United States Patent	12385070
Kind Code	B2
Date of Patent	August 12, 2025
Inventor(s)	Scharenberg; Andrew et al.

Homology directed repair compositions for the treatment of hemoglobinopathies

Abstract

The present disclosure provides improved compositions for the homology directed repair of the human globin locus for the prevention, treatment, or amelioration of at least one symptom of a hemoglobinopathy.

Inventors: **Scharenberg; Andrew (Seattle, WA), Jacoby; Kyle (Seattle, WA), Kiem; Hans-Peter (Seattle, WA), Rawlings; David J. (Seattle, WA), Lux; Christopher (Bremerton, WA), Pattabhi; Sowmya (Bellevue, WA), Humbert; Olivier M. (Seattle, WA)**

Applicant: **Seattle Children's Hospital** (Seattle, WA); **Fred Hutchinson Cancer Center** (Seattle, WA)

Family ID: **63918645**

Assignee: **Seattle Children's Hospital (Seattle, WA); Fred Hutchinson Cancer Center (Seattle, WA)**

Appl. No.: **18/344732**

Filed: **June 29, 2023**

Prior Publication Data

Document Identifier	Publication Date
US 20240124896 A1	Apr. 18, 2024

Related U.S. Application Data

continuation parent-doc US 16608182 US 12110499 WO PCT/US2018/029235 20180424 child-doc US 18344732
us-provisional-application US 62488927 20170424

Publication Classification

Int. Cl.: **C12N15/90** (20060101); **A61K35/28** (20150101); **A61P9/00** (20060101); **C07K14/805** (20060101); **C12N15/11** (20060101); **C12N15/86** (20060101)

U.S. Cl.:

CPC **C12N15/907** (20130101); **A61K35/28** (20130101); **A61P9/00** (20180101); **C07K14/805** (20130101); **C12N15/11** (20130101); **C12N15/86** (20130101); C12N2310/20 (20170501); C12N2750/14143 (20130101); C12N2800/80 (20130101)

Field of Classification Search

CPC: C12N (15/907); C12N (15/11); C12N (15/86); C12N (2310/20); C12N (2750/14143); C12N (2800/80); C12N (15/102); C12N (15/10); A61K (35/28); A61K (48/00); A61P (9/00); C07K (14/805)

References Cited

U.S. PATENT DOCUMENTS

Patent No.	Issued Date	Patentee Name	U.S. Cl.	CPC
6051402	12/1999	LeBoulch et al.	N/A	N/A
7901671	12/2010	Leboulch et al.	N/A	N/A
9017967	12/2014	Bonas et al.	N/A	N/A
9068199	12/2014	Leboulch et al.	N/A	N/A
2014/0080216	12/2013	Cost et al.	N/A	N/A
2015/0133528	12/2014	Krieg et al.	N/A	N/A
2015/0166969	12/2014	Takeuchi et al.	N/A	N/A
2020/0255857	12/2019	Gori et al.	N/A	N/A

FOREIGN PATENT DOCUMENTS

Patent No.	Application Date	Country	CPC
WO2014036219	12/2013	WO	N/A
WO2017115268	12/2016	WO	N/A
WO2017218948	12/2016	WO	N/A
WO2002088346	12/2021	WO	N/A

OTHER PUBLICATIONS

Office Action for European Application No. 18789938.0, Dated May 21, 2024, 4 pages. cited by applicant

Balazs and Godbey, “Liposomes for use in gene delivery,” J. Drug Deliv., vol. 2011, No. 326497, 2011, 12 pages. cited by applicant

Choi, et al., “Optimization of AAV expression cassettes to improve packaging capacity and transgene expression on neurons”, Molecular brain, vol. 7, No. 17, 2014, 10 pages. cited by applicant

Cong, et al., “Multiplex genome engineering using CRISPR/Cas systems,” Science, vol. 339, No. 6121, 2013, pp. 819-823. cited by applicant

Dever, et al., “CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells,” Nature, vol. 539, No. 7629, 2016, pp. 384-389. cited by applicant

Office Action Dated Dec. 20, 2021 for European Application No. 18789938, 5 pages. cited by applicant
European Office Action mailed Feb. 2, 2023, for European Patent Application No. 18789938, a foreign counterpart to U.S. Appl. No. 16/608,182, 6 pages. cited by applicant
Finotti, et al., "Recent trends in the gene therapy of B-thalassemia", J. Blood Med., vol. 6, 2015, pp. 69-85. cited by applicant
Gilman, et al., "Distal CCAAT box deletion in the A gamma globin gene of two black adolescents with elevated fetal A gamma globin," Nucleic Acids Res., vol. 16, No. 22, 1988, pp. 10635-10642. cited by applicant
Hoban, et al., "Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells," Blood, vol. 125, No. 17, 2015, pp. 2597-29604. cited by applicant
Huang and Yen, "Role of the hepatitis B virus posttranscriptional regulatory element in export of intronless transcripts," Mol. Cell. Biol., vol. 15, No. 7, 1995, pp. 3864-3869. cited by applicant
Jinek, et al., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," Science, vol. 337, No. 6096, 2012, pp. 816-821. cited by applicant
Jinek, et al., "RNA-programmed genome editing in human cells," eLife 2:e00471, 2013, 9 pages. cited by applicant
Kiem, et al., Abstract "Novel Gene Editing Approaches for Hemoglobinopathies," National Institutes of Health Grant No. HL 136135 (Funding Start Date Jan. 17, 2017). cited by applicant
Liu and Mertz, "HnRNP L binds a cis-acting RNA sequence element that enables intron-dependent gene expression," Genes Dev., vol. 9, No. 14, 1995, pp. 1766. cited by applicant
Liu, et al., "Poly(cationic lipid)-mediated in vivo gene delivery to mouse liver," Gene Therapy., vol. 10, No. 2, 2003, pp. 180-187. cited by applicant
Mali, et al., "RNA-guided human genome engineering via Cas9," Science, vol. 339, No. 6121, 2013, pp. 823-826. cited by applicant
Manca, et al., "Disorders of the Synthesis of Human Fetal Hemoglobin," IUBMB Life, vol. 60, No. 2, 2008, pp. 94-111. cited by applicant
Office Action Dated Aug. 8, 2022 for U.S. Appl. No. 16/608,182, 13 Pages. cited by applicant
Pattabhi, et al., "In Vivo Outcome of Homology-Directed Repair at the HBB Gene in HSC Using Alternative Donor Template Delivery Methods," Mol. Ther. Nucleic Acids, vol. 17, 2019, pp. 277-288. cited by applicant
Qi, et al., "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression," Cell, vol. 152, No. 5, 2013, pp. 1173-1183. cited by applicant
Ran, et al., "Genome engineering using the CRISPR-Cas9 system," Nature Protocols, vol. 8, No. 11, 2013, pp. 2281-2308. cited by applicant
Segal, "Bacteria herald a new era of gene editing," eLife 2:e00563, 2013, 3 pages. cited by applicant
Search Report and Written Opinion Dated Oct. 2, 2018 in International Application No. PCT/US2018/029235, 12 pages. cited by applicant
Wall, et al., "The human B-globin gene 3′ enhanver contains multiple binding sites for an erythroid-specific protein," Gen. Dev., vol. 2, 1988, 1089-1100. cited by applicant
Zetsche, et al., "Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system," Cell, vol. 163, No. 3, 2015, pp. 759-771. cited by applicant
Zufferey, et al., "Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors," J. Virol., vol. 73, No. 4, 1999, pp. 2886-2892. cited by applicant

---

Primary Examiner: Sajjadi; Fereydoun G

Assistant Examiner: Nicol; Alexander W

Attorney, Agent or Firm: Lee & Hayes PC

Background/Summary

RELATED APPLICATIONS (1) This application is a continuation of U.S. patent application Ser. No. 16/608,182, filed on Oct. 24, 2019, which is a national stage filing under 35 U.S.C. 371 of International Application No. PCT/US2018/029235, filed Apr. 24, 2018, which claims the benefit of U.S. Provisional Application No. 62/488,927, filed on Apr. 24, 2017. The entire teachings of the above applications are incorporated herein by reference.

BACKGROUND
Technical Field
(1) The present disclosure relates to improved compositions for use in homology directed repair of the human globin locus. More particularly, the disclosure relates to improved donor repair templates for editing the human globin locus for the prevention, treatment, or amelioration of at least one symptom of a hemoglobinopathy.
Description of the Related Art
(2) Hemoglobinopathies are a diverse group of inherited monogenetic blood disorders that result from variations in the structure and/or synthesis of hemoglobin. The most common hemoglobinopathies are sickle cell disease (SCD), α-thalassemia, and β-thalassemia. Approximately 5% of the world's population carries a globin gene mutation. The World Health Organization estimates that more than 300,000 infants are born each year with major hemoglobin disorders. Hemoglobinopathies manifest highly variable clinical manifestations that range from mild hypochromic anemia to moderate hematological disease to severe, lifelong, transfusion-dependent anemia with multiorgan involvement.
(3) The only potentially curative treatment available for hemoglobinopathies is allogeneic hematopoietic stem cell transplantation. However, it is estimated that HLA-compatible HSC transplants are available to less than 20% of affected individuals and long term toxicities are substantial. In addition, HSC transplants are also associated with significant mortality and morbidity in subjects that have SCD or severe thalassemias. The significant mortality and morbidity is due in part to pre-HSC transplantation transfusion-related iron overload, graft-versus-host disease (GVHD), and high doses of chemotherapy/radiation required for pre-transplant conditioning of the subject, among others.
(4) Supportive treatments for hemoglobinopathies include periodic blood transfusions for life, combined with iron chelation, and in some cases splenectomy. Additional treatments for SCD include analgesics, antibiotics, ACE inhibitors, and hydroxyurea. However, the side effects associated with hydroxyurea treatment include cytopenia, hyperpigmentation, weight gain, opportunistic infections, azoospermia, hypomagnesemia, and cancer.
(5) At best, patients treated with existing methods have a projected lifespan of 50 to 60 years.
BRIEF SUMMARY
(6) The present disclosure generally relates, in part, to improved donor repair templates used for editing a human γ-globin gene.
(7) In particular embodiments, a DNA donor repair template is contemplated comprising: a 5′ homology arm and a 3′ homology arm, wherein the donor repair template comprises a polynucleotide sequence within at least about 1 kb, at least about 1.5 kb, or at least about 2 kb upstream of the transcription start site of a human gamma globin gene and further comprises a deletion of Chr11: 5249959-5249971.

(8) In particular embodiments, a DNA donor repair template is contemplated comprising: a 5′ homology arm; a selection cassette; an erythroid expression control sequence; and a 3′ homology arm.
(9) In certain embodiments, the length of the 5′ homology arm is at least about 100 bp, at least about 200 bp, at least about 300 bp, at least about 400 bp, at least about 500 bp, at least about 600 bp, at least about 700 bp, at least about 800 bp, at least about 900 bp, or at least about 1000 bp and wherein the 5′ homology arm comprises a polynucleotide sequence within at least about 1 kb, at least about 1.5 kb, or at least about 2 kb upstream of the transcription start site of a human gamma globin gene.
(10) In some embodiments, the 5′ homology arm is at least about 100 bp, at least about 200 bp, at least about 300 bp, at least about 400 bp, at least about 500 bp, at least about 600 bp, at least about 700 bp, at least about 800 bp, at least about 900 bp, or at least about 1000 bp and wherein the 5′ homology arm comprises a polynucleotide sequence within 1 kb upstream of the transcription start site of a human gamma globin gene and the 5′ homology arm comprises a deletion in the region of Chr11: 5249957-5249977.
(11) In certain embodiments, the deletion in the region of Chr11: 5249957-5249977 is a deletion associated with hereditary persistence of fetal hemoglobin.
(12) In some embodiments, the deletion in the region of Chr11: 5249957-5249977 is a deletion associated with derepression of gamma globin expression.
(13) In particular embodiments, the deletion in the region of Chr11: 5249957-5249977, comprises a deletion of the polynucleotide sequence of any one of SEQ ID NOs: 1-6.
(14) In particular embodiments, the deletion in the region of Chr11: 5249957-5249977, comprises a deletion of the polynucleotide sequence of Chr11: 5249959-5249971.
(15) In further embodiments, the selection cassette comprises a ubiquitous promoter, a constitutive promoter, an inducible promoter, or hematopoietic stem cell promoter, operably linked to a polynucleotide sequence encoding a selectable marker, and one or more post-transcription regulatory elements.
(16) In certain embodiments, the promoter is selected from the group consisting of: a cytomegalovirus (CMV) immediate early promoter, a viral simian virus 40 (SV40) promoter, a Moloney murine leukemia virus (MoMLV) LTR promoter, a Rous sarcoma virus (RSV) LTR promoter, a herpes simplex virus (HSV) (thymidine kinase) promoter, a H5, P7.5, or P11 vaccinia virus promoter, a short elongation factor 1-alpha (EF1a-short) promoter, a long elongation factor 1-alpha (EF1a-long) promoter, an early growth response 1 (EGR1) promoter, a ferritin H (FerH) promoter, a ferritin L (FerL) promoter, a Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) promoter, a eukaryotic translation initiation factor 4A1 (EIF4A1) promoter, a heat shock 70 kDa protein 5 (HSPA5) promoter, a heat shock protein 90 kDa beta, member 1 (HSP90B1) promoter, a heat shock protein 70 kDa (HSP70) promoter, a β-kinesin (β-KIN) promoter, a human ROSA 26 promoter, a Ubiquitin C (UBC) promoter, a phosphoglycerate kinase-1 (PGK) promoter, a cytomegalovirus enhancer/chicken β-actin (CAG) promoter, a β-actin promoter and a myeloproliferative sarcoma virus enhancer, negative control region deleted, dl587rev primer-binding site substituted (MND) promoter.
(17) In particular embodiments, the selectable marker is selected from the group consisting of: a hygromycin-B phosphotransferase (HPH) gene, an amino 3′-glycosyl phosphotransferase (NEO) gene, a dihydrofolate reductase (DHFR) gene, an adenosine deaminase (ADA) gene, a multi-drug resistance (MDR) gene, an O.sup.6-methylguanine-DNA-methyltransferase (MGMT) gene, a bleomycin (BLE) gene, and a blasticidin-S deaminase (BSR) gene.
(18) In certain embodiments, the one or more post-transcription regulatory elements are selected from the group consisting of: a woodchuck hepatitis virus post-transcriptional response element (WPRE) or variant thereof, a hepatitis B virus post-transcriptional response element (HPRE) or variant thereof, and a polyadenylation sequence.

