTTTATAGACATAATTTATTAGCATGCATGAGCAAAATTAAGA AAAACAACAACAAATGAATGCATATATATGTATATGTATGTGTGTA TATATACACACATATATATATATATTTTTTCTTTTCTTACCAGAAGG TTTTAATCCAAATAAGGAGAAGATATGCTTAGAACCGAGGTAGAGT TTTTCATCCATTCTGTCTGTAAAGTATTTTGCATATCTGGAGACGCA GGAAGAGATCCATCTACATATCCCAAAGCTGAATTATGGTAGACAA AACTCTTCCACTTTTAGTGCATCAACTCTTATTGTGTATAAAGAA AATTGGGAAAACGATCTTCAATATGCTTACCAAGCTGTGATTCCAA ATATTACGTAAATACACTTGCAAGGAGGATGTTTTAGTAGCAAT TTGTACTGATGGTATGGGGCCAGAGATATATCTTAGAGGGAGGGC TGAGGGTTGAAGTCCAACCTCTAAGCCAGTGCCAGAAGAGCCAAG GACAGGTACGGCTGTCATCACTTAGACCTCACCTGTGGAGCCACA CCCTAGGGTTGGCCAAATCTACTCCAGGAGCAGGGAGGGCAGGAG CCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA TTTGCTTCTGACACAACTGTGTCTACTAGCAACCTCAAACAGACACC ATGGTGA
50	Donor DNA	E7 to E7V AAV.304	TCTGACTCCTGTCGAGAAGTCTGCAGTCACTGCTCTATGGGGGAAA
51	Donor DNA	HBB locus RHA (used for E7 to E7V AAV.304)	GTGAACGTGGATGAAGTTGGTGGTGAAGCCCTGGGCAGGTTGGTAT CAAGGTTACAAGACAGGTTTAAAGGAGACCAATAGAAACTGGGCAT GTGGAGACAGAGAAGACTCTTGGGTTTCTGATAGGCACCTGACTCTC TCTGCCATTGGTCTATTTTCCACCCTTAGGCTGCTGGTGGTCTAC CCTTGGACCCAGAGGTTCTTTGAGTCCCTTGGGGATCTGTCCACTCC TGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAA GTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCA AGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCA CGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGACGCTTGATGT TTTCTTTCCCTTCTTTTCTATGGTTAAGTTTATGTATAGGAAGGG GATAAGTAACAGGGTACAGTTTGAATGGGAAACAGACGAATGAT TGCATCAGTGTGGAAGTCTCAGGATCGTTTTAGTTTCTTTTATTTGTC TGTTTATAACAATTTGTTTCTTTTGTGTTAATCTTGTCTTCTTTTTTT TCTTCTCCGCAATTTTACTATTATACTTAATGCCTTAACATTGTGTA TAACAAAAGGAAATATCTCTGAGATACATTAAGTAACCTAAAAAA AACTTTACACAGTCTGCCTAGTACATTACTATTGGAATATATGTGT

-continued

SEQ ID NO:	Identifier	Name/ Description	Sequence
			<p>GCTTATTTGCATATTCATAATCTCCCTACTTTATTTCTTTTATTTTAA ATTGATACATAATCATTATACATATTTATGGGTTAAAGTGTAATGTT TTAATATGTGTACACATATTGACCAAATCAGGGTAATTTTGCAATTTG TAATTTTAAAAAATGCTTCTTCTTTTAAATATACTTTTGTGTTATCT TATTTCTAATACTTTCCCTAATCTCTTCTTTCAGGGCAATAATGATA CAATGTATCATGCCCTCTTGCACCATTCTAAAGAATAACAGTGATA ATTTCTGGGTTAAGCAATAGCAATATCTCTGCATATAAATATTTCT GCATATAAATTTGTAAGTGTGAAGAGGTTTCATATTGCTAATAGC AGCTACAATCCAGCTACCATTTCTGCTTTTATTTTATGGTTGGGATAA GGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTT CATACCTCTTATCTTCCCTCCACAGCTCCTGGGCAACGTGCTGGTCT GTGTGCTGGCCCATCACTTTGGCAAGAAATTCACCCACAGTGCA GGCTGCCATCAGAAAGTGGTGGCTGGTGGCTAATGCCCTGGCC CACAAGTATCACTAAGCTCGCTTTCTGCTGTCCAATTTCTATTAAA GGTTCCTTTGTTCCCTAAGTCCAACCTACTAACTGGGGGATATTATG AAGGCCCTTGAGCATCTGGATTCTGCCATAAAAAACATTTATTTT CATTGCAATGATGATTTTAAATATTTCTGAATATTTTACTAAAAAG GGAATGTGGGAGGTCAAGTGCATTTAAACATAAAGAAATGAAGAG CTAGTTCAAACCTTGGGAAAATACACTATATCTTAACTCCATGAA AGAAGGTGAGGCTGCAACAGCTAATGCACATTGGCAACAGCCCT GATGCATATGCCTTATTCATCCCTCAGAAAAGGATTCAGTAGAGG CTTGATTTGGAGGTTAAAGTTTGCTATGCTGTATTTTACATTACTT ATTGTTTTAGCTGTCTCATGAATGTCTTTTCACTACCCATTTGCTTA TCCTGCATCTCTCAGCCTTGACTCCACTCAGTCTCTTGTCTAGAGA TACCACCTTTCCCTGAAGTGTCTTCCATGTTTTACGGCGAGATG GTTTTCTCTCGCCTGGCCACTCAGCCTTAGTTGTCTCTGTTGCTTAT AGAGGCTTACTTGAAGAAGGAAAAACAGGGGTCATGGTTTGACTGT CCTGTGAGCCCTTCTCCCTGCCTCCCCACTCACAGTGACCCGGAA TCTGCAGTGCTAGTCTCCCGAATATCACTCTTTACAGTCTGCTT TGGAAGGACTGGGCTTAGTATGAAAAGTTAGGACTGAGAAGAATTT GAAAAGCGGCTTTTGTAGCTTGATATTCACACTGTCTTATTACCC TGTC</p>
52	Donor DNA	HBB locus LHA (used for E7 to E7V AAV.307)	<p>CTTGCTTTGACAATTTTGGTCTTTCAGAATACTATAAATATAACCTA TATTATAATTTTATAAAGTCTGTGCATTTTCTTTGACCCAGGATATT TGCAAAAGACATATTCAAACCTCCGCAGAACCTTTATTTACATAT ACATGCCCTCTTATATCAGGGATGTGAAACAGGGTCTTGAAAACGT CTAAATCTAAAACAATGCTAATGCAGGTTTAAATTTAATAAAATAA AATCCAAAATCTAACAGCCAAGTCAAATCTGCATGTTTAAACATT AAAATATTTTAAAGACGTCTTTTCCAGGATTCAACATGTGAAATCT TTTCTCAGGGATACACGTGTGCCTAGATCCTCATTTGCTTTAGTTTTT ACAGAGGAATGAATATAAAAGAAAATACTTAAATTTTATCCCTCT TACCTCTATAATCATACATAGGCATAATTTTTTAACCTAGGCTCCAG ATAGCCATAGAAGAACCAACACTTTCTGCGTGTGTGAGAATAATC AGAGTGAGATTTTTTACAAAGTACCTGATGAGGGTTGAGACAGGTA GAAAAGTGAGAGATCTCTATTTATTAGCAATAATAGAGAAAGCA TTTAAGAGAATAAAGCAATGGAAATAAGAAAATTTGTAATTTTCTT CTGATAACTAGAAAATAGAGGATCCAGTTTCTTTTGGTTAACTTAAAT TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT TAATCGTAGTTTCAGAGTGTAGAGCTGAAAGGAAGAAGTAGGAG AAACATGCAAGTAAAGTATAACACTTTCCTTACTAAACCGACAT GGGTTTCCAGGTAGGGGAGGATTGAGGATGACTGACAGGGCCCTT AGGGAACTGAGACCTACGCTGACCTCATAAATGCTTGCTACCT TTGCTGTTTTAATTACATCTTTAATAGCAGGAAGCAGAACTCTGCA CTTCAAAAGTTTTTCTCACCTGAGGAGTTAATTTAGTACAAGGGG AAAAAGTACAGGGGATGGGAGAAAGGCGATCACGTTGGGAAGCT ATAGAGAAAGAGAGTAAATTTTAGTAAGGAGGTTTAAACAAAC AAAATATAAAGAGAAATAGGAACCTGAATCAAGGAAATGATTTTA AAACGCAGTATCTTAGTGGACTAGAGGAAAAAATAATCTGAGCC AAGTAGAAGACCTTTTCCCTCCTACCCCTACTTTCTAAGTACAGAA GCTTTTGTGTTCCCGACACTCTTGACAGATTAGTCCAGGCAGAAA CAGTTAGATGTCCCGAGTTAACTCCTATTTGACACCACTGATTACC CCATTGATAGTCACTTTGGGTTGTAAGTGACTTTTATTTATTTGT ATTTTTGACTGCATTAAGAGGCTCTAGTTTTTATCTCTTGTGTTCC AAAACCTAATAAGTAACATAATGCACAGAGCACATTGATTGTATTT ATTCTATTTTTAGACATAATTTATTAGCATGCATGAGCAAAATAAGA AAAACAACAACAAATGAATGCATATATATGTATATGTATGTGTGTA TATATACACACATATATATATATTTTTTCTTTTCTTACCAGAGG TTTTAATCCAAATAAGGAGAAGATATGCTTAGAACCGAGGTAGAGT TTTCATCCATTCTGTCTGTAAGTATTTTGCATATTTGGAGACGCA GGAAGAGATCCATCTACATATCCCAAGCTGAATTTGGTAGACAA AACTCTTCCACTTTTAGTGCATCACTCTTATTTGTGTAATAAGAA AATTGGGAAAACGATCTTCAATATGCTTACCAAGCTGTGATTCCAA</p>

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
			ATATTACGTAAATACACTTGCAAAGGAGGATGTTTTAGTAGCAAT TTGTACTGATGGTATGGGGCCAAAGAGATATATCTTAGAGGGAGGGC TGAGGGTTTGAAGTCCAACCTCCTAAGCCAGTGCCAGAAAGAGCCAAG GACAGGTACGGCTGTCATCACTTAGACCTCACCCCTGTGGAGCCACA CCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAG CCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA TTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAAACAGACACC ATGGTGCA
53	Donor DNA	E7 to E7V AAV.307	TCTGACTCCTGTGCAAAAATCCGCTGTCAACGCCCTCTGGGGCAAG
54	Donor DNA	HBB locus RHA (used for E7 to E7V AAV.307)	GTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGTTGGTAT CAAGGTTACAAGACAGGTTTAAAGGAGACCAATAGAAAAGTGGGCAT GTGGAGACAGAGAAGACTCTGGGTTTCTGATAGGCACTGACTCTC TCTGCCTATTGGTCTATTTCCCAACCTTAGGCTGCTGGTGGTCTAC CCTTGGACCCAGAGGTTCTTTGAGTCCCTTGGGGATCTGTCCACTCC TGATGCTGTTATGGGCAACCTAAGGTGAAGGCTCATGGCAAGAAA TGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCA AGGGCACCTTTGCCACACTGAGTGAGCTGCACCTGTGACAAGCTGCA CGTGATCCTGAGAAGTTCAGGGTGAGTCTATGGGACGCTTGATGT TTTCTTTCCCTTCTTTCTATGGTTAAGTTCATGTCATAGGAAGGG GATAAGTAACAGGGTACAGTTTGAATGGGAAACAGACGAATGAT TGCATCAGTGTGGAAGTCTCAGGATCGTTTTAGTTTCTTTTATTGTC TGTTCAACAATTTGTTTCTTTTGTTTAATTCTTGCTTCTTTTCTTTT TCTTCTCCGCAATTTTACTATTATACTTAATGCCTTAACATTGTGTA TAACAAAAGGAAATATCTCTGAGATACATTAAGTAACATAAAAAA AACTTTACACAGTCTGCCTAGTACATTACTATTGGAATATATGTGT GCTTATTTGCATATTCATAATCTCCCTACTTTATTTTCTTTTATTTT ATTGATACATAATCATTATACATATTTATGGGTAAAGTGTAATGTT TTAATATGTGTACACATATTGACCAATCAGGGTAATTTTGCAATTG TAATTTTAAAAAATGCTTCTTCTTTTAAATATACTTTTGTGTTATCT TATTTCTAATACTTTCCCTAATCTCTTCTTTCAGGGCAATATGATA CAATGTATCATGCCTCTTGCACCATTTCAAAGAATAACAGTGATA ATTTCTGGGTAAAGGCAATAGCAATATCTCTGCATATAAATATTTCT GCATATAAATTGTAAGTGTGTAAGAGGTTTCAATTTGCTAATAGC AGCTACAATCCAGCTACCATCTGCTTTTATTTTATGTTGGGATAA GGCTGGATTATTTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTT CATACCTCTTATCTTCTCCACAGCTCCTGGGCAACGTGCTGGTCT GTGTGCTGGCCATCACTTTGGCAAGAATTCAACCCACAGTGCA GGCTGCCTATCAGAAAGTGGTGGCTGGTGGCTAATGCCCTGGCC CACAAGTATCACTAAGCTCGCTTTCTGCTGTCCAATTTCTATTAAA GGTTCTTTGTTCCCTAAGTCCAACCTACTAACTGGGGATATTATG AAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTT CATTGCAATGATGATTTAAATATTTCTGAATATTTTACTAAAAAG GGAATGTGGGAGTCAAGTGCATTTAAACATAAAGAAATGAAGAG CTAGTTCAAACCTTGGGAAAATACACTATATCTTAACTCCATGAA AGAAGGTGAGGCTGCAACAGCTAATGCACATTTGGCAACAGCCCT GATGCATATGCCTTATTCATCCCTCAGAAAAGGATTCAAGTAGAGG CTTGATTTGGAGGTTAAAGTTTGTCTATGCTGATTTTACATTACTT ATTGTTTGTAGCTGCTCCTCATGAATGCTTTTCACTACCCATTGCTTA TCCTGCATCTCTCAGCCTTGACTCCACTCAGTTCTCTTGCTTAGAGA TACCACCTTTCCCTGAAGTGTTCCTCCATGTTTACGGCGAGATG GTTTCTCTCGCTGGCCACTCAGCCTTAGTTGTCTCTGTTGTCTTAT AGAGGTCCTTGAAGAAGGAAAAACAGGGGTCATGGTTTGACTGT CCTGTGAGCCCTTCTCCCTGCCTCCCCACTCACAGTGACCCGGAA TCTGCAGTGCTAGTCTCCCGAAGTATCACTCTTTCACAGTCTGCTT TGAAGGACTGGGCTTAGTATGAAAAGTTAGGACTGAGAAGAATT GAAAGGCGGCTTTTGTAGCTTGATATTCACACTGTCTTATTACCC TGTC
55	Donor DNA	HBB locus LHA (used for GFP AAV)	GACTGCATTAAGAGGTCTCTAGTTTTTACCTCTTGTTTCCCAAAAC CTAATAAGTAAGTAAATGCACAGAGCACATTGATTTGTATTTATTCTA TTTTATGACATAATTTATTAGCATGCATGAGCAAATTAAGAAAAAC AACAACAAATGAATGCATATATATGTATATGTATGTGTGTACATAT ACACATATATATATATTTTTTCTTTTCTTACCAGAGGTTTAAATC CAAATAAGGAGAAGATATGCTTAGAAGTGAAGTAGAGTTTTCATCC ATTCTGTCTGTAAAGTATTTTGCATATTTCTGGAGACGAGGAAGAG ATCCATCTACATATCCCAAGCTGAATTATGGTAGACAAAACCTCTT CCACCTTTAGTGCATCAATTTCTTATTTGTGTAATAAGAAAATTGGG AAAACGATCTTCAATATGCTTACCAAGCTGTGATTCCAATATTAC GTAATAACACTTGCAAAGGAGGATGTTTTAGTAGCAATTTGTACT GATGATATGGGGCCAAAGAGATATCTTAGAGGGAGGGCTGAGGG

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
			TTTGAAGTCCAACCTCCTAAGCCAGTGCCAGAAGAGCCAAGGACAGG TACGGCTGTCACTACTTAGACCTCACCCTGTGGAGCCACACCTAG GGTTGGCCAATCTACTCCAGGAGCAGGGAGGGCAGGAGCCAGGG CTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTGCT TCTGACACAACCTGTGTTCACTAGCAACCTCAAACAGACACCATGGT GCATCTGACTCCTGAGGA
56	Donor DNA	GFP	ATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGGGTGGTGCCCATCC TGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCGTAA GTTTCATCTGCACACCGGCAAGCTGCCGTGCCCTGGCCACCCCTC GTGACCACCTGACCTACGGCGTGCACTGCTTCAGCCGCTACCCCG ACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAG CTACGTCCAGGAGCGCACCCTCTTCTTCAAGGACGACGGCAACTAC AAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAAC CGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATC CTGGGGCACAAGCTGGAGTACAACCTACAACAGCCACAACGTCTATA TCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGA TCCGCCACAACATCGAGGACGGCAGCGTGCACTCGCCGACCACTA CCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGAC AACCCTACCTGAGCACCAGTCCGCCCTGAGCAAAGACCCCAACG AGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCG GATCACTCTCGGCATGGACGAGCTGTACAAGTAA
57	Donor DNA	HBB locus RHA (used for GFP AAV)	CTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAAGC CCTGGGCAGGTGGTATCAAGGTTACAAGACAGGTTTAAGGAGACC AATAGAACTGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCTG ATAGGCACTGACTCTCTGCTTATGGTCTATTTTCCACCCTTAG GCTGCTGGTGGTCTACCTTGGACCCAGAGGTTCTTTGAGTCTTTG GGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCTAAGGTGA GGCTCATGGCAAGAAAGTGCTCGGTGCCCTTAGTGATGGCCTGGCT CACCTGGACAACCTCAAGGGCACCTTTGCCCACTGAGTGAGCTGC ACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCT ATGGGACGCTTGATGTTTCTTTCCCTTCTTTTCTATGGTTAAGTTC ATGTATAGGAAGGGGATAAGTAACAGGGTACAGTTTAGAATGGG AAACAGACGAATGATGTCATCAGTGTGAAGTCTCAGGATCGTTTT AGTTTCTTTTATTTGCTGTTTATAACAATGTTTCTTTTGTGTTAAT CTTGCTTTCTTTTCTTCTCTCCGCAATTTTACTATTATACTTAAT GCCTTAACATTGTGTATAACAAAAGGAATATCTCTGAGATACATT AAGTAACCTAAAAAAACTTTACACAGTCTGCCTAGTACATTACT ATTTGGAATATATGTGTGCTTATTTGCATATTCATAATCTCCCTACTT TATTTTCTTTTATTTTAAATGATACATAATCATTATACATATTATG GGTTAAAGTGTAATGTTTAAATATGTGTACACATATTGACCAAATCA GGGTAATTTTGCAATTTGTAATTTTAAAAAATGC
58		MND promoter	GGCCGCCAGTGTGATGGATATCTGCAGAATTGCCCCTTATGGGGAT CCGAACAGAGAGACAGCAGAATATGGGCCAACAGGATATCTGTG GTAAGCAGTTCTGCCCCGGCTCAGGGCCAGAACAGTTGGAACAG CAGAATATGGGCCAACAGGATATCTGTGGTAAGCAGTTCTTGCC CGGCTCAGGGCCAGAACAGATGGTCCCAGATGCGGTCCCCGCCCT CAGCAGTTCTAGAGAACCATCAGATGTTTCAGGGTGCCCCAAGG ACCTGAAATGACCCTGTGCCTTATTGAACTAACCAATCAGTTCGCT TCTCGCTTCTGTTGCGCGCTTCTGCTCCCCGAGCTCTATATAAGCA GAGCTCGTTTAGTGAACCGTCAGATCGCTGGAGACGCCATCCAG CTGTTTGGACCTCCATAGAAGACACCGACTCTAGAG
59		ER a promoter	GGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCC CCGAGAAGTTGGGGGGAGGGTTCGGCAATTGAACCGGTGCCTAGA GAAGGTGGCGCGGGGTAACCTGGGAAAGTGATGTCGTGTACTGGCT CCGCCCTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCACTA TTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACAC AGGTAAGTGCCGTGTGTGGTTCCCGCGGGCTGGCCTCTTTACGGG TTATGGCCCTTGCGTGCCCTTGAATTACTTCCACTGGCTGCAGTACGT GATTCCTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCC AGGCCCTTGCGCTTAAGGAGCCCTTCGCCCTCGTGCTTGAGTTGAG CCTGGCCTGGGCGCTGGGGCCGCGCGTGCGAATCTGGTGGCACCT TCGCGCCTGTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAAT TTTGATGACCTGCTGCGACGCTTTTTCGCAAGATAGTCTTGTA AATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCG CGGGCGGCGACGGGGCCCGTGCCTCCAGCGCACATGTTCCGCGAG CGGGGGCTGCGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTC TCAAGCTGGCCGGCTGCTCTGGTGCCTGGCCTCGCGCCGCGGT ATCGCCCCGCTTGGGCGCAAGGCTGGCCCGGTGCGCACAGTTG

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
			CGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTC AAAATGGAGGACGCGGCGCTCGGAGAGCGGGCGGGTGAGTCACC CACACAAAGGAAAAGGCGCTTCCGCTCTCAGCCGTCGCTTCATGT GACTCCACGGAGTACCGGGCGCGTCCAGGCACCTCGATTAGTTCT CGAGCTTTTGGAGTACGTCGCTTTAGGTTGGGGGGAGGGGTTTTAT GCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGAAGTTAGGC CAGCTTGGCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTGAG TTTGATCTTGTTTATTCTCAAGCCTCAGACAGTGGTTCAAAGTTT TTTTCTTCATTTCAGGTGTCGTGA
60		SFFV promoter	GTAACGCCATTTTGCAAGGCATGGAAAAATACCAACCAAGAATA GAGAAGTTCAGATCAAGGCGGGTACATGAAATAGCTAACGTTG GGCCAAACAGGATATCTGCGGTGAGCAGTTTCGGCCCCGGCCGGG GCCAAGAACAGATGGTCACCGCAGTTTCGGCCCCGGCCGAGGCCA AGAACAGATGGTCCCCAGATATGGCCCAACCTCAGCAGTTTCTTA AGACCATCAGATGTTTCCAGGCTCCCCAAGGACCTGAATGACC CTGCGCCTTATTTGAATTAACCAATCAGCCTGCTTCTCGCTTCTGT CGCGCGCTTCTGCTTCCGAGCTCTATAAAGAGCTCACAACCCCTC ACTCGGCGCGCAGTCTCTCCGACAGACTGAGTCG
61		2A peptide from porcine tescho virus	GCCACGAACCTTCTCTGTAAAGCAAGCAGGAGACGTGGAAGAA AACCCCGGTCCT
62		Synthetic poly(A) signal	AATAAAATCGCTATCCATCGAAGATGGATGTGTGTGGTTTTTGT GTG
63	PAM	Canonical PAM	N _x NRG (N = any nucleotide; R = A or G; x = 19-21)
64	PAM	SpCas9 PAM	NRG (N = any nucleotide, R = A or G)
65	Nuclear locali- zation signal (NLS)	SV40 NLS 1	PKKKRKV
66	NLS	SV40 NLS 2	PKKKRRV
67	NLS	Nucleoplasmin NLS	KRPAATKKAGQAKKKK
68	I53	i53 (DNA)	ATGCTGATCTTCGTGAAGACCTGACCGCAAGACCATCACCTGG AGGTGGAGCCAGCGACACCATCGAGAACGTGAAGGCCAAGATCC AGGACAAGGAGGCGATCCCCCGACAGCAGAGGCTGGCCTTCG CCGGCAAGAGCCTGGAGGACGGCAGGACCTGAGCGACTACAACA TCCTGAAGGACAGCAAGCTGCACCCCTGCTGAGGCTGAGGTGA
69	I53 mRNA	i53 (RNA)	AUGCUGAUCUUCUGAAGACCCUGACCGCAAGACCAUACCCUG GAGGUGGAGCCAGCGACACCAUCGAGAACGUGAAGGCCAAGAU CCAGGACAAGGAGGGCAUCCCCCGACAGCAGAGGCUGGCCUU CGCCGGCAAGAGCCUGGAGGACGGCAGGACCCUGAGCGACUACAA CAUCCUGAAGGACAGCAAGCUGACCCCCUGCUGAGGCUAGGUG A
70	I53	i53 (aa)	MLIFVKLTLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLAFAGKSL EDGRTLSDYNIILKDSKLHPLLRLR
71	DM mRNA	I53-DM (RNA)	AUGUUGAUUUUCUGUAAAACCCUACCGGAAAACCAUACCCUC GAGGUUGAACCCUCGGAUACGAUAGAAAUGUAAAGGCCAAGAU CCAGGAUAAGGAAGGAUUUCCUCCUGAUCAGCAGAGACUGGCCU UUGCUGGCAAAUCGCGUGGAAGAUGGACGUACUUUGUCUGACUAC AAUUAUUCUAAAGGACUCUAAACUUCUAGUGUUGAGACUUCG U
72		A10 (DNA)	ATGCAGATTTCAGTGAAGACCTTGGCCGGAAGCCATCACCTCG AGGTTGAACCTCGGATACGATAGAAAATGTAAAGGCCAAGATCC AGGATAAGGAAGGAATTCCTCCTGATCAGCAGCGACTGATCTTGC TGAAATGCGGCTGGAAGATGGACGTACTTTGTCTGACTACAATATT AAAAACGACTCTACTCTTTTCTTGTGTGAAAAATAGTGTACT

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
73	A10	(RNA)	AUGCAGAUUUACGUGAAGACCUUUGCCCGGAAGCCCAUCACCCUC GAGGUUGAACCCUCGGAUACGAUAGAAAUGUAAAGGCCAAGAU CCAGGAUAAGGAAGGAUUCUCCUGAUCAGCAGCAGACUGAUCU UUGCUGAAAUGCGCUGGAAGAUGGACGUACUUUGUCUGACUAC AAUAUUAAAACGACUCUACUCUUUUUCUUGUGUUGAAAAUAG UGUUACU
74	A10	(aa)	MQIYVKTFARKPITLEVEPSDTIENVKAKIQDKEGIPPDQQRILFAEMRL EDGRTLSDYNIKNDSTLFLVLKNSVT
75	A11	(DNA)	ATGCTGATTTTCGTGACCACCGATATGGGGATGACAATCTCACTCG AGGTTGAACCCCTCGGATACGATAGAAAATGTAAAGGCCAAGATCC AGGATAAGGAAGGAATTCCTCCTGATCAGCAGAGACTGATCTTTGG TGACAAGGATCTGGAAGATGGACGTACTTTGTCTGACTACAATATT CAAAGGAGTCTAGCCTTAATCTTGTGCTGAAACTTCGTGGTGGT
76	A11	(RNA)	AUGCUGAUUUUCGUGACCACCGAUUUGGGGAUGACAAUCUCACU CGAGGUUGAACCCUCGGAUACGAUAGAAAUGUAAAGGCCAAGA UCCAGGAUAAGGAAGGAUUCUCCUGAUCAGCAGAGACUGAUC UUUGGUGACAAGGAUCUGGAAGAUGGACGUACUUUGUCUGACUA CAAUUUCAAAGGAGUCUAGCCUUAUUCUUGUGUGAAACUUC GUGGUGGU
77	A11	(aa)	MLIFVTTDMGMTISLEVEPSDTIENVKAKIQDKEGIPPDQQRILFGDKDL EDGRTLSDYNIQKESLNLVLKLRGG
78	C08	(DNA)	ATGCAGATTTTCGTGACCACCGATATGTGGATGAGAATCTCACTCG AGGTTGAACCCCTCGGATACGATAGAAAATGTAAAGGCCAAGATCC AGGATAAGGAAGGAATTCCTCCTGATCAGCAGAGACTGATCTTTGG TGACAAGGATCTGGAAGATGGACGTACTTTGTCTGACTACAATATT CAAAGGAGTCTAGCCTTAATCTTGTGCTGAACCTTCGTGGTGGT
79	C08	(RNA)	AUGCAGAUUUUCGUGACCACCGAUUUGGGAUGAGAAUCUCACU CGAGGUUGAACCCUCGGAUACGAUAGAAAUGUAAAGGCCAAGA UCCAGGAUAAGGAAGGAUUCUCCUGAUCAGCAGAGACUGAUC UUUGGUGACAAGGAUCUGGAAGAUGGACGUACUUUGUCUGACUA CAAUUUCAAAGGAGUCUAGCCUUAUUCUUGUGUGAACCUUC GUGGUGGU
80	C08	(aa)	MQIFVTTDMWMRISLEVEPSDTIENVKAKIQDKEGIPPDQQRILFGDKD LEDGRTLSDYNIQKESLNLVLNLRGG
81	G08	(DNA)	ATGTTGATTTTCGTGAAAACCTTACCGGGAACCATCACCTCG AGGTTGAACCCCTCGGATACGATAGAAAATGTAAAGGCCAAGATCC AGGATAAGGAAGGAATTCCTCCTGATCAGCAGAGACTGATCTTTGC TGGCAAATCGCTGGAAGATGGACGTACTTTGTCTGACTACAATATT CTAAAGGACTCTAACTTCATCCTCTGTTGAGACTTCGTGGTGGT
82	G08	(RNA)	AUGUUGAUUUUCGUGAAAACCUUACCGGGAACCAUCACCCUC GAGGUUGAACCCUCGGAUACGAUAGAAAUGUAAAGGCCAAGAU CCAGGAUAAGGAAGGAUUCUCCUGAUCAGCAGAGACUGAUCU UUGCUGGCAAAUCGUGGAAGAUGGACGUACUUUGUCUGACUAC AAUAUUCAAAGGACUUAACUUAUCCUCUGUGAGACUUCG UGGUGGU
83	G08	(aa)	MLIFVKTLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQQRILFAGKSLE DGRRTLSDYNIKDSKLHPLLRGG
84	H04	(DNA)	ATGCGAATTATCGTGAACCTTTATGCGGAAGCCGATCACGCTCG AGGTTGAACCCCTCGGATACGATAGAAAATGTAAAGGCCAAGATCC AGGATAAGGAAGGAATTCCTCCTGATCAGCAGAGACTGTATTTGC GGCCAGTCAGCTGGAAGATGGACGTACTTTGTCTGACTACAATATT CAAAGGAGTCTACTCTTCTTCTTGTGTTAAGGCTGCTCCGCGTT
85	H04	(RNA)	AUGCAGAUUUUCGUGAAAACCUUUGCGGAAGCCGAUCACGCU CGAGGUUGAACCCUCGGAUACGAUAGAAAUGUAAAGGCCAAGA UCCAGGAUAAGGAAGGAUUCUCCUGAUCAGCAGAGACUGUAU UUUGCGCCAGUCAGCUGGAAGAUGGACGUACUUUGUCUGACUA CAAUUUCAAAGGAGUCUACUCUUCUUCUUGUGUAAGGCUGC UCCGCGUU
86	H04	(aa)	MRIIVKTFMRKPITLEVEPSDTIENVKAKIQDKEGIPPDQQRILYFAASQL EDGRTLSDYNIQKESLTLVVRLLRV