(19) In some embodiments, the polyadenylation sequence is selected from the group consisting of: an ideal poly(A) sequence, an SV40 poly(A) sequence, a bovine growth hormone (BGH) poly(A) sequence, and a rabbit β-globin poly(A) sequence.
(20) In some embodiments, the erythroid expression control sequence comprises a human β-globin LCR responsive promoter.
(21) In particular embodiments, the erythroid expression control sequence comprises an ankyrin gene promoter, an α-spectrin gene promoter, a β-spectrin gene promoter, or a β-globin gene promoter, optionally in combination with an HPFH-2 enhancer, an HS40 enhancer, or a β-globin gene 3′ enhancer.
(22) In further embodiments, the erythroid expression control sequence is positioned to be operably linked to an endogenous gamma globin gene when the DNA donor repair template is integrated into the human genome.
(23) In certain embodiments, the endogenous gamma globin gene is the A-gamma globin gene (HBGA; HBG1).
(24) In further embodiments, the endogenous gamma globin gene is the G-gamma globin gene (HBGG; HBG2).
(25) In additional embodiments, DNA donor repair template is integrated into the human genome at both the HBG1 locus and the HBG2 locus and the erythroid expression control sequence of the DNA donor repair template integrated at the HBG1 locus is positioned to be operably linked to the endogenous HBG1 gene and the erythroid expression control sequence of the DNA donor repair template integrated at the HBG2 locus is positioned to be operably linked to the endogenous HBG2 gene.
(26) In particular embodiments, the length of the 3′ homology arm is at least about 100 bp, at least about 200 bp, at least about 300 bp, at least about 400 bp, at least about 500 bp, at least about 600 bp, at least about 700 bp, at least about 800 bp, at least about 900 bp, or at least about 1000 bp and wherein the 3′ homology arm comprises a polynucleotide sequence downstream of the 5′ homology arm and upstream of the start codon of the human gamma globin gene.
(27) In particular embodiments, a DNA donor repair template is contemplated comprising: a 5′ homology arm; a polynucleotide encoding a therapeutic globin and one or more post-transcriptional control elements; a selection cassette; and a 3′ homology arm.
(28) In particular embodiments, the length of the 5′ homology arm is at least about 100 bp, at least about 200 bp, at least about 300 bp, at least about 400 bp, at least about 500 bp, at least about 600 bp, at least about 700 bp, at least about 800 bp, at least about 900 bp, or at least about 1000 bp and wherein the 5′ homology arm comprises a polynucleotide sequence within 1 kb upstream of the transcription start site of a human gamma globin gene and the 5′ homology arm comprises a deletion in the region of Chr11: 5249957-5249977.
(29) In some embodiments, the deletion in the region of Chr11: 5249957-5249977 is a deletion associated with hereditary persistence of fetal hemoglobin.
(30) In certain embodiments, the deletion in the region of Chr11: 5249957-5249977 is a deletion associated with derepression of gamma globin expression.
(31) In some embodiments, the deletion in the region of Chr11: 5249957-5249977, comprises a deletion of the polynucleotide sequence of any one of SEQ ID NOs: 1-6.
(32) In particular embodiments, the deletion in the region of Chr11: 5249957-5249977, comprises a deletion of the polynucleotide sequence of Chr11: 5249959-5249971.
(33) In further embodiments, the endogenous gamma globin promoter is operably linked to the polynucleotide encoding the therapeutic globin.
(34) In particular embodiments, the therapeutic globin is γ-globin, β-globin, δ-globin, or an anti-sickling β-globin.
(35) In certain embodiments, the anti-sickling β-globin is selected from the group consisting of: β-globin.sup.A-T87Q, β-globin.sup.A-T87Q/K120E/K95E, and β-globin.sup.A-T87Q/G16D/E22A.

(36) In additional embodiments, the one or more post-transcription regulatory elements are selected from the group consisting of: a woodchuck hepatitis virus post-transcriptional response element (WPRE) or variant thereof, a hepatitis B virus post-transcriptional response element (HPRE) or variant thereof, and a polyadenylation sequence.
(37) In particular embodiments, the polyadenylation sequence is selected from the group consisting of: an ideal poly(A) sequence, an SV40 poly(A) sequence, a bovine growth hormone (BGH) poly(A) sequence, and a rabbit β-globin poly(A) sequence.
(38) In particular embodiments, the DNA donor repair template further comprises an erythroid enhancer that enhances the expression of the polynucleotide encoding the therapeutic globin.
(39) In certain embodiments, the erythroid enhancer is selected from an HPFH-2 enhancer, an HS40 enhancer, or a β-globin gene 3′ enhancer.
(40) In some embodiments, the selection cassette comprises a ubiquitous promoter, a constitutive promoter, an inducible promoter, or hematopoietic stem cell promoter, operably linked to a polynucleotide sequence encoding a selectable marker, and optionally one or more post-transcription regulatory elements.
(41) In some embodiments, the promoter is selected from the group consisting of: a cytomegalovirus (CMV) immediate early promoter, a viral simian virus 40 (SV40) promoter, a Moloney murine leukemia virus (MoMLV) LTR promoter, a Rous sarcoma virus (RSV) LTR promoter, a herpes simplex virus (HSV) (thymidine kinase) promoter, a H5, P7.5, or P11 vaccinia virus promoter, a short elongation factor 1-alpha (EF1a-short) promoter, a long elongation factor 1-alpha (EF1a-long) promoter, an early growth response 1 (EGR1) promoter, a ferritin H (FerH) promoter, a ferritin L (FerL) promoter, a Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) promoter, a eukaryotic translation initiation factor 4A1 (EIF4A1) promoter, a heat shock 70 kDa protein 5 (HSPA5) promoter, a heat shock protein 90 kDa beta, member 1 (HSP90B1) promoter, a heat shock protein 70 kDa (HSP70) promoter, a β-kinesin (β-KIN) promoter, a human ROSA 26 promoter, a Ubiquitin C (UBC) promoter, a phosphoglycerate kinase-1 (PGK) promoter, a cytomegalovirus enhancer/chicken β-actin (CAG) promoter, a β-actin promoter and a myeloproliferative sarcoma virus enhancer, negative control region deleted, dl587rev primer-binding site substituted (MND) promoter.
(42) In particular embodiments, the selectable marker is selected from the group consisting of: a hygromycin-B phosphotransferase (HPH) gene, an amino 3-glycosyl phosphotransferase (NEO) gene, a dihydrofolate reductase (DHFR) gene, an adenosine deaminase (ADA) gene, a multi-drug resistance (MDR) gene, an O.sup.6-methylguanine-DNA-methyltransferase (MGMT) gene, a bleomycin (BLE) gene, and a blasticidin-S deaminase (BSR) gene.
(43) In certain embodiments, the one or more post-transcription regulatory elements are selected from the group consisting of: a woodchuck hepatitis virus post-transcriptional response element (WPRE) or variant thereof, a hepatitis B virus post-transcriptional response element (HPRE) or variant thereof, and a polyadenylation sequence.
(44) In some embodiments, the polyadenylation sequence is selected from the group consisting of: an ideal poly(A) sequence, an SV40 poly(A) sequence, a bovine growth hormone (BGH) poly(A) sequence, and a rabbit β-globin poly(A) sequence.
(45) In particular embodiments, the selection cassette comprises a ubiquitous promoter, a constitutive promoter, an inducible promoter, or hematopoietic stem cell promoter, operably linked to a polynucleotide sequence encoding a selectable marker, and a ribosomal skipping sequence or viral self-cleaving peptide.
(46) In certain embodiments, the length of the 3′ homology arm is at least about 100 bp, at least about 200 bp, at least about 300 bp, at least about 400 bp, at least about 500 bp, at least about 600 bp, at least about 700 bp, at least about 800 bp, at least about 900 bp, or at least about 1000 bp and wherein the 3′ homology arm comprises a polynucleotide sequence downstream of the start codon of the gamma globin gene.

(47) In certain embodiments, the gamma globin gene is the A-gamma globin gene (HBGA; HBG1).
(48) In various embodiments, a viral vector comprises a DNA donor repair template contemplated herein.
(49) In particular embodiments, the viral vector is a recombinant adeno-associated viral vector (rAAV) or a retrovirus.
(50) In some embodiments, the rAAV has one or more ITRs from AAV2.
(51) In further embodiments, the rAAV has a serotype selected from the group consisting of: AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, and AAV10.
(52) In certain embodiments, the rAAV has an AAV6 serotype.
(53) In additional embodiments, the retrovirus is a lentivirus.
(54) In particular embodiments, the lentivirus is an integrase deficient lentivirus.
(55) In some embodiments, a cell comprises a DNA donor repair template or a viral vector contemplated herein.
(56) In some embodiments, the DNA donor repair template has been inserted into a human gamma globin gene target site by homology directed repair.
(57) In further embodiments, the target site is an engineered nuclease target site set forth in SEQ ID NO: 7 or SEQ ID NO: 8.
(58) In particular embodiments, the cell is a hematopoietic cell.
(59) In certain embodiments, the cell is CD34.sup.+ cell.
(60) In particular embodiments, the cell is CD133.sup.+ cell.
(61) In various embodiments, a method for increasing gamma globin expression in a hematopoietic stem or progenitor cell comprising introducing one or more engineered nucleases that cleave a target site set forth in SEQ ID NO: 9 and a DNA donor repair template contemplated herein into the cell, whereby the DNA donor repair template is inserted into the cell genome by homology directed repair at a double strand break introduced by the one or more engineered nucleases.
(62) In particular embodiments, a method for increasing therapeutic globin expression in a hematopoietic stem or progenitor cell comprises introducing one or more engineered nucleases that cleave a target site set forth in SEQ ID NO: 7 or SEQ ID NO: 8 and a DNA donor repair template contemplated herein into the cell, whereby the DNA donor repair template is inserted into the cell genome by homology directed repair at a double strand break introduced by the one or more engineered nucleases.
(63) In some embodiments, a genome edited cell produced by a HDR with a donor repair template contemplated herein is provided.
(64) In various embodiments, a composition comprises a DNA donor repair template, a viral vector, or a cell contemplated herein.
(65) In various embodiments, a composition comprises a physiologically acceptable excipient and a DNA donor repair template, a viral vector, or a cell contemplated herein.
(66) In further embodiments, a method of treating a hemoglobinopathy in a subject comprises administering the subject a cell or composition contemplated herein.
(67) In various embodiments, a method of ameliorating at least one symptom, of a hemoglobinopathy in a subject comprises administering the subject a cell or composition contemplated herein.
(68) In particular embodiments, the β-globin alleles of the subject are β.sup.E/β.sup.0, β.sup.C/β.sup.0, β.sup.0/β.sup.0, β.sup.E/β.sup.E, β.sup.C/β.sup.+, β.sup.E/β.sup.+, β.sup.0/β.sup.+, β.sup.+/β.sup.+, β.sup.C/β.sup.C, β.sup.E/β.sup.S, β.sup.0/β.sup.S, β.sup.C/β.sup.S, β.sup.+/β.sup.S, or β.sup.S/β.sup.S.
(69) In additional embodiments, a method of treating a thalassemia in a subject comprises administering the subject an effective amount of a cell or composition contemplated herein.
(70) In certain embodiments, the thalassemia is a β-thalassemia.
(71) In particular embodiments, the β-globin alleles of the subject are β.sup.E/β.sup.0, β.sup.C/

β.sup.0, β.sup.0/β.sup.0, β.sup.C/β.sup.C, β.sup.E/β.sup.E, β.sup.E/β.sup.+, β.sup.C/β.sup.E, β.sup.C/β.sup.+, β.sup.0/β.sup.+, or β.sup.+/β.sup.+.
(72) In certain embodiments, a method of treating sickle cell disease in a subject comprises administering the subject an effective amount of a cell or composition contemplated herein.
(73) In particular embodiments, the β-globin alleles of the subject are β.sup.E/β.sup.S, β.sup.0/β.sup.S, β.sup.C/β.sup.S, β.sup.+/β.sup.S or β.sup.S/β.sup.S.
(74) In some embodiments, a method of treating a β-thalassemia in a subject comprises administering the subject an effective amount of a cell or composition contemplated herein.
(75) In particular embodiments, the β-globin alleles of the subject are β.sup.E/β.sup.0, β.sup.C/β.sup.0, β.sup.0/β.sup.0, β.sup.C/β.sup.C, β.sup.E/β.sup.E, β.sup.E/β.sup.+, β.sup.C/β.sup.E, β.sup.C/β.sup.+, β.sup.0/β.sup.+, or β.sup.+/β.sup.+.

Description

BRIEF DESCRIPTION OF SEVERAL VIEWS OF THE DRAWINGS
(1) FIG. **1** shows an upstream region of the γ-globin promoter, including the 13 bp sequence responsible for repression of γ-globin gene expression (top panel) and a nuclease target site strategy for disruption of the 13 bp sequence.
(2) FIG. **2** shows a diagram of the various illustrative donor repair templates that were integrated into the γ-globin locus by homology directed repair.
(3) FIG. **3** shows the flow cytometric analysis of HbF expression by HPLC following erythroid differentiation.
(4) FIGS. **4**A-**4**C show TALEN design and screening. FIG. **4**A shows a diagram of the human β hemoglobin locus on Chromosome 11 highlighting the Aγ(HBG1) promoter structure. FIG. **4**B provides a schematic of repressive elements that bind the γ-hemoglobin promoter including the CCAAT Displacement Protein (CDP/CUX1), DRED Complex (in conjunction with COUP-TFII), BCL11a (associated with the NURD complex which also partners with LRF) and the NF-Y binding sites. Negative transcription factors are shown in red. Positive transcriptions factors are shown in green. The putative BCL11a binding sequence (TGACCA) is underlined in red. The distal and proximal CCAAT boxes are underlined in green. Green * indicate the location of published HPFH SNPs. The bracketed green line highlights the 13 bp HPFH deletion as labeled. FIG. **4**C shows TALENs selected in silico and tested. Blue boxes represent Repeat Variable Diresidues (RVDs) and their corresponding nucleotide is listed below. Scissors represent FokI endonuclease and the dotted line and indicated bp numbers represent the spacer length between the TALEN pairs.
(5) FIGS. **5**A-**5**G demonstrate optimizing TALEN editing conditions to maximize efficiency. FIG. **5**A provides an experimental timeline for TALEN transfection of hPBSCs. FIG. **5**B shows cell viability (right axis) and total cell number (left axis) assessed at two recovery temperatures (37° C. and 30° C.) and over increasing TALEN mRNA doses measured at 24 hours post transfection. (n=6/condition, p-values *<0.05, **<0.005, ***<0.0005)). TALEN mRNA doses higher than 1 ug had a significant negative impact on transfected cells and the effect was more pronounced with a 30° C. recovery step. FIG. **5**C provides schematic of a novel ddPCR fall off assay designed to detect NHEJ events at both the HBG1 and HBG2 promoters. A common set of primers (green) are used to amplify the HBG1 and HGB2 alleles. A common NHEJ ddPCR Probe linked to HEX bind over the TALEN target cut site and is designed to fall off if single base insertion or deletion is present in the binding region. Unique HBG1 and HBG2 Ref ddPCR Probes linked to FAM bind to a region with three unique nucleotides allowing for editing of HBG1 and HBG2 to be assessed independently. FIG. **5**D shows ddPCR results for hPBSCs edited with TALEN mRNA at increasing concentrations and two recovery temperatures. All cells transfected with TALEN mRNA result in significant indel generation. 30° C. recovery results in significantly higher rates of NHEJ than 37°

C. but decreases with doses of TALEN mRNA over 1 ug. HBG1 NHEJ editing rates detected by this assay are roughly 50% of HBG2 due to the presence of large intergenic deletions. FIG. **5**E shows rates of NHEJ following 1 ug mRNA transfection and detected by ddPCR do not vary by donor (n=3 donors). FIG. **5**F shows NHEJ assessment by Next Gen Sequencing (Illumina MiSeq) detects editing rates lower than the ddPCR assay but demonstrate a consistent doubling in editing following 30° C. recovery. FIG. **5**G shows indel frequency by size (bp) demonstrating higher editing rates following 30° C. recovery (bottom, blue) also result in higher rates of 13 bp deletion as well as an increased frequency of larger (5-8 bp) deletions than 37° C. (above, black) recovery where 1 bp deletions predominate. Solid bars represent HBG1, Outlined bars represent HBG2.
(6) FIGS. **6**A-**6**D demonstrate TALEN-induced ds breaks drive fetal hemoglobin expression in differentiated hPBSCs. FIG. **6**A provides an experimental timeline for erythroid differentiation following transfection. FIG. **6**B shows representative flow at 14 days of differentiation comparing the erythroid progeny of mock and TALEN edited hPBSCs. The left panel demonstrates that the overall CD235a+ staining profile is nearly identical between mock (82.5%, blue) and TALEN edited cells (88.5%, orange). Isotype control is shown in brown. The right panel shows HbF expression is significantly higher in TALEN edited cells (61.1%) compared to mock (27.4%). FIG. **6**C shows combined analysis of HbF expression at increasing doses of TALEN mRNA and at different recovery temperatures demonstrating a significant increase in HbF expression in all TALEN transfected cells compared to mock ($p<0.0005$) and a greater increase seen at 30° C. recovery (outlined blue circles) compared to 37° C. (filled black circles). Cells transfected with 1 ug TALEN result in significantly more HbF expression with 37° C. recovery resulting in 44±3% HbF and 30° C. recovery resulting in 57±3% HbF compared to mock HbF expression of 24±2 (30° C.) or 24±5 (37° C.). FIG. **6**D shows hemoglobin protein expression detected by HPLC demonstrates a significant increase in overall HbF expression at 1 ug TALEN mRNA. There is a 2.4 fold increase (26±3% total HbF protein) seen with 37° C. culture following 1 ug transfection. 30° C. cold shock treatment resulted in a 4.6 fold increase in HbF protein expression to 41±8% when transfected with 1 ug mRNA ($p<0.005$).
(7) FIGS. **7**A-**7**H demonstrates sustained multi-lineage engraftment of TALEN edited hPBSCs in recipient W41 mice. FIG. **7**A provides experimental timeline for editing of hPBSCs followed by both primary and secondary transplants. FIG. **7**B shows human engraftment (% hCD45 positive) at sac for both the primary and secondary transplants. Shapes correspond to experimental cohorts. Primary and secondary transplant average engraftment was not significantly different following TALEN editing (Primary: mock 67±3% n=4, TALEN 71±8% n=9. Secondary: mock 3.6±1.9% n=2, TALEN 2.8±1.4% n=5). FIG. **7**C provides summary of FACS analysis of primary transplants at sac (Transplant 1 at 16 weeks, Transplant 2 at 24 weeks) with multilineage engraftment in mice transfused with either mock or TALEN edited cells. CD19+ engraftment is more robust at 24 weeks with population profiles otherwise similar. FIG. **7**D shows NHEJ rates detected by ddPCR at sac for both the primary and secondary transplants. Transplant 1 had a higher input editing rate (HBG1=56%, HBG2=46%) than transplant 2 (HBG1=20%, HBG2=28%). Editing rates drop by approximately 50% post transplant but are maintained following secondary transplant. FIG. **7**E shows indel frequency measured by Next Gen Sequencing demonstrates the edits maintained post transplant tend to be in the 2-7 bp range with a relative decrease in the 13 bp HPFH deletion. FIG. **7**F shows comparison of percent modification (deletion) seen at each nucleotide indicates that the majority of deletions occur over the BCL11a binding site (red line) and distal CCAAT box (green line) with a relative decrease in retained deletions post transplant of the 3′ end of the 13 bp HPFH deletion. FIG. **7**G shows flow analysis at the time of harvest shows that human HbF is significantly upregulated in mice that receive TALEN edited hPBSCs (Mock=21±1%, TALEN 33±15%). FIG. **7**H shows flow analysis following ex vivo liquid differentiation demonstrates significantly higher HbF in animals that received TALEN edited hPBSCs (Mock=59±2%, TALEN=70±6%).
(8) FIG. **8** demonstrates ribonucleoprotein delivery with Cas9 and sgRNA has been optimized for

editing at the HBG1 and G2 loci. The panel on the left shows ˜26% overall editing at the HbG1 and HbG2 loci by TIDE analysis. The panel on the right shows a T7 endonuclease assay that shows, no Indels in the electroporation only samples and the presence of Indels in the RNP and RNP+AAV samples.

(9) FIG. **9** provides templates that rely on large genomic deletions to induce fetal hemoglobin.

(10) FIG. **10** provides templates designed to express T87Q using the HBG1 promoter or drive the endogenous HBG1 gene using the d13 HPFH promoter or the HBB promoter.

(11) FIG. **11** provides 'Round 3' repair templates designed to express T87Q using the HBG1 promoter or drive the endogenous HBG1 gene using the d13 HPFH promoter or the HBB promoter and some containing an MGMT chemo selection cassette.

(12) FIG. **12** provides a summary of Rhesus related constructs.

(13) FIG. **13** shows construct 1263—GFP Control repair template. This template is used to assess HDR rates at the HGB1 locus. The data demonstrates efficient HDR within this locus. The HDR rates are comparable to our own and others published results at other genetic loci in human CD34+ HSC. This construct is used to compare to other HDR constructs that create deletions and/or have smaller HR arms, etc, and this acts as a benchmark for editing.

(14) FIG. **14** shows when HR arms bind to sequences distant from each other in the genome, they require a deletion event to occur and decrease HR efficiency.

(15) FIG. **15** provides alternative constructs using the HBGd13 Promoter driving HBB T87Q anti-sickling globin.

(16) FIG. **16** demonstrates GFP Positive cells generated using construct 1235/60 also express T87Q as well as increased levels of HBF.

(17) FIG. **17** demonstrates GFP Positive cells express both T87Q as well as increased levels of HBF

(18) FIG. **18** demonstrates the HDR templates V3E6 (with either 600 or 400 bp homology arms) are capable of expressing T87Q following HDR-editing. Fetal hemoglobin also increases in these cells likely due to NHEJ events in other alleles.

(19) FIG. **19** demonstrates Construct 1324 is a rAAV construct that can drive homology-dependent repair into the HBG1 locus. The donor template introduces a d13 promoter that drives gamma-1 globin. MND-GFP is in the reverse orientation. It has an alternate HDR site due to incidental homology.

(20) FIG. **20** demonstrates Construct 1325 is a rAAV construct that can drive homology-dependent repair into the HBG1 locus. The donor template introduces a d13 promoter that drives gamma-1 globin. This is a deletional construct that has a 240 bp deletion that may lower HDR rates. MND-GFP is in the reverse orientation.

(21) FIG. **21** shows that HDR occurs at HBG1 locus following with co-delivery of RNP and indicated AAV donor. The GFP+ population seen in the boxes are from day 6 post-editing and shows that all 3 AAV donors support HDR-based editing of HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples. Irrespective of the Cas9 vendor used there is integration of AAV into the HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples. Construct 1345 is shown on FIG. **27**.

(22) FIG. **22** shows that HDR occurs at HBG1 locus following with co-delivery of RNP and indicated AAV donor. The GFP+ population are from day 6 post-editing and shows that all 3 AAV donors support HDR-based editing of HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples. Irrespective of the Cas9 vendor used there is integration of AAV into the HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples.

(23) FIG. **23** shows that HDR occurs at HBG1 locus following with co-delivery of RNP and indicated AAV donor. The GFP+ population seen in the boxes are from day 13 post-editing and shows that all 3 AAV donors support HDR-based editing of HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples. Irrespective of the Cas9 vendor used there is integration of AAV

into the HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples.
(24) FIG. **24** shows that HDR occurs at HBG1 locus following with co-delivery of RNP and indicated AAV donor. The GFP+ population are from day 13 post-editing and shows that all 3 AAV donors support HDR-based editing of HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples. Irrespective of the Cas9 vendor used there is integration of AAV into the HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples.
(25) FIG. **25** shows that HDR drives increase in fetal hemoglobin induction in erythroid cells. All 3 AAV donors support HDR-based editing and lead to increased fetal hemoglobin induction following integration into the HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples. Irrespective of the Cas9 vendor used there is increased fetal hemoglobin induction with integration of AAV into the HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples.
(26) FIG. **26** shows that HDR drives increase in fetal hemoglobin induction in erythroid cells. All 3 AAV donors support HDR-based editing and lead to increased fetal hemoglobin induction following integration into the HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples. Irrespective of the Cas9 vendor used there is increased fetal hemoglobin induction with integration of AAV into the HBG1 locus in RNP+1324, RNP+1325, RNP+1345 samples.
(27) FIG. **27** demonstrates Construct 1345 is a rAAV construct that can drive homology-dependent repair into the HBG1 locus. The donor template introduces a HBB promoter that drives T87Q globin.
(28) FIG. **28** demonstrates Construct 1347 is a rAAV construct that can drive homology-dependent repair into the Rhesus HBG1 locus. The donor template introduces a HBB promoter that drives T87Q globin.
(29) FIG. **29** demonstrates Construct 1333 is a rAAV construct that can drive homology-dependent repair into the human HBG1 locus. The donor template introduces a HBB promoter that drives Gamma globin expression and allows for chemo therapeutic selection, as it has a MND-promoter driving P140K MGMT expression. Construct 1336 is a rAAV construct that can drive homology-dependent repair into the human HBG1 locus. The donor template introduces a HBG1 d13 promoter that drives T87Q globin expression and allows for chemo therapeutic selection, as it has a PGK-promoter driving P140K MGMT expression.
(30) FIG. **30** demonstrates Construct 1343 is a rAAV construct that can drive homology-dependent repair into the human HBG1 locus. The donor template introduces a HBB promoter that drives T87Q globin expression and allows for chemo therapeutic selection, as it has a PGK-promoter driving P140K MGMT expression. Construct 1346 is a rAAV construct that can drive homology-dependent repair into the human HBG1 locus. The donor template introduces a HBG1 d13 promoter that drives T87Q globin expression and allows for chemo therapeutic selection, as it has a PGK-promoter driving P140K MGMT expression.
(31) FIG. **31** demonstrates Construct 1347 is a rAAV construct that can drive homology-dependent repair into the human HBG1 locus. The donor template introduces a HBB promoter that drives T87Q globin expression and allows for chemo therapeutic selection, as it has a PGK-promoter driving P140K MGMT expression.
(32) FIG. **32** provides an example of an HBB promoter driving HBG1 expression.
(33) FIG. **33** demonstrates use of multiple HDR templates (numbers listed under flow plots) and TALEN co-delivery demonstrating increased HbF expression in GFP+ cells.
(34) FIG. **34** demonstrates MGMT chemoselection HDR-cassettes (Constructs 1333, 1336) are designed to drive fetal hemoglobin expression as well as the ability to expand engrafted cells pre- or post-transplant via chemoselection.
(35) FIG. **35** demonstrates HDR editing and chemoselection using 1333 MGMT HDR donor cassette. Following HDR-editing of CD34+ HSC, chemoselection in vitro allows for 5-fold expansion of the edited population over non-edited cells. Edited cells are tracked using ddPCR based assay.