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
87		I53 alt (DNA)	ATGTTGATTTTCGTGAAAACCCCTTACCGGAAAACCATCACCCCTCG AGGTTGAACCCCTCGGATACGATAGAAAATGTAAAGGCCAAGATCC AGGATAAGGAAGGAATTCCTCCTGATCAGCAGAGACTGGCCTTTGC TGGCAAATCGCTGGAAGATGGACGTACTTTGTCTGACTACAATATT CTAAAGGACTCTAAACTTCATCCTCTGTTGAGACTTCGT
88		I53 alt (RNA)	AUGUUGAUUUUCUGUAAAACCCUUAACCGGAAAACCAUCACCCUC GAGGUUGAACCCUCGGAUACGAUAGAAAUGUAAAGGCCAAGAU CCAGGAUAAGGAAGGAUUUCUCCUGAUCAGCAGAGACUGGCCU UUGCUGGCAAAUCGCUUGGAAGAUGGACGUACUUUGUCUGACUAC AAUAUUCUAAAGGACUCUAAACUUCUCCUGUUGAGACUUCG U
89		FLAG-tagged i53 DNA	ATGGACTACAAAGACGATGACGATAAAGCCGCCAGTTTAAACGGC GCGCCATTAAATTAAGGATCCAATGTTGATTTTCGTGAAAACCCCTTAC CGGAAAACCATCACCCCTCGAGGTTGAACCCCTCGGATACGATAGAA AATGTAAGGCCAAGATCCAGGATAAGGAAGGAATTCCTCCTGATC AGCAGAGACTGGCCTTTGCTGGCAAATCGCTGGAAGATGGACGTAC TTTGTCTGACTACAATATTCTAAAGGACTCTAACTTCATCCTCTGT TGAGACTTCGTTGA
90		FLAG-tagged i53 RNA	AUGGACUACAAAGACGAUGACGAUAAAGCCGCCAGUUUAAACGG CGCGCCAUUAAUUAAGGAUCCAAUGUUGAUUUUCGUGAAAACCC UUACCGGAAAACCAUCACCUCGAGGUUGAACCCUCGGAUACGA UAGAAAUGUAAAGGCCAAGAUCAGGAUAAGGAAGGAUUCCU CCUGAUCAGCAGAGACUGGCCUUUGCUGGCAAAUCGCUUGGAAGA UGGACGUACUUUGUCUGACUACAAUAUUCUAAAGGACUCUAAAC UUCAUCCUCUGUUGAGACUUCGUUGA
91		Linker (DNA)	GCCGCCAGTTTAAACGGCGCGCCATTAAATTAAGGATCCA
92		Linker (RNA)	GCCGCCAGUUUAAACGGCGCGCCAUUAAUUAAGGAUCCA
93		Linker	AASLNGAPLIKDP
94	Protein tag	6xHis	HHHHHH
95	Protein tag	Flag	MDYKDDDDK
96	Protein tag	FLAG (DNA)	GACTACAAAGACGATGACGATAAA
97	Protein tag	FLAG (RNA)	GACUACAAAGACGAUGACGAUAAA
98	Donor DNA	HBB locus LHA to RHA (used for E6V to E6, AAV.323)	CTTGCTTTTGACAATTTTGGTCTTTCAGAATACTATAAATATAACCTA TATTATAATTTTCATAAAGTCTGTGCATTTTCTTTGACCCAGGATATT TGCAAAAGACATATTCAAACCTCCGCAGAACCTTTATTTCACATAT ACATGCCCTCTTATATCAGGGATGTGAAACAGGGTCTTGAAAACGT CTAAATCTAAAACAATGCTAATGCAGGTTTAAATTTAATAAAATAA AATCCAAAATCTAACAGCCAAGTCAAATCTGCATGTTTAAACATTT AAAATATTTTAAAGACGTCTTTTCCAGGATTCAACATGTGAAATCT TTTCTCAGGGATACACGTGTGCCTAGATCCTCATTGCTTTAGTTTTTT ACAGAGGAATGAATATAAAAAGAAAATACTTAAATTTTATCCCTCT TACCTCTATAATCATACATAGGCATAATTTTTTAACCTAGGCTCCAG ATAGCCATAGAAGAACCAACACTTTCTGCGTGTGTGAGAATAATC AGAGTGAGATTTTTTCAAGTACCTGATGAGGGTTGAGACAGGTA GAAAAGTGAGAGATCTCTATTATTAGCAATAATAGAGAAAGCA TTTAAGAGAATAAAGCAATGGAAATAAGAAATTTGTAAATTTCCCTT CTGATAACTAGAAATAGAGGATCCAGTTTCTTTTGGTTAACTAAAT TTTATTTTCATTTTATTGTTTTATTTTATTTTATTTTATTTTGTG TAATCGTAGTTTCAGAGTGTAGAGCTGAAAGGAAGAAGTAGGAG AAACATGCAAGTAAAAGTATAACACTTTCTTACTAAACCGACAT GGGTTTCCAGGTAGGGGAGGATTGAGGATGACTGACAGGGCCCTT AGGGAACACTGAGACCTTACGCTGACCTCATAAATGCTTGCTACCT TTGCTGTTTTAATTACATCTTTTAATAGCAGGAAGCAGAACTCTGCA CTTCAAAGTTTTTCTCTACCTGAGGAGTTAATTTAGTACAAGGGG AAAAAGTACAGGGGATGGGAGAAAGGCGATCACGTTGGGAAGCT ATAGAGAAAGAGAGTAATTTTAGTAAGGAGGTTTAAACAAAC AAAATATAAGAGAAATAGGAACCTGAATCAAGGAATGATTTTA

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
			AAACGCAGTATTCTTAGTGGACTAGAGGAAAAAATAATCTGAGCC AAGTAGAAGACCTTTTCCCTCCTACCCCTACTTTCTAAGTACAGAG GGCTTTTGTTCCTCCAGACACTCTTGACAGATTAGTCCAGGCAGAAA CAGTTAGATGTCCTCCAGTTAACCCTCTATTGACACCACTGATTACC CCATTGATAGTCACACTTTGGGTTGTAAGTGACTTTTATTATTGTTG ATTTTGTAGTCATTAAGAGGTCTCTAGTTTTTATCTCTGTTTCC AAAACCTAATAAGTAACATAATGCACAGAGCATTGATTGTATT ATTCTATTTTACACATAATTTATTAGCATGCATGAGCAAAATTAAGA AAAACAACAACAAATGAATGCATATATATGTATATGTATGTGTGTA TATATACACACATATATATATATATTTTCTTTTCTTACCAGAAGG TTTTAATCCAAATAAGGAGAAGATATGCTTAGAACCGAGGTAGAGT TTTCATCCATTCTGTCTGTAAAGTATTTTGCATATCTGGAGACGCA GGAAAGAGATCCATCTACATATCCCAAGCTGAATTATGGTAGACAA AACTCTTCCACTTTTAGTGCATCAACTCTTATTGTGTAATAAGAA AATTGGGAAAAACGATCTTCAATATGCTTACCAAGCTGTGATTCCAA ATATTACGTAAATACACTTGCAAGGAGGATGTTTTAGTAGCAAT TTGTACTGATGGTATGGGGCCAAGAGATATATCTTAGAGGGAGGGC TGAGGGTTGAAGTCCAACCTCTAAGCCAGTGCCAGAAGAGCCAAG GACAGGTACGGCTGTCATCACTAGACCTCACCTGTGGAGCCACA CCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAG CCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA TTTGCTTCTGACACAACCTGTGTTCACTAGCAACCTCAACACAGACC ATGGTGCATCTGACTCCTGAAGAAAAATCCGCTGTCACTGCCCTGT GGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCA GGTTGGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAA ACTGGGCATGTGGAGACAGAGAAGACTCTGGGTTTCTGATAGGCA CTGACTCTCTCTGCCTATTGGTCTATTTCCACCCCTAGGCTGCTGG TGGTCTACCCTTGGACCCAGAGGTTCTTTAGTCCCTTGGGGATCTG TCCACTCCTGATGCTGTTATGGGCAACCTAAGGTGAAGGCTCATG GCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCTGGCTCACCTGGA CAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGAC AAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGAGC CTTGATGTTTTCTTTCCCTCTTTTCTATGGTTAAGTTCATGTCATA GGAAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAACAGAC GAATGATTGCATCAGTGTGAAGTCTCAGGATCGTTTTAGTTCTTT TATTTGCTGTTTATAACAATTGTTTTCTTTGTTTAAATCTTGCTTCT TTTTTTTCTTCTCCGAATTTTACTATTATACTAATGCCTTAACA TTGTGTATAACAAAAGGAATATCTCTGAGATACATTAAGTAACTT AAAAAAAACCTTACACAGTCTGCCTAGTACATTACTATTGGAAT ATATGTGTGCTTATTTGCATATTCATAATCTCCCTACTTTATTTTCTT TTATTTTAAATGTATACATAATCATTATACATATTATGGGTTAAAG TGTAATGTTTTAATATGTGTACACATATTGACCAATCAGGGTAATT TTGCATTGTAAATTTAAAAAATGCTTTCTTCTTTAATATACTTTT TGTTTTATCTTATTCTAATACTTTCCCTAATCTCTTCTTTCAGGGCA ATAATGATACAATGTATCATGCCTCTTTGCACCATTTCAAAGAATAA CAGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCTGCATATA AATATTTCTGCATATAAATTGTAACCTGATGTAAGAGGTTTCATATTG CTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTATGGT TGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTA ATCATGTTTCATACCTCTTATCTTCTCCACAGCTCCTGGGCAACGT GCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAGAATTCACCCAC CAGTGCAGGCTGCCATCAGAAAGTGGTGGCTGGTGTGGCTAATGC CCTGGCCCAAGATATCACTAAGCTCGCTTTCTTGCTGTCCAATTT TATTAAAGGTTCTTTGTTCCCTAAGTCCAACCTACTAACTGGGGGA TATTATGAAGGGCTTTGAGCATCTGGATTCTGCCTAATAAAAAACA TTTATTTTCAATTGCAATGATGATTTAAATTTTCTGAATATTTTAC TAAAAAGGGAATGTGGAGGTCAGTGCATTTAAACATAAAGAAA TGAAGAGCTAGTTCAAACCTGGGAAAAATACACTATATCTTAACT CCATGAAGAAGGTGAGGCTGCAACAGCTAATGCACATTGGCAA CAGCCCCGTGATGCATATGCCCTTATTCATCCCTCAGAAAAGGATTCA AGTAGAGGCTTGATTGGAGGTTAAAGTTTGTCTATGCTGTATTTTA CATTACTTATTGTTTTAGCTGTCCCTCATGAATGCTTTTCACTACCCA TATTGTTATCCTGCATCTCTCAGCCTTGACTCCACTCAGTTCTCTTG TTAGAGATACCACCTTTCCCTGAAGTGTCTTCCATGTTTTACGG CGAGATGGTTTCTCCTCGCTGGCCACTCAGCCTTAGTTGTCTCTGT TGTCTTATAGAGGCTACTTTGAAGAAGGAAAAACAGGGGTCATGGT TTGACTGTCCTGTGAGCCCTTCTTCCCTGCCCTCCCCACTCACAGTG ACCCGGAATCTGAGTGCTAGTCTCCCGAATATCACTCTTTCACA GTCTGCTTTGAAGGACTGGGCTTAGTATGAAAAGTTAGGACTGAG AAGAATTTGAAGGCGGCTTTTGTAGCTTGATATTCACACTACTGTCT TATTACCCTGTC

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
99	Donor DNA	HBB locus LHA (used for E6V to E6, AAV.323)	CTTGCTTTGACAATTTTGGTCTTTCAGAATACTATAAAATATAACCTA TATTATAATTTTATAAAGTCTGTGCATTTTCTTTGACCCAGGATATT TGCAAAAGACATATTCAAACCTCCGCAGAACACTTTATTTCACATAT ACATGCCTCTTATATCAGGGATGTGAAACAGGGTCTTGAAAACGT CTAAATCTAAAACAATGCTAATGCAGGTTTAAATTTAATAAAATAA AATCCAAATCTAACAGCCAGTCAAATCTGCATGTTTAAACATT AAAATATTTTAAAGACGTCTTTCCAGGATTCAACATGTGAAATCT TTTCTCAGGGATACACGTGTGCCTAGATCCTCATTGCTTTAGTTTTT ACAGAGGAATGAATATAAAAAGAAATACTTAAATTTTATCCCTCT TACCTCTATAATCATACATAGGCATAATTTTAAACCTAGGCTCCAG ATAGCCATAGAAGAACCAACACTTTCTGCGTGTGTGAGAAATAATC AGAGTGAGATTTTTCACAAGTACCTGATGAGGGTTGAGACAGGTA GAAAAAGTGAGAGATCTCTATTATTAGCAATAATAGAGAAAGCA TTTAAGAGAAATAAGCAATGGAAATAAGAAATTTGTAATTTTCCTT CTGATACTAGAAATAGAGGATCCAGTTTCTTTTGGTTAACCTAAAT TTTATTTTCATTTTATTGTTTTATTTTATTTTATTTTATTTTGTG TAATCGTAGTTTCAGAGTGTAGAGCTGAAAGGAAGAAGTAGGAG AAACATGCAAGTAAAAGTATAACACTTTCCTACTAAACCGACAT GGGTTTCCAGGTAGGGGCAGGATTGAGGATGACTGACAGGGCCCTT AGGGAACACTGAGACCTACGCTGACCTCATAAATGCTTGCTACCT TTGCTGTTTAAATTACATCTTTTAAAGCAGGAAGCAGAACTCTGCA CTTCAAAAGTTTTTCTCACCTGAGGAGTTAATTTAGTACAAGGGG AAAAAGTACAGGGGATGGGAGAAAGGCATCACGTTGGGAAGCT ATAGAGAAAGAGAGTAAATTTAGTAAAGGAGGTTAAACAACAA AAAATATAAAGAGAAATAGGAACCTGAATCAAGGAAATGATTTTA AAACGCAGTATCTTCTAGTGGACTAGAGGAAAAAATAATCTGAGCC AAGTAGAAGACCTTTCCCTCCTACCCCTACTTTCTAAGTCACAGA GGCTTTTGTTCCTCCAGACACTCTTGACAGATTAGTCCAGGCAGAAA CAGTTAGATGTCCCAAGTAAACCTCCTATTTGACACCACTGATTACC CCATTGATAGTCACACTTTGGGTGTAGTGACTTTTATTATTGT ATTTTGTAGTGCATTAAAGAGTCTCTAGTTTTTATCTCTGTTTCCC AAAACCTAATAAGTAACTAATGCACAGAGCATTGATTGTATT ATTTCTATTTTAGACATAATTTATTAGCATGCATGAGCAAAATAGA AAAACAACAACAATGAATGCATATATATGTATATGTATGTGTGTA TATATACACATATATATATATATTTTTCTTTCTTACCAGAAAGG TTTTAATCCAAATAAGGAGAAGATATGCTTAGAACCGAGGTAGAGT TTTCATCCATTCTGCTCTGAAGTATTTTGCATATTCTGGAGACGCA GGAAGAGATCCATCTACATATCCCAAGCTGAATTATGGTAGACAA AACTCTCCACTTTTAGTGCATCAACTTCTTTTGTGTAATAAGAA AATTGGGAAAACGATCTCAATATGCTTACCAAGCTGTGATTCCAA ATATTACGTAATAACTGCAAGGAGGATGTTTTAGTAGCAAT TTGTACTGATGGTATGGGGCAGAGATATATCTTAGAGGGAGGGC TGAGGGTTTGAAGTCCAACCTCCTAAGCCAGTGCCAGAAAGCCAA GACAGGTACGGTGTCTCACTTAGACCTCACCTGTGGAGGCCACA CCCTAGGGTTGGCCAATCTACTCCAGGAGCAGGGAGGGCAGGAG CCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA TTTGCTTCTGACACAACCTGTCTCTAGCAACCTCAACAGACACC ATGTGTCATCTGACTCCT
100	Donor DNA	HBB locus RHA (used for E6V to E6, AAV.323)	ACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGG CCCTGGGCAGGTTGGTATCAAGGTTACAAGACAGGTTTAAAGGAGAC CAATAGAAACTGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCT GATAGGCACTGACTCTCTGCTCATTGGTCTATTTTCCACCCCTTA GGCTGCTGGTGGTCTACCTTGGACCCAGAGGTTCTTTGAGTCCTTT GGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGA AGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCTGGC TCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTG CACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTC TATGGGACGCTTGATGTTTTCTTTCCCTTCTTTCTATGGTTAAGTT CATGTCATAGGAAGGGGATAAGTAACAGGGTACAGTTTAGAATGG GAAACAGACGAATGATTGCATCAGTGTGGAAGTCTCAGGATCGTTT TAGTTTCTTTTATTTGCTGTTTATAACAATTTGTTTCTTTTGTAAAT CTTGCTTTCTTTTTTTCTTCTCCGCAATTTTACTATTATACTTAAT GCCTTAACATTGTGTATAACAAAAGGAATATCTCTGAGATACATT AAGTAACTTAAAAAAAACCTTACACAGTCTGCCTAGTACATTACT ATTTGGAATATATGTGTGCTTATTGTCATATTCATAATCTCCCTACTT TATTTTCTTTTATTTTAAATTGATACATAATCATTATACATATTATG GGTTAAAGTGTAATGTTTAAATATGTGTACACATATTGACCAATCA GGGTAATTTGCATTGTAAATTTAAAAAATGCTTCTTCTTTTAAATA TACTTTTTTGTATTATCTTATTCTAATACTTTCCCTAATCTCTTCTTT CAGGGCAATAATGATACAAATGATCATGCCTCTTGCACCAATCTAA AGAATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCT

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
			GCATATAAATATTCTGCATATAAATTGTAAGTATGTAAGAGGTTT CATATTGCTAATAGCAGCTACAATCCAGTACCATTCTGCTTTTATT TTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCC TTTTGCTAATCATGTTCATACCTCTTATCTTCTCCACAGCTCCTGG GCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATT CACCCACAGTGACAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTG GCTAATGCCCTGGCCACAAGTATCACTAAGCTCGCTTCTTGCTGT CCAATTTCTATTAAAGGTTCTTTGTTCCTAAGTCCAACACTACTAAA CTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAAT AAAAAACATTTATTTTCATTGCAATGATGATTAAATTATTCTGA ATATTTTACTAAAAAGGGAATGTGGGAGGTCAGTGCATTAAAAACA TAAAGAAATGAAGAGCTAGTTCAAACCTTGGGAAAATACACTATAT CTAAACTCCATGAAGAAGGTGAGGCTGCAACAGCTAATGCACA TTGGCAACAGCCCCGTGATGCATATGCCTTATTCATCCCTCAGAAAA GGATTCAAGTAGAGGCTTGATTGGAGGTTAAAGTTTGTCTATGCT GTATTTTACATTACTTATGTGTTTAGCTGCTCATGAATGTCTTTTC ACTACCCATTTGCTTATCCTGCATCTCTCAGCCTTGACTCCACTCAG TTCTCTTGCTTAGAGATACCACCTTCCCTGAAGTGTCTCTTCCAT GTATTTACGGCGAGATGGTTTCTCCTCGCCTGGCCACTCAGCCTTAGT TGTCTCTGTGCTTATAGAGGTCTACTGAAGAAGGAAAAACAGG GGTCTAGGTTTACTGCTCTGTGAGCCCTTCTTCCCTGCCTCCCCCA CTCACAGTGACCCGAATCTGCAGTGCTAGTCTCCCGAAGTATCA CTCTTTACAGTCTGCTTTGGAAGGACTGGGCTTAGTATGAAAAGTT AGGACTGAGAAGAAATTGAAAGGCGGCTTTTGTAGCTTGATATTC ACTACTGTCTTATTACCTGTG
101	Donor DNA	HBB exons 1-3	ATGGTGCATCTGACTCCTGAAGAAAAATCCGCTGTCACTGCCTGT GGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCA GGTTGGTATCAAGGTTACAAGACAGGTTTAAAGAGACCAATAGAA ACTGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCTGATAGGCA CTGACTCTCTCTGCCCTATTGGTCTATTTTCCACCCTTAGGCTGCTG TGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCCTTGGGGATCTG TCCACTCTGTATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATG GCAAGAAAGTGCTCGGTGCCCTTAGTGATGGCTGGCTCACCTGGA CAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGAC AAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGACG CTTGATGTTTTCTTTCCCTTCTTTCTATGGTTAAGTTCATGTGATA GGAAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAAACAGAC GAATGATTGCATCAGTGTGGAAGTCTCAGGATCGCTTTAGTTTCTTT TATTTGCTGTTTATAACAATTGTTTTCTTTGTTTAAATCTTGCTTTCT TTTTTTTCTTCTCCGCAATTTTACTATTATACTTAATGCCTTAACA TTGTGTATAACAAAAGGAAATATCTCTGAGATACATTAAAGTAACCT AAAAAAAACTTTACACAGTCTGCCCTAGTACATTACTATTGGAAT ATATGTGTGCTTATTTGCATATTCATAATCTCCCTACTTTATTTTCTT TTATTTTAAATGTATACATAATCATTATACATATTATGGGTTAAAG TGTAAATGTTTTAATATGTGTACACATATTGACCAATCAGGTAATT TTGCATTTGTAATTTAAAAAATGCTTTCTTCTTTAATATACTTTTT TGTTTATCTTATTCTAATACTTTCCCTAATCTCTTTCTTTAGGGCA ATAATGATACAAATGTATCATGCCTCTTTGCACCATTCTAAAGAAATA CAGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCTGCATATA AATATTTCTGCATATAAATTGTAAGTATGTAAGAGGTTTCATATTG CTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGT TGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTA ATCATGTTTCATACCTCTTATCTCTCCACAGCTCCTGGGCAACGT GCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTACCCCCAC CAGTGCAGGCTGCCATCAGAAAGTGGTGGCTGGTGTGGCTAATGC CCTGGCCCAAGTATCAC
102	Donor DNA	E6V to E6 Insert (PAM underlined)	GAAGAAAAATCCGCTGTC
103	Donor template	Full AAV-323	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGC GTCGGGCGACCTTTGGTGCGCCGGCCTCAGTGAGCGAGCGAGCGCG CAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTGCGGCCGC ACGCGTCTTGCTTTGACAATTTGGTCTTTTCTCAGAATACTATAAATAT AACCTATATTATAATTTATAAAGTCTGTGCAATTTCTTTGACCCAG GATATTTGCAAAAGACATATTCAAACCTCCGAGAACACTTTATTTTC ACATATACATGCCTCTTATATCAGGGATGTGAAACAGGGTCTTGAA AACTGTCTAAATCTAAACAATGCTAATGCAGGTTTAAATTTAATA AAATAAAATCCAAATCTAACAGCCAAGTCAAATCTGCATGTTTTA ACATTTAAATATTTTAAAGACGTCTTTTCCAGGATTCAACATGTG AAATCTTTTCTCAGGATACAGTGTGCTTAGATCCTCATTGCTTTA

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
			<p> GTTTTTACAGAGGAATGAATATAAAAAGAAAATACTTAAATTTTA TCCCTCTTACCTCTATAATCATACATAGGCATAATTTTTTAACCTAG GCTCCAGATAGCCATAGAAGAACCAACACTTCTGCGTGTGTGAG AATAATCAGAGTGAGATTTTTTCAAGTACCTGATGAGGGTTGAG ACAGGTAGAAAAAGTGAGAGATCTCTATTTATTTAGCAATAATAGA GAAAGCATTTAAGAGAATAAGCAATGGAAATAAGAAATTTGTAA ATTTCTCTCTGATAACTAGAAATAGAGGATCCAGTTTCTTTTGGTTA ACCTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT ATTTTGTGTAATCGTAGTTTTCAGAGTGTTAGAGCTGAAAGGAAGAA GTAGGAGAAACATGCAAGTAAAGTATAACACTTCTTACTATAA CCGACATGGGTTTCCAGGTAGGGGAGGATTGAGGATGACTGACAG GGCCCTTAGGGAACTGAGACCTTACGCTGACCTCATAAATGCTT GCTACCTTGCTGTTTAAATTACATCTTTAATAGCAGGAAGCAGAA CTCTGCACTTCAAAAGTTTTCTCACCTGAGGAGTTAATTTAGTAC AAGGGGAAAAAGTACAGGGGGATGGGAGAAAGGCGATCACGTTGG GAAGCTATAGAGAAAGAGTAAATTTTAGTAAAGGAGGTTAA ACAAACAAAATATAAGAGAAATAGGAACCTGAATCAAGGAAATG ATTTTAAACCGCAGTATTCTTAGTGGACTAGAGGAAAAAATAATC TGAGCCAGTAGAAGACCTTTCCCTCTACCCCTACTTTCTAAGT CACAGAGGCTTTTGTCTCCCGACACTCTTGCAGATTAGTCCAGG CAGAAACAGTTAGATGTCCTCAGTTAACCTCCTATTTGACACCACT GATTACCCCATGTAGTACACTTGGGTTGTAAGTGACTTTTAT TTATTTGTATTTTGTACTGCATTAAGAGGTCTCTAGTTTTTATCTCT TGTTTTCCAAAACCTAATAAGTAACTAATGCACAGAGCACATTGAT TTGATTTATTCTATTTTATAGACATAATTATAGCATGCATGAGCA AATTAAGAAAAACAACAACAAATGAATGCATATATATGTATATGTA TGTGTGTATATATACACATATATATATATATTTTTCTTTTCTTAC CAGAAGGTTTTAATCCAATAAGGAGAAGATATGCTTAGAACCGAG GTAGAGTTTTTATCCATTCTGTCTGTAAGTATTTGCATATTTCTGG AGACGAGGAGAGATCCATCTACATATCCCAAGCTGAATATATGG TAGACAAAACCTCTCCACTTTTAGTGACATCAACTCTTATTTGTGTA ATAAGAAAATTGGGAAAACGATCTTCAATATGCTTACCAAGCTGTG ATTCAAAATATACGTAATAACACTTGCAAGGAGGAGTGTTTTATG TAGCAATTTGTACTGATGGTATGGGGCAAGAGATATATCTTAGAG GGAGGGCTGAGGGTTGAAGTCCAACCTTAAGCCAGTGCCAGAAG AGCCAAGGACAGGTACGGCTGTCTACCTTAGACCTCACCTGTGG AGCCACACCTTAGGGTTGGCCTACTCTCCAGGAGCAGGGAGGG CAGGAGCCAGGGCTGGGCATAAAAGTACGGGCAGAGCCATCTATT GCTTACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAA AGACACCATGGTGATCTGACTCCTGAAGAAAAATCCGCTGTCACT GCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCC TGGGCAGGTTGGTATCAAGGTTACAGGAGGTTTAAAGGAGACCAA TAGAAACTGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCTGAT AGGCACTGACTCTCTCTGCTTATTGGTCTATTTTCCACCTTAGGC TGCTGGTGGTCTACCTTGGACCCAGAGGTTCTTTGAGTCTTTTGGG GATCTGTCCACTCCTGATGCTGTTATGGGCAACCTAAGGTGAAG CTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCTGGCTCA CCTGGACAACTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCAC TGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTAT GGGACGCTTGATGTTTTCTTCTCCCTCTTTTCTATGGTTAAGTTTAT GTCATAGGAAGGGGATAAGTAAACAGGGTACAGTTTAGAATGGGAA ACAGACGAATGATTGCATCAGTGTGAAGTCTCAGGATCGTTTTAG TTTCTTTTATTGCTGTTTCATAACAATTGTTTTCTTTTGTTAATCTT GCTTTCTTTTCTTCTCCGCAATTTTACTATTATACCTAATGCTT TTAACATTGTGTATAACAAAAGGAAATATCTCTGAGATACATTAAG TAACTTAAAAAAAACCTTTACACAGTCTGCCAGTACATTACTATTT GGAATATATGTGTGCTTATTTGCATATTCATAATCTCCCTACTTTATT TTCTTTTATTTTAAATGATACATAATCATTATACATATTTATGGGTT AAAGTGTAATGTTTTAATATGTGTACACATATTGACCAAAATCAGGG TAATTTTGCAATTTGAATTTTAAAAAATGCTTCTCTTTTAAATATAC TTTTTTGTTTATCTTATTTCTAATACTTTCCCTAATCTCTTTCTTCAG GGCAATAATGATACAAATGTATCATGCCTCTTGCACCATTTCTAAAG AATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCTGC ATATAAATATTTCTGCATATAAATGTAACTGATGTAAGAGGTTTCA TATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTT ATGGTTGGGATAAGGCTGGATTATCTGAGTCCAAGCTAGGCCCTT TTGCTAATCATGTTTACATCTTATCTCTCCACAGCTCCTGGG CAACGTGCTGGTCTGTGCTGGCCATCACTTTGGCAAGAATTTC ACCCCAACAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGG CTAATGCCCTGGCCACAAGTATCACTAAGCTCGCTTTCTTGTCTGC CAATTTCTATTAAAGGTTCTTTGTTCCCTAAGTCCAACCTACTAAAC TGGGGGATATTATGAAGGCTTGAGCATCTGGATTCTGCCCTAATA AAAAACATTATTTTATTTGCAATGATGATTTAAATTATTTCTGAA </p>

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
			<p> TATTTTACTAAAAAGGGAATGTGGGAGGTCAGTGCATTTAAACAT AAAGAAATGAAGAGCTAGTTCAAACCTTGGGAAAATACACTATATC TTAAATCCATGAAAGAAGGTGAGGCTGCAAACAGCTAATGCACAT TGCCAACAGCCCCGTATGCATATGCCTTATTCATCCCTCAGAAAAG GATTCAAGTAGAGGCTTGATTTGGAGGTTAAAGTTTGTCTATGCTGT ATTTTACATTACTTATTGTTTGTAGCTGTCTCATGAATGCTCTTTTCAC TACCCATTGCTTATCCTGCATCTCTCAGCCTTGACTCCACTCAGTTC TCTTGCTTAGAGATACCACCTTTCCCTGAAGTGTTCCTTCCATGTTT TACGGCAGATGGTTTCTCCTCGCCTGGCCACTCAGCCTTAGTTGTCT TCTGTGTCTTATAGAGTCTACTTGAAGAAGGAAAAACAGGGGTC ATGGTTTGACTGTCTGTGAGCCCTTCTTCCCTGCCTCCCCACTCA CAGTGACCCGGAATCTCAGTGTCTAGTCTCCCGGAACATCACTCTT TCACAGTCTGCTTGAAGGACTGGGCTTAGTATGAAAGTTAGGA CTGAGAAGAATTTGAAAGCGGCTTTTGTAGCTTGATATTCATAC TGTCTTATTACCTGTCCGTAACACGTGCGCCGAGGCTGCAGCG TCGTCTCCCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTC TGCGCGCTCGCTCGCTCACTGAGGCGGGCGACAAAGGTCGCCCG ACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGC AGCTGCCCTGACGGGCGCCTGATGCGGTATTTCTCCTTACGCATCT GTGCGGTATTTACACCCGCATACGTCAAAGCAACCATAGTACGCGC CCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAG CGTGACCGCTACACTTGCACGCGCCTAGCGCCCGCTCCTTTCGCTT TCTTCCCTTCTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTC TAAATCGGGGGCTCCCTTAGGGTTCCGATTAGTGCTTTACGGCAGC CTCGACCCCAAAAACTTGATTTGGGTGATGGTTACGTAGTGGG CATCGCCCTGATAGACGGTTTTCGCCCTTTGACGTTGGAGTCCACG TTCTTTAATAGTGGACTCTTGTTCAAACTGGAACAACACTCAACCC TATCTCGGGCTATCTTTGATTTATAAGGGATTTTGCCGATTTCGG CCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAA TTTTAACAAAAATTAACGTTTACAAATTTATGGTGCATCTCAGTA CAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCGACACCCGCG AACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGCATCC GCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGA GGTTTTACCGTCAACACGAAACGCGCAGACGAAAGGCGCTCGT GATACGCCTATTTTATAGGTTAATGTGCATGAACAATAAACTGTCT GCTTACATAAACAGTAATACAAGGGGTGTATGAGCCATATTCAAC GGGAAACGTGAGGCGCGGATTAAATTCCAACATGGATGCTGATT ATATGGGTATAAATGGGCTCGCGATAATGTGCGGCAATCAGGTGCG ACAATCTATCGCTGTATGGGAAGCCGATGCGCCAGAGTTGTTTCT GAAACATGGCAAAGGTAGCGTTGCCAATGATGTACAGATGAGATG GTCAGACTAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCA AGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCCTGCG ATCCCGGAAAAACAGCATTCAGGTATTAGAAGAATATCCTGATT CAGGTGAAAAATATTGTTGATGCGCTGGCAGTGTCTCTGCGCCGGT GCATTGATTCCTGTTGTAATGTCTTTTAAACAGCGATCGCGTAT TTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGA TGCGAGTGATTTTGTGACGAGCGTAATGGCTGGCCTGTTGAACAA GTCTGGAAGAAATGCATAAACTTTTGCATTCTCACCAGATTACG TCGTCACTCATGGTGATTCTCACTTGATAACCTTATTTTGACGAG GGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAG ACCGATACCAGGATCTTGCCATCCTATGGAACCTGCCTCGGTGAGTTT TCTCCTTCATTACAGAAACGGCTTTTCAAAAATATGGTATTGATAA TCCTGATATGAATAAATGCAGTTTCATTGTATGCTCGATGAGTTT TCTAATCTCATGACCAAAATCCCTAACGTGAGTTTTCGTTCCACTG AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCT TTTTTTCTGCGCGTAATCTGCTGCTTGCAACAAAAAACCCACCGCT ACCAGCGGTGGTTTGTGTCGGATCAAGAGCTACCAACTCTTTTTC CGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTCCT TCTAGTGTAGCCGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCA CCGCCATACATCCTCGCTCTGCTAATCTGTTACCAAGTGGCTGCTGC TAGTGGCGATAAGTCTGTCTTACCGGTTGGACTCAAGACGATAG TTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCTGTGA CACAGCCAGCTTGAGCGCAACGACCTACACCGAACTGAGATACCT ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAA GGCGGACAGGTATCCGTAAGCGGCGAGGTCGGAACAGGAGAGCG CACGAGGGAGCTTCCAGGGGGAACGCTTGGTATCTTTATAGTCTCT GTCCGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTC GTCAGGGGGCGGAGCCTATGGAACAAACGCAACGCGGCTT TTTACGGTTCTTGGCCTTTTGTGCGCTTTTGTCTCACATGT </p>
104	AAV 5' ITR		<p> CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGC GTCGGGCGACCTTTGGTCGCGCGGCTCAGTGAGCGAGCGAGCGCG CAGAGAGGAGTGGCCAACTCCATCACTAGGGGTTCTT </p>