(36) FIG. **36** demonstrates HDR editing and chemoselection using 1333 MGMT HDR donor cassette. Cells edited with an MGMT cassette are able to be chemoselected as shown here by flow cytometry expansion of MGMT Hi cells post selection.
(37) FIG. **37** demonstrates 20-30% INDEL rate at both loci using γ1 (HBG1) and γ2 (HBG2) specific probes with ddPCR. This increases to 50% with a 30 C recovery step.
(38) FIG. **38** provides a T7 analysis showing del13 TALEN pair transfection induces INDELs in human CD34 cells at both the γ1 (HBG1) and γ2 (HBG2) locus.
(39) FIG. **39** shows confirmed editing rates (50% in this example) via Next Gen Sequencing. There is an overrepresentation of the 13 bp deletion likely the result of microhomology in the region.
(40) FIG. **40** shows results of HbF induction by flow cytometry. TALEN editing of peripheral blood CD34 cells followed by erythroid differentiation results in significantly increased number of F-cells.
(41) FIG. **41** shows results of HbF induction by HPLC. TALEN editing of peripheral blood CD34 cells followed by erythroid differentiation results in significantly increased HbF protein expression. The difference in protein expression is greater than the percent F-cells suggesting that the increased F-cells express higher levels of HbF than control F-cells.
(42) FIG. **42** shows results of human engraftment at week 24. No significant difference in percent human engraftment between control and edited cells is seen.
(43) FIG. **43** shows engrafted edited CD34 cells generate all human hematopoietic lineages. Human erythroid, CD34+ Lymphoid and Myeloid, and CD19+ cells were all identified and sorted from harvested marrows following transplant (W24).
(44) FIG. **44** shows TALEN edited CD34 cells produce more F-cells. At sac (A) there is a significantly higher rate of human F-Cells detected in the marrow. Differentiated CD34 cells from the marrow produce more F-Cells (B).
(45) FIG. **45** shows sorted cells from all lineages retain INDELs from TALEN editing. T7 Analysis demonstrating INDELS are present in vitro.
(46) FIG. **46** provides a transplants summary for various animals tested.
(47) FIGS. **47**A-**47**I show results of recapitulating the 13-nucleotide HPFH deletion by CRISPR/Cas9 gene editing. FIG. **47**A shows recapitulation of 13-nucleotide HPFH deletion by CRISPR/Cas9 gene editing. FIG. **47**B shows immunophenotypic separation of HSPC subsets from bone marrow enriched-CD34+ cells (A17117) by surface antibody staining. FIG. **47**C shows HbG gene editing efficiency measured at 24 h post CRISPR RNPs electroporation in sorted HSPCs subset. Results are means and standard deviations from 2 donors. FIG. **47**D shows proportion of 13-nt HPFH deletion relative to all other deletions in reactions from FIG. **47**C. * denotes statistically significant decrease (t-test, $P<0.05$) in 13-nt deletion in CD90+ subset as compared to CD34+ cells. FIG. **47**E shows colony-forming cells in CD34+ and HSPC subsets plated on methylcellulose media at 24 h post mock electroporation. FIG. **47**F shows colony-forming cells in CD34+ and HSPC subsets plated on methylcellulose media at 24 h post CRISPR/Cas9 RNPs electroporation. Results are from the same donor. FIG. **47**G shows deletion profile in CD34+-edited cells (A17117) at 4 days post electroporation. Each color box shows deletion with a frequency higher than 1% and the white portion at the bottom shows combined deletions that contributes less than 1%. FIG. **47**H provides sequences of the most common deletions (color-coded) from FIG. **47**G. FIG. **47**I shows in vitro HbF expression (defined as ratio of HbF/HbA measured by flow cytometry) in differentiated erythroblast as function of HbG editing levels (measured by TIDE). Results are from 4 different donors and HbF/HbA ratio was normalized to mock-treated samples in each donor for comparison.
(48) FIGS. **48**A-**48**D show hematopoietic recovery in all transplanted animals by measuring neutrophil count (cells per μL) (FIG. **48**A), platelet count (cells per μL) (FIG. **48**B), lymphocyte count (cells per μL) (FIG. **48**C), and monocyte count (cells per μL) (FIG. **48**D) in view of days post transplantation.

(49) FIGS. **49**A-**49**D show tracking of HbG editing in vivo in all transplanted animals. FIG. **49**A shows HbG editing efficiency by Miseq analysis measured over days post transplantation. FIG. **49**B shows percent HPFH 13-nt deletion by Miseq analysis measured over days post transplantation. FIGS. **49**A-**49**D show mutant frequency for animals A17114 and A17116.
(50) FIGS. **50**A-**50**E show fetal hemoglobin production in all transplanted animals. Percent F-cells (flow cytometry) (FIG. **50**A) and percent γ-globin relative to β-like globin (HPLC) (FIG. **50**B) was measured over days post transplantation for all transplanted animals. Percent F-cells (flow cytometry) was measured in view of percent γ-globin (HPLC) for four of the transplanted animals (FIG. **50**B). Percent γ-globin and percent F-cells were then measured in view of HbG editing (measured by TIDE) (FIG. **50**D). For four animal models the percent total of gamma-1, gamma-2, beta, alpha-2, and alpha-1 was measured over days post transplantation.
(51) FIGS. **51**A-**51**G shows multilineage engraftment of CRISPR/Cas9-edited HSPCs in bone marrow of transplanted animals at 6 months post infusion. FIGS. **51**A-**51**B show immunophenotypic separation of HSPC subsets from bone marrow enriched-CD34+ cells by surface antibody staining, as well as showing colony-forming cells in CD34+ and HSPC subsets for animal models A17114 (FIG. **51**A) and A17116 (FIG. **51**B). FIG. **51**C shows HbG editing efficiency (measured by TIDE) for three animal models measuring CD34+ cells, CD45RA+ cells, CD50− cells, and CD90+ cells. FIGS. **51**D and **51**E provide flow cytometry sorting of HSPCs subsets. FIG. **51**F shows HbG editing efficiency (measured by TIDE) for three animal models measuring T cells, B cells, Gran cells, mono cells, 71+ erythroid cells. FIG. **51**G shows mutant frequency for animal model A17114 in multiple cell types.
(52) FIG. **52** provides quantification of large deletion events and off-target sites.
(53) FIGS. **53**A-**53**F show HbG editing efficiency in different HSPC subsets. FIG. **53**A shows titration of molar ratios of Cas9 protein to gRNA for optimization of HbG editing efficiency (determined by Surveyor assay). Circles show individual data points and bar shows mean. Results are from 3 different NHP donors. FIG. **53**B shows size distribution of HbG deletions in edited NHP CD34+ cells 4 days post treatment. Results are from Miseq analysis using 1 donor and deletion frequency was normalized to 100%. FIG. **53**C shows flow cytometry sorting of HSPCs subsets from A17114 after CD34+ enrichment. FIG. **53**D shows HbG editing efficiency in HSPCs sorted subsets from FIG. **53**C determined by Miseq analysis. FIG. **53**E shows flow cytometry sorting of HSPCs subsets from A17117 after CD34+ enrichment. FIG. **53**F shows HbG editing efficiency in HSPCs sorted subsets from FIG. **53**E determined by Miseq analysis.
(54) FIGS. **54**A-**54**D demonstrates validation of CD90+ editing approach. FIGS. **54**A-**54**D show results pre-editing and post-editing, including flow cytometry sorting (FIG. **54**A), showing colony forming cells (FIG. **54**B), measuring percent indels by Miseq analysis (FIG. **54**C), and measuring HbG editing efficiency (measured by TIDE) (FIG. **54**D).
(55) FIGS. **55**A-**55**B provide flow cytometry analysis of infusion product pre- (FIG. **55**A) and post- (FIG. **55**B) CRISPR/Cas9 editing.
(56) FIGS. **56**A-**56**B provides colony-forming assays of infusion product pre- (FIG. **56**A) and post- (FIG. **56**B) editing.
(57) FIG. **57** shows quantitative PCR measurement of hemoglobin expression in peripheral blood of transplanted animals.
(58) FIG. **58** shows validation of RP-HPLC approach for analysis of hemoglobin expression from NHP peripheral blood.
(59) FIG. **59** shows representative HPLC profiles of peripheral blood from experimental animals pre- and post-transplant.
(60) FIGS. **60**A-**60**F provides validation of sorting approach from bone marrow analysis of animal models A17114 and A17116.
(61) FIGS. **61**A-**61**B shows. FIG. **61**A shows longitudinal analysis of gamma globin (HbG) editing in peripheral blood of transplanted animals as determined by TIDE analysis. FIG. **61**B shows

frequency of circulating F-cells in transplanted animals as compared to control transplants (grey) and to an untransplanted control (black).
(62) FIGS. **62**A-**62**C demonstrate targeted integration of Globin T87Q/GFP donor cassette by HDR in Rhesus CD34+ HSPCs. FIG. **62**A provides a schematic of targeted integration strategy using combination of CRISPR/Cas9 and AAV donor delivery. FIG. **62**B shows HDR time course experiment in rhesus CD34+ cells treated with or without CRISPR/Cas9 RNPs and with 2 different doses of AAV donor. FIG. **62**C provides representative flow plots of treated cells (day 7) showing GFP positive cells as surrogate for HDR events.
BRIEF DESCRIPTION OF THE SEQUENCE IDENTIFIERS
(63) SEQ ID NOs: 1-6 set forth polynucleotide sequences in the human γ-globin promoter, that when disrupted, are associated with HPFH.
(64) SEQ ID NOs: 7-8 set forth polynucleotide sequence of the human γ-globin gene and its reverse complement, both of which can be used to target donor repair templates contemplated herein.
(65) SEQ ID NO: 9 sets forth a polynucleotide sequence of 1 kb upstream of the transcriptional start site of the human γ-globin gene.
(66) SEQ ID NOs: 10-57 set forth various AAV donor templates.
(67) SEQ ID NOs: 58-59 set forth TALEN plasmid sequences.
(68) SEQ ID NOs: 60-69 set forth polynucleotide sequences for exemplary 2A sites.
(69) SEQ ID NOs: 70-80 set forth linker sequences.
DETAILED DESCRIPTION
A. Overview
(70) The present disclosure generally relates to, in part, improved genome editing compositions for use in homology directed repair (HDR). Without wishing to be bound by any particular theory, the DNA donor repair templates contemplated herein are used to increase the amount of a therapeutic globin in a cell and to select the cells. The therapeutic cells can be used to treat, prevent, or ameliorate at least one symptom associated with a hemoglobinopathy.
(71) Normal adult hemoglobin comprises a tetrameric complex of two alpha-(α) globin proteins and two beta-(β-) globin proteins. In development, the fetus produces fetal hemoglobin (HbF), which comprises two gamma-(γ) globin proteins instead of the two β-globin proteins. At some point during perinatal development, a “globin switch” occurs; erythrocytes down-regulate γ-globin expression and switch to predominantly producing β-globin. This switch results primarily from decreased transcription of the γ-globin genes and increased transcription of β-globin genes.
(72) There is a segment of the human population that has deletions in various regions of the globin locus that lead to a condition known as Hereditary Persistence of Fetal Hemoglobin (HPFH). The deletions associated with HPFH are associated with increases in HbF in adulthood and are referred to herein collectively as HPFH deletions, a number of which are described herein, and others that are known in the art. HPFH is not associated with any significant clinical manifestations, even when 100% of the individual's hemoglobin is HbF. Thus, individuals that have a hemoglobinopathy that also have HPFH, have increased HbF expression, which can lessen the severity of the disease.
(73) In particular preferred embodiments, the genome editing compositions contemplated herein are used to engineer the globin locus to phenocopy a 13 bp deletion in the γ-globin promoter associated with HPFH (Gilman et al. *Nucleic Acids Res.* 1988 Nov. 25; 16(22): 10635-10642). Without wishing to be bound by any particular theory, it is contemplated that the DNA donor repair templates contemplated herein can be used to derepress the γ-globin locus to drive expression of a therapeutic globin gene and to select genome edited cells. The DNA donor repair templates contemplated herein are also advantageous because they can be designed to alter both the A-γ-globin gene and the G-γ-globin gene. A further advantage is that the engineered nuclease(s) used to introduce a DSB into the locus to facilitate HDR in the presence of a DNA donor repair template, can still lead to therapeutic editing in the absence of a DNA donor repair template because cleavage

of the 13 bp target sequence will be repaired by NHEJ, thereby producing indels at the target site, disrupting the repressive function of the intact 13 bp sequence, and derepressing γ-globin expression.