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
105		AAV 3' ITR	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAGGTCGCCGACGCCCGGGC TTTGCCCGGCGGCTCAGTGAGCGAGCGAGCGCGAGCTGCCTGC AGG
106	Donor	HBB Locus LHA to RHA template (AAV.304)	CTTGCTTTGACAATTTTGGTCTTTCAGAATACTATAAATATAACCTA TATTATAATTTCTAAAGTCTGTGCATTTTCTTTGACCCAGGATATT TGCAAAAGACATATTCAAACCTCCGCAGAACACTTTATTTCACATAT ACATGCCCTCTTATATCAGGGATGTGAAACAGGGTCTTGAAAACGT CTAAATCTAAAACAATGCTAATGCAGGTTTAAATTTAATAAAATAA AATCCAAAATCTAACAGCCAAAGTCAAATCTGCATGTTTAAACATTT AAAATATTTTAAAGACGTCTTTTCCAGGATTCAACATGTGAAATCT TTTCTCAGGGATACACGTGTGCCTAGATCCTCATTGCTTTAGTTTTTT ACAGAGGAATGAATATAAAAGAAAATACTTAAATTTTATCCCTCT TACCTCTATAATCATACATAGGCATAATTTTTTAACTAGGCTCCAG ATAGCCATAGAAGAACCAACACTTTCTGCGTGTGTGAGAATAATC AGAGTGAGATTTTTCACAAGTACCTGATGAGGGTTGAGACAGGTA GAAAAGTGAGAGATCTCTATTATTAGCAATAATAGAGAAAGCA TTTAAGAGAATAAAGCAATGGAAATAAGAAATTTGTAATTTCTCT CTGATAACTAGAAATAGAGGATCCAGTTCTTTTGGTTAACTAAAT TTTATTTTCTTTTATTTGTTTTATTTTATTTTATTTTATTTTATTTG TAATCGTAGTTTCAGAGTGTAGAGCTGAAAGGAAGAAGTAGGAG AAACATGCAAGTAAAGTATAACACTTTCCTTACTAAACCGACAT GGGTTTCCAGGTAGGGGCGAGGATCAGGATGACTGACAGGGCCCTT AGGGAACACTGAGACCCCTACGCTGACCTCATAAATGCTTGCTACCT TTGCTGTTTTAATTACATCTTTAATAGCAGGAAGCAGAACTCTGCA CTTCAAAAGTTTTTCTCACCTGAGGAGTTAATTTAGTACAAGGG AAAAAGTACAGGGGATGGGAGAAAGGCGATCACGTTGGGAAGCT ATAGAGAAAGAGAGTAAATTTTAGTAAAGGAGGTTTAAACAAAC AAAATATAAAGAGAAATAGGAACTGAATCAAGGAAATGATTTTA AAACGCAGTATTTCTAGTGGACTAGAGGAAAAAATAATCTGAGCC AAGTAGAAGACCTTTTCCCTCCTACCCCTACTTTCTAAGTACAGAG GGCTTTTGTTCCTCCAGACACTTTGACAGATTAGTCCAGGCAGAAA CAGTTAGATGTCCCGAGTTAACCTCCTATTTGACACCACTGATTACC CCATTGATAGTACACTTTGGGTTGTAAGTGACTTTTATTATTATTGT ATTTTGTACTGCATTAAAGAGTCTCTAGTTTTTTATCTCTTGTTCCT AAAACCTAATAAGTAACTAATGCACAGACACATTGATTGTATTT ATTCTATTTTGTAGACATAATTTATTAGCATGCATGAGCAAAATTAAGA AAAACAACAACAATGAATGCATATATGTATATGTATGTGTGTAT TATATACACACATATATATATATATTTTTTCTTTTCTTACCAGAAGG TTTTAATCCAAATAAGGAGAGATATGCTTAGAACCGAGGTAGAGT TTTCATCCATTCTGTCTGTAAAGTATTTGTCATATTTGAGAGACGA GGAAGAGATCCATCTACATATCCCAAGCTGAATTATGGTAGACAA AACTCTTCCACTTTTAGTGATCAACTTCTTATTGTGTAATAAGAA AATTGGGAAAACGATCTTCAATATGCTTACCAAGCTGTGATTCCAA ATATTACGTAATACTACTGCAAGGAGGATGTTTTAGTAGCAAT TTGTAAGTATGATGTTGGGCGCAAGAGATATCTTAGAGGGAGGGC TGAGGGTTGAAGTCCAACCTCTAAGCCAGTGCCAGAGAGCCAAAG GACAGGTACGGCTGTCTCACTTAGACCTCACCTGTGGAGCCACA CCCTAGGGTTGGCCAATCTACTCCAGGAGCAGGAGGGCAGGAG CCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA TTTGCTTCTGACACAACCTGTGTTCTACTAGCAACCTCAACAGACACC ATGGTGATCTGACTCCTGTGAGAAAGTCTGCAGTCACTGCTCTATG GGGAAAGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCG GTTGGTATCAAGGTTACAAGACAGGTTTAAAGGAGACCAATAGAAAC TGGGCATGTGGAGACAGAGAAGACTCTGGGTTCTGATAGGCACT GACTCTCTCTGCTTATTGGTCTATTTTCCACCCTTAGGCTGCTGGT GGTCTACCTTTGGACCCAGAGGTTCTTTGAGTCCCTTTGGGGATCTGT CCACTCCTGATGCTGTTATGGGCAACCTTAAGGTGAAGGCTCATGG CAAGAAAGTGTCTCGGTGCCCTTAGTGATGGCTGGCTCACCTGGAC AACCTCAAGGGCACCTTTGCCCACTGAGTGAGCTGCACTGTGACA AGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGACGC TTGATGTTTTCTTTCCCTCTTTTCTATGGTTAAGTTCATGTCATAG GAAGGGGATAAGTAACAGGGTACAGTTTGAATGGGAAACAGACG AATGATTGCATCAGTGTGGAAGTCTCAGGATCGTTTTAGTTTCTTTT ATTTGCTGTTTCAACAATTTGTTTTCTTTGTTAATCTTGTCTTCTT TTTTTTTCTTCTCCGAATTTTACTATTATACTTAATGCCTTAACAT TGTGTATAACAAAAGGAAATATCTCTGAGATACATTAAGTAACTTA AAAAAAACTTTACACAGTCTGCCTAGTACATTACTATTGGAAATA TATGTGTGCTTATTGTGATATTCATAATCTCCCTACTTTATTTCTTT TATTTTTAATTGATACATAATCATTATACATATTTATGGGTTAAAGT GTAATGTTTTAATATGTGTACACATATTGACCAATCAGGGTAATTT

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
107	Donor	Full AAV (AAV. 304) template	<p> TGCATTGTAAATTTAAAAAATGCTTCTCTCTTTAATATACTTTTT GTTTATCTTATTTCTAATACTTTCCCTAATCTCTTTCTTCAGGGCAA TAATGATACAAATGTATCATGCTCTTTGCACCACTTAAAGAATAAC AGTGATAATTTCTGGGTAAAGCAATAGCAATATCTCTGCATATAA ATATTTCTGCATATAAATTGTAACGTAGTGAAGAGGTTTCATATTGC TAATAAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTATGGTT GGGATAAGGCTGGATTATTTCTGAGTCCAAGCTAGGCCCTTTTGCTA ATCATGTTCCATACCTCTTATCTTCTCCACAGCTCCTGGGCAACGT GCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAAATCACCCAC CAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGC CCTGGCCCAAGTATCACTAAGCTCGCTTTCTTGCTGTCCAATTTT TATTAAGGTTCTTTGTTCCTTAAGTCCAACCTAACTAGGGGA TATTATGAAGGCTTGAGCATCTGGATTCTGCCTAATAAAAAACA TTTATTTTCATTGCAATGATGTATTTAAATTATTTCTGAATATTTTAC TAAAAAGGGAATGTGGGAGGTGAGTGCATTTAAAAACATAAGAAA TGAAGAGCTAGTTCAAACCTGGGAAAAATACACTATATCTTAACT CCATGAAAGAGGTGAGGCTGCAACAGCTAATGCACATTTGGCAA CAGCCCTGTATGCATATGCCTTATTCATCCCTCAGAAAAGGATTCA AGTAGAGGCTTGATTGGAGGTAAAGTTTGTCTATGCTGTATTTTA CATTACTTATTGTTTTAGCTGTCTCATGAATGTCTTTTCACTACCCA TTTGCTTATCTGCATCTCTCAGCCTTGACTCCACTCAGTTCTCTTGC TTAGAGATACCACCTTCCCTGAAGTGTCTCTTCCATGTTTTACGG CGAGATGGTTTTCTCTCGCTGGCCACTCAGCCTTAGTTGTCTCTGT TGTCTTATAGAGGTCTACTTGAAGAAGGAAAAACAGGGGTCTAGGT TTGACTGTCTGTGAGCCCTTCTTCCCTGCCTCCCCACTCACAGTG ACCCGGAATCTGCAGTGTAGTCTCCCGAATCATCTCTTTCACA GTCTGCTTTGGAAGGACTGGGCTTAGTATGAAAAGTTAGGACTGAG AAGAATTTGAAAGCGGCTTTTGTAGCTTGATATTCATACTGTCT TATTACCTGTCT </p> <p> CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCGGGC AAAGCCCGGGCGTGGGCGACCTTTGGTCCGCCGGCCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGAGTGGCCAACTCCATCACTAGGGGT TCCTGCGGCCGCACGCGTCTTGCTTTGACAATTTTGGTCTTTAGAA TACTATAAATAAACCTATATTATAATTTTCAAAAGTCTGTGCATTT TCTTTGACCAGGATATTGCAAAAGACATATTTCAAATCTCCGCGAG AACACTTTATTTACATATACATGCCTCTTATATCAGGGATGTGAAA CAGGGTCTTGAAAACGTCTAAATCTAAACAATGCTAATGCAGGT TTAAATTTAATAAAATAAAATCCAAATCTAACAGCCAAAGTCAAT CTGCATGTTTTAACATTTAAAAATTTTAAAGACGTCTTTCCAGG ATTCAACATGTGAAATCTTTTCTCAGGGATACACGTGTGCTAGATC CTCATTGCTTTAGTTTTTACAGAGGAATGAATATAAAAAGAAAAT ACTTAAATTTTATCCCTCTTACCTCTATAATCATACATAGGCATAAT TTTTTAACCTAGGCTCCAGATAGCCATAGAAGAACCAAACTTTT TGCGTGTGTGAGAAATCAGAGTGAGATTTTTTACAAAGTACCTG ATGAGGTTGAGACAGGTAGAAAAAGTGAGAGATCTCTATTATTT AGCAATAATAGAGAAAGCATTTAAGAGAAATAAGCAATGGAAATA AGAAATTTGTAAATTTCTTCTGATAACTAGAAATAGAGGATCCAG TTTCTTTTGGTTAACTAAATTTTATTTTCTTTTATTTTATTTAT TTTATTTTATTTTATTTTGTGTAATCGTAGTTTTCAGAGTGTAGAGCT GAAAGGAAGAGTAGGAGAAACATGCAAGTAAAAGTATAACACT TTCTTACTAAACCGACATGGGTTTCCAGGTAGGGCGAGGATTCAG GATGACTGACAGGGCCCTTAGGGAACACTGAGACCTACGCTGACC TCATAAATGCTTGCTACCTTTGCTGTTTAAATTACATCTTTAATAGC AGGAAGCAGAACTCTGCACCTCAAAGTTTTTCTCACCTGAGGAG TTAATTTAGTACAAGGGGAAAAAGTACAGGGGGATGGGAGAAAGG CGATCACGTTGGGAAGCTATAGAGAAAGAGTAAATTTTAGTAA AGGAGGTTTAAACAACAAAAATATAAGAGAAATAGGAACCTGAA TCAAGGAAATGATTTTAAACGCAGTATTCTTAGTGGACTAGAGGA AAAAATAATCTGAGCCAAGTAGAAGACCTTTCCCTCTTACCCC TACTTTCTAAGTCAAGAGGCTTTTGTTCCTCCACAGCACTCTTGCA GATTAGTCCAGGCAGAAACAGTTAGATGTCCCAAGTTAACCTCCTA TTTGACACCACTGATTACCCCATGATAGTCACACTTGGGGTGTAA GTGACTTTTTATTTATTTGTATTTTGTGACTGCATTAAGAGGTCTCTAG TTTTTATCTCTTGTTTCCAAAACTAATAAGTAACATAATGCACAG AGCACATTGATTGTATTTATTTCTATTTTATAGACATAATTTATAGC ATGCATGAGCAAAATTAAGAAAAACAACAAATGAATGCATATA TATGTATATGTATGTGTATATATACACACATATATATATATATTT TTTCTTTTCTTACCAGAAGGTTTTAATCCAAATAAGGAGAGATATG CTTAGAACCGAGGTAGAGTTTTTATCCATTCTGTCTGTAAGTATTT TGCATATTTCTGAGACGCGAGGAAGAGATCCATCTACATATCCCAA GCTGAATATGGTAGACAAAACCTTCCACTTTTAGTGCATCAACTT CTTATTTGTGTAATAAGAAAATGGGAAACGATCTTCAATATGCTT </p>

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
			ACCAAGCTGTGATTCCAAATATTACGTAAATACACTTGCAAAGGAG GATGTTTTTAGTAGCAATTTGTACTGATGGTATGGGGCCAAGAGAT ATATCTTAGAGGGAGGGCTGAGGGTTTGAAGTCCAACCTCTAAGCC AGTGCCAGAAGAGCCAAGGACAGGTACGGCTGTCACTACTTAGACC TCACCCCTGTGGAGCCACACCTAGGGTTGGCCAATCTACTCCCAGG AGCAGGGAGGGCAGGAGCCAGGGCTGGGCATAAAAGTCAGGGCAG AGCCATCTATTGCTTACATTTGCTTCTGACACAACCTGTGTTCACTAG CAACCTCAAACAGACACCATGGTGCATCTGACTCCTGTCGAGAAGT CTGCAGTCACTGCTCTATGGGGGAAAGTGAACGTGGATGAAGTTGG TGGTGAGGCCCTGGGCAGGTGGTATCAAGGTTACAAGACAGGTTT AAGGAGACCAATAGAACTGGGCATGTGGAGACAGAGAAGACTCT TGGGTTTCTGATAGGCACCTGACTCTCTGCTCATTGGTCTATTTTCC ACCCCTTAGGCTGCTGGTGGTCTACCCCTGGACCCAGAGGTTCTTTG AGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCCT AAGGTGAAGGCTCATGGCAAGAAAGTGCCTGGTGCCTTTAGTGATG CTGCTGGCTCACCTGGACACCTCAAGGGCACCTTTGCCACACTGAG TGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGG GTGAGTCTATGGGACGCTTGATGTTTTCTTCCCTTCTTTTCTATGG TTAAGTTCAATGTCATAGGAAGGGATAAGTAACAGGGTACAGTTTA GAATGGGAAACAGACGAATGATTGCATCAGTGTGGAAGTCTCAGG ATCGTTTTAGTTTCTTTTATTTGCTGTTTATAACAATGTTTTCTTTT GTTTAATCTTGCTTCTTTTCTTTTCTTCTCCGCAATTTTACTATTA TACTTAATGCCCTAACATTGTGTATAACAAAAGGAAATATCTCTGA GATACATTAAAGTAACCTAAAAAACTTTACACAGTCTGCCTAGT ACATTACTATTTGGAATATATGTGTGCTTATTGTCATATTCATAATC TCCCTACTTTATTTCTTTTATTTTAAATGATACATAATCATTATAC ATATTTATGGGTTAAAGTGTAATGTTTTAATATGTGTACACATATTG ACCAATCAGGGTAATTTGCAATTTGTAATTTTAAAAATGCTTTCT TCTTTTAAATATACTTTTTGTATTATCTTATTCTAATACTTTCCCTAAT CTCTTTCTTTCAGGGCAATAATGATACATATGTCATGCTCTTTGTC ACCATTTCTAAAGAATAACAGTGATAATTTCTGGGTTAAGGCAATAG CAATATCTCTGCATATAAATATTTCTGCATATAAATTGTAAGTATG TAAGAGGTTTCTATATTGCTAATAGCAGCTACAATCCAGCTACCATTC TGCTTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAA GCTAGGCCCTTTTGCTAATCATGTTTACATACCCTTATCTCTCTCCAC AGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGC AAAGAATTCACCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGG CTGGTGTGGCTAATGCCCTGGCCCAAGTATCACTAAGCTCGCTTT CTTGCTGTCCAATTTCTATTAAAGGTTCTTTGTTCCCTAAGTCCAA CTACTAACTGGGGATATTATGAAGGCTTGGAGCATCTGGATTCT TGCCTAATAAAAAACATTTATTTTCAATGCAATGATGATTTTAAAT ATTTCTGAATATTTTACTAAAAAGGGAATGTGGGAGGTCAGTGCAT TTAAAAACATAAAGAAATGAAGAGCTAGTTCAAACTTGGGAAAAAT ACACTATATCTTAACTCCATGAAAGAAGGTGAGGCTGCAACAGC TAATGCACATTGGCAACAGCCCTGATGCATATGCCTTATTATCTCC TCAGAAAAGGATTCAAGTAGAGGCTTGATTTGGAGGTAAAGTTTT GCTATGCTGATTTTACATTACTTATTGTTTATGCTGCTCCTCATGAAT GTCTTTTCACTACCATTTGCTTATCCTGCATCTCTCAGCCTTGACTC CACTCAGTTCTCTGCTTAGAGATACACCTTTCCCTGAAGTGTTC CTTCCATGTTTTACGGCGAGATGGTTTCTCCTCGCCTGGCCACTCAG CCTTAGTTGTCTCTGTTGTCTATAGAGGCTACTTGAAGAAGGAAA AACAGGGGTATGGTTTGACTGTCTGTGAGCCCTTCTTCCCTGCCT CCCCCACTCACAGTGACCCGGAATCTGCAGTGTCTAGTCTCCCGGAA CTATCACTCTTTCACAGTCTGCTTTGGAAGGACTGGGCTTAGTATGA AAAGTTAGGACTGAGAAGAATTGAAAGGCGGCTTTTGTAGCTTG ATATCACTACTGTCTTATTACCTGTGCGTAACCACGTGCGGCCGA GGCTGCAGCGTGTCTCTCTAGGAACCCCTAGTGATGGAGTTGGC CACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGCGACCA AAGGTCGCGCGACGCCGGGCTTTGCCCGGGCGGCTCAGTGAGCG AGCGAGCGCGCAGCTGCCGTCAGGGGCGCCTGATGCGGTATTTCT CCTTACGCATCTGTGCGGTATTTACACCGCATACGTCAAAGCAAC CATAGTACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGTGTGGT GGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCC GCTCCTTTCTGCTTTCTTCCCTTCTTCTCGCCACGTTCCGCGGCTTT CCCCGTCAAGCTCTAAATCGGGGGCTCCCTTAGGGTTCCGATTTAG TGCTTTACGGCACCTCGACCCCAAAAACTTGATTTGGGTGATGGTT CAGCTAGTGGGCGCATCGCCCTGATAGAGGTTTTTTCGCCCTTTGACG TTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCTAACTGGAAC AACACTCAACCTATCTCGGGCTATTCTTTGATTTATAAGGGATTT TGCCGATTTTCGGCTATTGGTTAAAAAATGAGCTGATTTAAACAAA ATTTAACGCGAATTTTAAACAAAATATTAACTTTTACAATTTTATGGT GCACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGCC CCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTG

-continued

SEQ ID NO:	Identifi- fier	Name/ Description	Sequence
			CTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT GCATGTGTCAGAGGTTTTCCCGTCATCACCGAAACGCGCGAGACG AAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGAAC AATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTATG AGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAATTCACA TGGATGCTGATTATATGGGTATAAATGGGCTCGCGATAATGTCGG GCAATCAGGTGCGACAATCTATCGCTTGATGGGAAGCCCGATGCG CCAGAGTTGTTTCTGAAACATGGCAAAGTAGCGTTGCCAATGATG TTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGT TACTCACCACTGCGATCCCGGAAAAACAGCATTCCAGGTATTAGA AGAATATCCTGATTCAGGTGAAAAATTGTTGATGCGCTGGCAGTG TTCTGCGCGGTTGCATTGATTCCTGTTTGTAAATTGCTTTTAAAC AGCGATCGCGTATTTCTGCTCGCTCAGGCGCAATCACGAATGAATA ACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTG GCCTGTTGAACAAGTCTGAAAGAAATGCATAAACTTTTGCCATT TCACCGGATTGAGTCGTCCTCATGGTGATTTCTCACTTGATAACCT TATTTTGGACGAGGGGAAATTAATAGGTTGATTTGATGTTGGACGA TCGGAATCGCAGACCGATACCAGGATCTGCCATCCTATGGAACT GCCTCGGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTCAAAAA TATGTTATGATAAATCCTGATATGAATAAATGTCAGTTTCATTGAT GCTCGATGAGTTTCTAATCTCATGACCAAATCCCTTAACGTGAG TTTTCGTCCACTGAGCGTCAGACCCGTAGAAAAGATCAAAGGAT CTTCTTGAGATCCTTTTTCTGCGCGTAATCTGCTGCTTGCAAAACA AAAAAACCAACCGCTACCAAGCGGTGTTTGTGTTGCGGATCAAGAGC TACCAACTCTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGAT ACCAAATACTGCTCTCTAGTGAGCCGTAGTTAGGCCACCACTTCA AGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTA CCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGTTTGG ACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAA CGGGGGTTCTGTCACACAGCCAGCTTGAGCGAACGACCTACAC CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTT CCCGAAGGGAGAAAGCGGCAGGATATCCGGTAAGCGGCGAGGTC GGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCGCTGG TATCTTTATAGTCTGTGCGGTTTCGCCACCCTTGACTTGAGCGTCG ATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCC AGCAACGCGGCTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGC TCACATGT
108	Donor	HBB Locus LHA to RHA template (AAV.307)	CTTGCTTTGACAATTTTGGTCTTTCAGAATACTATAAATATAACCTA TATTATAATTTCAATAAGTCTGTGCATTTTCTTTGACCCAGGATATT TGCAAAAGACATATTCAAACCTCCGCAGAACCTTTATTTACATAT ACATGCCCTTTATATCAGGGATGTGAAACAGGGTCTTGAAAACGT CTAATCTAAAACAATGCTAATGCAGGTTTAAATTTAATAAAATAA AATCCAAAATCTAACAGCCAAAGTCAAATCTGCATGTTTAAACATT AAAATATTTTAAAGAGCTCTTTTCCAGGATTCAACATGTGAAATCT TTTCTCAGGGATACACGTGTGCCTAGATCCTCATTGCTTTAGTTTTT ACAGAGGAATGAATATAAAAGAAAATACTTAAATTTATCCCTCT TACCTCTATAATCATACATAGGCATAATTTTTTAACCTAGGCTCCAG ATAGCCATAGAAGAACCAACACTTTCTGCGTGTGTGAGAATAATC AGAGTGAGATTTTTCACAAGTACCTGATGAGGGTTGAGACAGGTA GAAAAAGTGAGAGATCTCTATTATTAGCAATAATAGAGAAAGCA TTTAAGAGAATAAAGCAATGGAAATAAGAAATTTGTAATTTCCCTT CTGATAACTAGAAATAGAGGATCCAGTTTCTTTTGGTTAACCTAAAT TTTATTTTATTTTATTTGTTTTATTTTATTTTATTTTATTTTATTTTGTG TAATCGTAGTTTCAGAGTGTTAGAGCTGAAAGGAAGAAGTAGGAG AAACATGCAAGATAAAGTATAACACTTTCCTTACTAAACCGACAT GGTTTTCAGGTAGGGGAGGATTCAGGATGACTGACAGGGCCCTT AGGGAACTAGAGACCTACGCTGACCTCATAAATGCTTGCTACCT TTGCTGTTTAAATTACATCTTTAATAGCAGGAAGCAGAACTCTGCA CTTCAAAAGTTTTTCTCACCTGAGGAGTTAATTTAGTACAAGGGG AAAAAGTACAGGGGATGGGAGAAAGCGATCACGTTGGGAAGCT ATAGAGAAAGAGAGTAAATTTTAGTAAGGAGGTTTAAACAAAC AAAATATAAAGAAATAGGAACCTGAATCAAGGAAATGATTTTA AAACGCAGTATTCCTAGTGGACTAGAGGAAAAAATAATCTGAGCC AAGTAGAAGACCTTTCCCTCTACCCCTACTTTCTAAGTCACAGA GGCTTTTGTTCCTCCAGACACTCTTGAGATTAGTCCAGGCAGAAA CAGTTAGATGTCCCGATTAACTCCTATTTGACACCACTGATTACC CCATTGATAGTCACACTTTGGGTTGTAAGTGACTTTTTATTTATTTGT ATTTTGTAGTCATTAAAGAGTCTCTAGTTTTTTATCTCTTGTTTCCC AAAACCTAATAAGTAACATAATGCACAGACACATTGATTGTATTT ATTTCTATTTTAGACATAATTTATTAGCATGCATGAGCAAAATTAAGA AAAACAACAACAATGAATGCATATATATGTATATGTATGTGTGA

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
109	Donor	Full AAV (AAV.307) template	<p> TATATACACACATATATATATATATTTTTTCTTTTCTTACCAGAAGG TTTTAATCCAAATAAGGAGAAGATATGCTTAGAACCGAGGTAGAGT TTTCATCCATTCTGTCTGTAAGTATTTTGCATATCTGGAGACGCA GGAAAGAGATCCATCTACATATCCCAAAGCTGAATTATGGTAGACAA AACTCTTCCACTTTTAGTGCATCAACTCTTATTTGTGTAATAAGAA AATTGGGAAAAACGATCTTCAATATGCTTACCAAGCTGTGATTCCAA ATATTACGTAAATACACTTGCAAGGAGGATGTTTTAGTAGCAAT TTGTACTGATGGTATGGGGCCAAGAGATATATCTTAGAGGGAGGGC TGAGGGTTTGAAGTCCAACCTCTAAGCCAGTGCCAGAAAGAGCCAAG GACAGGTACGGCTGTCATCACTTAGACCTCACCCCTGTGGAGCCACA CCCTAGGGTTGGCCAACTACTCTCCAGGAGCAGGGAGGGCAGGAG CCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA TTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAACACAGACCC ATGGTGCATCTGACTCCTGTGCAAAAATCCGCTGTCAACGCCCTCTG GGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAG GTTGGTATCAAGGTTACAAGACAGGTTTAAAGGAGACCAATAGAAAC TGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCTGATAGGCACT GACTCTCTCTGCTATTGGTCTATTTCCCAACCTTAGGCTGCTGGT GGTCTACCCCTGGACCCAGAGGTTCTTTGAGTCTTTGGGGATCTGT CCACTCCTGATGCTGTTATGGGCAACCTAAGGTGAAGGCTCATGG CAAGAAAGTGCTCGGTGCTTTAGTGATGGCTGGCTCACCTGGAC AACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCATGTGACA AGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGACGC TTGATGTTTTCTTTCCCTTCTTTCTATGGTTAAGTTTATGTCATAG GAAGGGGATAAGTAACAGGGTACAGTTTGAATGGGAAACAGACG AATGATTGCATCAGTGTGAAGTCTCAGGATCGTTTTAGTTTCTTTT ATTTGCTGTTTATAACAATTGTTTTCTTTGTTAATTCTTGTCTTCTT TTTTTTTCTTCTCCGAATTTTACTATTATACTTAATGCCTTAACAT TGTGTATAACAAAAGGAATATCTCTGAGATACATTAAGTAACCTTA AAAAAAACTTTACACAGTCTGCCTAGTACATTACTATTGGAAATA TATGTGTGCTTATTTGCATATTATAATCTCCCTACTTTATTTCTTT TATTTTAAATTGATACATAATCATTATACATATTTATGGGTTAAAGT GTAATGTTTAAATATGTGTACACATATTGACCAATCAGGGGTAATTT TGCAATTTGTAATTTTAAAAAATGCTTTCTTTTAAATATACTTTT GTTTATCTTATTTCTAATACTTTCCCTAATCTCTTTCTTTCAGGGCAA TAATGATACAAATGATATCGCTCTTTGACCACTTCTAAAGAAATAAC AGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCTGCATATAA ATATTTCTGCATATAAATGTAACTGATGTAAGAGGTTTCATATTGC TAATAGCAGCTACAATCCAGCTACCACTCTGCTTTTATTTTATGGTT GGGATAAGGCTGGATTATCTGAGTCCAAGCTAGGCCCTTTTGCTA ATCATGTTTCATACCTCTTATCTTCTCCACAGCTCCTGGGCAACGT GCTGCTCTGTGTGCTGGCCCATCACTTTGGCAAGAAATCACCCAC CAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGC CCTGGCCCAAGTATCACTAAGCTCGCTTTCTTGCTGTCCAATTTT TATTAAGGTTCTTTGTTCCCTAAGTCCACTACTAACTGGGGGA TATTATGAAGGGCTTGAGCATCTGGATTCTGCCTAATAAAAAACA TTTATTTTCATGCAATGATGATTTAAATTATTTCTGAATATTTTAC TAAAAAGGGAATGTGGGAGGTGAGTGCATTTAAACATAAAGAAA TGAAGAGCTAGTTCAAACCTGGGAAAAATACACTATATCTTAACT CCATGAAAGAAGGTGAGGCTGCAACAGCTAATGCACATTGGCAA CAGCCCTTGATGCATATGCCTTATTCATCCCTCAGAAAGGATTCA AGTAGAGGCTTGATTGGAGGTTAAAGTTTGTCTATGCTGTATTTTA CATTACTTATTGTTTTAGCTGTCTCATGAATGCTTTTCACTACCCA TTTGCTTATCCTGCATCTCTCAGCCTTGACTCCACTCAGTTCTCTTG TTAGAGATACCACCTTTCCCTGAAGTGTCTCTCCATGTTTTACGG CGAGATGGTTTCTCTCGCTGGCCACTCAGCCTTAGTTGTCTCTGT TGTCTTATAGAGGTCTACTTGAAGAAGGAAAAACAGGGGTCTAGGT TTGACTGTCTGTGAGCCCTTCTTCCCTGCTCCCACTCAGGTG ACCCGGAATCTGCAGTGTAGTCTCCGGAATATCACTCTTTCACA GTCTGCTTTGAAAGGACTGGGCTTAGTATGAAAGTTAGGACTGAG AAGAATTTGAAAGGCGGCTTTTGTAGCTTGATATCACTACTGTCT TATTACCTGTCT </p>

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
			<p> ATTCAACATGTGAAATCTTTTCTCAGGGATACACGTGTGCCTAGATC CTCATTGCTTTAGTTTTTTTACAGAGGAATGAATATAAAAAGAAAAT ACTTAAATTTTATCCCTCTTACCTCTATAATCATACATAGGCATAAT TTTTTAACCTAGGCTCCAGATAGCCATAGAAGAACCAACACTTTC TGCGTGTGTGAGAATAATCAGAGTGAGATTTTTTACAAAGTACCTG ATGAGGGTTGAGACAGGTAGAAAAAGTGAGAGATCTCTATTATTAT AGCAATAATAGAGAAAGCATTTAAGAGATAAAGCAATGGAATA AGAAATTTGTAAATTTCTTTCTGATAACTAGAAATAGAGGATCCAG TTTCTTTTGGTTAACCTAAATTTTATTTTCAATTTATTGTTTTATTAT TTTATTTTATTTTATTTTGTGTAACTGTAGTTTCAGAGTGTTAGAGCT GAAAGGAAGAAGTAGGAGAAACATGCAAGTAAAAGTATAACACT TTCCTTACTAAACCGACATGGGTTTCCAGGTAGGGGCAGGATTTCAG GATGACTGACAGGGCCCTTAGGGAACACTGAGACCTACGCTGACC TCATAAATGCTTGCTACCTTTGCTGTTTTAATTACATCTTTAATAGC AGGAAGCAGAACTCTGCACCTCAAAAGTTTTTCTCACCTGAGGAG TTAATTTAGTACAAGGGGAAAAAGTACAGGGGGATGGGAGAAAG CGATCACGTTGGGAAGCTATAGAGAAAGAAGAGTAAATTTAGTAA AGGAGGTTTAAACAACAAAAATATAAGAGAAATAGGAACTTGAA TCAAGGAAATGATTTTAAACGCAGTATCTTAGTGAGCTAGAGGA AAAAAATAATCTGAGCCAAGTAGAAGACCTTTTCCCCCTCCTACCCC TACTTTCTAAGTCACAGAGGCTTTTGTTCCTCCAGACACTCTTGCA GATTAGTCCAGGCAGAAACAGTTAGATGTCCCGAGTTAACCTCTA TTTGACACCACTGATTACCCCATTTGATAGTCACACTTTGGGTTGTAA GTGACTTTTTATTTATTTGTATTTTTGACTGCATTAAGAGGCTCTAG TTTTTTATCTCTTGTTCCTCCAAACCTAATAAGTAACATATGCACAG AGCACATTGATTTGTATTTATTTCTATTTTACACATAATTTATTAGC ATGCATGAGCAAAATTAAGAAAAACACACAAATGAATGCATATA TATGTATATGTATGTGTATATATACACACATATATATATATATTT TTTCTTTTCTTACCAGAAGGTTTTAATCCAATAAGGAGAAGATATG CTTAGAACCCAGGTAGAGTTTTTCATCCATTCGTCTCTGTAAGTATTT TGCAATATTCTGGAGACGCAGGAAGAGATCCATCTACATATCCCAA GCTGAATTATGGTAGACAAAACCTTCCACTTTTAGTGCATCAACTT CTTATTTGTGTAATAAGAAAAATGGGAAAACGATCTTCAATATGCTT ACCAAGCTGTGATTCCAATATTAAGTAAATACACTTGCAAGGAG GATGTTTTTAGTAGCAATTTGTACTGATGGTATGGGGCCAAGAGAT ATATCTTAGAGGGAGGGCTGAGGGTTTGAAGTCCAACCTCTAAGCC AGTGCCAGAAGAGCCAAGGACAGGTACGGCTGTCTCACTTAGACC TCACCCCTGTGGAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGG AGCAGGGAGGGGAGGAGCCAGGGCTGGGCATAAAAGTCAGGGCAG AGCCATCTATTGCTTACATTTGCTTCTGACCAACTGTGTTCACTAG CAACCTCAAACAGACACCATGGTGATCTGACTCCTGTGCGAAAAAT CCGCTGTCAACGCCCTCTGGGGCAAGGTGAACGTGGATGAAGTTGG TGGTGAGGCCCTGGGCAGGTGGTATCAAGGTTACAAGACAGGTTT AAGGAGACCAATAGAACTGGGCATGTGGAGACAGAGAAGACTCT TGGGTTTTCTGATAGGCACGTACTCTCTGCTATTTGGTCTATTTTCC CACCTTAGGCTGCTGGTGGTCTACCTTGGACCCAGAGGTTCTTTG AGTCTTTTGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCTT AAGGTGAAGGCTCATGGCAAGAAAGTGTCTGGTGCCTTTAGTGATG GCCTGGCTCACCTGGACAACCTCAAGGGCACCTTGGCCACTGAG TGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGG GTGAGTCTATGGGACGCTTGATGTTTTCTTTCCCTTCTTTCTATGG TTAAGTTTATGTCATAGGAAGGGGATAAGTAACAGGGTACAGTTTA GAATGGGAAACAGACGAATGATTGCATCAGTGTGAAGTCTCAGG ATCGTTTTAGTTTTCTTTTATTTGCTGTTTATAACAATGTTTTCTTTT GTTTAATTTCTGCTTTCTTTTTTTTTCTTCTCCGCAATTTTACTATTA TACTTAATGCCTTAACATTGTGTATAACAAAAGGAAATATCTCTGA GATACATTAAAGTAACCTAAAAAAAACCTTACACAGTCTGCCTAGT ACATTACTATTTGGAATATATGTGTCTTATTTGCATATTCATAATC TCCCTACTTTATTTCTTTTATTTTAAATGATACATAATCATTATAC ATATTTTAGGGTTAAAGTGTAATGTTTTAATATGTGTACACATATTG ACCAATCAGGGTAATTTTGCAATTTGTAAATTTTAAAAATGCTTTCT TCTTTTAAATATACTTTTTTGTATCTTATTCTAATACTTTCCCTAAT CTCTTTCTTTTCAAGGCAATAATGATACATATGATCATGCTCTTTGTC ACCATTCATAAGAAATAACAGTGATAATTTCTGGGTAAAGGCAATAG CAATATCTCTGCATATAAATATTTCTGCATATAAATGTAACTGATG TAAGAGGTTTCTATATTGCTAATAGCAGCTACAACTCCAGCTACCATTC TGCTTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAA GCTAGGCCCTTTTGCTAATCATGTTTATACCTCTTATCTCTCTCCAC AGCTCCTGGGCACGCTGTGGTCTGTGTGCTGGCCCATCACTTTGGC AAAGAAATCACCCACCAAGTGAGGCTGCCTATCAGAAAGTGGTG CTGGTGTGGCTAATGCCCTGGCCCAAGTATCACTAAGCTCGCTTT CTTGCTGTCCAATTTCTATTAAAGGTTCTTTGTTCCCTAAGTCCAA CTACTAACTGGGGATATTATGAAGGCTTGGAGCATCTGAGTTT TGCCATAATAAAAAACATTTATTTTCAATTGCAATGATGATTAAAT </p>

-continued

SEQ ID NO:	Identi- fier	Name/ Description	Sequence
			<p> ATTTCTGAATATTTTACTAAAAAGGGAATGTGGGAGGTCAGTGCAT TTAAAAACATAAAGAAATGAAGAGCTAGTTCAAACCTTGGGAAAAAT ACACTATATCTTAAACTCCATGAAAGAAGGTGAGGCTGCAACACAGC TAATGCACATTGGCAACAGCCCTGATGCATATGCCTTATTCATCCC TCAGAAAAGGATTCAAGTAGAGGCTTGATTTGGAGGTTAAAGTTTT GCTATGCTGTATTTACATTACTTATTGTTTAGCTGTCTCATGAAT GTCTTTTCACTACCCATTGCTTATCCTGCATCTCTCAGCCTTGACTC CACTCAGTTCTCTTGCTTAGAGATACCACCTTTCCCTGAAGTGTTCT CTTCATGTTTTACGGCGAGATGGTTTCTCCTCGCTGGCCACTCAG CCTTAGTTGTCCTGTTGCTTATAGAGGTCCTTGAAGAAGGAAA AACAGGGGTCATGGTTTGACTGTCTGTGAGCCCTTCTCCCTGCCT CCCCCACTCACAGTGACCGGAATCTGCAGTGCTAGTCTCCCGGAA CTATCACTCTTTCACAGTCTGCTTGGAAAGGACTGGGCTTAGTATGA AAAGTTAGGACTGAGAAGAATTGAAAGCGGCTTTTGTAGCTTG ATATTCACTACTGTCTTATTACCTGTGCGTAACCACGTGCGGCCGA GGCTGCAGCGTCGTCTCCCTAGGAACCCCTAGTGATGGAGTTGGC CACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGCGACCA AAGGTCGCCCCGACGCCCGGCTTTCGCCGGCGGCTCAGTGAGCG AGCGAGCGCGCAGCTGCCGTCAGGGGCGCTGATGCGGTATTTTCT CCTTACGCATCTGTGCGGTATTTACACCCGCATACGTCAAAGCAAC CATAGTACGCGCCCTGTAGCGCGCATTAAAGCGCGCGGGTGTGT GGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCC GCTCCTTTTCGCTTCTTCCCTTCCCTTTCGCCACGTTCCGCCGGCTT CCCCCTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCCGATTTAG TGCTTTACGGCACCTCGACCCAAAAAATCTGATTGGGTGATGGTT CACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACG TTGGAGTCCACGTTCTTTAATAGTGGACTCTGTTCCAACTGGAAC AACACTCAACCCTATCTCGGGCTATTCTTTGATTATAAGGGATT TGCCGATTTTCGCCCTATTGGTTAAAAATGAGCTGATTTAACAATA ATTTAACCGCAATTTTAAACAAATATTAACGTTTACAATTTTATGGT GCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCC CCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGCTCTG CTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT GCATGTGTGAGAGGTTTTACCGTCACTACCGAAACGCGCGAGACG AAAGGGCTCGTGATACGCCATTATTTATAGGTAAATGTCATGAAC AATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATG AGCCATATTCAACGGGAACGTCGAGGCGCGATTAAATTCACAAC TGGATGCTGATTATATAGGGTATAAATGGGCTCGCGATAATGTCGG GCAATCAGGTGCGACAATCTATCGCTTGATGGGAAGCCGATGCG CCAGAGTTGTTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATG TTACAGATGAGATGGTCAGACTAACTGGCTGACGGAATTTATGCC TCTTCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGT TACTCACTACTGCGATCCCGGAAAAACAGCATTCAGGTATTAGA AGAATATCCTGATTCAAGTGAATAATTTGATGCGCTGGCAGTG TTCTGCGCGGTTGCATTGCTGTTTGTGTAATTGTCCTTTTAAC AGCGATCGCGTATTTCTGCTCGCTCAGGCGCAATCACGAATGAATA ACGGTTTGGTTGATGCGAGTGATTTTATGAGCAGCGTAATGGCTG GCCTGTTGAACAAGTCTGGAAGAAATGCATAAACTTTTGCCATT TCACCGGATTCAAGTCTGCTCATGGTGATTTCTCACTTGATAACCT TATTTTTGACGAGGGGAAATTAATAGGTGTATTGATGTTGACGA GTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAAC GCCTCGGTGAGTTTTCTCCTTCAATACAGAAACGGCTTTTCAAAAA TATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTGAT GCTCGATGAGTTTTTCTAATCTCATGACCAAAATCCCTTAACGTGAG TTTTCGTTCACTGAGCGTCAGACCCGTCAGAAAAGATCAAAGGAT CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAACA AAAAAACACCGCTACCGCGGTGGTTGTTTGCCGGATCAAGAGC TACCAACTCTTTTCCGAAGGTAATGGCTTCAGCAGAGCGCAGAT ACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCA AGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCTGTGTA CCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGTTGG ACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAA CGGGGGGTTCTGTGCACACAGCCAGCTTGAGCGGAACGACCTACAC CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTT CCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGCGAGGGTC GGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCTGG TATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTGC ATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCC AGCAACGCGGCTTTTACGGTTCTGCGCTTTTGTGCGCTTTTGTG TCACATGT </p>

-continued

SEQ ID NO:	Identifier	Name/Description	Sequence
110	Donor	5' ITR template	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGC AAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGGT TCCT

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 116

<210> SEQ ID NO 1
 <211> LENGTH: 114
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: gRNA-related, sgRNA1
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (1)..(30)
 <223> OTHER INFORMATION: n is any nucleotide; n is present at least 17
 and up to 30 times
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (107)..(114)
 <223> OTHER INFORMATION: u is present at least 1 time and up to 8 times
 <400> SEQUENCE: 1
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau 60
 aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuuu 114

<210> SEQ ID NO 2
 <211> LENGTH: 119
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: gRNA-related, SpCas9 sgRNA
 <220> FEATURE:
 <221> NAME/KEY: modified_base
 <222> LOCATION: (1)..(6)
 <223> OTHER INFORMATION: every c and u is a 2' O-methyl
 phosphorothioate nucleotide
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (2)..(2)
 <223> OTHER INFORMATION: n is a phosphorothioate nucleotide
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (4)..(4)
 <223> OTHER INFORMATION: n is a phosphorothioate nucleotide
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (6)..(6)
 <223> OTHER INFORMATION: n is a phosphorothioate nucleotide
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (7)..(36)
 <223> OTHER INFORMATION: n is any nucleotide and repeats 17 times and
 up to 30 times
 <220> FEATURE:
 <221> NAME/KEY: modified_base
 <222> LOCATION: (113)..(118)
 <223> OTHER INFORMATION: every u is a 2' O-methyl phosphorothioate
 nucleotide
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (114)..(114)

-continued

<223> OTHER INFORMATION: n is a phosphorothioate nucleotide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (116)..(116)
<223> OTHER INFORMATION: n is a phosphorothioate nucleotide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (118)..(118)
<223> OTHER INFORMATION: n is a phosphorothioate nucleotide

<400> SEQUENCE: 2

cnunnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnguuu uagagcuaga aaugcaagu 60

uaaaauaagg cuaguccguu aucaacuuga aaaaguggca ccgagucggu gcununu 119

<210> SEQ ID NO 3
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, AAVS1 target sequence A

<400> SEQUENCE: 3

ggggccacta gggacaggat 20

<210> SEQ ID NO 4
<211> LENGTH: 20
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, AAVS1 sgRNA spacer A

<400> SEQUENCE: 4

ggggccacua gggacaggau 20

<210> SEQ ID NO 5
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, AAVS1 target sequence B

<400> SEQUENCE: 5

gccagtagcc agccccgtcc 20

<210> SEQ ID NO 6
<211> LENGTH: 20
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, AAVS1 sgRNA spacer B

<400> SEQUENCE: 6

gccaguagcc agccccgucc 20

<210> SEQ ID NO 7
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, BFP target sequence

<400> SEQUENCE: 7

-continued

tgaagcactg cagccat	18
 <210> SEQ ID NO 8 <211> LENGTH: 18 <212> TYPE: RNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic: gRNA-related, BFP sgRNA spacer <400> SEQUENCE: 8 ugaagcacug cagccau	
	18
 <210> SEQ ID NO 9 <211> LENGTH: 20 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic: gRNA-related, GFP target sequence A <400> SEQUENCE: 9 gctgaagcac tgcacgccgt	
	20
 <210> SEQ ID NO 10 <211> LENGTH: 20 <212> TYPE: RNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic: gRNA-related, GFP sgRNA spacer A <400> SEQUENCE: 10 gcugaagcac ugcacgccgu	
	20
 <210> SEQ ID NO 11 <211> LENGTH: 20 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic: gRNA-related, GFP target sequence B <400> SEQUENCE: 11 ctcgtgacca ccctgacct	
	20
 <210> SEQ ID NO 12 <211> LENGTH: 20 <212> TYPE: RNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic: gRNA-related, GFP sgRNA spacer B <400> SEQUENCE: 12 cucgugacca cccugaccua	
	20
 <210> SEQ ID NO 13 <211> LENGTH: 20 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic: gRNA-related, GSD1a target sequence <400> SEQUENCE: 13 tcttttgaca gcgtccatac	
	20
 <210> SEQ ID NO 14	

-continued

<211> LENGTH: 20
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, GSD1a Ch32 gRNA spacer

<400> SEQUENCE: 14

ucuuuggaca gcguccauac 20

<210> SEQ ID NO 15
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, HBB target sequence

<400> SEQUENCE: 15

cttgccccac agggcagtaa 20

<210> SEQ ID NO 16
<211> LENGTH: 20
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, HBB RO2 sgRNA spacer

<400> SEQUENCE: 16

cuugccccac agggcaguaa 20

<210> SEQ ID NO 17
<211> LENGTH: 106
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: gRNA-related, HBB sgRNA

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)..(6)
<223> OTHER INFORMATION: every c and u is a 2' O-methyl
phosphorothioate nucleotide

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2)..(2)
<223> OTHER INFORMATION: n is a phosphorothioate nucleotide

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: n is a phosphorothioate nucleotide

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (6)..(6)
<223> OTHER INFORMATION: n is a phosphorothioate nucleotide

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (100)..(104)
<223> OTHER INFORMATION: every u is a 2' O-methyl phosphorothioate
nucleotide

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (101)..(101)
<223> OTHER INFORMATION: n is a phosphorothioate nucleotide

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (103)..(103)
<223> OTHER INFORMATION: n is a phosphorothioate nucleotide

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (105)..(105)
<223> OTHER INFORMATION: n is a phosphorothioate nucleotide

-continued

<400> SEQUENCE: 17

cnunungccc cacagggcag uaaguuuag agcuagaaa agcaaguuaa auaaggcua 60

guccguuauac aacuugaaaa aguggcaccg agucggugcu nununu 106

<210> SEQ ID NO 18

<211> LENGTH: 20

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: gRNA-related, CFTR target sequence

<400> SEQUENCE: 18

tctgtatcta tattcatcat 20

<210> SEQ ID NO 19

<211> LENGTH: 20

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: gRNA-related, CFTR sgRNA spacer

<400> SEQUENCE: 19

ucuguauca uauucauau 20

<210> SEQ ID NO 20

<211> LENGTH: 23

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: gRNA-related, HBB Target Sequence

<400> SEQUENCE: 20

cttgccccac agggcagtaa cgg 23

<210> SEQ ID NO 21

<211> LENGTH: 127

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN1 (Ht-CR282)

<400> SEQUENCE: 21

gccacctaag gcaagctgac cctgaagttc atctgcacca cgggcaagct gcccggtgcc 60

tggcccaccc tcgtgaccac cctgacgtac ggcgtgcagt gcttcagccg ctaccccgac 120

cacatga 127

<210> SEQ ID NO 22

<211> LENGTH: 127

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN2 (Hn-CR283)

<400> SEQUENCE: 22

tcacatgtgtc ggggtagcgg ctgaagcact gcacgccgta cgtcaggggtg gtcacgaggg 60

tgggccaggg cacgggcagc ttccgggtgg tgcagatgaa cttcaggggtc agcttgccgt 120

aggtggc 127

<210> SEQ ID NO 23

-continued

```

<211> LENGTH: 127
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN3 (Hn-39-88)

<400> SEQUENCE: 23
cgctcctgga cgtagccttc gggcatggcg gacttgaaga agtcgtgctg cttcatgtgg      60
tcggggtagc ggctgaagca ctgcacgccg tacgtcaggg tggtcacgag ggtgggccag      120
ggcacgg                                           127

<210> SEQ ID NO 24
<211> LENGTH: 152
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN4 (Ht-91-61)

<400> SEQUENCE: 24
gccacctacg gcaagctgac cctgaagttc atctgcacca ccggcaagct gcccggtgcc      60
tggccacccc tcgtgaccac cctgacgtac ggcgtgcagt gcttcagccg ctaccccgac      120
cacatgaagc agcacgactt cttcaagtcc gc                                           152

<210> SEQ ID NO 25
<211> LENGTH: 152
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN (Hn-91-61)

<400> SEQUENCE: 25
gcggaactga agaagtcgtg ctgcttcatt tggtcggggg agcggtgaa gcaactgcacg      60
ccgtacgtca ggggtggtcag gagggtaggc cagggcacgg gcagcttgcc ggtggtgcag      120
atgaacttca gggtcagctt gccgtaggtg gc                                           152

<210> SEQ ID NO 26
<211> LENGTH: 127
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN (Hn-91-36)

<400> SEQUENCE: 26
tcattgtgtc ggggtagcgg ctgaagcact gcacgccgta cgtcaggggt gtcacgaggg      60
tgggccaggg cacgggcagc ttgccggtgg tgcagatgaa cttcaggggtc agcttgccgt      120
aggtggc                                           127

<210> SEQ ID NO 27
<211> LENGTH: 152
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN (Ht-91-61)

<400> SEQUENCE: 27
gccacctacg gcaagctgac cctgaagttc atctgcacca ccggcaagct gcccggtgcc      60
tggccacccc tcgtgaccac cctgacgtac ggcgtgcagt gcttcagccg ctaccccgac      120

```

-continued

cacatgaagc agcacgactt cttcaagtcc gc 152

<210> SEQ ID NO 28
<211> LENGTH: 127
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN (Ht-39-88)

<400> SEQUENCE: 28

cgctcctgga cgtagccttc gggcatggcg gacttgaaga agtcgtgctg cttcatgtgg 60
tcggggtagc ggctgaagca ctgcacgccg tacgtcaggg tggtcacgag ggtgggccag 120
ggcacgg 127

<210> SEQ ID NO 29
<211> LENGTH: 127
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN 1067

<400> SEQUENCE: 29

gccacctaag gcaagctgac cctgaagttc atctgcacca ccggcaagct gcccgtgccc 60
tggccacccc tcgtgaccac cctgagccac ggggtgcagt gcttcagccg ctaccccgac 120
cacatga 127

<210> SEQ ID NO 30
<211> LENGTH: 127
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN 1068

<400> SEQUENCE: 30

tcattgtggtc ggggtagcgg ctgaagcact gcaccccggtg gctcaggggtg gtcacgaggg 60
tggggccaggg cacgggcagc ttgccgggtgg tgcagatgaa cttcaggggtc agettgccgt 120
aggtggc 127

<210> SEQ ID NO 31
<211> LENGTH: 127
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN 1069

<400> SEQUENCE: 31

tcattctgcac caccgggaag ctgcccgtgc cctggcccac cctcgtgacc accctgagcc 60
acgggggtgca gtgcttcagc cgctaccccg accacatgaa gcagcacgac ttcttcaagt 120
ccgcat 127

<210> SEQ ID NO 32
<211> LENGTH: 127
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN 1070

<400> SEQUENCE: 32

-continued

atggcggact tgaagaagtc gtgctgcttc atgtggtcgg ggtagcggt gaagcactgc	60
accccggtggc tcagggtggt cagcaggggtg ggccaggga cgggcagctt gccgggtggtg	120
cagatga	127

<210> SEQ ID NO 33
 <211> LENGTH: 127
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN 1061

<400> SEQUENCE: 33

ctcctggacg tagccttcgg gcatggcgga ctgaagaag tcgtgctgct tcatgtggtc	60
ggggtagcgg ctgaagcact gcaccccggtg gctcaggggtg gtcacgaggg tggggcaggg	120
cacgggc	127

<210> SEQ ID NO 34
 <211> LENGTH: 127
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN 1062

<400> SEQUENCE: 34

gcccgtgccc tggcccaccc tcgtgaccac cctgagccac ggggtgcagt gcttcagccg	60
ctaccccgac cacatgaagc agcagcactt cttcaagtcc gccatgcccg aaggctacgt	120
ccaggag	127

<210> SEQ ID NO 35
 <211> LENGTH: 127
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN 1063

<400> SEQUENCE: 35

gacttgaaga agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca ctgcaccccg	60
tgggtcaggg tggtcacgag ggtgggccag ggcacgggca gcttgccggt ggtgcagatg	120
aacttca	127

<210> SEQ ID NO 36
 <211> LENGTH: 127
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Donor DNA, ssODN 1064

<400> SEQUENCE: 36

tgaagtcat ctgcaccacc ggcaagctgc ccgtgccctg gccaccctc gtgaccacc	60
tgagccacgg ggtgcagtgc ttcagccgct accccgacca catgaagcag cagcattct	120
tcaagtc	127

<210> SEQ ID NO 37
 <211> LENGTH: 50
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:

-continued

<223> OTHER INFORMATION: Synthetic: Donor DNA, 50-0 dsDNA

<400> SEQUENCE: 37

cccagaaact tgttctgttt ttccatagga ttctctttgg acagtgcctt 50

<210> SEQ ID NO 38

<211> LENGTH: 150

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, 150-0 dsDNA

<400> SEQUENCE: 38

agggcactgt ccaagagaa tcctatggaa aaacagaaca agtttctggg gttactgaat 60

gaatgctttt gccc aaagcc tacaccttca agaagagtgt agcctgagaa ggatttcaca 120

tgttgcctct agaagggaga actgggtggc 150

<210> SEQ ID NO 39

<211> LENGTH: 142

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, 93-50 ssODN

<400> SEQUENCE: 39

tcttgaaggt gtaggctttg ggcaaaagca ttcattcagt aacccagaa acttgttctg 60

tttttccata ggattctctt tggacagtgc cttactggtt gggctctgga tactgactac 120

tacagcaaca cttccgtgcc ct 142

<210> SEQ ID NO 40

<211> LENGTH: 125

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, 25-100 ssODN

<400> SEQUENCE: 40

taggattctc tttggacagt gcccttactg gtgggtcctg gatactgact actacagcaa 60

cacttccgtg cccctgataa agcagttccc tgtaacctgt gagactggac caggtaagcg 120

tccca 125

<210> SEQ ID NO 41

<211> LENGTH: 125

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, H3-95-30 ssODN

<400> SEQUENCE: 41

taagcacagt ggaagaatct cattctgttc tcagttttcc tggattatgc ctggcaccat 60

taaagaaaat atcataagct ttgggtgtttg ctatgatgaa tatagataca gaagcgatcat 120

caaag 125

<210> SEQ ID NO 42

<211> LENGTH: 125

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

-continued

<223> OTHER INFORMATION: Synthetic: Donor DNA, N1-95-30 ssODN

<400> SEQUENCE: 42

```

aattaagcac agtgaagaa ttctattctg ttctcagttt tctggatta tgctggcac    60
cattaaagaa aatatcatct ttggtgtttg ctagcatgat gaatatagat acagaagcgt    120
catca                                              125

```

<210> SEQ ID NO 43

<211> LENGTH: 1000

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA , AAVS1 locus LHA (used for BFP donor)

<400> SEQUENCE: 43

```

ccccagctct tctctgttca gccctaagaa tctgggtcc agccctctct actctagccc    60
ccaacccctc agccactaag gcaattgggg tgcaggaatg ggggcagggt accagcctca    120
ccaagtgtt gataaaccca cgtgggttac cctaagaact tgggaacagc cacagcaggg    180
gggcgatgct tggggacctg cctggagaag gatgcaggac gaaaaacaca gcccaggtg    240
gagaaactgg ccgggaatca agagtacccc agagacagtg accaaccatc cctgttttcc    300
taggactgag ggtttcagtg ctaaaactag gctgtcctgg gcaaacagca taagctggtc    360
acccacacac cagacctgac ccaaacccag ctccctctct tcttggccac gtaacctgag    420
aagggaatcc ctctctctct aaccccagcc caccccaatg ctccaggcct cctgggatac    480
cccgaaagag gagtttgcca agcagtcacc ccacagttgg aggagaatcc acccaaaagg    540
cagcctggta gacagggctg ggggtggcctc tcgtggggtc caggccaagt aggtggcctg    600
gggcctcttg gggatgcagg ggaaggggga tgcaggggaa cggggatgca ggggaacggg    660
gctcagctct aagagcagag ccaggaaccc ctgtagggaa ggggcaggag agccaggggc    720
atgagatggt ggacgaggaa gggggacagg gaagcctgag cgcctctcct gggttgcca    780
aggactcaaa ccagaagcc cagagcaggg ccttagggaa gcgggaccct gctctgggcg    840
gaggaatatg tcccagatag cactggggac tctttaagga aagaaggatg gagaaagaga    900
aaggagtag aggcggccac gacctggtga acacctagga cgcaccattc tcacaaaggg    960
agttttccac acggacaccc cctctctcac cacagccctg    1000

```

<210> SEQ ID NO 44

<211> LENGTH: 717

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, BFP locus donor

<400> SEQUENCE: 44

```

atggtgagca agggcgagga gctgttcacc ggggtgggtc ccacctcgtt cgagctggac    60
ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac    120
ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc    180
ctcgtgacca ccctgaccca tggcgtgcag tgcttcagcc gctacccoga ccacatgaag    240
cagcacgact tcttcaagtc cgccatgcc gaaggetacg tccaggagcg caccatcttc    300
ttcaaggacg acggcaacta caagaccgc gccaggtga agttcgaggg cgacaccctg    360

```

-continued

gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac	420
aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac	480
ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc	540
gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac	600
tacctgagca cccagtgcaa gctgagcaaa gaccccaacg agaagcgcgga tcacatggtc	660
ctgctggagt tcgtgaccgc cgccgggata actctcgga tggacgagct gtacaag	717

<210> SEQ ID NO 45
 <211> LENGTH: 1000
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Donor DNA, AAVS1 locus RHA (used for BFP donor)

<400> SEQUENCE: 45

actgtggggt ggaggggaca gataaaagta cccagaacca gagccacatt aaccggccct	60
gggaatataa ggtggtccca gctcggggac acaggatccc tggaggcagc aaacatgctg	120
tcctgaagtg gacatagggg ccggggttgg aggaagaaga ctagctgagc tctcggaacc	180
ctggaagatg ccatgacagg gggctggaag agctagcaca gactagagag gtaagggggg	240
taggggagct gccc aaatga aaggagttag aggtgacccg aatccacagg agaacggggt	300
gtccaggcaa agaaagcaag aggatggaga ggtggctaaa gccagggaga cgggggtactt	360
tggggttgtc cagaaaaacg gtgatgatgc aggcctacaa gaaggggagg cgggacgcaa	420
gggagacatc cgtcggagaa ggccatccta agaaacgaga gatggcacag gcccagaag	480
gagaagggaa agggaaaccca gcgagtgaag acggcatggg gttgggtgag ggaggagaga	540
tgcccgaga ggaccagac acggggagga tccgctcaga ggacatcacg tggtcagcg	600
ccgagaagga agtgctccg aaagagcatc ctggggcagc aacacagcag agagcaagg	660
gaagaggag tggaggaaga cggaacctga aggaggcggc agggaggat ctgggccagc	720
cgtagagggt acccaggcca caagctgcag acagaaagcg gcacaggccc aggggagaga	780
atgcaggtca gagaagcag gacctgcctg ggaaggggaa acagtgggcc agaggcgcg	840
cagaagccag tagagctcaa agtggtccg actcaggaga gagacggcag cgttagagg	900
cagagtccg gcggcacagc aagggcactc gggggcgaga ggagggcagc gcaaagtgc	960
aatggccagg gccaggcaga tagaccagac tgagctatgg	1000

<210> SEQ ID NO 46
 <211> LENGTH: 1000
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Donor DNA, AAVS1 locus LHA (used for GFP donor)

<400> SEQUENCE: 46

ccccagctct tctctgttca gccctaagaa tcctgggtcc agccctcct actctagccc	60
ccaacccct agccactaag gcaattggg tgcaggaatg ggggcagggt accagcctca	120
ccaagtgggt gataaaccca cgtggggtac cctaagaact tgggaacagc cacagcagg	180
gggcgatgct tggggacctg cctggagaag gatgcaggac gagaaacaca gcccagggtg	240

-continued

```

gagaaactgg ccggaatca agagtcaccc agagacagtg accaaccatc cctgttttcc 300
taggactgag ggtttcagtg ctaaaactag gctgtcctgg gcaaacagca taagctggtc 360
acccacacacc cagacctgac ccaaaccag ctccctgct tcttgccac gtaacctgag 420
aagggaatcc ctctctctg aacccagcc caccccaatg ctccaggcct cctgggatac 480
cccgaagagt gagtttgcca agcagtcacc ccacagttgg aggagaatcc acccaaaagg 540
cagcctggta gacagggtg ggttgccctc tcgtggggc caggccaagt aggtggcctg 600
gggcctctgg gggatgcagg ggaaggggga tgcaggggaa cggggatgca ggggaacggg 660
gctcagctg aagagcagag ccaggaaccc ctgtagggaa ggggcaggag agccaggggc 720
atgagatggt ggacgaggaa gggggacagg gaagcctgag cgctctcct gggcttgcca 780
aggactcaaa ccagaagcc cagagcaggg ccttagggaa gcgggacct gctctgggcg 840
gaggaatatg tccagatag cactggggac tctttaagga aagaaggatg gagaaagaga 900
aagggagtag aggcggccac gacctggtga acacctagga cgcaccattc tcacaaaggg 960
agttttccac acggacaccc cctctctcac cacagccctg 1000

```

```

<210> SEQ ID NO 47
<211> LENGTH: 720
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, GFP donor to AAVS1 locus