(74) In other particular preferred embodiments, the genome edited compositions contemplated herein are used to increase expression of endogenous γ-globin and select for edited cells. Without wishing to be bound by any particular theory, it is contemplated that the DNA donor repair templates contemplated herein can be used for selection and to introduce a β-globin LCR responsive promoter in operable linkage to an endogenous γ-globin gene to drive expression of endogenous γ-globin. The DNA donor repair templates contemplated herein are also advantageous because they can be designed to alter both the A-γ-globin gene and the G-γ-globin gene. A further advantage is that the engineered nuclease(s) used to introduce a DSB into the locus to facilitate HDR in the presence of a DNA donor repair template, can still lead to therapeutic editing in the absence of a DNA donor repair template because cleavage of the 13 bp target sequence will be repaired by NHEJ, thereby producing indels at the target site, disrupting the repressive function of the intact 13 bp sequence, and derepressing γ-globin expression.

(75) The engineered nucleases contemplated herein can be used to introduce a double-strand break in a target polynucleotide sequence, which may be repaired by non-homologous end joining (NHEJ) in the absence of a polynucleotide template, e.g., a donor repair template, or by homology directed repair (HDR), i.e., homologous recombination, in the presence of a donor repair template. Nuclease variants contemplated in certain embodiments, can also be designed as nickases, which generate single-stranded DNA breaks that can be repaired using the cell's base-excision-repair (BER) machinery or homologous recombination in the presence of a donor repair template. NHEJ is an error-prone process that frequently results in the formation of small insertions and deletions that disrupt gene function. Homologous recombination requires homologous DNA as a template for repair and can be leveraged to create a limitless variety of modifications specified by the introduction of donor DNA containing the desired sequence at the target site, flanked on either side by sequences bearing homology to regions flanking the target site.

(76) Genome edited cells engineered by HDR with a DNA donor repair template and one or more engineered nucleases are contemplated in particular embodiments.

(77) Genome edited cells and compositions comprising the same are also contemplated for use in the treatment, prevention, and/or amelioration of at least one symptom of a hemoglobinopathy.

(78) Accordingly, the methods and compositions contemplated herein represent a quantum improvement compared to existing gene editing strategies for the treatment of hemoglobinopathies.

(79) The practice of the particular embodiments will employ, unless indicated specifically to the contrary, conventional methods of chemistry, biochemistry, organic chemistry, molecular biology, microbiology, recombinant DNA techniques, genetics, immunology, and cell biology that are within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See e.g., Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (3rd Edition, 2001); Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); Maniatis et al., *Molecular Cloning: A Laboratory Manual* (1982); Ausubel et al., *Current Protocols in Molecular Biology* (John Wiley and Sons, updated July 2008); *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, Greene Pub. Associates and Wiley-Interscience; Glover, *DNA Cloning: A Practical Approach*, vol. I & II (IRL Press, Oxford, 1985); Anand, *Techniques for the Analysis of Complex Genomes*, (Academic Press, New York, 1992); *Transcription and Translation* (B. Hames & S. Higgins, Eds., 1984); Perbal, *A Practical Guide to Molecular Cloning* (1984); Harlow and Lane, *Antibodies*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1998) *Current Protocols in Immunology* Q. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach and W. Strober, eds., 1991); *Annual Review of Immunology*; as well as monographs in journals such as *Advances in Immunology.*

B. Definitions

(80) Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of particular embodiments, preferred embodiments of compositions, methods and materials are described herein. For the purposes of the present disclosure, the following terms are defined and in some cases elaborated on below.

(81) The articles “a,” “an,” and “the” are used herein to refer to one or to more than one (i.e., to at least one, or to one or more) of the grammatical object of the article. By way of example, “an element” means one element or one or more elements.

(82) The use of the alternative (e.g., “or”) should be understood to mean either one, both, or any combination thereof of the alternatives.

(83) The term “and/or” should be understood to mean either one, or both of the alternatives.

(84) As used herein, the term “about” or “approximately” refers to a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that varies by as much as 15%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2% or 1% to a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length. In one embodiment, the term “about” or “approximately” refers a range of quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length ±15%, ±10%, ±9%, ±8%, ±7%, ±6%, ±5%, ±4%, ±3%, ±2%, or ±1% about a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length.

(85) In one embodiment, a range, e.g., 1 to 5, about 1 to 5, or about 1 to about 5, refers to each numerical value encompassed by the range. For example, in one non-limiting and merely illustrative embodiment, the range “1 to 5” is equivalent to the expression 1, 2, 3, 4, 5; or 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, or 5.0; or 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, or 5.0.

(86) As used herein, the term “substantially” refers to a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that is 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or higher compared to a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length. In one embodiment, “substantially the same” refers to a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that produces an effect, e.g., a physiological effect, that is approximately the same as a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length.

(87) Throughout this specification, unless the context requires otherwise, the words “comprise”, “comprises” and “comprising” will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements. By “consisting of” is meant including, and limited to, whatever follows the phrase “consisting of” Thus, the phrase “consisting of” indicates that the listed elements are required or mandatory, and that no other elements may be present. By “consisting essentially of” is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase “consisting essentially of” indicates that the listed elements are required or mandatory, but that no other elements are present that materially affect the activity or action of the listed elements.

(88) Reference throughout this specification to “one embodiment,” “an embodiment,” “a particular embodiment,” “a related embodiment,” “a certain embodiment,” “an additional embodiment,” or “a further embodiment” or combinations thereof means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment.

Thus, the appearances of the foregoing phrases in various places throughout this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments. It is also understood that the positive recitation of a feature in one embodiment, serves as a basis for excluding the feature in a particular embodiment.

(89) The term “ex vivo” refers generally to activities that take place outside an organism, such as experimentation or measurements done in or on living tissue in an artificial environment outside the organism, preferably with minimum alteration of the natural conditions. In particular embodiments, “ex vivo” procedures involve living cells or tissues taken from an organism and cultured or modulated in a laboratory apparatus, usually under sterile conditions, and typically for a few hours or up to about 24 hours, but including up to 48 or 72 hours, depending on the circumstances. In certain embodiments, such tissues or cells can be collected and frozen, and later thawed for ex vivo treatment. Tissue culture experiments or procedures lasting longer than a few days using living cells or tissue are typically considered to be “in vitro,” though in certain embodiments, this term can be used interchangeably with ex vivo.

(90) The term “in vivo” refers generally to activities that take place inside an organism. In one embodiment, cellular genomes are engineered, edited, or modified in vivo.

(91) By “enhance” or “promote” or “increase” or “expand” or “potentiate” refers generally to the ability of a DNA donor repair template, genome editing composition, or genome edited cell contemplated herein to produce, elicit, or cause a greater response (i.e., physiological response) compared to the response caused by either vehicle or control. A measurable response may include an increase in γ-globin expression, HbF expression, and/or an increase in transfusion independence, among others apparent from the understanding in the art and the description herein. An “increased” or “enhanced” amount is typically a “statistically significant” amount, and may include an increase that is 1.1, 1.2, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30 or more times (e.g., 500, 1000 times) (including all integers and decimal points in between and above 1, e.g., 1.5, 1.6, 1.7. 1.8, etc.) the response produced by vehicle or control.

(92) By “decrease” or “lower” or “lessen” or “reduce” or “abate” or “ablate” or “inhibit” or “dampen” refers generally to the ability of a DNA donor repair template, genome editing composition, or genome edited cell contemplated herein to produce, elicit, or cause a lesser response (i.e., physiological response) compared to the response caused by either vehicle or control. A measurable response may include a decrease in endogenous β-globin, transfusion dependence, RBC sickling, and the like. A “decrease” or “reduced” amount is typically a “statistically significant” amount, and may include a decrease that is 1.1, 1.2, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30 or more times (e.g., 500, 1000 times) (including all integers and decimal points in between and above 1, e.g., 1.5, 1.6, 1.7. 1.8, etc.) the response (reference response) produced by vehicle, or control.

(93) By “maintain,” or “preserve,” or “maintenance,” or “no change,” or “no substantial change,” or “no substantial decrease” refers generally to the ability of a DNA donor repair template, genome editing composition, or genome edited cell contemplated herein to produce, elicit, or cause a substantially similar or comparable physiological response (i.e., downstream effects) in as compared to the response caused by either vehicle or control. A comparable response is one that is not significantly different or measurably different from the reference response.

(94) A “target site” or “target sequence” is a chromosomal or extrachromosomal nucleic acid sequence that defines a portion of a nucleic acid to which a binding molecule will bind and/or cleave, provided sufficient conditions for binding and/or cleavage exist. When referring to a polynucleotide sequence or SEQ ID NO that references only one strand of a target site or target sequence, it would be understood that the target site or target sequence bound and/or cleaved by a nuclease variant is double-stranded and comprises the reference sequence and its complement. In a preferred embodiment, the target site is a sequence in the human gamma globin gene that when

disrupted, is associated with an HPFH phenotype.
(95) “Recombination” refers to a process of exchange of genetic information between two polynucleotides, including but not limited to, donor capture by non-homologous end joining (NHEJ) and homologous recombination. For the purposes of this disclosure, “homologous recombination (HR)” refers to the specialized form of such exchange that takes place, for example, during repair of double-strand breaks in cells via homology-directed repair (HDR) mechanisms. This process requires nucleotide sequence homology, uses a “donor” molecule as a template to repair a “target” molecule (i.e., the one that experienced the double-strand break), and is variously known as “non-crossover gene conversion” or “short tract gene conversion,” because it leads to the transfer of genetic information from the donor to the target. Without wishing to be bound by any particular theory, such transfer can involve mismatch correction of heteroduplex DNA that forms between the broken target and the donor, and/or “synthesis-dependent strand annealing,” in which the donor is used to resynthesize genetic information that will become part of the target, and/or related processes. Such specialized HR often results in an alteration of the sequence of the target molecule such that part or all of the sequence of the donor polynucleotide is incorporated into the target polynucleotide.
(96) “NHEJ” or “non-homologous end joining” refers to the resolution of a double-strand break in the absence of a donor repair template or homologous sequence. NHEJ can result in insertions and deletions at the site of the break. NHEJ is mediated by several sub-pathways, each of which has distinct mutational consequences. The classical NHEJ pathway (cNHEJ) requires the KU/DNA-PKcs/Lig4/XRCC4 complex, ligates ends back together with minimal processing and often leads to precise repair of the break. Alternative NHEJ pathways (altNHEJ) also are active in resolving dsDNA breaks, but these pathways are considerably more mutagenic and often result in imprecise repair of the break marked by insertions and deletions. While not wishing to be bound to any particular theory, it is contemplated that modification of dsDNA breaks by end-processing enzymes, such as, for example, exonucleases, e.g., Trex2, may bias repair towards an altNHEJ pathway.
(97) “Cleavage” refers to the breakage of the covalent backbone of a DNA molecule. Cleavage can be initiated by a variety of methods including, but not limited to, enzymatic or chemical hydrolysis of a phosphodiester bond. Both single-stranded cleavage and double-stranded cleavage are possible. Double-stranded cleavage can occur as a result of two distinct single-stranded cleavage events. DNA cleavage can result in the production of either blunt ends or staggered ends. In certain embodiments, polypeptides and nuclease variants, e.g., homing endonuclease variants, megaTALs, etc. contemplated herein are used for targeted double-stranded DNA cleavage. Endonuclease cleavage recognition sites may be on either DNA strand.
(98) An “exogenous” molecule is a molecule that is not normally present in a cell, but that is introduced into a cell by one or more genetic, biochemical or other methods. Exemplary exogenous molecules include, but are not limited to small organic molecules, protein, nucleic acid, carbohydrate, lipid, glycoprotein, lipoprotein, polysaccharide, any modified derivative of the above molecules, or any complex comprising one or more of the above molecules. Methods for the introduction of exogenous molecules into cells are known to those of skill in the art and include, but are not limited to, lipid-mediated transfer (i.e., liposomes, including neutral and cationic lipids), electroporation, direct injection, cell fusion, particle bombardment, biopolymer nanoparticle, calcium phosphate co-precipitation, DEAE-dextran-mediated transfer and viral vector-mediated transfer.
(99) An “endogenous” molecule is one that is normally present in a particular cell at a particular developmental stage under particular environmental conditions. Additional endogenous molecules can include proteins, for example, endogenous globins.
(100) A “gene,” refers to a DNA region encoding a gene product, as well as all DNA regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent

to coding and/or transcribed sequences. A gene includes, but is not limited to, promoter sequences, enhancers, silencers, insulators, boundary elements, terminators, polyadenylation sequences, post-transcription response elements, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, replication origins, matrix attachment sites, and locus control regions.

(101) “Gene expression” refers to the conversion of the information, contained in a gene, into a gene product. A gene product can be the direct transcriptional product of a gene (e.g., mRNA, tRNA, rRNA, antisense RNA, ribozyme, structural RNA or any other type of RNA) or a protein produced by translation of an mRNA. Gene products also include RNAs which are modified, by processes such as capping, polyadenylation, methylation, and editing, and proteins modified by, for example, methylation, acetylation, phosphorylation, ubiquitination, ADP-ribosylation, myristilation, and glycosylation.

(102) As used herein, the term “genetically engineered” or “genetically modified” refers to the chromosomal or extrachromosomal addition of extra genetic material in the form of DNA or RNA to the total genetic material in a cell. Genetic modifications may be targeted or non-targeted to a particular site in a cell's genome. In one embodiment, genetic modification is site specific. In one embodiment, genetic modification is not site specific.

(103) As used herein, the term “genome editing” refers to the substitution, deletion, and/or introduction of genetic material at a target site in the cell's genome, which restores, corrects, disrupts, and/or modifies expression of a gene or gene product. Genome editing contemplated in particular embodiments comprises introducing one or more nuclease variants into a cell to generate DNA lesions at or proximal to a target site in the cell's genome, optionally in the presence of a donor repair template.

(104) As used herein, the term “gene therapy” refers to the introduction of extra genetic material into the total genetic material in a cell that restores, corrects, or modifies expression of a gene or gene product, or for the purpose of expressing a therapeutic polypeptide. In particular embodiments, introduction of genetic material into the cell's genome by genome editing that restores, corrects, disrupts, or modifies expression of a gene or gene product, or for the purpose of expressing a therapeutic polypeptide is considered gene therapy.

(105) As used herein, the terms “polynucleotide” or “nucleic acid” refer to deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and DNA/RNA hybrids. Polynucleotides may be single-stranded or double-stranded and either recombinant, synthetic, or isolated. Polynucleotides include, but are not limited to: pre-messenger RNA (pre-mRNA), messenger RNA (mRNA), RNA, short interfering RNA (siRNA), short hairpin RNA (shRNA), microRNA (miRNA), ribozymes, genomic RNA (gRNA), plus strand RNA (RNA(+)), minus strand RNA (RNA(−)), tracrRNA, crRNA, single guide RNA (sgRNA), synthetic RNA, synthetic mRNA, genomic DNA (gDNA), PCR amplified DNA, complementary DNA (cDNA), synthetic DNA, or recombinant DNA. Preferably, polynucleotides of the invention include polynucleotides or variants having at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to any of the reference sequences described herein (see, e.g., SEQ ID NOs: 1-9), typically where the variant maintains at least one biological activity of the reference sequence. In various illustrative embodiments, viral vector and transfer plasmid polynucleotide sequences and compositions comprising the same are contemplated. In particular embodiments, polynucleotides encoding one or more therapeutic polypeptides and/or other genes of interest are contemplated.

(106) As used herein, the terms “polynucleotide variant” and “variant” and the like refer to polynucleotides displaying substantial sequence identity with a reference polynucleotide sequence or polynucleotides that hybridize with a reference sequence under stringent conditions that are defined hereinafter. These terms include polynucleotides in which one or more nucleotides have been added or deleted, or replaced with different nucleotides compared to a reference polynucleotide. In this regard, it is well understood in the art that certain alterations inclusive of

mutations, additions, deletions and substitutions can be made to a reference polynucleotide whereby the altered polynucleotide retains the biological function or activity of the reference polynucleotide.

(107) As used herein, the term “isolated” means material, e.g., a polynucleotide, a polypeptide, a cell, that is substantially or essentially free from components that normally accompany it in its native state. In particular embodiments, the term “obtained” or “derived” is used synonymously with isolated. For example, an “isolated polynucleotide,” as used herein, refers to a polynucleotide that has been purified from the sequences which flank it in a naturally-occurring state, e.g., a DNA fragment that has been removed from the sequences that are normally adjacent to the fragment.

(108) The recitations “sequence identity” or, for example, comprising a “sequence 50% identical to,” as used herein, refer to the extent that sequences are identical on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a “percentage of sequence identity” may be calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, I) or the identical amino acid residue (e.g., Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. Included are nucleotides and polypeptides having at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to any of the reference sequences described herein, typically where the polypeptide variant maintains at least one biological activity of the reference polypeptide.

(109) Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include “reference sequence,” “comparison window,” “sequence identity,” “percentage of sequence identity,” and “substantial identity”. A “reference sequence” is at least 12 but frequently 15 to 18 and often at least 25 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (i.e., only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a “comparison window” to identify and compare local regions of sequence similarity. A “comparison window” refers to a conceptual segment of at least 6 contiguous positions, usually about 50 to about 100, more usually about 100 to about 150 in which a sequence is compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. The comparison window may comprise additions or deletions (i.e., gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (i.e., resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul et al., 1997, Nucl. Acids Res. 25:3389. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons Inc., 1994-1998, Chapter 15.

(110) Terms that describe the orientation of polynucleotides include: 5′ (normally the end of the polynucleotide having a free phosphate group) and 3′ (normally the end of the polynucleotide having a free hydroxyl (OH) group). Polynucleotide sequences can be annotated in the 5′ to 3′ orientation or the 3′ to 5′ orientation.

(111) The terms “complementary” and “complementarity” refer to polynucleotides (i.e., a sequence

of nucleotides) related by the base-pairing rules. For example, the complementary strand of the DNA sequence 5′ A G T C A T G 3′ is 3′ T C A G T A C 5′. The latter sequence is often written as the reverse complement with the 5′ end on the left and the 3′ end on the right, 5′ C A T G A C T 3′. A sequence that is equal to its reverse complement is said to be a palindromic sequence. Complementarity can be “partial,” in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there can be “complete” or “total” complementarity between the nucleic acids.

(112) The term “nucleic acid cassette” or “expression cassette” as used herein refers to genetic sequences within the vector which can express a polynucleotide. In one embodiment, the nucleic acid cassette contains a polynucleotide(s)-of-interest. In another embodiment, the nucleic acid cassette contains one or more expression control sequences, e.g., a promoter, enhancer, post-transcriptional regulatory element, poly(A) sequence, and a polynucleotide(s)-of-interest. Vectors may comprise one, two, three, four, five or more nucleic acid cassettes. The nucleic acid cassette is positionally and sequentially oriented within the vector such that the nucleic acid in the cassette can be transcribed into RNA. Preferably, the cassette has its 3′ and 5′ ends adapted for ready insertion into a vector, e.g., it has restriction endonuclease sites at each end. In a preferred embodiment, the nucleic acid cassette one or more expression control sequences operably linked to a polynucleotide encoding a therapeutic RNA, e.g., a shmiR, and/or a polypeptide, that can be used to treat, prevent, or ameliorate a genetic disorder. The cassette can be removed and inserted into a plasmid or viral vector as a single unit.

(113) As used herein, the term “polynucleotide(s)-of-interest” refers to one or more polynucleotides, e.g., a polynucleotide encoding a polypeptide (i.e., a polypeptide-of-interest), inserted into an expression vector that is desired to be expressed. In preferred embodiments, polynucleotides comprise one or more polynucleotides-of-interest that encode one or more therapeutic globins. In particular embodiments, the polynucleotide-of-interest is a transgene that encodes a polypeptide that provides a therapeutic function for the treatment of a hemoglobinopathy, e.g., α-globin, β-globin, β-globin.sup.A-T87Q, anti-sickling globins, γ-globin, and δ globin.

(114) Polynucleotides, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters and/or enhancers, untranslated regions (UTRs), Kozak sequences, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, internal ribosomal entry sites (IRES), recombinase recognition sites (e.g., LoxP, FRT, and Att sites), termination codons, transcriptional termination signals, and polynucleotides encoding self-cleaving polypeptides, epitope tags, as disclosed elsewhere herein or as known in the art, such that their overall length may vary considerably. It is therefore contemplated that a polynucleotide fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol.

(115) The term “expression control sequence” refers to a polynucleotide sequence that comprises one or more promoters, enhancers, or other transcriptional control elements or combinations thereof that are capable of directing, increasing, regulating, or controlling the transcription or expression of an operatively linked polynucleotide. In particular embodiments, vectors of the invention comprise one or more expression control sequences that are specific to particular erythroid cells, erythroid cell types, or erythroid cell lineages. In preferred embodiments, vectors comprise one or more expression control sequences specific to erythroid cells, e.g., an erythroid specific expression control sequence.

(116) An “endogenous” expression control sequence is one which is naturally linked to a given gene in the genome. An “exogenous” expression control sequence is one which is placed in juxtaposition to a gene by means of genetic manipulation (i.e., molecular biological techniques) such that transcription of that gene is directed by the linked enhancer/promoter. A “heterologous” expression control sequence is an exogenous sequence that is from a different species than the cell being genetically manipulated. A “synthetic” expression control sequence may comprise elements

of one or more endogenous and/or exogenous sequences, and/or sequences determined in vitro or in silico that provide optimal promoter and/or enhancer activity for the particular gene therapy. In particular embodiments, a vector comprises exogenous, endogenous, or heterologous expression control sequences such as promoters and/or enhancers.

(117) The term “promoter” as used herein refers to an expression control sequence that comprises a recognition site of a polynucleotide (DNA or RNA) to which an RNA polymerase binds. The term “enhancer” refers to an expression control sequence that comprises a segment of DNA which contains sequences capable of providing enhanced transcription and in some instances can function independent of their orientation relative to another control sequence. An enhancer can function cooperatively or additively with promoters and/or other enhancer elements. The term “promoter/enhancer” refers to a segment of DNA which contains sequences capable of providing both promoter and enhancer functions.

(118) The term “operably linked”, refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. In one embodiment, the term refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, and/or enhancer or other expression control sequence) and a second polynucleotide sequence, e.g., a polynucleotide-of-interest, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

(119) The terms “polypeptide” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues and to variants and synthetic analogues of the same. Thus, these terms apply to amino acid polymers in which one or more amino acid residues are synthetic non-naturally occurring amino acids, such as a chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally-occurring amino acid polymers. Illustrative examples of polypeptides include, but are not limited to globin polypeptides, suitable for use in the compositions and methods of particular embodiments. Also, see, e.g., U.S. Pat. Nos. 6,051,402; 7,901,671; and 9,068,199, the full disclosure and claims of which are specifically incorporated herein by reference in their entireties.

(120) Particular embodiments contemplated herein, also include polypeptide “variants.” The recitation polypeptide “variant” refers to polypeptides that are distinguished from a reference polypeptide by the addition, deletion, truncations, modifications, and/or substitution of at least one amino acid residue, and that retain a biological activity. In certain embodiments, a polypeptide variant is distinguished from a reference polypeptide by one or more substitutions, which may be conservative or non-conservative, as known in the art. In certain embodiments, a variant polypeptide includes an amino acid sequence having at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity or similarity to a corresponding sequence of a reference polypeptide. In certain embodiments, amino acid additions or deletions occur at the C-terminal end and/or the N-terminal end of the reference polypeptide.

(121) Polypeptides contemplated in particular embodiments include fusion polypeptides. “Fusion polypeptides” or “fusion proteins” refer to a polypeptide having at least two, three, four, five, six, seven, eight, nine, or ten polypeptide segments, including, but not limited to one or more linker and/or self-cleaving polypeptides.

(122) A peptide “linker” sequence refers to a polypeptide sequence that may be employed to separate any two or more polypeptide components by a distance sufficient to ensure that each polypeptide folds into its appropriate secondary and tertiary structures so as to allow the polypeptide domains to exert their desired functions. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues

that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., Gene 40:39-46, 1985; Murphy et al., Proc. Natl. Acad. Sci. USA 83:8258-8262, 1986; U.S. Pat. Nos. 4,935,233 and 4,751,180. Linker sequences are not required when a particular fusion polypeptide segment contains non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference. Preferred linkers are typically flexible amino acid subsequences which are synthesized as part of a recombinant fusion protein. Linker polypeptides can be between 1 and 200 amino acids in length, between 1 and 100 amino acids in length, or between 1 and 50 amino acids in length, including all integer values in between. Exemplary linkers include, but are not limited to the following amino acid sequences: glycine polymers (G)n; glycine-serine polymers (G1-5S1-5)n, where n is an integer of at least one, two, three, four, or five; glycine-alanine polymers; alanine-serine polymers; GGG (SEQ ID NO: 70); DGGGS (SEQ ID NO: 71); TGEKP (SEQ ID NO: 72) (see e.g., Liu et al., PNAS 5525-5530 (1997)); GGRR (SEQ ID NO: 73) (Pomerantz et al. 1995, supra); (GGGGS)n wherein n=1, 2, 3, 4 or 5 (SEQ ID NO: 74) (Kim et al., PNAS 93, 1156-1160 (1996); EGKSSGSGSESKVD (SEQ ID NO: 75) (Chaudhary et al., 1990, Proc. Natl. Acad. Sci. U.S.A. 87:1066-1070); KESGSVSSEQLAQFRSLD (SEQ ID NO: 76) (Bird et al., 1988, Science 242:423-426), GGRRGGGS (SEQ ID NO: 77); LRQRDGERP (SEQ ID NO: 78); LRQKDGGGSERP (SEQ ID NO: 79); LRQKD(GGGS)2ERP (SEQ ID NO: 80). Alternatively, flexible linkers can be rationally designed using a computer program capable of modeling both DNA-binding sites and the peptides themselves (Desjarlais & Berg, PNAS 90:2256-2260 (1993), PNAS 91:11099-11103 (1994) or by phage display methods.