```

```

<400> SEQUENCE: 47
atggtgagca agggcgagga gctgttcacc ggggtgggtg ccatectggt cgagctggac 60
ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120
ggcaagctga cctgaagtt catctgcacc accggcaagc tgcccggtgc ctggcccacc 180
ctcgtgacca cctgacctc cgcgctgcag tgettcagcc gctacccga ccacatgaag 240
cagcacgact tctcaagtc cgccatgcc gaaggetacg tccaggagcg caccatcttc 300
ttcaaggacg acggcaacta caagaccgc gccgagggtg agttcgaggg cgacaccctg 360
gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420
aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac 480
ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540
gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600
tacctgagca ccagtcctgc cctgagcaaa gacccaacg agaagcgcg tcacatggtc 660
ctgctggagt tcgtgaccgc cgccgggata actctcgca tggacgagct gtacaagtaa 720

```

```

<210> SEQ ID NO 48
<211> LENGTH: 1000
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, AAVS1 locus RHA (used
for GFP donor)

```

```

<400> SEQUENCE: 48
actgtggggt ggaggggaca gataaaagta ccagaaacca gagccacatt aaccggccct 60
gggaatataa ggtggtccca gctcggggac acaggatccc tggaggcagc aaacatgctg 120

```

-continued

tcctgaagtg gacatagggg cccgggttgg aggaagaaga ctagctgagc tctcggaccc	180
ctggaagatg ccatgacagg gggctggaag agctagcaca gactagagag gtaagggggg	240
taggggagct gccc aaatga aaggagttag aggtgacccg aatccacagg agaacggggt	300
gtccaggcaa agaaagcaag aggatggaga ggtggctaaa gccagggaga cgggggtactt	360
tggggttgtc cagaaaaacg gtgatgatgc aggcctacaa gaaggggagg cgggacgcaa	420
gggagacatc cgtcggagaa ggccatccta agaaacgaga gatggcacag gcccagaaag	480
gagaaggaaa agggaaccca gcgagtgaag acggcatggg gttgggtgag ggaggagaga	540
tgcccggaga ggaccagac acggggagga tccgctcaga ggacatcacg tgggtgcagc	600
ccgagaagga agtgcctcgg aaagagcatc cttgggcagc aacacagcag agagcaagg	660
gaagaggag tggaggaaga cggaacctga aggaggcggc agggaaggat ctgggccagc	720
cgtagagggtg acccaggcca caagctgcag acagaaagcg gcacaggccc aggggagaga	780
atgcaggtea gagaagcag gacctgctg ggaaggggaa acagtgggccc agaggcggcg	840
cagaagccag tagagctcaa agtgggtccg actcaggaga gagacggcag cgtttagagg	900
cagagtccg gcgacacagc aagggcactc gggggcgaga ggagggcagc gcaaagtga	960
aatggccagg gccaggcaga tagaccagac tgagctatgg	1000

<210> SEQ ID NO 49

<211> LENGTH: 2189

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, HBB locus LHA (used for E7 to E7V AAV.304)

<400> SEQUENCE: 49

cttgctttga caatttttgt ctttcagaat actataaata taacctatat tataatttca	60
taaagtctgt gcattttctt tgaccagga tatttgcaa agacatatc aaacttccgc	120
agaacacttt atttcacata tacatgctc ttatatcagg gatgtgaaac agggctctga	180
aaactgtcta aatctaaac aatgctaag caggttttaa ttaataaaa taaaatccaa	240
aatctaacag ccaagtcaaa tctgcatgtt ttaacattta aaatatttta aagacgtctt	300
ttcccaggat tcaacatgtg aaatcttttc tcagggatac acgtgtgcct agatcctcat	360
tgcttttagt ttttacagag gaatgaatat aaaaagaaaa tacttaaat ttatccctct	420
tacctctata atcacataa ggcataattt tttaacctag gctccagata gccatagaag	480
aacaaacac tttctgcgtg tgtgagaata atcagagtga gattttttca caagtaacctg	540
atgaggggtg agacaggtag aaaaagttag agatctctat ttatttagca ataatagaga	600
aagcatttta gagaataaag caatggaaat aagaaatttg taaatttcct tctgataact	660
agaaatagag gatccagttt cttttgggta acctaaattt tatttcattt tattgtttta	720
ttttatttta ttttatttta ttttggtgaa tcgtagtctc agagtgttag agctgaaagg	780
aagaagtagg agaaacatgc aaagtaaaag tataacactt tccttactaa accgacatgg	840
gtttccagggt aggggcagga ttcaggatga ctgacagggc ccttagggaa cactgagacc	900
ctacgctgac ctcataaatg cttgctacct ttgctgtttt aattacatct tttaatagca	960
ggaagcagaa ctctgcactt caaaagtttt tcctcacctg aggagttaat ttagtacaag	1020
gggaaaaagt acagggggat gggagaaagg cgatcacgtt gggaagctat agagaaagaa	1080

-continued

```

gagtaaattt tagtaaagga ggtttaaaca aacaaaatat aaagagaaat aggaacttga 1140
atcaaggaaa tgattttaaa acgcagtatt cttagtggac tagaggaaaa aaataatctg 1200
agccaagtag aagacctttt cccctcctac ccctactttc taagtcacag aggctttttg 1260
ttccccaga cactcttgca gattagtcca ggcagaaaca gttagatgtc ccagttaac 1320
ctcctatttg acaccactga ttacccatt gatagtcaca ctttgggttg taagtgactt 1380
tttatttatt tgtatttttg actgcattaa gaggtctcta gttttttatc tcttgtttcc 1440
caaaacctaa taagtaacta atgcacagag cacattgatt tgtatttatt ctatttttag 1500
acataattta ttagcatgca tgagcaaatt aagaaaaaca acaacaaatg aatgcatata 1560
tatgtatatg tatgtgtgta tatatacaca catatatata tatatttttt cttttcttac 1620
cagaaggttt taatccaaat aaggagaaga tatgcttaga accgaggtag agttttcatc 1680
cattctgtcc tgtaagtatt ttgcatattc tggagacgca ggaagagatc catctacata 1740
tcccaaagct gaattatggg agacaaaact cttccacttt tagtgcacat acttcttatt 1800
tgtgtataaa gaaaattggg aaaacgatct tcaatatgct taccaagctg tgattccaaa 1860
tattacgtaa atacacttgc aaaggaggat gtttttagta gcaatttgta ctgatgggat 1920
ggggccaaga gatatatctt agagggaggg ctgagggttt gaagtccaac tcctaagcca 1980
gtgccagaag agccaaggac aggtacggct gtcacactt agacctcacc ctgtggagcc 2040
acaccctagg gttggccaat ctactcccag gagcaggagc ggcaggagcc agggctgggc 2100
ataaaagtea gggcagagcc atctattgct tacatttgct tctgacacaa ctgtgttcac 2160
tagcaacctc aaacagacac catggtgca 2189

```

```

<210> SEQ ID NO 50
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, E7 to E7V AAV.304

<400> SEQUENCE: 50

```

```

tctgactcct gtcgagaagt ctgcagtcac tgctctatgg gggaaa 46

```

```

<210> SEQ ID NO 51
<211> LENGTH: 2200
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, HBB locus RHA (used for
E7 to E7V AAV.304)

```

```

<400> SEQUENCE: 51

gtgaacgtgg atgaagtgg tggtagggcc ctgggcaggt tggatatcaag gttacaagac 60
agggttaagg agaccaatag aaactgggca tgtggagaca gagaagactc ttgggtttct 120
gataggcact gactctctct gcctattggg ctattttccc acccttaggc tgctgggtgt 180
ctacccttgg acccagaggt tctttgagtc ctttggggat ctgtccactc ctgatgctgt 240
tatgggcaac cctaagggtga aggctcatgg caagaaagtg ctcggtgcct ttagtgatgg 300
cctggctcac ctggacaacc tcaagggcac ctttgccaca ctgagtgagc tgcactgtga 360
caagctgcac gtggatcctg agaacttcag ggtgagteta tgggacgctt gatgttttct 420

```

-continued

ttcccccttct tttctatggt taagttcatg tcataggaag gggataagta acaggggtaca	480
gtttagaatg ggaacacagac gaatgattgc atcagtgtgg aagtctcagg atcgtttttag	540
tttcttttat ttgctgttca taacaattgt tttcttttgt ttaattcttg ctttcttttt	600
ttttcttctc cgcaattttt actattatac ttaatgcctt aacattgtgt ataacaaaag	660
gaaatatctc tgagatacat taagtaactt aaaaaaaaaa tttacacagt ctgcctagta	720
cattactatt tggaatatat gtgtgcttat ttgcatattc ataactctcc tactttattt	780
tcttttattt ttaattgata cataatcatt atacatattt atgggttaaa gtgtaatgtt	840
ttaatatgtg tacacatatt gaccaaatac gggtaatttt gcatttgtaa ttttaaaaaa	900
tgctttcttc ttttaataata cttttttgtt tatcttattt ctaatacttt ccctaacttc	960
tttctttcag ggcaataatg atacaatgta tcatgcctct ttgcaccatt ctaaagaata	1020
acagtgtata tttctggggt aaggcaatag caatatctct gcataataat atttctgcat	1080
ataaattgta actgatgtaa gaggtttcat attgctaata gcagctacaa tccagctacc	1140
attctgcttt tattttatgg ttgggataag gctggattat tctgagtcca agctaggccc	1200
ttttgctaata catgttcata cctcttatct tctctccaca gctcctgggc aacgtgctgg	1260
tctgtgtgct ggcccatcac ttgggcaaag aattcacccc accagtgcag gctgcctatc	1320
agaaagtggg ggctgggtg gctaattgcc tggccacaaa gtatcactaa gctcgctttc	1380
ttgctgtcca atttctatta aaggttcctt tgttccctaa gtccaactac taaactgggg	1440
gatattatga agggccttga gcatctggat tctgcctaata aaaaaacatt tattttcatt	1500
gcaatgatgt atttaaatga tttctgaata ttttactaaa aagggaatgt gggaggtcag	1560
tgcattttaa acataaagaa atgaagagct agttcaaac tggggaaaat aactatatc	1620
ttaaaactcca tgaagaagg tgaggctgca aacagctaata gcacattggc aacagcccct	1680
gatgcatatg ccttattcat ccctcagaaa aggattcaag tagaggcttg atttgagggt	1740
taaagttttg ctatgtgtga ttttacatta cttattgttt tagctgtcct catgaatgtc	1800
ttttcactac ccatttgcct atcctgcac tctcagcctt gactccactc agttctcttg	1860
cttagagata ccacctttcc cctgaagtgt tcttccatg ttttacggcg agatgggttc	1920
tctcgcctg gccactcagc cttagtgtgc tctgtgtgt tatagaggtc tacttgaaga	1980
aggaaaaaca ggggtcatgg ttgactgtc ctgtgagccc ttcttccctg cctccccac	2040
tcacagtgc cgggaatctg cagtgtagt ctcccgaac tatcactctt tcacagtctg	2100
ctttggaagg actgggctta gtatgaaaag ttaggactga gaagaatttg aaaggcggct	2160
ttttgtagct tgatattcac tactgtctta ttacctgtc	2200

<210> SEQ ID NO 52

<211> LENGTH: 2189

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, HBB locus LHA (used for E7 to E7V AAV.307)

<400> SEQUENCE: 52

cttgccttga caattttggt ctttcagaat actataaata taacctatat tataatttca	60
taaagtctgt gcattttctt tgaccagga tatttgcaa agacatattc aaacttccgc	120
agaacacttt atttcacata tacatgctc ttatatcagg gatgtgaaac agggctctga	180

-continued

aaactgtcta aatctaaaac aatgctaatag cagggtttaa ttttaataaaa taaaatccaa	240
aatctaacag ccaagtcaaa tctgcatgtt ttaacattta aaatatattta aagacgtctt	300
ttcccaggat tcaacatgtg aaatcttttc tcaggagatc acgtgtgcct agatcctcat	360
tgcttttagtt ttttacagag gaatgaatat aaaaagaaaa tacttaaaatt ttatccctct	420
tacctctata atcacataata ggcatatatt ttttaacctag gctccagata gccatagaag	480
aaccaaacac tttctgcgtg tgtgagaata atcagagtga gattttttca caagtacctg	540
atgaggggtg agacaggtag aaaaagttag agatctctat ttatttagca ataatagaga	600
aagcatttaa gagaataaag caatggaaat aagaaatttg taaatttctt tctgataact	660
agaaatagag gatccagttt ctttttggtta acctaaattt tatttcattt tattgtttta	720
ttttatttta ttttatttta ttttgtgtaa tcgtagtctt agagtgttag agctgaaagg	780
aagaagtagg agaacatgc aaagtaaaag tataacactt tccttactaa accgacatgg	840
gtttccaggt aggggcagga ttcaggatga ctgacagggc ccttagggaa cactgagacc	900
ctacgctgac ctcataaatg cttgtctact ttgtgtttt aattacatct tttaatagca	960
ggaagcagaa ctctgcactt caaaagtttt tcctcacctg aggagttaat ttagtacaag	1020
gggaaaaagt acagggggat ggagagaaag cgatcacgtt gggaagctat agagaaagaa	1080
gagtaaatat tagtaaaaga ggtttaaaca aacaaaatat aaagagaaat aggaacttga	1140
atcaaggaaa tgattttaaa acgcagtatt cttagtggac tagaggaaaa aaataatctg	1200
agccaagtag aagacctttt cccctcctac ccttactttc taagtcacag aggccttttg	1260
ttccccaga cactcttgca gattagtcca ggcagaaaca gttagatgac ccaggttaac	1320
ctcctatttg acaccactga ttacccattt gatagtcaca ctttgggttg taagtactt	1380
tttatttatt tgtatttttg actgcattaa gaggtctcta gttttttatc tctgttttc	1440
caaaaacctaa taagtaacta atgcacagag cacattgatt tgtatttatt ctatttttag	1500
acataattta ttagcatgca tgagcaaatt aagaaaaaca acaacaaatg aatgcatata	1560
tatgtatatg tatgtgtgta tatatacaca catatatata tatatttttt cttttcttac	1620
cagaaggttt taatccaaat aaggagaaga tatgcttaga accgaggtag agttttcatc	1680
cattctgtcc tgtaagtatt ttgcatttc tggagacgca ggaagagatc catctacata	1740
tccccaaagt gaattatggt agacaaaact cttccacttt tagtgcata acttcttatt	1800
tgtgtaataa gaaaattggg aaaacgatct tcaatatgct taccaagctg tgattccaaa	1860
tattacgtaa atacacttgc aaaggaggat gtttttagta gcaatttgta ctgatggat	1920
ggggccaaga gatatatctt agagggaggg ctgagggttt gaagtccaac tcctaagcca	1980
gtgccagaag agccaaggac aggtacggct gtcacactt agacctcacc ctgtggagcc	2040
acaccctagg gttggccaat ctactcccag gagcaggagag ggcaggagcc agggctgggc	2100
ataaaagtca gggcagagcc atctattgct tacatttgct tctgacacaa ctgtgttcac	2160
tagcaacctc aaacagacac catggtgca	2189

<210> SEQ ID NO 53

<211> LENGTH: 46

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, E7 to E7V AAV.307

-continued

<400> SEQUENCE: 53

tctgactcct gtcgaaaaat ccgctgtcac cgcctctctgg ggcaag 46

<210> SEQ ID NO 54

<211> LENGTH: 2200

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, HBB locus RHA (used for E7 to E7V AAV.307)

<400> SEQUENCE: 54

gtgaacgtgg atgaagttgg tggtagggcc ctgggcaggt tggatcaag gttacaagac 60

aggtttaagg agaccaatag aaactgggca tgtggagaca gagaagactc ttgggtttct 120

gataggcact gactctctct gcctatttgg ctattttccc acccttaggc tgctggtggt 180

ctacccttgg acccagaggt tctttgagtc ctttggggat ctgtccactc ctgatgctgt 240

tatgggcaac cctaaggtga aggcctcatgg caagaaagtg ctcggtgcct ttagtgatgg 300

cctggctcac ctggacaacc tcaaggggcac ctttgccaca ctgagtgagc tgcactgtga 360

caagctgcac gtggatcctg agaacttcag ggtgagtcta tgggacgctt gatgttttct 420

ttcccttctt tttctatggt taagtctatg tcataggaag gggataagta acagggatca 480

gtttagaatg ggaacacagac gaatgattgc atcagtgtgg aagtctcagg atcgttttag 540

tttcttttat ttgctgttca taacaattgt tttcttttgt ttaattcttg ctttcttttt 600

ttttcttctc cgcaattttt actattatac ttaatgcctt aacatttgtt ataacaaaag 660

gaaatatctc tgagatacat taagtaactt aaaaaaaaaac tttacacagt ctgcctagta 720

cattactatt tggaaatatat gtgtgcttat ttgcatatcc ataactctcc tactttattt 780

tcttttattt ttaattgata cataatcatt atacatattt atgggttaaa gtgtaatgtt 840

ttaatatgtg tacacatatt gaccaaatca gggtaatttt gcattttgtaa ttttaaaaaa 900

tgctttcttc ttttaataata cttttttgtt tatcttattt ctaatacttt ccctaacttc 960

tttctttcag ggcaataatg atacaatgta tcatgcctct ttgcaccatt cttaaagaata 1020

acagtataaa tttctgggtt aaggcaatag caatatctct gcataataat atttctgcat 1080

ataaattgta actgatgtaa gaggtttcat attgctaata gcagctacaa tccagctacc 1140

attctgcttt tattttatgg ttgggataag gctggattat tctgagtcca agctaggccc 1200

ttttgctaata catgttcata cctcttatct tctctccaca gctcctgggc aacgtgctgg 1260

tctgtgtgct ggcccatcac tttggcaaag aattcaccct accagtgcag gctgcctatc 1320

agaaagtggt ggctggtgtg gctaattgcc tggccacaa gtatcactaa gctcgctttc 1380

tgctgtcca atttctatta aagggttcct tgttccctaa gtccaactac taaactgggg 1440

gatattatga agggccttga gcactctggat tctgcctaat aaaaaacatt tattttcatt 1500

gcaatgatgt atttaaatga tttctgaata ttttactaaa aagggaatgt gggaggtcag 1560

tgcatttaaa acataaagaa atgaagagct agttcaaacc ttgggaaaaa aactatatc 1620

ttaaactcca tgaaagaagg tgaggctgca aacagctaata gcacattggc aacagcccct 1680

gatgcatatg ccttattcat cctcagaaa aggattcaag tagaggcttg atttgagggt 1740

taaagtttgg ctatgctgta ttttacatta cttattgttt tagctgtcct catgaatgtc 1800

-continued

```

ttttcactac ccatttgett atcctgcate tctcagcett gactccacte agttctcttg 1860
cttagagata ccacctttcc cctgaagtgt tcttccatg ttttacggcg agatgggttc 1920
tctctgectg gccactcagc cttagtgtgc tctgttgtct tatagaggtc tacttgaaga 1980
aggaaaaaca ggggtcatgg ttgactgtc ctgtgagccc ttcttccctg cctcccccac 2040
tcacagtgc ccggaatctg cagtgcctagt ctcccgaac tatcactctt tcacagtctg 2100
ctttggaagg actgggctta gtatgaaaag ttaggactga gaagaatttg aaaggcggct 2160
ttttgtagct tgatattcac tactgtctta ttaccctgtc 2200

```

```

<210> SEQ ID NO 55
<211> LENGTH: 803
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, HBB locus LHA (used for
GFP AAV)

```

```

<400> SEQUENCE: 55
gactgcatta agaggctctct agttttttac ctcttggttc ccaaaccta ataagtaact 60
aatgcacaga gcacattgat ttgtatttat tctattttta gacataatth attagcatgc 120
atgagcaaat taagaaaaac aacaacaaat gaatgcatat atatgtatat gtatgtgtgt 180
acatatacac atatatatat attttttttc ttttcttacc agaaggtttt aatccaaata 240
aggagaagat atgcttagaa ctgaggtaga gttttcatcc attctgtcct gtaagtatth 300
tgcatattct ggagacgcag gaagagatcc atctacatat cccaaagctg aattatggta 360
gacaaaaactc ttccactttt agtgcacaa tttcttattt gtgtaataag aaaattggga 420
aaacgatctt caatatgett accaagctgt gattccaaat attacgtaa tacacttgca 480
aaggaggatg ttttttagtag caatttgtag tgatggtagt gggccaagag atatatctta 540
gaggaggaggc tgagggtttg aagtccaact cctaagccag tgccagaaga gccagggaca 600
ggtagggctg tcatcactta gacctcacc tgtaggagcca caccctaggg ttggccaatc 660
tactcccagg agcaggaggg gcaggagcca gggctgggca taaaagtcag ggcagagcca 720
tctattgett acatttgett ctgacacaac tgtgttccat agcaacctca aacagacacc 780
atggtgcacg tgactcctga gga 803

```

```

<210> SEQ ID NO 56
<211> LENGTH: 720
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, GFP

```

```

<400> SEQUENCE: 56
atggtgagca agggcgagga gctgttcacc ggggtgggtc ccactctggc cgagctggac 60
ggcgacgtaa acggccacaa gtccagcgtg tccggcgagg gcgagggcga tgccacctac 120
ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180
ctcgtgacca ccctgacctc cgccgtgcag tgcttcagcc gctaccccga ccacatgaag 240
cagcagcact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc 300
ttcaaggacg acggcaacta caagaccgc gccgaggtga agttcgaggc cgacaccctg 360
gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420

```

-continued

aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac	480
ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc	540
gaccactacc agcagaacac ccccatcggc gacggccccg tgetgctgcc cgacaaccac	600
tacctgagca ccagtcgcg cctgagcaaa gaccccaacg agaagcgcg tcatatggtc	660
ctgctggagt tcgtgaccgc cgccgggac actctcggca tggacgagct gtacaagtaa	720

<210> SEQ ID NO 57
 <211> LENGTH: 920
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Donor DNA, HBB locus RHA (used for GFP AAV)

<400> SEQUENCE: 57

ctgccctgtg gggcaagggt aacgtggatg aagttggtgg tgaggccctg ggcaggttgg	60
tatcaaggtt acaagacagg ttaagggaga ccaatagaaa ctgggcatgt ggagacagag	120
aagactcttg ggtttctgat aggcactgac tctctctgcc tattggteta ttttcccacc	180
cttaggctgc tgggtgteta cctttggacc cagaggttct ttgagtcctt tggggatctg	240
tccactcctg atgctgttat gggcaaccct aaggtgaagg ctcatggcaa gaaagtgtct	300
ggtgccttta gtgatggcct ggctcacctg gacaacctca agggcacctt tgccacactg	360
agtgagctgc actgtgacaa gctgcacgtg gatcctgaga acttcagggt gagtctatgg	420
gacgcttgat gttttcttct ccttctttt ctatggttaa gttcatgtca taggaagggg	480
ataagtaaca gggtagctt tagaatggga aacagacgaa tgattgcatc agtggtgaag	540
tctcaggatc gttttagttt cttttatttg ctgttcataa caattgtttt cttttgttta	600
attcttgctt tctttttttt tcttctcgc aatttttact attatactta atgcctaac	660
attgtgtata acaaaaggaa atatctctga gatacatata gtaacttaaa aaaaaacttt	720
acacagctcg cctagtacat tactatttgg aatatatgtg tgcttatttg catattcata	780
atctccctac tttattttct tttattttta attgatacat aatcattata catatttatg	840
ggttaaagtg taatgtttta atatgtgtac acatattgac caaatcaggg taattttgca	900
tttgaattt taataaatgc	920

<210> SEQ ID NO 58
 <211> LENGTH: 451
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: MND promoter

<400> SEQUENCE: 58

ggccgccagt gtgatggata tctgcagaat tcgcccttat ggggatccga acagagagac	60
agcagaatat gggccaaaca ggatatctgt ggtaagcagt tcctgccccg gctcagggcc	120
aagaacagtt ggaacagcag aatatgggcc aaacaggata tctgtggtaa gcagttcctg	180
ccccggctca gggccaagaa cagatgggtc ccagatgcgg tcccgccctc agcagtttct	240
agagaacat cagatgttct cagggtgccc caaggacctg aaatgacct gtgccttatt	300
tgaactaacc aatcagttcg cttctcgctt ctgttcgcgc gcttctgctc cccgagctct	360

-continued

atataagcag agctcggtta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt	420
ttgacctcca tagaagacac cgactctaga g	451

<210> SEQ ID NO 59
 <211> LENGTH: 1178
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: EF1 alpha promoter

<400> SEQUENCE: 59

ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtcgccga gaagttgggg	60
ggaggggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa ctgggaaagt	120
gatgtcgtgt actggctccg cctttttccc gaggtgggg gagaacgta tataagtga	180
gtagtcgccc tgaacgttct ttttcgcaac gggtttgccc ccagaacaca ggtaagtgcc	240
gtgtgtggtt cccgcgggccc tggcctcttt acgggttatg gcccttgctg gcttgaatt	300
acttccactg gctgcagtac gtgattcttg atcccgagct tcgggttgga agtgggtggg	360
agagttcgag gccttgctgct taaggagccc cttcgcctcg tgcttgagtt gaggcctggc	420
ctgggcgctg gggccgcgcg gtgcgaatct ggtggcacct tcgcgcctgt ctgctgctt	480
tcgataagtc tctagccatt taaaattttt gatgacctgc tgcgacgctt tttttctggc	540
aagatagtct tgtaaatgcg ggccaagatc tgcacactgg tatctcggtt ttggggccg	600
cgggcggcga cggggcccggt gcgtcccagc gcacatgttc ggcgaggcgg ggctgcgag	660
cgcggccacc gagaatcgga cgggggtagt ctcaagctgg ccggcctgct ctggtgctg	720
gcctcgccgc gccgtgtatc gccccgccct gggcggcaag gctggcccg tcggcaccag	780
ttgctgagc ggaagatgg ccgcttccc gcccgtctgc agggagctca aaatggagga	840
cgcggcgcctc gggagagcgg gcgggtgagt caccacaca aaggaaaagg gcctttccgt	900
cctcagccgt cgcttcatgt gactccacgg agtacgggc gccgtccagg cactcgatt	960
agttctcgag cttttggagt acgtcgtctt taggttgggg ggaggggttt tatgcatgg	1020
agtttcccca cactgagtgg gtggagactg aagttaggcc agcttggcac ttgatgtaat	1080
tctccttgga atttgccctt ttgagtttg gatcttggtt cattotcaag cctcagacag	1140
tggttcaaag tttttttctt ccatttcagg tgctgtga	1178

<210> SEQ ID NO 60
 <211> LENGTH: 402
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: SFFV promoter

<400> SEQUENCE: 60

gtaacgccat tttgcaaggc atggaaaaat accaaaccaa gaatagagaa gttcagatca	60
agggcgggta catgaaaata gctaacgttg ggccaaacag gatattctgcg gtgagcagtt	120
tcggccccgg cccggggcca agaacagatg gtcaccgcag tttcggcccc ggcccaggc	180
caagaacaga tgggtcccag atatggccca accctcagca gtttcttaag acccatcaga	240
tgtttccagg ctcccccaag gacctgaaat gacctgcgc cttatttgaa ttaaccaatc	300
agcctgcttc tcgcttctgt tcgcgcgctt ctgcttccc agctctataa aagagctcac	360

-continued

aacccctcac tcggcgcgcc agtctctccga cagactgagt cg 402

<210> SEQ ID NO 61
<211> LENGTH: 57
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: 2A peptide from porcine teschovirus

<400> SEQUENCE: 61

gccacgaact tctctctgtt aaagcaagca ggagacgtgg aagaaaaccc cggtcct 57

<210> SEQ ID NO 62
<211> LENGTH: 49
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Synthetic poly(A) signal

<400> SEQUENCE: 62

aataaaatcg ctatccatcg aagatggatg tgtgttggtt tttgtgtg 49

<210> SEQ ID NO 63
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Canonical PAM
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(21)
<223> OTHER INFORMATION: n is any nucleotide and repeats 19 times and
up to 21 times
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (22)..(22)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 63

nnnnnnnnnn nnnnnnnnnn nng 24

<210> SEQ ID NO 64
<211> LENGTH: 3
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: SpCas9 PAM
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(1)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 64

nrg 3

<210> SEQ ID NO 65
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: SV40 NLS 1

<400> SEQUENCE: 65

Pro Lys Lys Lys Arg Lys Val
1 5

-continued

<210> SEQ ID NO 66
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: SV40 NLS 2

<400> SEQUENCE: 66

Pro Lys Lys Lys Arg Arg Val
1 5

<210> SEQ ID NO 67
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Nucleoplasmin NLS

<400> SEQUENCE: 67

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> SEQ ID NO 68
<211> LENGTH: 225
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: i53 (DNA)

<400> SEQUENCE: 68

atgctgatct tcgtgaagac cctgaccggc aagaccatca ccctggaggt ggagcccagc 60
gacaccatcg agaacgtgaa ggccaagatc caggacaagg agggcatccc ccccgaccag 120
cagaggctgg ccttcgccgg caagagcctg gaggaaggca ggaccctgag cgactacaac 180
atcctgaagg acagcaagct gcacccctcg ctgaggctga ggtga 225

<210> SEQ ID NO 69
<211> LENGTH: 225
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: I53 mRNA

<400> SEQUENCE: 69

augcugaucu ucgugaagac ccugaccggc aagaccauca ccugggaggu ggagcccagc 60
gacaccaucg agaacgugaa ggccaagauc caggacaagg agggcaucac ccccgaccag 120
cagaggcugg ccuucgccgg caagagccug gaggaaggca ggaccugag cgacuacaac 180
auccugaagg acagcaagcu gcaccccccug cugaggcuga gguga 225

<210> SEQ ID NO 70
<211> LENGTH: 74
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: i53 (aa)

<400> SEQUENCE: 70

Met Leu Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp

-continued

	20	25	30	
Lys	Glu	Gly	Ile	Pro
	35		40	
Pro	Asp	Gln	Gln	Arg
				Leu
Ala	Phe	Ala	Gly	Lys
	45			
Ser	Leu	Glu	Asp	Gly
	50		55	
Arg	Thr	Leu	Ser	Asp
				Tyr
Asn	Ile	Leu	Lys	Asp
				60
Ser	Lys	Leu	His	Pro
	65		70	
Leu	Arg	Leu	Arg	

<210> SEQ ID NO 71
 <211> LENGTH: 222
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: I53-DM (RNA)

<400> SEQUENCE: 71

auguugauuu	ucgugaaaac	ccuuaccggg	aaaaccauca	cccucgaggu	ugaacccucg	60
gauacgauag	aaaau guaaa	ggccaagauc	caggauaagg	aaggauuucc	uccugaucag	120
cagagacugg	ccuuugcugg	caaaucgcug	gaagauggac	guacuuuguc	ugacuacaau	180
auucuaaagg	acucuaaaacu	ucaucuagug	uugagacuuc	gu		222

<210> SEQ ID NO 72
 <211> LENGTH: 228
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: A10 (DNA)

<400> SEQUENCE: 72

atgcagattt	acgtgaagac	ctttgcccgg	aagcccatca	ccctcgaggt	tgaaccctcg	60
gatacgatag	aaaatgtaaa	ggccaagatc	caggataagg	aaggaattcc	tcttgatcag	120
cagcgactga	tctttgtctg	aatgcggctg	gaagatggac	gtactttgtc	tgactacaat	180
attaaaaacg	actctactct	ttttcttgtg	ttgaaaaata	gtgttact		228

<210> SEQ ID NO 73
 <211> LENGTH: 228
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: A10 (RNA)

<400> SEQUENCE: 73

augcagauuu	acgugaagac	cuuugcccgg	aagcccauca	cccucgaggu	ugaacccucg	60
gauacgauag	aaaau guaaa	ggccaagauc	caggauaagg	aaggauuucc	uccugaucag	120
cagcgacuga	ucuuugcuga	aaugcggcug	gaagauggac	guacuuuguc	ugacuacaau	180
auuaaaaaacg	acucucacucu	uuuucuuugug	uugaaaaaua	guguuacu		228

<210> SEQ ID NO 74
 <211> LENGTH: 76
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: A10 (aa)

<400> SEQUENCE: 74

Met	Gln	Ile	Tyr	Val	Lys	Thr	Phe	Ala	Arg	Lys	Pro	Ile	Thr	Leu	Glu
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

-continued

1	5	10	15
---	---	----	----

Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
 20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Glu Met
 35 40 45

Arg Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Lys Asn Asp
 50 55 60

Ser Thr Leu Phe Leu Val Leu Lys Asn Ser Val Thr
 65 70 75

<210> SEQ ID NO 75
 <211> LENGTH: 228
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: All (DNA)

<400> SEQUENCE: 75

atgctgattt tcgtgaccac cgatatgggg atgacaatct cactcgaggt tgaaccctcg	60
gatacgatag aaaatgtaaa ggccaagatc caggataagg aaggaattcc tcctgatcag	120
cagagactga tctttgtgta caaggatctg gaagatggac gtactttgtc tgactacaat	180
attcaaaagg agtctagcct taatcttggt ctgaaacttc gtggtggt	228

<210> SEQ ID NO 76
 <211> LENGTH: 228
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: All (RNA)

<400> SEQUENCE: 76

augcugauuu ucgugaccac cgauaugggg augacaauc cacucgaggu ugaaccucg	60
gauacgauag aaaauguaaa ggccaagauc caggauaagg aaggaauucc uccugaucag	120
cagagacuga ucuuugguga caaggauaug gaagauggac guacuuuguc ugacuacaau	180
auucaaaagg agucuagccu uaaucuugug cugaaacuuc gugguggu	228

<210> SEQ ID NO 77
 <211> LENGTH: 76
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: All (aa)

<400> SEQUENCE: 77

Met Leu Ile Phe Val Thr Thr Asp Met Gly Met Thr Ile Ser Leu Glu	
1 5 10 15	

Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
 20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Gly Asp Lys
 35 40 45

Asp Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
 50 55 60

Ser Ser Leu Asn Leu Val Leu Lys Leu Arg Gly Gly
 65 70 75

```

<210> SEQ ID NO 78
<211> LENGTH: 228
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: C08 (DNA)

<400> SEQUENCE: 78

atgcagattt tctgtaccac cgatatgtgg atgagaatct cactcgaggt tgaaccctcg      60
gatacgatag aaaatgtaaa ggccaagatc caggataagg aaggaattcc tctgatcag      120
cagagactga tctttgtgta caaggatctg gaagatggac gtactttgtc tgactacaat      180
attcaaaagg agtctagcct taatcttggt ctgaaccttc gtggtggt      228

<210> SEQ ID NO 79
<211> LENGTH: 228
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: C08 (RNA)

<400> SEQUENCE: 79

augcagauuu ucgugaccac cgauaugugg augagaaucu cacucgaggu ugaaccucug      60
gauacgauag aaaauguaaa ggccaagauc caggauaagg aaggaauucc uccugaucag      120
cagagacuga ucuuugguga caaggaucug gaagauggac guacuuuguc ugacuacaau      180
auucaaaagg agucuagccu uaaucuugug cugaaccuuc gugguggu      228

<210> SEQ ID NO 80
<211> LENGTH: 76
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: C08 (aa)

<400> SEQUENCE: 80

Met Gln Ile Phe Val Thr Thr Asp Met Trp Met Arg Ile Ser Leu Glu
1             5             10            15

Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
                20            25            30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Gly Asp Lys
            35            40            45

Asp Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
            50            55            60

Ser Ser Leu Asn Leu Val Leu Asn Leu Arg Gly Gly
65             70             75

<210> SEQ ID NO 81
<211> LENGTH: 228
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: G08 (DNA)

<400> SEQUENCE: 81

atgttgattt tctgtaaaac ccttaccggg aaaacatca ccctcgaggt tgaaccctcg      60
gatacgatag aaaatgtaaa ggccaagatc caggataagg aaggaattcc tctgatcag      120
cagagactga tctttgtctg caaatcgctg gaagatggac gtactttgtc tgactacaat      180

```

-continued

attctaaagg actctaaact tcctcctctg ttgagacttc gtggtggt 228

<210> SEQ ID NO 82
 <211> LENGTH: 228
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: G08 (RNA)

<400> SEQUENCE: 82

```

auguugauuu ucgugaaaac ccuuaacggg aaaaccauca cccucgaggu ugaacccucg      60
gauacgauag aaaauguaaa ggccaagauc caggauaagg aaggaauucc uccugaucag      120
cagagacuga ucuuugcugg caaaucgcug gaagauggac guacuuuguc ugacuacaau      180
auucuaaagg acucuaaaacu ucauccucug uugagacuuc gugguggu                228

```

<210> SEQ ID NO 83
 <211> LENGTH: 76
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: G08 (aa)

<400> SEQUENCE: 83

```

Met Leu Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1             5             10            15
Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
                20            25            30
Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
                35            40            45
Ser Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Leu Lys Asp
                50            55            60
Ser Lys Leu His Pro Leu Leu Arg Leu Arg Gly Gly
65              70              75

```

<210> SEQ ID NO 84
 <211> LENGTH: 228
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: H04 (DNA)

<400> SEQUENCE: 84

```

atgcgaatta tcgtgaaaac ctttatgctg aagccgatca cgctcgaggt tgaaccctcg      60
gatacgatag aaaatgtaaa ggccaagatc caggataagg aaggaattcc tcctgatcag      120
cagagactgt attttgctgc cagtcagctg gaagatggac gtactttgtc tgactacaat      180
attcaaaagg agtctactct tcttcttctg gtaaggctgc tccgcgtt                228

```

<210> SEQ ID NO 85
 <211> LENGTH: 228
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: H04 (RNA)

<400> SEQUENCE: 85

```

augcgaaaua ucgugaaaac cuuuaugcgg aagccgauca cgcucgaggu ugaacccucg      60
gauacgauag aaaauguaaa ggccaagauc caggauaagg aaggaauucc uccugaucag      120

```

-continued

cagagacugu auuuugcggc cagucagcug gaagauggac guacuuuguc ugacuacaau 180

auucaaaagg agucucucuc ucuucuugug guaaggcugc uccgcguu 228

<210> SEQ ID NO 86

<211> LENGTH: 76

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: H04 (aa)

<400> SEQUENCE: 86

Met Arg Ile Ile Val Lys Thr Phe Met Arg Lys Pro Ile Thr Leu Glu
1 5 10 15Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
20 25 30Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Tyr Phe Ala Ala Ser
35 40 45Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
50 55 60Ser Thr Leu Leu Leu Val Val Arg Leu Leu Arg Val
65 70 75

<210> SEQ ID NO 87

<211> LENGTH: 222

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: I53 alt (DNA)

<400> SEQUENCE: 87

atgttgattt tcgtgaaaac ccttaccggg aaaaccatca ccctcgaggt tgaaccctcg 60

gatacgatag aaaatgtaaa ggccaagatc caggataagg aaggaattcc tcctgatcag 120

cagagactgg cctttgctgg caaatcgctg gaagatggac gtactttgtc tgactacaat 180

attctaaagg actctaaact tcatcctctg ttgagacttc gt 222

<210> SEQ ID NO 88

<211> LENGTH: 222

<212> TYPE: RNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: I53 alt (RNA)

<400> SEQUENCE: 88

auguugauuu ucgugaaaac ccuaccggg aaaaccauca ccucgaggu ugaaccucg 60

gauacgauag aaaauguaaa ggccaagauc caggauaagg aaggauuucc uccugaucag 120

cagagacugg ccuucgugg caaaucgcug gaagauggac guacuuuguc ugacuacaau 180

auucuaaagg acucuaaaacu ucauccucug uugagacuuc gu 222

<210> SEQ ID NO 89

<211> LENGTH: 291

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: FLAG-tagged i53 DNA

<400> SEQUENCE: 89

-continued

atggactaca aagacgatga cgataaagcc gccagtttaa acggcgcgcc attaattaag	60
gatccaatgt tgattttcgt gaaaaccctt accgggaaaa ccatcaccct cgaggttgaa	120
ccctcgata cgatagaaaa tgtaaaggcc aagatccagg ataaggaagg aattcctcct	180
gatcagcaga gactggcctt tgctggcaaa tcgctggaag atggacgtac tttgtctgac	240
tacaatatc taaaggactc taaacttcat cctctgttga gacttcgttg a	291

<210> SEQ ID NO 90
 <211> LENGTH: 291
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: FLAG-tagged i53 RNA

<400> SEQUENCE: 90

auggacuaca aagacgauga cgauaaagcc gccaguuaa acggcgcgcc auuaauuaag	60
gauccaaugu ugauuuucgu gaaaaccuu accgggaaaa ccaucacccu cgagguugaa	120
cccucggaua cgauagaaaa uguaaaggcc aagauccagg auaaggaagg aaauccuccu	180
gaucagcaga gacuggccuu ugcuggcaaa ugcuggaag auggacguac uuugucugac	240
uacaauauuc uaaaggacuc uaaacuucan ccucuguuga gacuucguug a	291

<210> SEQ ID NO 91
 <211> LENGTH: 39
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Linker (DNA)

<400> SEQUENCE: 91

gccgccagtt taaacggcgc gccattaatt aaggatcca	39
--	----

<210> SEQ ID NO 92
 <211> LENGTH: 39
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Linker (RNA)

<400> SEQUENCE: 92

gccgccaguu uaaacggcgc gccauuaauu aaggaucca	39
--	----

<210> SEQ ID NO 93
 <211> LENGTH: 13
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Linker

<400> SEQUENCE: 93

Ala Ala Ser Leu Asn Gly Ala Pro Leu Ile Lys Asp Pro
1 5 10

<210> SEQ ID NO 94
 <211> LENGTH: 6
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic: Protein tag, 6xHis

<400> SEQUENCE: 94

-continued

His His His His His His
1 5

<210> SEQ ID NO 95
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Protein tag, Flag

<400> SEQUENCE: 95

Met Asp Tyr Lys Asp Asp Asp Lys
1 5

<210> SEQ ID NO 96
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Protein tag, FLAG (DNA)

<400> SEQUENCE: 96

gactacaaag acgatgacga taaa 24

<210> SEQ ID NO 97
<211> LENGTH: 24
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Protein tag, FLAG (RNA)

<400> SEQUENCE: 97

gacuacaaag acgaugacga uaaa 24

<210> SEQ ID NO 98
<211> LENGTH: 4435
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor DNA, HBB locus LHA to RHA
(used for E6V to E6, AAV.323)

<400> SEQUENCE: 98

cttgtcttga caatttttgt ctttcagaat actataaata taacctatat tataatttca 60
taaagtctgt gcattttctt tgaccagga tatttgcaaa agacatattc aaacttcgc 120
agaacacttt atttcacata tacatgcctc ttatatcagg gatgtgaaac agggctctga 180
aaactgtcta aatctaaaac aatgctaag caggtttaaa ttaataaaa taaaatccaa 240
aatctaacag ccaagtcaaa tctgcatgtt ttaacattta aaatatttta aagacgtctt 300
ttcccaggat tcaacatgtg aaatcttttc tcaggagatc acgtgtgcct agatcctcat 360
tgcttttagtt ttttacagag gaatgaatat aaaaagaaaa tacttaaatt ttatccctct 420
tacctctata atcacacata ggcataattt tttaacctag gctccagata gccatagaag 480
aaccaaaacac tttctgcgtg tgtgagaata atcagagtga gattttttca caagtacctg 540
atgaggggtg agacaggtag aaaaagttag agatctctat ttatttagca ataatagaga 600
aagcatttaa gagaataaag caatggaaat aagaaatttg taaatttcct tctgataact 660
agaaatagag gatccagttt cttttgggta acctaaattt tatttcattt tattgtttta 720

-continued

ttttatttta	ttttatttta	ttttgtgtaa	tcgtagtttc	agagtgttag	agctgaaagg	780
aagaagtagg	agaacatgc	aaagtaaaag	tataacactt	tccttactaa	accgacatgg	840
gtttccaggt	aggggcagga	ttcaggatga	ctgacagggc	ccttagggaa	cactgagacc	900
ctacgctgac	ctcataaatg	cttgctacct	ttgctgtttt	aattacatct	tttaatagca	960
ggaagcagaa	ctctgcactt	caaaagtttt	tcctcacctg	aggagttaat	ttagtacaag	1020
gggaaaaagt	acagggggat	gggagaaaag	cgatcacgtt	gggaagctat	agagaaagaa	1080
gagtaaat	tagtaaaagga	ggtttaaaca	aacaaaatat	aaagagaaat	aggaacttga	1140
atcaaggaaa	tgatttttaa	acgcagtatt	cttagtggac	tagaggaaaa	aaataatctg	1200
agccaagtag	aagacctttt	ccctcctac	ccctactttc	taagtcacag	aggctttttg	1260
ttcccccaga	cactcttgca	gattagtcca	ggcagaaaca	gtagatgtc	cccagttaac	1320
ctcctatttg	acaccactga	ttccccatt	gatagtcaca	ctttgggttg	taagtgactt	1380
tttatttatt	tgtatttttg	actgcattaa	gaggtctcta	gttttttate	tcttgtttcc	1440
caaaacctaa	taagtaacta	atgcacagag	cacattgatt	tgtatttatt	ctatttttag	1500
acataattta	ttagcatgca	tgagcaaat	aagaaaaaca	acaacaaatg	aatgcatata	1560
tatgtatatg	tatgtgtgta	tatatacaca	catatatata	tatatTTTTT	cttttcttac	1620
cagaagggtt	taatccaaat	aaggagaaga	tatgcttaga	accgaggtag	agttttcatc	1680
cattctgtcc	tgtaagtatt	ttgcatattc	tggagacgca	ggaagagatc	catctacata	1740
tcccaaagct	gaattatggt	agacaaaact	cttccacttt	tagtgcacga	acttcttatt	1800
tgtgtaataa	gaaaattggg	aaaacgatct	tcaatatgct	taccaagctg	tgattccaaa	1860
tattacgtaa	atacacttgc	aaaggaggat	gttttttagta	gcaatttgta	ctgatgggat	1920
ggggccaaga	gatatatctt	agagggaggg	ctgagggttt	gaagtccaac	tcctaagcca	1980
gtgccagaag	agccaaggac	aggtagggct	gtcatcactt	agacctcacc	ctgtggagcc	2040
acaccctagg	gttggccaat	ctactcccag	gagcagggag	ggcaggagcc	agggtggggc	2100
ataaaagtca	gggcagagcc	atctattgct	tacatttgct	tctgacacaa	ctgtgttcac	2160
tagcaacctc	aaacagacac	catggtgcat	ctgactcctg	aagaaaaatc	cgctgtcact	2220
gccctgtggg	gcaagggtgaa	cgtggatgaa	gttgggtgtg	aggccctggg	cagggtggta	2280
tcaaggttac	aagacaggtt	taaggagacc	aatagaaact	gggcagtgtg	agacagagaa	2340
gactcttggg	tttctgatag	gcactgactc	tctctgccta	ttggtctatt	ttcccacct	2400
taggtctgtg	gtggtctacc	cttggaccca	gaggttcttt	gagtcctttg	gggatctgtc	2460
cactcctgat	gctgttatgg	gcaaccctaa	ggtgaaggct	catggcaaga	aagtgtctcg	2520
tgcctttagt	gatggcctgg	ctcacctgga	caacctcaag	ggcacctttg	ccacactgag	2580
tgagctgcac	tgtgacaagc	tgacagtgga	tcctgagaac	ttcaggggtga	gtctatggga	2640
cgcttgatgt	tttctttccc	cttcttttct	atgggttaagt	tcatgtcata	ggaaggggat	2700
aagtaacagg	gtacagttta	gaatgggaaa	cagacgaatg	attgcatcag	tgtggaagtc	2760
tcaggatcgt	tttagtttct	tttatttgct	gttcataaca	attgttttct	ttgtttaat	2820
tcttgcttct	tttttttttc	ttctccgcaa	tttttactat	tatacttaat	gccttaacat	2880
tgtgtataac	aaaaggaaat	atctctgaga	tacattaagt	aacttaaaaa	aaaactttac	2940
acagtctgcc	tagtacatta	ctatttggaa	tatatgtgtg	cttatttgca	tattcataat	3000

-continued

ctccctactt tattttcttt tatttttaaat tgatacataa tcattatata tatttatggg	3060
ttaaagtgtg atgttttaaat atgtgtacac atattgacca aatcagggtg attttgcatt	3120
tgtaatttta aaaaatgctt tcttctttta atatactttt ttgtttatct tatttctaata	3180
actttcccta atctctttct ttcagggtcaa taatgataca atgtatcatg cctctttgca	3240
ccattotaaa gaataacagt gataatttct gggttaaggc aatagcaata tctctgcata	3300
taaatatttc tgcataataa ttgtaactga tgtaagaggt ttcattatgc taatagcagc	3360
tacaatccag ctaccattct gcttttattt tatggttggg ataaggctgg attattctga	3420
gtccaagcta ggcctttttg ctaatcatgt tcataacctc tatcttctc ccacagctcc	3480
tgggcaacgt gctggtctgt gtgctggccc atcactttgg caaagaattc accccaccag	3540
tgagggtgc ctatcagaaa gtgggtggctg gtgtggctaa tgccctggcc cacaagtatc	3600
actaagctcg ctttcttctgt gtccaatttc tattaagggt tcctttgttc cctaagtcca	3660
actactaaac tgggggatat tatgaagggc ctgagcctc tggattctgc ctaataaaaa	3720
acatttattt tcattgcaat gatgtattta aattatttct gaattattta ctaaaaaggg	3780
aatgtgggag gtcagtgcatt ttaaaacata aagaaatgaa gagctagtcc aaacctggg	3840
aaaatacact atacttataa ctccatgaaa gaaggtgagg ctgcaaacag ctaatgcaca	3900
ttggcaacag cccctgatgc atatgcctta ttcctccctc agaaaaggat tcaagtagag	3960
gcttgatttg gaggttaaag ttttgctatg ctgtatttta cattacttat tgttttagct	4020
gtcctcatga atgtcttttc actaccatt tgccttacct gcctctctca gccttgactc	4080
cactcagttc tcttgcttag agataccacc tttccctga agtgctcctt ccattgttta	4140
cggcgagatg gtttctctc gcctggccac tcagccttag ttgtctctgt tgtcttatag	4200
aggtctactt gaagaaggaa aaacaggggt catggtttga ctgtctctgt agcccttctt	4260
ccctgcctcc cccactcaca gtgaccgga atctgcagtg ctagtctccc ggaactatca	4320
ctctttcaca gtctgctttg gaaggactgg gcttagtatg aaaagttagg actgagaaga	4380
atttgaaagg cggtcttttg tagcttgata ttcactactg tcttattacc ctgtc	4435

<210> SEQ ID NO 99

<211> LENGTH: 2199

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, HBB locus LHA (used for E6V to E6, AAV.323)

<400> SEQUENCE: 99

cttgctttga caattttggg ctttcagaat actataaata taacctatat tataatttca	60
taaagtctgt gcattttctt tgaccagga tatttgcaa agacatattc aaacttcgc	120
agaacacttt atttcacata tacatgctc ttatatcagg gatgtgaaac agggctttga	180
aaactgtcta aatctaaaac aatgctaagt caggtttaaa ttaataaaa taaaatccaa	240
aatctaacag ccaagtcaaa tctgcatgtt ttaacattta aaatatttta aagacgtctt	300
ttcccaggat tcaacatgtg aaatcttttc tcagggtatc acgtgtgcct agatcctcat	360
tgcttttagtt ttttacagag gaatgaatat aaaaagaaaa tacttaaat ttatccctct	420
tacctctata atcatacata ggcatattt tttaacctag gctccagata gccatagaag	480
aaccaaaac tttctgcgtg tgtgagaata atcagagtga gattttttca caagtacctg	540

-continued

atgaggggtg agacaggtag aaaaagttag agatctctat ttatttagca ataatagaga	600
aagcatttaa gagaataaag caatggaaat aagaaatttg taaatttcct tctgataact	660
agaaatagag gatccagttt cttttggtta acctaaattt tatttcattt tattgtttta	720
ttttatttta ttttatttta ttttgtgtaa tcgtagtctc agagtgttag agctgaaag	780
aagaagtagg agaacatgc aaagtaaaag tataaacatt tccttactaa accgacatgg	840
gtttccaggt aggggcagga ttcaggatga ctgacagggc ccttagggaa cactgagacc	900
ctacgctgac ctcataaatg cttgctacct ttgctgtttt aattacatct tttaatagca	960
ggaagcagaa ctctgcactt caaaagtttt tcctcacctg aggagttaat ttagtacaag	1020
gggaaaaagt acagggggat gggagaaagg cgtcacggtt gggaagctat agagaagaa	1080
gagtaatttt tagtaaggga ggtttaaaca aacaaaatat aaagagaaat aggaacttga	1140
atcaaggaaa tgattttaaa acgcagtatt cttagtggac tagaggaaaa aaataatctg	1200
agccaagtag aagacctttt cccctcctac cctacttttc taagtcacag aggccttttg	1260
ttccccaga cactcttgca gattagtcca ggcagaaaca gttagatgac ccagttaac	1320
ctcctatttg acaccactga ttacccatt gatagtcaca ctttgggttg taagtactt	1380
tttattttatt tgtatttttg actgcattaa gaggtctcta gttttttatc tcttgtttcc	1440
caaaacctaa taagtaacta atgcacagag cacattgatt tgtattttatt ctatttttag	1500
acataattta ttagcatgca tgagcaaatt aagaaaaaca acaacaaatg aatgcatata	1560
tatgtatatg tatgtgtgta tatatacaca catatatata tatatttttt cttttcttac	1620
cagaaggttt taatccaaat aaggagaaga tatgcttaga accgaggtag agttttcacc	1680
cattctgtcc tgtaagtatt ttgcatattc tggagacgca ggaagagatc catctacata	1740
tcccaaagct gaattatggg agacaaaact cttccacttt tagtgcatca acttcttatt	1800
tgtgtaataa gaaaattggg aaaacgatct tcaatatgct taccaagctg tgattccaaa	1860
tattacgtaa atacacttgc aaaggaggat gtttttagta gcaatttgta ctgatgggat	1920
ggggccaaga gatatatctt agagggaggg ctgaggggtt gaagtccaac tcctaagcca	1980
gtgccagaag agccaaggac aggtacggct gtcacactt agacctcacc ctgtggagcc	2040
acaccctagg gttggccaat ctactcccag gagcaggagc ggcaggagcc agggctgggc	2100
ataaaagtca gggcagagcc atctattgct tacatttgct tctgacacaa ctgtgttcac	2160
tagcaacctc aaacagacac catggtgcat ctgactcct	2199

<210> SEQ ID NO 100

<211> LENGTH: 2218

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA , HBB locus RHA (used for E6V to E6, AAV.323)

<400> SEQUENCE: 100

actgccctgt ggggcaaggt gaacgtggat gaagttggtg gtgaggccct gggcagggtg	60
gtatcaaggt tacaagacag gtttaaggag accaatagaa actgggcatg tgagacaga	120
gaagactctt gggtttctga taggcactga ctctctctgc ctattggtct attttccac	180
ccttaggctg ctggtggtct acccttgac ccagaggctc tttgagctct ttggggatct	240

-continued

gtccactcct gatgctgtta tgggcaaccc taaggtgaag gctcatggca agaaagtgt	300
cgggtgccttt agtgatggcc tggtcacct ggacaaccc aagggcacct ttgccacact	360
gagtgagctg cactgtgaca agctgcacgt ggatcctgag aacttcaggg tgagtctatg	420
ggacgcttga tgttttcttt ccccttcttt tctatggta agttcatgtc ataggaaggg	480
gataagtaac agggtagagt ttagaatggg aaacagacga atgattgcat cagtgtggaa	540
gtctcaggat cgttttagtt tcttttattt gctgttcata acaattgttt tcttttgttt	600
aattcttget ttcttttttt ttcttctcgg caatttttac tattatactt aatgccttaa	660
cattgtgtat aacaaaagga aatatctctg agatacatta agtaacttaa aaaaaaactt	720
tacacagtct gcctagtaca ttactatttg gaatatatgt gtgcttattt gcatattcat	780
aatctcccta ctttattttt ttttattttt aattgataca taatcattat acatatttat	840
gggttaaagt gtaatgtttt aatatgtgta cacatattga ccaaatcagg gtaattttgc	900
atttgaatt ttaaaaaatg ctttctctt ttaataact ttttgttta tcttatttct	960
aatactttcc ctaatctctt tcttcaggg caataatgat acaatgtatc atgcctcttt	1020
gcaccattct aaagaataac agtgataatt tctgggttaa ggcaatagca atatctctgc	1080
atataaatat ttctgcatat aaattgtaac tgatgtaaga ggtttcatat tgctaatagc	1140
agctacaatc cagctaccat tctgctttta ttttatggtt gggataaggc tggattattc	1200
tgagtccaag ctaggccctt ttgctaataca tgttcatacc tcttatcttc ctcccacagc	1260
tcctgggcaa cgtgctggtc tgtgtgctgg cccatcactt tggcaaagaa ttcaccccac	1320
cagtgcaggc tgcctatcag aaagtgggtg ctggtgtggc taatgccctg gccacaagt	1380
atcactaagc tcgctttctt gctgtccaat ttctattaaa gggtcctttg ttccctaagt	1440
ccaactacta aactggggga tattatgaag ggccttgagc atctggattc tgctaataa	1500
aaaacattta ttttcattgc aatgatgtat ttaaatatt tctgaatatt ttactaaaa	1560
gggaatgtgg gaggtcagtg catttaaac ataaagaaat gaagagctag ttcaaacctt	1620
gggaaaatac actatatctt aaactccatg aaagaagggt aggctgcaa cagctaattgc	1680
acattggcaa cagcccctga tgcataatgc ttattcatcc ctcaaaaag gattcaagta	1740
gaggcttgat ttggagggtta aagttttgct atgctgtatt ttacattact tattgtttta	1800
gctgtctca tgaatgtctt ttactaccc atttgcttat cctgcatctc tcagccttga	1860
ctccactcag ttctcttget tagagatacc acctttcccc tgaagtgttc ctccatgtt	1920
ttacggcgag atgggttctc ctgcctggc cactcagcct tagttgtctc tgttgtctta	1980
tagagggtcta cttgaagaag gaaaaacagg ggtcatggtt tgactgtcct gtgagccctt	2040
cttcctgcc tccccactc acagtgacct ggaatctgca gtgctagtct cccggaacta	2100
tcactcttct acagtctgct ttggaaggac tgggcttagt atgaaaagtt aggactgaga	2160
agaatttgaa aggcggcttt ttgtagcttg atattcacta ctgtcttatt acctgtc	2218

<210> SEQ ID NO 101

<211> LENGTH: 1421

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, HBB exons 1-3

<400> SEQUENCE: 101

-continued

atggtgcatc tgactcctga agaaaaatcc gctgtcactg ccctgtgggg caaggtgaac	60
gtggatgaag ttggtgggga ggccctgggc aggttggat caaggttaca agacaggttt	120
aaggagacca atagaaactg ggcatgtgga gacagagaag actcttgggt ttctgatagg	180
cactgactct ctctgcctat tggctctatt tcccaccctt aggctgctgg tggctaccc	240
ttggaccag aggttctttg agtcctttgg ggatctgtcc actcctgatg ctgttatggg	300
caaccctaag gtgaaggctc atggcaagaa agtgctcggt gccttttagt atggcctggc	360
tcacctggac aacctcaagg gcacctttgc cacactgagt gagctgcact gtgacaagct	420
gcacgtggat cctgagaact tcagggtgag tctatgggac gcttgatgtt ttctttcccc	480
ttcttttcta tggtaagtt catgtcatag gaaggggata agtaacaggg tacagtttag	540
aatgggaaac agacgaatga ttgcatcagt gtggaagtct caggatcggt ttagtttctt	600
ttatttgctg ttcataacaa ttgttttctt ttgtttaatt ctgctttct tttttttct	660
tctccgaat ttttactatt atacttaatg ccttaacatt gtgtataaca aaaggaaata	720
tctctgagat acattaagta acttaaaaaa aaactttaca cagtctgcct agtacattac	780
tatttggaat atatgtgtgc ttatttgcac attcataatc tccctacttt attttctttt	840
atttttaatt gatacataat cattatacat atttatgggt taaagtgtaa tgttttaata	900
tgtgtacaca tattgaccaa atcagggtaa ttttgcattt gtaattttaa aaaatgcttt	960
cttcttttaa tatacttttt tgtttatctt atttctaata ctttccctaa tctctttctt	1020
tcagggcaat aatgatacaa tgtatcatgc ctctttgcac cattctaaag aataacagt	1080
ataatttctg ggttaaggca atagcaatat ctctgcatat aaatatttct gcatataaat	1140
tgtaactgat gtaagagggt tcatattgct aatagcagct acaatccagc taccattctg	1200
cttttatttt atggttggga taaggctgga ttattctgag tccaagctag gcccttttgc	1260
taatcatgtt catacctott atcttcctcc cacagctcct gggcaacgtg ctggtctgtg	1320
tgctggccca tcactttggc aaagaattca cccaccagt gcaggctgcc tatcagaaag	1380
tgggtggctgg tgtggctaata gccctggccc acaagtatca c	1421

<210> SEQ ID NO 102

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor DNA, E6V to E6 Insert

<400> SEQUENCE: 102

gaagaaaaat ccgctgtc	18
---------------------	----

<210> SEQ ID NO 103

<211> LENGTH: 7107

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor template, Full AAV-323

<400> SEQUENCE: 103

cctgcaggca gctgcgcgct cgctcgtca ctgaggccgc ccgggcgtcg ggcgaccttt	60
ggtcgccccg cctcagttag cgagcgagcg cgcagagagg gagtggccaa ctccatcact	120
aggggttctt gcggccgcac gcgtcttgct ttgacaattt tggcttttca gaatactata	180

-continued

aatataacct	atattataat	ttcataaagt	ctgtgcattt	tctttgaccc	aggatatttg	240
caaaagacat	attcaaaactt	cgcagaaca	ctttatttca	catatacatg	cctcttatat	300
cagggatgtg	aaacagggtc	ttgaaaactg	tctaaatcta	aaacaatgct	aatgcagggt	360
taaatttaat	aaaataaaat	ccaaaatcta	acagccaagt	caaactctga	tgttttaaca	420
tttaaaatat	tttaagacg	tcttttccca	ggattcaaca	tgtgaaatct	tttctcaggg	480
atacacgtgt	gcctagatcc	tcattgcttt	agttttttac	agaggaatga	atataaaaag	540
aaaatactta	aattttatcc	ctcttacctc	tataatcata	cataggcata	attttttaac	600
ctaggctcca	gatagccata	gaagaaccaa	acactttctg	cgtgtgtgag	aataatcaga	660
gtgagatttt	ttcacaagta	cctgatgagg	gttgagacag	gtagaaaaag	tgagagatct	720
ctatttattt	agcaataata	gagaaagcat	ttaagagaat	aaagcaatgg	aaataagaaa	780
tttgtaaatt	tccttctgat	aactagaaat	agaggatcca	gtttcttttg	gttaacctaa	840
attttatttc	attttattgt	tttattttat	tttattttat	tttattttgt	gtaatcgtag	900
tttcagagt	ttagagctga	aaggaagaag	taggagaaac	atgcaaagta	aaagtataac	960
actttcctta	ctaaaccgac	atgggtttcc	aggtaggggc	aggattcagg	atgactgaca	1020
gggcccttag	ggaacactga	gacctacgc	tgacctcata	aatgcttgct	acctttgctg	1080
ttttaattac	atcttttaat	agcaggaagc	agaactctgc	acttcaaaag	tttttcctca	1140
cctgaggagt	taatttagta	caaggggaaa	aagtacaggg	ggatgggaga	aaggcgatca	1200
cgttgggaag	ctatagagaa	agaagagtaa	attttagtaa	aggaggttta	aacaaacaaa	1260
atataaagag	aataggaac	ttgaatcaag	gaaatgattt	taaaacgcag	tattcttagt	1320
ggactagagg	aaaaaaataa	tctgagccaa	gtagaagacc	ttttcccttc	ctaccctac	1380
tttctaagtc	acagaggctt	tttgttcccc	cagacactct	tgagatttag	tccaggcaga	1440
aacagttaga	tgtccccagt	taacctocta	tttgacacca	ctgattaccc	cattgatagt	1500
cacactttgg	gttgtaagtg	actttttatt	tatttgattt	tttgactgca	ttaagaggtc	1560
tctagttttt	tatctcttgt	ttcccaaaac	ctaataagta	actaatgcac	agagcacatt	1620
gatttgattt	tattctattt	ttagacataa	tttattagca	tgcatgagca	aattaagaaa	1680
aacaacaaca	aatgaatgca	tatatatgta	tatgtatgtg	tgtatatata	cacacatata	1740
tatatatatt	ttttcttttc	ttaccagaag	gttttaatcc	aaataaggag	aagatatgct	1800
tagaaccgag	gtagagtttt	catccattct	gtcctgtaag	tattttgcat	attctggaga	1860
cgcaggaaga	gatccatcta	catatcccaa	agctgaatta	tggtagacaa	aactcttcca	1920
cttttagtgc	atcaacttct	tatttggtga	ataagaaaat	tgggaaaacg	atcttcaata	1980
tgcttaccaa	gctgtgatcc	caaataattac	gtaaatacac	ttgcaaagga	ggatgttttt	2040
agtagcaatt	tgtactgatg	gtatggggcc	aagagatata	tcttagaggg	agggtgagg	2100
gtttgaagtc	caactcctaa	gccagtgcc	gaagagccaa	ggacaggtag	ggctgtcatc	2160
acttagacct	cacctgtgg	agccacaccc	tagggttggc	caatctactc	ccaggagcag	2220
ggagggcagg	agccagggct	gggcataaaa	gtcagggcag	agccatctat	tgcttacatt	2280
tgcttctgac	acaactgtgt	tcactagcaa	cctcaaacag	acaccatggt	gcatctgact	2340
cctgaagaaa	aatccgctgt	cactgccttg	tggggcaagg	tgaacgtgga	tgaagttggt	2400
ggtgaggccc	tgggcagggt	ggtatcaagg	ttacaagaca	ggtttaagga	gaccaataga	2460