C. Engineered Nucleases

(123) In particular embodiments, HDR at target sites comprises introducing one or more double strand breaks (DSB) at a target site using one or more engineered nucleases in the presence of a DNA donor repair template contemplated herein. The terms “reprogrammed nuclease,” “engineered nuclease,” or “nuclease variant” are used interchangeably and refer to a nuclease comprising one or more DNA binding domains and one or more DNA cleavage domains, wherein the nuclease has been designed and/or modified from a parental or naturally occurring nuclease, to bind and cleave a double-stranded DNA target sequence. The nuclease variant may be designed and/or modified from a naturally occurring nuclease or from a previous nuclease variant. Nuclease variants contemplated in particular embodiments may further comprise one or more additional functional domains, e.g., an end-processing enzymatic domain of an end-processing enzyme that exhibits 5′-3′ exonuclease, 5′-3′ alkaline exonuclease, 3′-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase, template-dependent DNA polymerase or template-independent DNA polymerase activity.

(124) Illustrative examples of nucleases that may be engineered to bind and cleave a target sequence include, but are not limited to homing endonucleases (meganucleases), megaTALs, transcription activator-like effector nucleases (TALENs), zinc finger nucleases (ZFNs), and clustered regularly-interspaced short palindromic repeats (CRISPR)/Cas nuclease systems.

(125) 1. Homing Endonucleases/Meganucleases

(126) In various embodiments, a plurality of homing endonucleases or meganucleases are introduced into a cell and engineered to bind to, and to introduce single-stranded nicks or double-strand breaks (DSBs) at a target site in a human γ-globin gene, e.g., a polynucleotide sequence as set for the in any one of SEQ ID NOs: 1-9. “Homing endonuclease” and “meganuclease” are used interchangeably and refer to naturally-occurring nucleases or engineered meganucleases that recognize 12-45 base-pair cleavage sites and are commonly grouped into five families based on sequence and structure motifs: LAGLIDADG, GIY-YIG, HNH, His-Cys box, and PD-(D/E)XK.

(127) Engineered HEs do not exist in nature and can be obtained by recombinant DNA technology or by random mutagenesis. Engineered HEs may be obtained by making one or more amino acid

alterations, e.g., mutating, substituting, adding, or deleting one or more amino acids, in a naturally occurring HE or previously engineered HE. In particular embodiments, an engineered HE comprises one or more amino acid alterations to the DNA recognition interface.

(128) Engineered HEs contemplated in particular embodiments may further comprise one or more linkers and/or additional functional domains, e.g., an end-processing enzymatic domain of an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. In particular embodiments, engineered HEs are introduced into a hematopoietic cell with an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. The HE and 3′ processing enzyme may be introduced separately, e.g., in different vectors or separate mRNAs, or together, e.g., as a fusion protein, or in a polycistronic construct separated by a viral self-cleaving peptide or an IRES element.

(129) A “DNA recognition interface” refers to the HE amino acid residues that interact with nucleic acid target bases as well as those residues that are adjacent. For each HE, the DNA recognition interface comprises an extensive network of side chain-to-side chain and side chain-to-DNA contacts, most of which is necessarily unique to recognize a particular nucleic acid target sequence. Thus, the amino acid sequence of the DNA recognition interface corresponding to a particular nucleic acid sequence varies significantly and is a feature of any natural or engineered HE. By way of non-limiting example, an engineered HE contemplated in particular embodiments may be derived by constructing libraries of HE variants in which one or more amino acid residues localized in the DNA recognition interface of the natural HE (or a previously engineered HE) are varied. The libraries may be screened for target cleavage activity against each predicted target site using cleavage assays (see e.g., Jarjour et al., 2009. *Nuc. Acids Res.* 37(20): 6871-6880).

(130) LAGLIDADG homing endonucleases (LHE) are the most well studied family of meganucleases, are primarily encoded in archaea and in organellar DNA in green algae and fungi, and display the highest overall DNA recognition specificity. LHEs comprise one or two LAGLIDADG catalytic motifs per protein chain and function as homodimers or single chain monomers, respectively. Structural studies of LAGLIDADG proteins identified a highly conserved core structure (Stoddard 2005), characterized by an αββαββα fold, with the LAGLIDADG motif belonging to the first helix of this fold. The highly efficient and specific cleavage of LHE's represent a protein scaffold to derive novel, highly specific endonucleases. However, engineering LHEs to bind and cleave a non-natural or non-canonical target site requires selection of the appropriate LHE scaffold, examination of the target locus, selection of putative target sites, and extensive alteration of the LHE to alter its DNA contact points and cleavage specificity, at up to two-thirds of the base-pair positions in a target site.

(131) Illustrative examples of LHEs from which engineered LHEs may be designed include, but are not limited to I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-Ncrl, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, and I-Vdi141I.

(132) Other illustrative examples of LHEs from which engineered LHEs may be designed include, but are not limited to I-CreI and I-SceI.

(133) In one embodiment, the engineered LHE is selected from the group consisting of: I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, and SmaMI.

(134) In one embodiment, the engineered LHE is I-OnuI.

(135) In a particular embodiment, the engineered I-OnuI LHE comprises one or more amino acid substitutions in the DNA recognition interface. In particular embodiments, the I-OnuI LHE comprises at least 70%, at least 71%, at least 72%, at least 73%, at least 74%, at least 75%, at least

76%, at least 77%, at least 78%, at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity with the DNA recognition interface of I-OnuI (Taekuchi et al. 2011. *Proc Natl Acad Sci U.S.A.* 2011 Aug. 9; 108(32): 13077-13082) or an engineered variant of I-OnuI.

(136) In one embodiment, the I-OnuI LHE comprises at least 70%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 95%, more preferably at least 97%, more preferably at least 99% sequence identity with the DNA recognition interface of I-OnuI (Taekuchi et al. 2011. *Proc Natl Acad Sci U.S.A.* 2011 Aug. 9; 108(32): 13077-13082) or an engineered variant of I-OnuI.

(137) In a particular embodiment, an engineered I-OnuI LHE comprises one or more amino acid substitutions or modifications in the DNA recognition interface, particularly in the subdomains situated from positions 24-50, 68 to 82, 180 to 203 and 223 to 240 of I-OnuI.

(138) In one embodiment, an engineered I-OnuI LHE comprises one or more amino acid substitutions or modifications at additional positions situated anywhere within the entire I-OnuI sequence. The residues which may be substituted and/or modified include but are not limited to amino acids that contact the nucleic acid target or that interact with the nucleic acid backbone or with the nucleotide bases, directly or via a water molecule. In one non-limiting example an engineered I-OnuI LHE contemplated herein comprises one or more substitutions and/or modifications, preferably at least 5, preferably at least 10, preferably at least 15, more preferably at least 20, even more preferably at least 25 in at least one position selected from the position group consisting of positions: 19, 24, 26, 28, 30, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 72, 75, 76 77, 78, 80, 82, 168, 180, 182, 184, 186, 188, 189, 190, 191, 192, 193, 195, 197, 199, 201, 203, 223, 225, 227, 229, 231, 232, 234, 236, 238, 240 of I-OnuI.

(139) 2. MegaTALs

(140) In various embodiments, one or more megaTALs are introduced into a hematopoietic cell and engineered to bind and introduce DSBs at a target site in a human γ-globin gene, e.g., a polynucleotide sequence as set for the in any one of SEQ ID NOs: 1-9. A “megaTAL” refers to an engineered nuclease comprising an engineered TALE DNA binding domain and an engineered meganuclease, and optionally comprise one or more linkers and/or additional functional domains, e.g., an end-processing enzymatic domain of an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. In particular embodiments, a megaTAL can be introduced into a T cell with an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. The megaTAL and 3′ processing enzyme may be introduced separately, e.g., in different vectors or separate mRNAs, or together, e.g., as a fusion protein, or in a polycistronic construct separated by a viral self-cleaving peptide or an IRES element.

(141) A “TALE DNA binding domain” is the DNA binding portion of transcription activator-like effectors (TALE or TAL-effectors), which mimics plant transcriptional activators to manipulate the plant transcriptome (see e.g., Kay et al., 2007. *Science* 318:648-651). TALE DNA binding domains contemplated in particular embodiments are engineered de novo or from naturally occurring TALEs, e.g., AvrBs3 from *Xanthomonas campestris* pv. *vesicatoria*, *Xanthomonas gardneri*, *Xanthomonas translucens*, *Xanthomonas axonopodis*, *Xanthomonas perforans*, *Xanthomonas alfalfa*, *Xanthomonas citri*, *Xanthomonas euvesicatoria*, and *Xanthomonas oryzae* and brg11 and hpx17 from *Ralstonia solanacearum*. Illustrative examples of TALE proteins for deriving and designing DNA binding domains are disclosed in U.S. Pat. No. 9,017,967, and references cited therein, all of which are incorporated herein by reference in their entireties.

(142) In particular embodiments, a megaTAL comprises a TALE DNA binding domain comprising one or more repeat units that are involved in binding of the TALE DNA binding domain to its corresponding target DNA sequence. A single "repeat unit" (also referred to as a "repeat") is typically 33-35 amino acids in length. Each TALE DNA binding domain repeat unit includes 1 or 2 DNA-binding residues making up the Repeat Variable Di-Residue (RVD), typically at positions 12 and/or 13 of the repeat. The natural (canonical) code for DNA recognition of these TALE DNA binding domains has been determined such that an HD sequence at positions 12 and 13 leads to a binding to cytosine (C), NG binds to T, NI to A, NN binds to G or A, and NG binds to T. In certain embodiments, non-canonical (atypical) RVDs are contemplated.
(143) Illustrative examples of non-canonical RVDs suitable for use in particular megaTALs contemplated in particular embodiments include, but are not limited to HH, KH, NH, NK, NQ, RH, RN, SS, NN, SN, KN for recognition of guanine (G); NI, KI, RI, HI, SI for recognition of adenine (A); NG, HG, KG, RG for recognition of thymine (T); RD, SD, HD, ND, KD, YG for recognition of cytosine (C); NV, HN for recognition of A or G; and H*, HA, KA, N*, NA, NC, NS, RA, S* for recognition of A or T or G or C, wherein (*) means that the amino acid at position 13 is absent. Additional illustrative examples of RVDs suitable for use in particular megaTALs contemplated in particular embodiments further include those disclosed in U.S. Pat. No. 8,614,092, which is incorporated herein by reference in its entirety.
(144) In particular embodiments, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 3 to 30 repeat units. In certain embodiments, a megaTAL comprises 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 TALE DNA binding domain repeat units. In a preferred embodiment, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 5-16 repeat units, more preferably 7-15 repeat units, more preferably 9-15 repeat units, and more preferably 9, 10, 11, 12, 13, 14, or 15 repeat units.
(145) In particular embodiments, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 3 to 30 repeat units and an additional single truncated TALE repeat unit comprising 20 amino acids located at the C-terminus of a set of TALE repeat units, i.e., an additional C-terminal half-TALE DNA binding domain repeat unit (amino acids −20 to −1 of the C-cap disclosed elsewhere herein, infra). Thus, in particular embodiments, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 3.5 to 30.5 repeat units. In certain embodiments, a megaTAL comprises 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 13.5, 14.5, 15.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, or 30.5 TALE DNA binding domain repeat units. In a preferred embodiment, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 5.5-13.5 repeat units, more preferably 7.5-12.5 repeat units, more preferably 9.5-15.5 repeat units, and more preferably 9.5, 10.5, 11.5, 12.5, 13.5, 14.5, or 15.5 repeat units.
(146) In particular embodiments, a megaTAL comprises an "N-terminal domain (NTD)" polypeptide, one or more TALE repeat domains/units, a "C-terminal domain (CTD)" polypeptide, and an engineered meganuclease.
(147) As used herein, the term "N-terminal domain (NTD)" polypeptide refers to the sequence that flanks the N-terminal portion or fragment of a naturally occurring TALE DNA binding domain. The NTD sequence, if present, may be of any length as long as the TALE DNA binding domain repeat units retain the ability to bind DNA. In particular embodiments, the NTD polypeptide comprises at least 120 to at least 140 or more amino acids N-terminal to the TALE DNA binding domain (0 is amino acid 1 of the most N-terminal repeat unit). In particular embodiments, the NTD polypeptide comprises at least about 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, or at least 140 amino acids N-terminal to the TALE DNA binding domain. In one embodiment, a megaTAL contemplated herein comprises an NTD polypeptide of at least about amino acids +1 to +122 to at least about +1 to +137 of a *Xanthomonas*

TALE protein (0 is amino acid 1 of the most N-terminal repeat unit). In particular embodiments, the NTD polypeptide comprises at least about 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, or 137 amino acids N-terminal to the TALE DNA binding domain of a *Xanthomonas* TALE protein. In one embodiment, a megaTAL contemplated herein comprises an NTD polypeptide of at least amino acids +1 to +121 of a *Ralstonia* TALE protein (0 is amino acid 1 of the most N-terminal repeat unit). In particular embodiments, the NTD polypeptide comprises at least about 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, or 137 amino acids N-terminal to the TALE DNA binding domain of a *Ralstonia* TALE protein.