-continued

aactgggcat	gtggagacag	agaagactct	tgggtttctg	ataggcactg	actctctctg	2520
cctattggtc	tattttccca	cccttaggct	gctgggtggtc	tacccttgga	cccagaggtt	2580
ctttgagtc	tttggggatc	tgtccactcc	tgatgctggt	atgggcaacc	ctaaggtgaa	2640
ggctcatggc	aagaaagtgc	tcggtgcctt	tagtgatggc	ctggctcacc	tggacaacct	2700
caagggcacc	tttgccacac	tgagttagct	gcactgtgac	aagctgcacg	tggatcctga	2760
gaacttcagg	gtgagtctat	gggacgcttg	atgttttctt	tccccttctt	ttctatgggt	2820
aagttcatgt	cataggaagg	ggataagtaa	cagggtagag	tttagaatgg	gaaacagacg	2880
aatgattgca	tcagtgtgga	agtctcagga	tcgttttagt	ttcttttatt	tgctgttcat	2940
aacaattggt	ttcttttggt	taattcttgc	tttctttttt	tttcttctcc	gcaattttta	3000
ctattatact	taatgcctta	acattgtgta	taacaaaagg	aaatatctct	gagatacatt	3060
aagtaactta	aaaaaaaaact	ttcacacagtc	tgccctagtag	attactatct	ggaatatatg	3120
tgtgcttatt	tgcatattca	taatctccct	actttatctt	cttttatctt	taattgatac	3180
ataatcatta	tacatattta	tgggttaaag	tgtaatgttt	taatatgtgt	acacatattg	3240
accaaatacag	ggtaattttg	catttgtaat	tttaaaaaat	gctttcttct	tttaataatac	3300
ttttttgttt	atcttatttc	taatacttcc	cctaactctct	ttctttcagg	gcaataatga	3360
tacaatgtat	catgcctctt	tgaccatttc	taaagaataa	cagtataaat	ttctgggtta	3420
aggcaatagc	aatatctctg	catataaata	tttctgcata	taaattgtaa	ctgatgtaag	3480
aggtttcata	ttgctaatag	cagctacaat	ccagctacca	ttctgctttt	attttatggg	3540
tgggataagg	ctggattatt	ctgagtcaca	gctaggccct	tttgctaata	atgttcatac	3600
ctcttatctt	cctccacacg	ctcctgggca	acgtgctggt	ctgtgtgctg	gcccacact	3660
ttggcaaaga	attcacccca	ccagtgcagg	ctgcctatca	gaaagtgggtg	gctgggtggtg	3720
ctaataccct	ggcccacaag	tatcactaag	ctcgctttct	tgctgtccaa	ttctatatta	3780
aggttccttt	gttccctaag	tccaactact	aaactggggg	atattatgaa	gggccttgag	3840
catctggatt	ctgcctaata	aaaaacattt	attttcattg	caatgatgta	tttaaatatt	3900
ttctgaatat	tttactaaaa	agggaatgtg	ggaggtcagt	gcatttaaaa	cataaagaaa	3960
tgaagagcta	gttcaaacct	tgggaaaata	cactatatct	taaactccat	gaaagaagggt	4020
gaggctgcaa	acagctaatg	cacattggca	acagccctg	atgcataatg	cttattcacc	4080
cctcagaaaa	ggattcaagt	agaggcttga	tttgagggtt	aaagttttgc	tatgctgtat	4140
tttacattac	ttattgtttt	agctgtcttc	atgaatgtct	tttactacc	catttgctta	4200
tctgcatct	ctcagccttg	actccactca	gttctcttgc	ttagagatac	cacctttccc	4260
ctgaagtgtt	ccttccatgt	tttacggcga	gatggtttct	cctcgccctg	ccactcagcc	4320
ttagtgtct	ctgttgtctt	atagaggtct	acttgaagaa	ggaaaaacag	gggtcatggt	4380
ttgactgtcc	tgtgagccct	tcttccctgc	ctccccact	cacagtgacc	cggaaatctgc	4440
agtgtatgct	tcccgaact	atcactcttt	cacagtctgc	tttggaaagga	ctgggcttag	4500
tatgaaaagt	taggactgag	aagaatttga	aaggcggtt	ttttagctt	gatattcact	4560
actgtcttat	tacctgtctg	gtaaccacgt	gcggccgagg	ctgcagcgtc	gtcctcccta	4620
ggaaccccta	gtgatggagt	tggccactcc	ctctctgcgc	gctcgctcgc	tcactgaggc	4680
cgggcgacca	aaggtcgccc	gacgccggg	ctttgcccg	gcggccctcag	tgagcgagcg	4740

-continued

agcgcgcagc	tgctgcagg	ggcgctgat	gcggtat	ctccttacgc	atctgtgcgg	4800
tatttcacac	cgcatacgtc	aaagcaacca	tagtacgcgc	cctgtagcgg	cgcattaagc	4860
gcgcggggtg	tggtggttac	gcgagcgtg	accgtacac	tgccagcgc	cctagcgccc	4920
gctcctttcg	ctttcttccc	ttcctttctc	gccacgttcg	cggctttcc	cgtcaagct	4980
ctaaatcggg	ggctcccttt	agggttcgca	tttagtgctt	tacggcacct	cgaccccaaa	5040
aaacttgatt	tggtgatgg	ttcacgtagt	gggccatcgc	cctgatagac	ggtttttcgc	5100
cctttgacgt	tggagtccac	gttctttaat	agtggactct	tggtccaaac	tggaacaaca	5160
ctcaacccta	tctcgggcta	ttcttttgat	ttataaggga	ttttgccgat	ttcggcctat	5220
tggttaaaaa	atgagctgat	ttaacaaaaa	tttaacgcga	attttaacaa	aatattaacg	5280
tttacaattt	tatggtgcac	tctcagtaca	atctgctctg	atgccgcata	gttaagccag	5340
ccccgacacc	cgccaacacc	cgtgacgcg	cctgacggg	cttgtctgct	cccgcatcc	5400
gcttacagac	aagctgtgac	cgtctccggg	agctgcattg	gtcagagggt	ttcacggtca	5460
tcaccgaaac	gcgcgagacg	aaagggcctc	gtgatacgc	tatttttata	ggtaaatgtc	5520
atgaacaata	aaactgtctg	cttacataaa	cagtaataca	aggggtgtta	tgagccatat	5580
tcaacgggaa	acgtcgaggc	cgcgattaaa	ttccaacatg	gatgctgatt	tatatgggta	5640
taaatgggct	cgcgataatg	tcgggcaatc	aggtgcgaca	atctatcgct	tgtatgggaa	5700
gcccgatgcg	ccagagtgtg	ttctgaaaca	tggcaaaggt	agcgttgcca	atgatgttac	5760
agatgagatg	gtcagactaa	actggctgac	ggaatttatg	cctcttccga	ccatcaagca	5820
ttttatccgt	actcctgatg	atgcatgggt	actcaccact	gcgatccccg	gaaaaacagc	5880
attccaggta	ttagaagaat	atcctgatcc	agggtgaaat	attgttgatg	cgctggcagt	5940
gttctcgcgc	cgggtgcatt	cgattcctgt	ttgtaattgt	ccttttaaca	gcgatcgctg	6000
atttcgtctc	gctcaggcgc	aatcacgaat	gaataacggg	ttggttgatg	cgagtgat	6060
tgatgacgag	cgtaaatggc	ggcctgttga	acaagtctgg	aaagaaatgc	ataaactttt	6120
gccattctca	ccggattcag	tcgtcactca	tggtgat	tcacttgata	accttatttt	6180
tgacgagggg	aaattaatag	gttgatttga	tggtggacga	gtcggaatcg	cagaccgata	6240
ccagatcctt	gccatcctat	ggaactgcct	cgggtgagttt	tctccttcat	tacagaaacg	6300
gctttttcaa	aaatatggta	ttgataatcc	tgatatgaat	aaattgcagt	ttcatttgat	6360
gctcgtgag	tttttcta	ctcatgacca	aaatccctta	acgtgagttt	tcgttccact	6420
gagcgtcaga	ccccgtagaa	aagatcaaag	gatcttcttg	agatcctttt	tttctgcgcg	6480
taatctgctg	cttgcaaaaca	aaaaaaccac	cgtaccagc	gggtggtttgt	ttgccggatc	6540
aagagctacc	aactcttttt	ccgaaggtaa	ctggcttcag	cagagcgcag	ataccaaata	6600
ctgtccttct	agtgtagccg	tagttaggcc	accacttcaa	gaactctgta	gcaccgccta	6660
catacctcgc	tctgctaate	ctgttaccag	tggtgctg	cagtggcgat	aagtcgtgtc	6720
ttaccgggtt	ggactcaaga	cgatagttac	cggataaggc	gcagcggtcg	ggctgaacgg	6780
ggggttcgtg	cacacagccc	agcttggagc	gaacgacctc	caccgaactg	agatacctac	6840
agcgtgagct	atgagaaagc	gccacgcttc	ccgaaggag	aaaggcggac	aggtatccgg	6900
taagcggcag	ggtcggaaca	ggagagcgca	cgagggagct	tccaggggga	aacgcctggt	6960
atctttatag	tcctgtcggg	tttcgccacc	tctgacttga	gcgtcgat	ttgtgatgct	7020

-continued

 cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta cggttcctgg 7080

ccttttgctg gccttttgct cacatgt 7107

<210> SEQ ID NO 104

<211> LENGTH: 130

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: AAV 5'ITR

<400> SEQUENCE: 104

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgctg ggcgaccttt 60

ggtcgccccg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120

aggggttcct 130

<210> SEQ ID NO 105

<211> LENGTH: 141

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: AAV 3'ITR

<400> SEQUENCE: 105

aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctactgagg 60

ccggcgacc aaaggtcgcc cgacgccccg gctttgcccg ggcgccctca gtgagcgagc 120

gagcgcgag ctgcctgcag g 141

<210> SEQ ID NO 106

<211> LENGTH: 4435

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor template, HBB Locus LHA to RHA (AAV.304)

<400> SEQUENCE: 106

cttgcttga caattttggt ctttcagaat actataaata taacctatat tataatttca 60

taaagtctgt gcattttctt tgaccagga tatttgcaa agacatatc aaacttccgc 120

agaacacttt atttcacata tacatgctc ttatatcagg gatgtgaaac agggcttga 180

aaactgtcta aatctaaac aatgctaag caggttttaa ttaataaaa taaaatccaa 240

aatctaacag ccaagtcaaa tctgcatgtt ttaacattta aaatatttta aagacgtctt 300

ttcccaggat tcaacatgtg aaatcttttc tcagggatac acgtgtgcct agatcctcat 360

tgcttttagtt ttttacagag gaatgaatat aaaaagaaaa tacttaaat ttatccctct 420

tacctctata atcacataaata ggacataatt ttaacctag gctccagata gccatagaag 480

aaccaaacac tttctgctg tgtgagaata atcagagtga gattttttca caagtacctg 540

atgaggggtg agacaggtag aaaaagttag agatctctat ttatttagca ataatagaga 600

aagcatttaa gagaataaag caatggaaat aagaaatttg taaatttcct tctgataact 660

agaaatagag gatccagttt ctttttggtta acctaaattt tatttcattt tattgtttta 720

ttttatttta ttttatttta ttttggttaa tcgtagtctc agagtgttag agctgaaagg 780

aagaagtagg agaacatgc aaagtaaaag tataacactt tccttactaa accgacatgg 840

gtttccaggt aggggcagga ttcaggatga ctgacagggc ctttagggaa cactgagacc 900

-continued

ctacgctgac	ctcataaatg	cttgcctac	ttgctgtttt	aattacatct	tttaatagca	960
ggaagcagaa	ctctgcactt	caaaagt	tttctcacctg	aggagttaat	ttagtacaag	1020
gggaaaaagt	acagggggat	gggagaaagg	cgatcacgtt	gggaagctat	agagaaagaa	1080
gagtaaat	tagtaagga	ggtttaaaca	aacaaaatat	aaagagaaat	aggaacttga	1140
atcaaggaaa	tgatttttaa	acgcagtatt	cttagtggac	tagaggaaaa	aaataatctg	1200
agccaagtag	aagacctttt	cccctcctac	ccctactttc	taagtcacag	aggctttttg	1260
ttccccaga	cactcttgca	gatttagtcca	ggcagaaaca	gttagatgtc	cccagttaac	1320
ctcctatttg	acaccactga	ttacccatt	gatagtcaca	cttgggttg	taagtactt	1380
tttatttatt	tgtatttttg	actgcattaa	gaggtctcta	gttttttctc	tctgttttc	1440
caaaacctaa	taagtaacta	atgcacagag	cacattgatt	tgtatttatt	ctatttttag	1500
acataattta	ttagcatgca	tgagcaaatt	aagaaaaaca	acaacaaatg	aatgcatata	1560
tatgtatatg	tatgtgtgta	tatatacaca	catatatata	tatatttttt	cttttcttac	1620
cagaaggttt	taatccaaat	aaggagaaga	tatgcttaga	accgaggtag	agttttctac	1680
cattctgtcc	tgtaagtatt	ttgcatttc	tgagacgca	ggaagagatc	catctacata	1740
tcccaaagct	gaattatgg	agacaaaact	cttccacttt	tagtgcac	acttcttatt	1800
tgtgtaataa	gaaaaatggg	aaaacgatct	tcaatatgct	taccaagctg	tgattccaaa	1860
tattacgtaa	atacacttgc	aaaggaggat	gttttttagta	gcaatttgta	ctgatgggat	1920
ggggccaaga	gatatatctt	agagggagg	ctgagggttt	gaagtccaac	tcctaagcca	1980
gtgccagaag	agccaaggac	aggtacggct	gtcatcactt	agacctcacc	ctgtggagcc	2040
acaccctagg	gttgcccaat	ctactcccag	gagcaggag	ggcaggagcc	agggtgggc	2100
ataaaagtca	gggcagagcc	atctattgct	tacatttgc	tctgacacaa	ctgtgttcac	2160
tagcaacctc	aaacagacac	catgggtgat	ctgactcctg	tcgagaagtc	tgacgtcact	2220
gctctatggg	ggaaagtga	cgtggatgaa	gttggtggtg	aggccctggg	cagggtggta	2280
tcaagggttac	aagacagggt	taaggagacc	aatagaaact	gggcatgtgg	agacagagaa	2340
gactcttggg	tttctgatag	gcactgactc	tctctgccta	ttggtctatt	ttccaccct	2400
taggctgctg	gtggtctacc	cttggaacca	gaggttcttt	gagtcctttg	gggatctgtc	2460
cactcctgat	gctgttatgg	gcaaccctaa	ggtgaaggct	catggcaaga	aagtgtctcg	2520
tgcccttagt	gatggcctgg	ctcacctgga	caacctcaag	ggcacctttg	ccacactgag	2580
tgagctgcac	tgtgacaagc	tgacgtgga	tcctgagaac	ttcagggtga	gtctatggga	2640
cgcttgatgt	tttctttccc	cttcttttct	atgggttaagt	tcatgtcata	ggaaggggat	2700
aagtaacagg	gtacagttta	gaatgggaaa	cagacgaatg	attgcatcag	tgtggaagtc	2760
tcaggatcgt	tttagtttct	tttatttgc	gttcataaca	attgttttct	ttgtttta	2820
tcttgcttct	tttttttttc	ttctccgcaa	tttttactat	tatactta	gccttaacat	2880
tgtgtataac	aaaaggaaat	atctctgaga	tacattaagt	aacttaaaaa	aaaactttac	2940
acagtctgcc	tagtacatta	ctatttggaa	tatatgtgtg	cttatttgc	tattcataat	3000
ctccctactt	tattttcttt	tattttta	tgatacata	tcattatata	tatttatggg	3060
ttaaagtgta	atgtttta	atgtgtacac	atattgacca	aatcagggtg	attttgcatt	3120
tgtaatttta	aaaaatgctt	tcttctttta	atatactttt	ttgtttatct	tatttcta	3180

-continued

```

actttcccta atctctttct ttcagggcaa taatgataca atgtatcatg cctctttgca 3240
ccattcctaaa gaataacagt gataatttct ggggtaaggc aatagcaata tctctgcata 3300
taaatatttc tgcataataa ttgtaactga tgtaagaggt tcatattgc taatagcagc 3360
tacaatccag ctaccattct gcttttattt tatgggtggg ataaggctgg attattctga 3420
gtccaagcta ggcccttttg ctaatcatgt tcatacctct tatcttcctc ccacagctcc 3480
tgggcaacgt gctggctgtg gtgctggccc atcactttgg caaagaattc accccaccag 3540
tgcaggctgc ctatcagaaa gtgggtggctg gtgtggctaa tgcctgggc cacaagtatc 3600
actaagctcg ctttcttgct gtccaatttc tattaaaggt tcctttgttc cctaagtcca 3660
actactaaac tgggggatat tatgaagggc cttgagcatc tggattctgc ctaataaaaa 3720
acattttatt tcattgcaat gatgtattta aattatttct gaatatttta ctaaaaaggg 3780
aatgtgggag gtcagtgc ataaaaata aagaaatgaa gagctagttc aaaccttggg 3840
aaaatacact atatcttaaa ctccatgaaa gaaggtgagg ctgcaaacag ctaatgcaca 3900
ttggcaacag ccctgatgc atatgcctta ttcacccctc agaaaaggat tcaagtagag 3960
gcttgatttg gaggttaaag ttttgctatg ctgtatttta cttacttat tgttttagct 4020
gtcctcatga atgtcttttc actaccatt tgccttatct gcctctctca gccttgactc 4080
cactcagttc tcttgcttag agataccacc ttccctctga agtgctcctt ccatgtttta 4140
cggcgagatg gtttctctc gcctggccac tcagccttag ttgtctctgt tgtcttatag 4200
aggtctactt gaagaaggaa aaacaggggt catggtttga ctgtcctgtg agcccttctt 4260
ccctgcctcc cccactcaca gtgaccgga atctgcagtg ctagtctccc ggaactatca 4320
ctctttcaca gtctgctttg gaaggactgg gcttagtatg aaaagttagg actgagaaga 4380
atttgaaagg cggtcttttg tagcttgata ttcactactg tcttattacc ctgtc 4435

```

<210> SEQ ID NO 107

<211> LENGTH: 7118

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor template, Full AAV (AAV.304)

<400> SEQUENCE: 107

```

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaa cccgggcgct 60
gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc tgcggcgca cgcgtcttgc tttgacaatt ttggtctttc 180
agaatactat aaataaacc tatattataa tttcataaag tctgtgcatt ttctttgacc 240
caggatattt gcaaaagaca tattcaaact tccgcagaac actttatttc acatatacat 300
gcctcttata tcagggatgt gaaacagggt cttgaaaact gtctaaatct aaaacaatgc 360
taatgcaggt ttaaatttaa taaaataaaa tccaaaatct aacagccaag tcaaactgca 420
atgttttaac atttaaaata ttttaagac gtcttttccc aggattcaac atgtgaaatc 480
ttttctcagg gatacacgtg tgcctagatc ctcatgtctt tagtttttta cagaggaatg 540
aatataaaaa gaaaataact aaattttatc cctcttacct ctataatcat acataggcat 600
aattttttaa cctaggctcc agatagccat agaagaacca aacactttct gcgtgtgtga 660
gaataatcag agtgagatth tttcacaagt acctgatgag gggtgagaca ggtagaaaaa 720

```

-continued

gtgagagatc tctatttatt tagcaataat agagaaagca ttttaagagaa taaagcaatg	780
gaaataagaa atttgtaaatt ttccttctga taactagaaa tagaggatcc agtttctttt	840
ggttaaccta aattttattt cattttattg ttttatttta ttttatttta ttttattttg	900
tgtaatcgta gtttcagagt gtttagactg aaaggaagaa gtaggagaaa catgcaaagt	960
aaaagtataa cactttcctt actaaaccga catgggtttc caggtagggg caggattcag	1020
gatgactgac agggccctta gggaacactg agaccctacg ctgacctcat aaatgcttgc	1080
tacctttgct gttttaatta catcttttaa tagcaggaag cagaactctg cacttcaaaa	1140
gtttttcttc acctgaggag ttaatttagt acaaggggaa aaagtacagg gggatgggag	1200
aaaggcgatc acgttgggaa gctatagaga aagaagagta aattttagta aaggagggtt	1260
aaacaaacaa aatataaaga gaaataggaa ctggaatcaa ggaaatgatt ttaaaacgca	1320
gtattcttag tggactagag gaaaaaata atctgagcca agtagaagac cttttccct	1380
cctaccctta ctttctaagt cacagaggct ttttgttccc ccagacactc ttgcagatta	1440
gtccaggcag aaacagttag atgtccccag ttaacctcct atttgacacc actgattacc	1500
ccattgatag tcacactttg ggttgtaagt gactttttat ttatttgtat tttgactgc	1560
attaagaggt ctctagtttt ttatctcttg tttcccaaaa cctaataagt aactaatgca	1620
cagagcacat tgatttgtat ttattctatt tttagacata atttattagc atgcatgagc	1680
aaattaagaa aaacaacaac aaatgaatgc atatatatgt atatgtatgt gtgtatatat	1740
acacacatat atatatatat tttttctttt cttaccagaa ggttttaatc caaataagga	1800
gaagatatgc ttagaaccga ggtagagttt tcatccattc tgtcctgtaa gtattttgca	1860
tattctggag acgcaggaag agatccatct acatatccca aagctgaatt atggtagaca	1920
aaactcttcc acttttagtg catcaacttc ttatttgtgt aataagaaaa ttgggaaaac	1980
gatcttcaat atgcttacca agctgtgatt ccaaatacta cgtaaatata cttgcaaagg	2040
aggatgtttt tagtagcaat ttgtactgat ggtatggggc caagagatat atcttagagg	2100
gagggtctgag gggttgaaagt ccaactccta agccagtgcc agaagagcca aggacaggta	2160
cggtctgcat cacttagacc tcacctgtg gagccacacc ctagggttgg ccaatctact	2220
cccaggagca gggagggcag gagccagggc tgggcataaa agtcagggca gagccatcta	2280
ttgcttacat ttgcttctga cacaactgtg ttcactagca acctcaaaca gacaccatgg	2340
tgcactctgac tcctgtcgag aagtctgcag tcaactgctc atgggggaaa gtgaacgtgg	2400
atgaagttgg tggtagaggc ctgggcagggt tggatcaag gttacaagac aggtttaagg	2460
agaccaatag aaactgggca tgtggagaca gagaagactc ttgggtttct gataggcact	2520
gactctctct gcctattggg ctattttccc acccttaggc tgcgtggtgg ctacccttgg	2580
accagagggt tctttgagtc ctttggggat ctgtccactc ctgatgctgt tatgggcaac	2640
cctaaggtga aggctcatgg caagaaagtg ctcggtgcct ttagtgatgg cctggctcac	2700
ctggacaacc tcaagggcac ctttgccaca ctgagtgagc tgcactgtga caagctgcac	2760
gtggatcctg agaacttcag ggtgagtcta tgggacgctt gatgttttct tccccctct	2820
tttctatggg taagtctcatg tcataggaag gggataagta acagggtaca gtttagaatg	2880
ggaaacagac gaatgattgc atcagtgtgg aagtctcagg atcgttttag tttcttttat	2940
ttgctgttca taacaattgt tttcttttgt ttaattcttg ctttcttttt tttcttctc	3000

-continued

cgcaattttt	actattatac	ttaatgcctt	aacatttgtt	ataacaaaag	gaaatatctc	3060
tgagatacat	taagtaactt	aaaaaaaaac	tttacacagt	ctgcctagta	cattactatt	3120
tggaatatat	gtgtgccttat	ttgcatattc	ataatctccc	tactttatct	tcttttattt	3180
ttaattgata	cataatcatt	atacatattt	atgggttaaa	gtgtaattgt	ttaatatgtg	3240
tacacataatt	gaccaaatac	gggtaatttt	gcatttgtaa	ttttaaaaaa	tgctttcttc	3300
ttttaataata	cttttttggt	tatcttattt	ctaatacttt	ccctaatactc	tttctttcag	3360
ggcaataatg	atacaatgta	tcatgcctct	ttgcaccatt	ctaaagaata	acagtataaa	3420
tttctgggtt	aaggcaatag	caatatctct	gcataataat	attcttgcac	ataaattgta	3480
actgatgtaa	gaggtttcat	attgctaata	gcagctacaa	tccagctacc	attctgcttt	3540
tatttttatgg	ttgggataag	gctggattat	tctgagtcca	agctaggccc	ttttgctaata	3600
catgttcata	cctcttatct	tctctccaca	gctcctgggc	aacgtgctgg	tctgtgtgct	3660
ggcccatcac	tttggaag	aattcacccc	accagtgcag	gctgcctatc	agaaagtggg	3720
ggctgggtgtg	gctaatagcc	tgcccccaca	gtatcaacta	gctcgtcttc	ttgctgtcca	3780
atttctatta	aaggttcctt	tgctccctaa	gtccaactac	taaactgggg	gatattatga	3840
agggccttga	gcactctggat	tctgcctaata	aaaaaacatt	tattttcatt	gcaatgatgt	3900
atttaaatga	tttctgaata	ttttactaaa	aagggaatgt	gggaggtcag	tgcaattaaa	3960
acataaagaa	atgaagagct	agttcaaacc	ttgggaaaat	acactatata	ttaaactcca	4020
tgaaagaagg	tgaggctgca	aacagctaata	gcacattggc	aacagccctc	gatgcataatg	4080
ccttattcat	ccctcagaaa	aggattcaag	tagaggcttg	atttgagggt	taaagtcttg	4140
ctatgctgta	ttttacatta	cttattgttt	tagctgtcct	catgaatgtc	ttttcactac	4200
ccatttgctt	atcctgcata	tctcagcctt	gactccactc	agttctcttg	cttagagata	4260
ccacctttcc	cctgaagtgt	tccttccatg	ttttacggcg	agatgggttc	tcctcgcctg	4320
gccactcagc	cttagttgtc	tctgttgtct	tatagaggtc	tacttgaaga	aggaaaaaca	4380
ggggctcatgg	tttgactgtc	ctgtgagccc	ttcttccctg	cctccccac	tcacagtgc	4440
ccggaatctg	cagtgtctag	ctcccgaac	tatcactctt	tcacagtctg	ctttggaagg	4500
actgggctta	gtatgaaaag	ttaggactga	gaagaatttg	aaaggcggct	ttttgtagct	4560
tgatattcac	tactgtctta	ttaccctgtc	ggtaaccacg	tgccggccgag	gctgcagcgt	4620
cgtctcctct	aggaaacctc	agtgtggag	ttggccactc	cctctctgcg	cgtctgctcg	4680
ctcactgagg	ccgggagacc	aaaggtcgcc	cgacgcccgg	gctttgcccg	ggcgccctca	4740
gtgagcgagc	gagcgcgagc	ctgcctgcag	ggcgccctga	tgccgtattt	tctccttacg	4800
catctgtgcg	gtatttcaca	ccgcatacgt	caaagcaacc	atagtacgag	ccctgtagcg	4860
gcgcattaag	cgcggcgggt	gtgggtggta	cgcgcagcgt	gaccgctaca	cttgccagcg	4920
ccctagcgcc	cgtccttttc	gctttcttcc	cttcctttct	cgccacgttc	gcgggctttc	4980
cccgtcaagc	tctaaatcgg	gggctccctt	tagggttccg	atttagtgct	ttacggcacc	5040
tcgaccccaa	aaaacttgat	ttgggtgatg	gttcacgtag	tgggccatcg	ccctgataga	5100
cgggtttttcg	ccctttgacg	ttggagtgca	cgttctttta	tagtggaactc	ttgttccaaa	5160
ctggaacaac	actcaacctc	atctcgggct	attcttttga	tttataaggg	attttgccga	5220
tttcggccta	ttgggttaaaa	aatgagctga	tttaacaaaa	atttaacgag	aattttaaca	5280

-continued

```

aaatattaac gtttacaatt ttatggtgca ctctcagtac aatctgctct gatgccgcat 5340
agttaagcca gccccgacac ccgccaacac ccgctgacgc gccctgacgg gcttgtctgc 5400
tccccggcatc cgcttacaga caagctgtga ccgtctccgg gagctgcatg tgtcagaggt 5460
tttcaccgtc atcaccgaaa cgcgcgagac gaaagggcct cgtgatacgc ctatttttat 5520
aggttaatgt catgaacaat aaaactgtct gcttacataa acagtaatac aaggggtgtt 5580
atgagccata ttcaacggga aacgtcgagg ccgcgattaa attccaacat ggatgctgat 5640
ttatatgggt ataaatgggc tcgcgataat gtcgggcaat caggtgcgac aatctatcgc 5700
ttgtatggga agcccgatgc gccagagttg tttctgaaac atggcaaagg tagcgttgcc 5760
aatgatgtta cagatgagat ggtcagacta aactggctga cggaatttat gcctcttccg 5820
accatcaagc attttatccg tactctgat gatgcattgt tactcaccac tgcgatcccc 5880
ggaaaaacag cattccaggt attagaagaa tatcctgatt caggtgaaaa tattgttgat 5940
gcgctggcag tgttcctgcg ccggttgcat tcgattcctg tttgtaattg tccttttaac 6000
agcgatcgcg tatttcgtct cgtcaggcg caatcacgaa tgaataacgg tttggttgat 6060
gcgagtgatt ttgatgacga gcgtaatggc tggcctgttg aacaagtctg gaaagaaatg 6120
cataaacttt tgccattctc accggattca gtcgtcactc atggtgatth ctcacttgat 6180
aaccttattt ttgacgaggg gaaattaata ggttgatttg atgttgagc agtcggaatc 6240
gcagaccgat accaggatct tgccatccta tggaaactgcc tcggtgagtt ttctccttca 6300
ttacagaaac ggctttttca aaaatatggt attgataatc ctgatatgaa taaattgcag 6360
tttcatttga tgctcgatga gtttttctaa tctcatgacc aaaatccctt aacgtgagtt 6420
ttcgttccac tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt 6480
ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg 6540
tttgccggat caagagctac caactctttt tccgaaggta actggcttca gcagagcgca 6600
gataccaaat actgtccttc tagttagacc gtagttaggc caccacttca agaactctgt 6660
agcaccgcct acatacctcg ctctgctaatt cctgttacca gtggctgctg ccagtggcga 6720
taagtcgtgt cttaccgggt tggactcaag acgatagtta ccgataagg cgcagcggtc 6780
gggctgaacg gggggttcgt gcacacagcc cagcttgag cgaacgacct acaccgaact 6840
gagataccta cagcgtgagc tatgagaaaag cgccacgctt ccggaaggga gaaaggcgga 6900
caggtatccg gtaagcgga gggtcggaac aggagagcgc acgagggagc ttccaggggg 6960
aaacgcctgg tatctttata gtctgtcgg gtttcgccac ctctgacttg agcgtcgatt 7020
tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt 7080
acggttctcg gccttttctg ggccttttgc tcacatgt 7118

```

<210> SEQ ID NO 108

<211> LENGTH: 4435

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic: Donor template, HBB Locus LHA to
RHA (AAV.307)

<400> SEQUENCE: 108

```

cttgctttga caattttggt ctttcagaat actataaata taacctatat tataatttca 60

```

-continued

taaagtctgt	gcattttctt	tgaccagga	tatttgcaa	agacatattc	aaacttcgc	120
agaacacttt	atttcacata	tacatgcctc	ttatatcagg	gatgtgaaac	agggtcttga	180
aaactgtcta	aactctaaac	aatgctaata	cagggtttaa	tttaataaaa	taaaatccaa	240
aacttaacag	ccaagtcaaa	tctgcatgtt	ttaacattta	aaatatttta	aagacgtctt	300
ttcccaggat	tcaacatgtg	aaatcttttc	tcagggatag	acgtgtgcct	agatcctcat	360
tgcttttagtt	ttttacagag	gaatgaatat	aaaaagaaaa	tacttaaat	ttatccctct	420
tacctctata	atcatacata	ggcataat	tttaacctag	gctccagata	gccatagaag	480
aaccaaacac	tttctgctg	tgtgagaata	atcagagtga	gattttttca	caagtacctg	540
atgaggggtg	agacaggtag	aaaaagtgag	agatctctat	ttatttagca	ataatagaga	600
aagcatttaa	gagaataaag	caatggaaat	aagaaatttg	taaatttcct	tctgataact	660
agaaatagag	gatccagttt	cttttggtta	acctaaat	tatttcattt	tattgtttta	720
ttttatttta	ttttatttta	ttttgtgtaa	tcgtagtctc	agagtgttag	agctgaaagg	780
aagaagtagg	agaacatgc	aaagtaaaag	tataacactt	tccttactaa	accgacatgg	840
gtttccaggt	aggggcagga	ttcaggatga	ctgacagggc	ccttagggaa	cactgagacc	900
ctacgctgac	ctcataaatg	cttgcctact	ttgctgtttt	aattacatct	tttaatagca	960
ggaagcagaa	ctctgcactt	caaaagt	tcctcacctg	aggagttaat	ttagtacaag	1020
gggaaaaagt	acagggggat	gggagaaagg	cgatcacgtt	gggaagctat	agagaaagaa	1080
gagtaaat	tagtaaaagg	gggttaaaaca	aacaaaatat	aaagagaaat	aggaacttga	1140
atcaaggaaa	tgattttaaa	acgcagtatt	cttagtgagc	tagaggaaaa	aaataatctg	1200
agccaagtag	aagacctttt	ccctcctctac	ccctactttc	taagtcacag	aggctttttg	1260
ttcccccaga	cactcttgca	gatttagtcca	ggcagaaaca	gttagatgtc	cccagttaac	1320
ctcctatttg	acaccactga	ttaccccat	gatagtcaca	ctttgggttg	taagtgactt	1380
tttattttatt	tgtatttttg	actgcattaa	gaggtctcta	gtttttttatc	tcttgtttcc	1440
caaaaacctaa	taagtaacta	atgcacagag	cacattgatt	tgtattttatt	ctatttttag	1500
acataattta	ttagcatgca	tgagcaaatt	aagaaaaaca	acaacaaatg	aatgcatata	1560
tatgtatatg	tatgtgtgta	tatatataca	catatatata	tatatTTTTT	cttttcttac	1620
cagaaggttt	taatccaaat	aaggagaaga	tatgcttaga	accgaggtag	agtttttcac	1680
cattctgtcc	tgtaagtatt	ttgcatattc	tgagagcgca	ggaagagatc	catctacata	1740
tccccaaagt	gaattatgg	agacaaaact	cttccacttt	tagtgcac	acttcttatt	1800
tgtgtaataa	gaaaattggg	aaaacgatct	tcaatatgct	taccaagctg	tgattccaaa	1860
tattacgtaa	atacacttgc	aaaggaggat	gttttttagta	gcaatttgta	ctgatgggat	1920
ggggccaaga	gatatatctt	agagggagg	ctgagggttt	gaagtccaac	tcctaagcca	1980
gtgccagaag	agccaaggac	aggtacggct	gtcatcactt	agacctcacc	ctgtggagcc	2040
acaccctagg	gttggccaat	ctactcccag	gagcagggag	ggcaggagcc	agggtggggc	2100
ataaaagtca	gggcagagcc	atctattgct	tacatttgct	tctgacacaa	ctgtgttcac	2160
tagcaacctc	aaacagacac	catggtgcat	ctgactcctg	tcgaaaaatc	cgctgtcacc	2220
gccctctggg	gcaaggtgaa	cgtggatgaa	gttggtgggtg	aggccctggg	cagggtggta	2280
tcaagggttac	aagacaggtt	taaggagacc	aatagaaaact	gggcatgtgg	agacagagaa	2340

-continued

gactcttggg	tttctgatag	gcactgactc	tctctgccta	ttggtctatt	ttcccacct	2400
taggctgctg	gtggtctacc	cttggaacca	gaggttcttt	gagtcctttg	gggatctgtc	2460
cactcctgat	gctgttatgg	gcaaccctaa	ggtgaaggct	catggcaaga	aagtgtctcg	2520
tgcctttagt	gatggcctgg	ctcacctgga	caacctcaag	ggcacctttg	ccacactgag	2580
tgagctgcac	tgtgacaagc	tgacacgtga	tcttgagaac	ttcaggggtga	gtctatggga	2640
cgcttgatgt	tttctttccc	cttcttttct	atggttaagt	tcatgtcata	ggaaggggat	2700
aagtaacagg	gtacagttta	gaatgggaaa	cagacgaatg	attgcacag	tgtggaagtc	2760
tcaggatcgt	tttagtttct	tttatttgct	gttcataaca	attgttttct	ttgttttaat	2820
tcttgctttc	tttttttttc	ttctccgcaa	tttttactat	tatacttaac	gccttaacat	2880
tgtgtataac	aaaaggaaat	atctctgaga	tacattaagt	aacttaaaaa	aaaactttac	2940
acagtctgcc	tagtacatta	ctatttgga	tatatgtgtg	cttatttgca	tattcataat	3000
ctccctactt	tattttcttt	tatttttaac	tgatacataa	tcattataca	tatttatggg	3060
ttaaagtgtg	atgttttaac	atgtgtacac	atattgacca	aatcagggtg	attttgcatt	3120
tgtaatttta	aaaaatgctt	tcttctttta	atatactttt	ttgtttatct	tattttctaat	3180
actttcccta	atctctttct	ttcagggcaa	taatgataca	atgtatcatg	cctctttgca	3240
ccattctaaa	gaataacagt	gataatttct	gggttaaggc	aatagcaata	tctctgcata	3300
taaatatttc	tgcatataaa	ttgtaactga	tgtaagagg	ttcatattgc	taatagcagc	3360
tacaatccag	ctaccattct	gcttttattt	tatggttggg	ataaggctgg	attattctga	3420
gtccaagcta	ggcccttttg	ctaactcatg	tcatacctct	tatcttcctc	ccacagctcc	3480
tgggcaacgt	gctggtctgt	gtgctggccc	atcactttgg	caaagaattc	accccaccag	3540
tgaggtctgc	ctatcagaaa	gtggtggctg	gtgtggctaa	tgccctggcc	cacaagtatc	3600
actaagctcg	ctttcttgct	gtccaatttc	tattaaagg	tcctttgttc	cctaagtcca	3660
actactaaac	tgggggatat	tatgaagggc	cttgagcctc	tggattctgc	ctaataaaaa	3720
acattttatt	tcattgcaat	gatgtattta	aattatttct	gaatatttta	ctaaaaaggg	3780
aatgtgggag	gtcagtgcac	ttaaaacata	aagaaatgaa	gagctagtcc	aaaccttggg	3840
aaaatacact	atatcttaaa	ctccatgaaa	gaaggtgagg	ctgcaaacag	ctaatagcaca	3900
ttggcaacag	cccctgatgc	atatgcctta	ttcatccctc	agaaaaggat	tcaagtagag	3960
gcttgatttg	gaggttaaag	ttttgctatg	ctgtatttta	cattacttat	tgtttttagct	4020
gtcctcatga	atgtcttttc	actaccattt	tgtttatcct	gcctctctca	gccttgactc	4080
cactcagttc	tcttgcttag	agataccacc	tttcccctga	agtgctcctt	ccatgtttta	4140
cggcgagatg	gtttctcttc	gcctggccac	tcagccttag	ttgtctctgt	tgtcttatag	4200
aggtctactt	gaagaaggaa	aaacaggggt	catggtttga	ctgtcctgtg	agcccttctt	4260
ccctgcctcc	cccactcaca	gtgaccgga	atctgcagtg	ctagtctccc	ggaactatca	4320
ctctttcaca	gtctgctttg	gaaggactgg	gcttagtatg	aaaagttagg	actgagaaga	4380
atttgaaagg	cggctttttg	tagcttgata	ttcactactg	tcttattacc	ctgtc	4435

<210> SEQ ID NO 109

<211> LENGTH: 7118

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

-continued

<223> OTHER INFORMATION: Synthetic: Donor template, Full AAV (AAV.307)

<400> SEQUENCE: 109

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc	60
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	120
actccatcac taggggttcc tgcggccgca cgcgtcttgc ttgacaatt ttggtctttc	180
agaatactat aaatataacc tatattataa tttcataaag tctgtgcatt ttctttgacc	240
caggatattt gcaaaagaca tattcaaact tccgcagaac actttatttc acatatacat	300
gcctcttata tcagggatgt gaaacagggt cttgaaaact gtctaaatct aaaacaatgc	360
taatgcagggt ttaaatTTaa taaaataaaa tccaaaatct aacagccaag tcaaactcgc	420
atgttttaac atttaaaata ttttaaagac gtcttttccc aggattcaac atgtgaaatc	480
ttttctcagg gatacacgtg tgccctagatc ctcatgtctt tagtttttta cagaggaatg	540
aatataaaaa gaaaaactt aaattttatc cctcttacct ctataatcat acataggcat	600
aatttttta cctaggctcc agatagccat agaagaacca aacactttct gcgtgtgtga	660
gaataatcag agtgagatTT tttcacaagt acctgatgag gggtgagaca ggtagaaaaa	720
gtgagagatc tctattttatt tagcaataat agagaaagca tttaaagaaa taaagcaatg	780
gaaataagaa atttgtaaat ttccttctga taactagaaa tagaggatcc agtttctttt	840
ggttaaccta aattttattt cattttattg ttttatttta ttttatttta ttttattttg	900
tgtaatcgta gtttcagagt gtttagactg aaaggaagaa gtaggagaaa catgcaaagt	960
aaaagtataa cactttcctt actaaaccga catgggtttc caggtagggg caggattcag	1020
gatgactgac agggccctta ggaacactg agaccctacg ctgacctcat aaatgcttgc	1080
tacctttgct gttttaatta catcttttaa tagcaggaag cagaactctg cacttcaaaa	1140
gtttttctc acctgaggag ttaatttagt acaaggggaa aaagtacagg gggatgggag	1200
aaaggcgatc acgttgaggaa gctatagaga aagaagagta aatttttagta aaggagggtt	1260
aaacaaacaa aatataaaga gaaataggaa ctggaatcaa ggaaatgatt ttaaaacgca	1320
gtattcttag tggactagag gaaaaaata atctgagcca agtagaagac cttttccctt	1380
cctaccctta ctttctaagt cacagaggct ttttgttccc ccagacactc ttgcagatta	1440
gtccaggcag aaacagttag atgtccccag ttaacctcct atttgacacc actgattacc	1500
ccattgatag tcacactttg ggttgtaagt gactttttat ttatttgtat tttgactgc	1560
attaagaggt ctctagtttt ttatctcttg tttcccaaaa cctaataagt aactaatgca	1620
cagagcacat tgatttgtat ttattctatt tttagacata atttattagc atgcatgagc	1680
aaattaagaa aaacaacaac aaatgaatgc atatatatgt atatgtatgt gtgtatatat	1740
acacacatat atatatatat tttttctttt cttaccagaa gggttttaatc caaataagga	1800
gaagatatgc ttagaacoga ggtagagttt tcatccattc tgtcctgtaa gtattttgca	1860
tattctggag acgcaggaag agatccatct acatatccca aagctgaatt atggtagaca	1920
aaactcttcc acttttagtg catcaacttc ttatttgtgt aataagaaaa ttgggaaaac	1980
gatcttcaat atgcttacca agctgtgatt ccaaatatta cgtaaataca cttgcaaagg	2040
aggatgtttt tagtagcaat ttgtactgat ggtatggggc caagagatat atcttagagg	2100
gagggtctgag gggttgaaat ccaactccta agccagtgcc agaagagcca aggacaggta	2160

-continued

cggctgtcat	cacttagacc	tcacctgtg	gagccacacc	ctagggttgg	ccaatctact	2220
cccaggagca	gggagggcag	gagccagggc	tgggcataaa	agtcagggca	gagccatcta	2280
ttgcttacat	ttgcttctga	cacaactgtg	ttcactagca	acctcaaaca	gacaccatgg	2340
tgcacttgac	tcctgtcgaa	aaatccgctg	tcaccgcctt	ctggggcaag	gtgaacgtgg	2400
atgaagttgg	tggtgaggcc	ctgggcaggt	tggatatcaag	gttacaagac	aggtttaagg	2460
agaccaatag	aaactgggca	tgtggagaca	gagaagactc	ttgggtttct	gataggcact	2520
gactctctct	gcctattggt	ctattttccc	acccttaggc	tgctggtggg	ctacccttgg	2580
accagaggt	tctttgagtc	ctttggggat	ctgtccactc	ctgatgctgt	tatgggcaac	2640
cctaagggtga	aggctcatgg	caagaaagtg	ctcgggtgcct	ttagtgatgg	cctggctcac	2700
ctggacaacc	tcaagggcac	ctttgccaca	ctgagtgagc	tgcactgtga	caagctgcac	2760
gtggatcctg	agaacttcag	ggtgagtcta	tgggacgctt	gatgttttct	ttccccctct	2820
tttctatggt	taagtctcatg	tcataggaag	gggataagta	acaggggtaca	gtttagaatg	2880
ggaaacagac	gaatgattgc	atcagtgtgg	aagtctcagg	atcgttttag	tttcttttat	2940
ttgctgttca	taacaattgt	tttcttttgt	ttaattcttg	ctttcttttt	tttcttcttc	3000
cgcaattttt	actattatac	ttaatgcctt	aacattgtgt	ataacaaaag	gaaatatctc	3060
tgagatacat	taagtaactt	aaaaaaaaac	tttacacagt	ctgcctagta	cattactatt	3120
tggatatatat	gtgtgcttat	ttgcatattc	ataatctccc	tactttatatt	tcttttatatt	3180
ttaattgata	cataatcatt	atacatattt	atgggttaaa	gtgtaatgtt	ttaatatgtg	3240
tacacataatt	gaccaaatac	gggtaatttt	gcatttgtaa	ttttaaaaaa	tgctttcttc	3300
ttttaatata	cttttttggg	tatcttatatt	ctaatacttt	ccctaatactc	tttctttcag	3360
ggcaataatg	atacaatgta	tcatgcctct	ttgcaccatt	ctaaagaata	acagtataaa	3420
tttctggggt	aaggcaatag	caatatctct	gcataataat	atttotgcat	ataaattgta	3480
actgatgtaa	gaggtttcat	attgctaata	gcagctacaa	tccagctacc	attctgcttt	3540
tattttatgg	ttgggataag	gctggattat	tctgagtcca	agctaggccc	ttttgctaatt	3600
catgttcata	cctcttatct	tcctcccaca	gctcctgggc	aacgtgctgg	tctgtgtgct	3660
ggcccatcac	tttggcaaa	aattcacccc	accagtgcag	gctgcctatc	agaaagtggg	3720
ggctggtgtg	gctaattgcc	tggcccacaa	gtatcactaa	gctcgtcttc	ttgctgtcca	3780
atttctatta	aaggttcctt	tgttccttaa	gtccaactac	taaactgggg	gatattatga	3840
agggccttga	gcactctggat	tctgcctaatt	aaaaaacatt	tattttcatt	gcaatgatgt	3900
atttaaatga	tttctgaata	ttttactaaa	aagggaatgt	gggaggtcag	tgcatttaaa	3960
acataaagaa	atgaagagct	agttcaaacc	ttgggaaaat	acactatata	ttaaactcca	4020
tgaaaagaagg	tgaggctgca	aacagctaatt	gcacattggc	aacagccctt	gatgcatatg	4080
ccttattcat	ccctcagaaa	aggattcaag	tagaggcttg	atttgagggt	taaagttttg	4140
ctatgctgta	ttttacatta	cttattgttt	tagctgtcct	catgaatgtc	ttttcactac	4200
ccatttgctt	atcctgcata	tctcagcctt	gactccactc	agttctcttg	cttagagata	4260
ccacctttcc	cctgaagtgt	tccttccatg	ttttacggcg	agatgggttc	tcctcgcctg	4320
gccactcagc	cttagttgtc	tctgttgtct	tatagaggtc	tacttgaaga	aggaaaaaca	4380
ggggctcatgg	tttgactgtc	ctgtgagccc	ttcttccctg	cctcccccac	tcacagtgc	4440

-continued

ccggaatctg cagtgcctagt ctcccgggaaac tatcactctt tcacagctctg ctttggaagg	4500
actgggctta gtatgaaaag ttaggactga gaagaatttg aaaggcggt tttttagct	4560
tgatattcac tactgtctta ttacctgtc ggtaaccacg tgccggcgag gctgcagcgt	4620
cgtcctccct aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg	4680
ctcactgagg ccggggcgacc aaaggctgcc cgacgcccgg gctttgcccc ggccggcctca	4740
gtgagcgagc gagcgcgag ctgcctgcag gggcgctga tgccgtatct tctccttacg	4800
catctgtgcg gtatttcaca ccgcatacgt caaagcaacc atagtacgag ccctgtagcg	4860
gcgcatatag ccggggcggt gtggtgggta ccgcagcgt gaccgctaca cttgccagcg	4920
ccctagcgcc cgctccttcc gctttcttcc ctctcttct cgccacgttc gccggcttcc	4980
cccgtcaagc tctaaatcgg gggctccctt taggggtccg atttagtgct ttacggcacc	5040
tcgaccccaa aaaacttgat ttgggtgatg gtccacgtag tgggccatcg ccctgataga	5100
cgggttttcc ccctttgagc ttggagtcga cgttctttaa tagtggaactc ttgttccaaa	5160
ctggaacaac actcaaccct atctcgggct attcttttga tttataaggg attttgccga	5220
tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgag aattttaaca	5280
aaatattaac gtttacaatt ttatggtgca ctctcagtag aatctgctct gatgcgcgat	5340
agttaagcca gcccgcacac ccgccaacac ccgctgacgc gccctgacgg gcttgtctgc	5400
tcccggcatc cgcttacaga caagctgtga ccgtctccgg gagctgcatg tgccagaggt	5460
tttcaccgtc atcaccgaaa ccgcgcgagc gaaaggccct cgtgatacgc ctatttttat	5520
aggttaatgt catgaacaat aaaactgtct gcttacataa acagtaatac aaggggtgtt	5580
atgagccata ttcaacggga aacgtcgagg ccgcgattaa attccaacat ggatgctgat	5640
ttatatgggt ataaatgggc tcgcgataat gtccggcaat caggtgcgac aatctatcgc	5700
ttgtatggga agcccgatgc gccagagttg tttctgaaac atggcaaagg tagcgttgcc	5760
aatgatgtta cagatgagat ggtcagacta aactggctga cggaatttat gcctcttccg	5820
accatcaagc attttatccg tactctgat gatgcaggt tactcaccac tgcgacccc	5880
ggaaaaacag cattccaggt attagaagaa tatcctgatt caggtgaaaa tattgttgat	5940
gcgctggcag tgttccctgc ccgggttgcgt tcgattcctg tttgtaattg tccctttaac	6000
agcgatcgcg tatttctgtc cgctcaggcg caatcacgaa tgaataacgg tttggttgat	6060
gcgagtgatt ttgatgacga gcgtaatggc tggcctgttg aacaagctcg gaaagaaatg	6120
cataaaacttt tgccattctc accggattca gtcgtcactc atggtgatct ctcacttgat	6180
aaccttattt ttgacgaggg gaaattaata ggttgatttg atgttgagc agtcggaatc	6240
gcagaccgat accaggatct tgccatccta tggaaactgcc tcggtgagtt ttctcctca	6300
ttacagaaac ggctttttca aaaatatggt attgataatc ctgatatgaa taaattgcag	6360
tttcatttga tgctcgatga gtttttctaa tctcatgacc aaaatccctt aacgtgagtt	6420
ttcgttccac tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt	6480
ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cgggtggttg	6540
tttgccggat caagagctac caactctttt tccgaaggta actggcttca gcagagcgca	6600
gataccaaat actgtccttc tagtgtagcc gtagttaggc caccacttca agaactctgt	6660
agcaccgcct acatacctcg ctctgctaatt cctgttacca gtggctgctg ccagtggcga	6720

-continued

```

taagtcgtgt cttaccgggt tggactcaag acgatagtta cgggataagg cgcagcggtc 6780
gggctgaacg ggggggttcgt gcacacagcc cagcttgag cgaacgacct acaccgaact 6840
gagataccta cagcgtgagc tatgagaaag cgccacgctt cccgaaggga gaaaggcgga 6900
caggatccg gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg 6960
aaacgcctgg tatctttata gtctgtcgg gtttcgccac ctctgacttg agcgtcgatt 7020
tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt 7080
acggttcctg gccttttgct ggccttttgc tcacatgt 7118

```

```

<210> SEQ ID NO 110
<211> LENGTH: 141
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Donor template, 5' ITR

```

```

<400> SEQUENCE: 110

```

```

cctgcaggca gctgcgcgct cgctcgtca ctgaggccgc cggggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccg gcctcagtga gcgagcgcgc gcgcagagag ggagtggcca 120
actccatcac taggggttcc t 141

```

```

<210> SEQ ID NO 111
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Wild-type HBB gene region
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2)..(2)
<223> OTHER INFORMATION: n = C or T

```

```

<400> SEQUENCE: 111

```

```

anctgactcc tgaggagaag tctgccgtta ctgccctgtg gggcaagg 48

```

```

<210> SEQ ID NO 112
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: Sickie HBB gene region
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2)..(2)
<223> OTHER INFORMATION: n = C or T

```

```

<400> SEQUENCE: 112

```

```

anctgactcc tggggagaag tctgccgtta ctgccctgtg gggcaagg 48

```

```

<210> SEQ ID NO 113
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: R02 gRNA

```

```

<400> SEQUENCE: 113

```

```

cttgccccac agggcagtaa 20

```

```

<210> SEQ ID NO 114

```

-continued

```

<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: AAV.307

<400> SEQUENCE: 114
atctgactcc tgtcgaaaaa tccgctgtca cgcacctctg gggcaagg           48

<210> SEQ ID NO 115
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: AAV.304

<400> SEQUENCE: 115
atctgactcc tgtcgagaag tctgcagtca ctgctctatg ggggaaaag           48

<210> SEQ ID NO 116
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic: B-thal HBB gene region
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2)..(2)
<223> OTHER INFORMATION: n = C or T

<400> SEQUENCE: 116
anctgactcc tgaggagaag tctgcagtta ctgccctgtg gggcaagg           48

```

1. A method for increasing homology directed repair (HDR) of a double-strand break (DSB) in a target gene in a cell or a population of cells, the method comprising contacting a cell or a population of cells with a 53BP1 inhibitor, wherein the cell or the population of cells is a quiescent human cell or a population of quiescent human cells that is induced to divide, and wherein the DSB is mediated by a site-directed nuclease, thereby increasing HDR of the DSB in the target gene in the cell or the population of cells.

2. (canceled)

3. The method of claim 1, wherein the 53BP1 inhibitor is (i) a 53BP1 binding polypeptide that inhibits 53BP1 recruitment to the DSB in the cell or the population of cells; or (ii) a 53BP1 binding polypeptide comprising an amino acid sequence selected from a group consisting of: SEQ ID NO: 70, SEQ ID NO: 74, SEQ ID NO: 77, SEQ ID NO: 80, SEQ ID NO: 83, or SEQ ID NO: 86.

4. (canceled)

5. The method of claim 1, wherein the 53BP1 inhibitor comprises (i) a nucleic acid comprising a nucleotide sequence encoding a 53BP1 binding polypeptide that inhibits 53BP1 recruitment to the DSB site in the cell or the population of cells; or (ii) a nucleic acid comprising a nucleotide sequence selected from a group consisting of: SEQ ID NO: 69, SEQ ID NO: 73, SEQ ID NO: 76, SEQ ID NO: 79, SEQ ID NO: 82, SEQ ID NO: 85, or SEQ ID NO: 88.

6. (canceled)

7. The method of claim 5, wherein the nucleic acid comprises (i) a vector comprising a nucleotide sequence encoding the 53BP1 binding polypeptide; or (ii) a vector

comprising a nucleotide sequence selected from a group consisting of: SEQ ID NO: 68, SEQ ID NO: 72, SEQ ID NO: 75, SEQ ID NO: 78, SEQ ID NO: 81, SEQ ID NO: 84, or SEQ ID NO: 87.

8. (canceled)

9. The method of claim 7, wherein the vector is an adeno associated viral vector (AAV).

10-11. (canceled)

12. The method of claim 1, wherein the cell or population of cells is isolated from a tissue sample obtained from a human donor, or a peripheral blood sample obtained from a human donor.

13. (canceled)

14. The method of claim 1, wherein the cell or population of cells is:

- (i) a hematopoietic stem or progenitor cell (HSPC) or a population of HSPCs;
- (ii) a long-term HSPC (LT-HSPC) or a population of LT-HSPCs;
- (iii) a CD34 expressing cell or a population of CD34 expressing cells;
- (iv) a white blood cell or a population of white blood cells;
- (v) in the G0 phase of the cell cycle; or
- (vi) a combination of (i)-(v).

15-18. (canceled)

19. The method of claim 1, wherein the cell or population of cells is induced to divide by treatment with one or more extrinsic signals, wherein the one or more extrinsic signals are selected from: an extrinsic signal that stimulates a cell signaling receptor, a mitogen, or an environmental factor.

20-23. (canceled)

24. The method of claim **1**, wherein the DSB is mediated by a Cas9 DNA endonuclease and one or more guide RNAs (gRNA), or a *Streptococcus pyogenes* Cas9 (SpCas9) and one or more gRNAs.

25. (canceled)

26. The method of claim **1**, further comprising contacting the cell or population of cells with a donor polynucleotide comprising a nucleotide sequence that corrects or induces a mutation when incorporated at the site of the DSB, wherein the donor polynucleotide comprises a double-stranded DNA (dsDNA) or a single-stranded oligonucleotide (ssODN).

27-29. (canceled)

30. The method of claim **26**, wherein the donor polynucleotide comprises a vector comprising the donor polynucleotide, wherein the vector comprises a recombinant single-stranded adeno-associated viral vector (rAAV).

31. (canceled)

32. The method of claim **30**, wherein the rAAV is serotype 6 (AAV6).

33. The method of claim **1**, further comprising contacting the cell or population of cells with an inhibitor of DNA-PK.

34. (canceled)

35. The method of claim **1**, wherein the cell or population of cells is induced to divide prior to contact with the 53BP1 inhibitor, is contacted with the 53BP1 inhibitor prior to being induced to divide, or is contacted with the 53BP1 inhibitor concurrent with being induced to divide.

36-37. (canceled)

38. The method of claim **1**, wherein the cell or population of cells is contacted ex vivo and the method further comprises administering the cell or population of cells to a subject in need thereof, or wherein the cell or population of cells is contacted with the 53BP1 inhibitor in vivo.

39-82. (canceled)

83. The method of claim **12**, wherein the tissue sample is obtained from a human donor, and wherein the human donor is:

- (i) administered an HSPC mobilizing agent prior to obtaining the tissue sample;
- (ii) administered (i), wherein the HSPC mobilizing agent is Plurixifor;
- (iii) administered a combination of HSPC mobilizing agents prior to obtaining the tissue sample; or
- (iv) administered (iii), wherein the combination of HSPC mobilizing agents comprises Plurixifor and granulocyte colony stimulating factor (G-CSF).

84-87. (canceled)

88. The method of claim **19**, wherein the extrinsic signal is:

- (i) one or more cytokines;
- (ii) one or more cytokines selected from interleukin-3 (IL-3), stem cell factor (SCF), Fms-like tyrosine kinase 3 (Flt3) ligand, and thrombopoietin;
- (iii) a mitogen;
- (iv) an environmental factor; or
- (v) hypoxia.

89-110. (canceled)

111. The method of claim **38**, wherein the population of cells is contacted ex vivo and the contacting results in an increase in HDR frequency in the cell population to 50% or more.

112. (canceled)

113. The method of claim **111**, wherein contacting results in:

- (i) a 1-2 fold increase in HDR frequency in the cell population, or wherein contacting results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2

fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population;

- (ii) a 1-2 fold decrease in indel frequency in the cell population, or wherein contacting results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold decrease in indel frequency in the cell population;
- (iii) an increase in engraftment in vivo following administration;
- (iv) an increase in chimerism in vivo following administration; or
- (v) a combination of (i)-(iv).

114-117. (canceled)

118. A method for increasing HDR of a DSB in a target gene in a population of human CD34+ cells, the method comprising contacting the population of human CD34+ cells with

- (i) an inhibitor of 53BP1, wherein the inhibitor is nucleic acid comprising a nucleotide sequence encoding a 53BP1 binding polypeptide comprising the amino acid sequence of SEQ ID NO: 69, and
- (ii) a Cas9 DNA endonuclease and one or more gRNAs to effect a DSB within the target gene in the population of cells.

119. The method of claim **118**, wherein the nucleic acid comprising a nucleotide sequence encoding the 53BP1 binding polypeptide is an RNA or a vector.

120. (canceled)

121. The method of claim **119**, wherein the vector is an AAV or an AAV of the DJ serotype (AAV-DJ).

122-123. (canceled)

124. The method of claim **118**, wherein the Cas9 DNA endonuclease is a polypeptide, and wherein the Cas9 DNA endonuclease is pre-complexed with one or more gRNAs.

125. (canceled)

126. The method of claim **124**, wherein the Cas9 DNA endonuclease and one or more gRNAs are electroporated into the cells, and wherein the cell population is contacted with the nucleic acid encoding the 53BP1 polypeptide prior to electroporation, is contacted with the nucleic acid encoding the 53BP1 polypeptide during electroporation, or is contacted with the nucleic acid encoding the 53BP1 polypeptide subsequent to electroporation.

127-129. (canceled)

130. The method of claim **118**, further comprising contacting the population of cells with a donor polynucleotide comprising a nucleotide sequence that corrects or induces a mutation when incorporated at the site of the DSB, wherein the donor polynucleotide comprises a dsDNA or a ssODN.

131-134. (canceled)

135. The method of claim **130**, wherein the donor polynucleotide comprises a vector comprising the donor polynucleotide, wherein the vector comprises rAAV or AAV6.

136-137. (canceled)

138. The method of claim **130**, wherein the donor polynucleotide is administered to the cells prior to electroporation or following electroporation.

139. (canceled)

140. The method of claim **118**, further comprising contacting the population of cells with an inhibitor of DNA-PK.

141. (canceled)

142. The method of claim **118**, wherein the cells are cultured with

- (i) one or more cytokines;
- (ii) one or more cytokines selected from: thrombopoietin, Flt3 ligand, SCF, IL-3; or

(iii) a combination of cytokines comprising: thrombopoietin, Flt3 ligand, SCF, and IL-3.

143-147. (canceled)

148. The method of claim **118**, wherein the cells are cultured under hypoxic conditions comprising atmospheric oxygen content less than 5%.

149. The method of claim **118**, wherein the population of cells is contacted ex vivo and the method further comprises administering the cells to a subject in need thereof, wherein contacting results in an increase in HDR frequency in the cell population to 50% or more.

150. (canceled)

151. The method of claim **149**, wherein contacting results in:

- (i) a 1-2 fold increase in HDR frequency in the cell population, or wherein contacting results in a 1.1-2 fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold increase in HDR frequency in the cell population;
- (ii) a 1-2 fold decrease in indel frequency in the cell population, or wherein contacting results in a 1.1-2

fold, 1.2-2 fold, 1.3-2 fold, 1.4-2 fold, 1.5-2 fold, 1.6-2 fold, 1.7-2 fold, 1.8-2 fold, or 1.9-2 fold decrease in indel frequency in the cell population;

(iii) an increase in engraftment in vivo following administration;

(iv) an increase in chimerism in vivo following administration; or

(v) a combination of (i)-(iv).

152-155. (canceled)

156. A method for increasing HDR of a DSB in a target gene in a population of human CD34+ cells, the method comprising contacting the population of human CD34+ cells with an effective amount of an inhibitor of the catalytic subunit of DNA-dependent protein kinase, and wherein the DSB is mediated by a site-directed nuclease, thereby increasing HDR of the DSB in the target gene in the population of human CD34+ cells.

157-228. (canceled)

* * * * *