(148) As used herein, the term "C-terminal domain (CTD)" polypeptide refers to the sequence that flanks the C-terminal portion or fragment of a naturally occurring TALE DNA binding domain. The CTD sequence, if present, may be of any length as long as the TALE DNA binding domain repeat units retain the ability to bind DNA. In particular embodiments, the CTD polypeptide comprises at least 20 to at least 85 or more amino acids C-terminal to the last full repeat of the TALE DNA binding domain (the first 20 amino acids are the half-repeat unit C-terminal to the last C-terminal full repeat unit). In particular embodiments, the CTD polypeptide comprises at least about 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 443, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, or at least 85 amino acids C-terminal to the last full repeat of the TALE DNA binding domain. In one embodiment, a megaTAL contemplated herein comprises a CTD polypeptide of at least about amino acids −20 to −1 of a *Xanthomonas* TALE protein (−20 is amino acid 1 of a half-repeat unit C-terminal to the last C-terminal full repeat unit). In particular embodiments, the CTD polypeptide comprises at least about 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acids C-terminal to the last full repeat of the TALE DNA binding domain of a *Xanthomonas* TALE protein. In one embodiment, a megaTAL contemplated herein comprises a CTD polypeptide of at least about amino acids −20 to −1 of a *Ralstonia* TALE protein (−20 is amino acid 1 of a half-repeat unit C-terminal to the last C-terminal full repeat unit). In particular embodiments, the CTD polypeptide comprises at least about 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acids C-terminal to the last full repeat of the TALE DNA binding domain of a *Ralstonia* TALE protein.

(149) In particular embodiments, a megaTAL contemplated herein, comprises a fusion polypeptide comprising a TALE DNA binding domain engineered to bind a target sequence, a meganuclease engineered to bind and cleave a target sequence, and optionally an NTD and/or CTD polypeptide, optionally joined to each other with one or more linker polypeptides contemplated elsewhere herein. Without wishing to be bound by any particular theory, it is contemplated that a megaTAL comprising TALE DNA binding domain, and optionally an NTD and/or CTD polypeptide is fused to a linker polypeptide which is further fused to an engineered meganuclease. Thus, the TALE DNA binding domain binds a DNA target sequence that is within about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 nucleotides away from the target sequence bound by the DNA binding domain of the meganuclease. In this way, the megaTALs contemplated herein, increase the specificity and efficiency of genome editing.

(150) In particular embodiments, a megaTAL contemplated herein, comprises one or more TALE DNA binding repeat units and an engineered LHE selected from the group consisting of: I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-CreI, I-SceI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-Ncrl, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, and I-Vdi141I, or preferably I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, and SmaMI, or more preferably I-OnuI.

(151) In particular embodiments, a megaTAL contemplated herein, comprises an NTD, one or more TALE DNA binding repeat units, a CTD, and an engineered LHE selected from the group

consisting of I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-CreI, I-SceI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-Ncrl, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, and I-Vdi141I, or preferably I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, and SmaMI, or more preferably I-OnuI.

(152) In particular embodiments, a megaTAL contemplated herein, comprises an NTD, about 9.5 to about 15.5 TALE DNA binding repeat units, and an engineered I-OnuI LHE selected from the group consisting of: I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-CreI, I-SceI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-Ncrl, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, and I-Vdi141I, or preferably I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, and SmaMI, or more preferably I-OnuI.

(153) In particular embodiments, a megaTAL contemplated herein, comprises an NTD of about 122 amino acids to 137 amino acids, about 9.5, about 10.5, about 11.5, about 12.5, about 13.5, about 14.5, or about 15.5 binding repeat units, a CTD of about 20 amino acids to about 85 amino acids, and an engineered I-OnuI LHE selected from the group consisting of: I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-CreI, I-SceI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-Ncrl, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, and I-Vdi141I, or preferably I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, and SmaMI, or more preferably I-OnuI.

(154) 3. Talens

(155) In various embodiments, a plurality of transcription activator-like effector nucleases (TALENs) are introduced into a cell and engineered to bind to, and to introduce single-stranded nicks or double-strand breaks (DSBs) at a target site in a human γ-globin gene, e.g., a polynucleotide sequence as set forth in any one of SEQ ID NOs: 1-9. A “TALEN” refers to an engineered nuclease comprising an engineered TALE DNA binding domain contemplated elsewhere herein and an endonuclease domain (or endonuclease half-domain thereof), and optionally comprise one or more linkers and/or additional functional domains, e.g., an end-processing enzymatic domain of an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. In particular embodiments, a TALEN can be introduced into a T cell with an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. The TALEN and 3′ processing enzyme may be introduced separately, e.g., in different vectors or separate mRNAs, or together, e.g., as a fusion protein, or in a polycistronic construct separated by a viral self-cleaving peptide or an IRES element.

(156) In one embodiment, targeted double-stranded cleavage is achieved with two TALENs, each comprising an endonuclease half-domain can be used to reconstitute a catalytically active cleavage domain. In another embodiment, targeted double-stranded cleavage is achieved using a single polypeptide comprising a TALE DNA binding domain and two endonuclease half-domains.

(157) TALENs contemplated in particular embodiments comprise an NTD, a TALE DNA binding domain comprising about 3 to 30 repeat units, e.g., about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 repeat units, and an endonuclease domain or half-domain.

(158) TALENs contemplated in particular embodiments comprise an NTD, a TALE DNA binding domain comprising about 3.5 to 30.5 repeat units, e.g., about 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5, 10.5,

11.5, 12.5, 13.5, 14.5, 15.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, or 30.5 repeat units, a CTD, and an endonuclease domain or half-domain.
(159) TALENs contemplated in particular embodiments comprise an NTD of about 121 amino acids to about 137 amino acids as disclosed elsewhere herein, a TALE DNA binding domain comprising about 9.5 to about 15.5 repeat units (i.e., about 9.5, about 10.5, about 11.5, about 12.5, about 13.5, about 14.5, or about 15.5 repeat units), a CTD of about 20 amino acids to about 85 amino acids, and an endonuclease domain or half domain.
(160) In particular embodiments, a TALEN comprises an endonuclease domain of a type restriction endonuclease. Restriction endonucleases (restriction enzymes) are present in many species and are capable of sequence-specific binding to DNA (at a recognition site), and cleaving DNA at or near the site of binding. Certain restriction enzymes (e.g., Type-IIS) cleave DNA at sites removed from the recognition site and have separable binding and endonuclease domains. In one embodiment, TALENs comprise the endonuclease domain (or endonuclease half-domain) from at least one Type-IIS restriction enzyme and one or more TALE DNA-binding domains contemplated elsewhere herein.
(161) Illustrative examples of Type-IIS restriction endonuclease domains suitable for use in TALENs contemplated in particular embodiments include endonuclease domains of the at least 1633 Type-IIS restriction endonucleases disclosed at “rebase.neb.com/cgi-bin/sublist?S.”
(162) Additional illustrative examples of Type-IIS restriction endonuclease domains suitable for use in TALENs contemplated in particular embodiments include those of endonucleases selected from the group consisting of Aar I, Ace III, Aci I, Alo I, Alw26 I, Bae I, Bbr7 I, Bbv I, Bbv II, BbvC I, Bcc I, Bce83 I, BceA I, Bcef I, Bcg I, BciV I, Bfi I, Bin I, Bmg I, Bpu10 I, BsaX I, Bsb I, BscA I, BscG I, BseR I, BseY I, Bsi I, Bsm I, BsmA I, BsmF I, Bsp24 I, BspG I, BspM I, BspNC I, Bsr I, BsrB I, BsrD I, BstF5 I, Btr I, Bts I, Cdi I, CjeP I, Drd II, EarI, Eci I, Eco31 I, Eco57 I, Eco57M I, Esp3 I, Fau I, Fin I, Fok I, Gdi II, Gsu I, Hga I, Hin4 II, Hph I, Ksp632 I, Mbo II, Mly I, Mme I, Mnl I, Pfl1108, I Ple I, Ppi I Psr I, RleA I, Sap I, SfaN I, Sim I, SspD5 I, Sth132 I, Sts I, TspDT I, TspGW I, Tth111 II, UbaP I, Bsa I, and BsmB I.
(163) In one embodiment, a TALEN contemplated herein comprises an endonuclease domain of the Fok I Type-IIS restriction endonuclease.
(164) In one embodiment, a TALEN contemplated herein comprises a TALE DNA binding domain and an endonuclease half-domain from at least one Type-IIS restriction endonuclease to enhance cleavage specificity, optionally wherein the endonuclease half-domain comprises one or more amino acid substitutions or modifications that minimize or prevent homodimerization.
(165) Illustrative examples of cleavage half-domains suitable for use in particular embodiments contemplated in particular embodiments include those disclosed in U.S. Patent Publication Nos. 20050064474; 20060188987, 20080131962, 20090311787; 20090305346; 20110014616, and 20110201055, each of which are incorporated by reference herein in its entirety.
(166) 4. Zinc Finger Nucleases
(167) In various embodiments, a plurality of zinc finger nucleases (ZFNs) are introduced into a cell and engineered to bind to, and to introduce single-stranded nicks or double-strand breaks (DSBs) at a target site in a human γ-globin gene, e.g., a polynucleotide sequence as set for the in any one of SEQ ID NOs: 1-9. A “ZFN” refers to an engineered nuclease comprising one or more zinc finger DNA binding domains and an endonuclease domain (or endonuclease half-domain thereof), and optionally comprise one or more linkers and/or additional functional domains, e.g., an end-processing enzymatic domain of an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. In particular embodiments, a ZFN can be introduced into a T cell with an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. The ZFN and 3′ processing enzyme may be introduced separately, e.g., in

different vectors or separate mRNAs, or together, e.g., as a fusion protein, or in a polycistronic construct separated by a viral self-cleaving peptide or an IRES element.
(168) In one embodiment, targeted double-stranded cleavage is achieved using two ZFNs, each comprising an endonuclease half-domain can be used to reconstitute a catalytically active cleavage domain. In another embodiment, targeted double-stranded cleavage is achieved with a single polypeptide comprising one or more zinc finger DNA binding domains and two endonuclease half-domains.
(169) In one embodiment, a ZNF comprises a TALE DNA binding domain contemplated elsewhere herein, a zinc finger DNA binding domain, and an endonuclease domain (or endonuclease half-domain) contemplated elsewhere herein.
(170) In one embodiment, a ZNF comprises a zinc finger DNA binding domain, and a meganuclease contemplated elsewhere herein.
(171) In particular embodiments, the ZFN comprises a zinc finger DNA binding domain that has one, two, three, four, five, six, seven, or eight or more zinc finger motifs and an endonuclease domain (or endonuclease half-domain). Typically, a single zinc finger motif is about 30 amino acids in length. Zinc fingers motifs include both canonical C.sub.2H.sub.2 zinc fingers, and non-canonical zinc fingers such as, for example, C.sub.3H zinc fingers and C.sub.4 zinc fingers.
(172) Zinc finger binding domains can be engineered to bind any DNA sequence. Candidate zinc finger DNA binding domains for a given 3 bp DNA target sequence have been identified and modular assembly strategies have been devised for linking a plurality of the domains into a multi-finger peptide targeted to the corresponding composite DNA target sequence. Other suitable methods known in the art can also be used to design and construct nucleic acids encoding zinc finger DNA binding domains, e.g., phage display, random mutagenesis, combinatorial libraries, computer/rational design, affinity selection, PCR, cloning from cDNA or genomic libraries, synthetic construction and the like. (See, e.g., U.S. Pat. No. 5,786,538; Wu et al., *PNAS* 92:344-348 (1995); Jamieson et al., *Biochemistry* 33:5689-5695 (1994); Rebar & Pabo, *Science* 263:671-673 (1994); Choo & Klug, *PNAS* 91:11163-11167 (1994); Choo & Klug, *PNAS* 91: 11168-11172 (1994); Desjarlais & Berg, *PNAS* 90:2256-2260 (1993); Desjarlais & Berg, *PNAS* 89:7345-7349 (1992); Pomerantz et al., *Science* 267:93-96 (1995); Pomerantz et al., *PNAS* 92:9752-9756 (1995); Liu et al., *PNAS* 94:5525-5530 (1997); Griesman & Pabo, *Science* 275:657-661 (1997); Desjarlais & Berg, *PNAS* 91:11-99-11103 (1994)).
(173) Individual zinc finger motifs bind to a three or four nucleotide sequence. The length of a sequence to which a zinc finger binding domain is engineered to bind (e.g., a target sequence) will determine the number of zinc finger motifs in an engineered zinc finger binding domain. For example, for ZFNs in which the zinc finger motifs do not bind to overlapping subsites, a six-nucleotide target sequence is bound by a two-finger binding domain; a nine-nucleotide target sequence is bound by a three-finger binding domain, etc. In particular embodiments, DNA binding sites for individual zinc fingers motifs in a target site need not be contiguous, but can be separated by one or several nucleotides, depending on the length and nature of the linker sequences between the zinc finger motifs in a multi-finger binding domain.
(174) In particular embodiments, ZNFs contemplated herein comprise, a zinc finger DNA binding domain comprising two, three, four, five, six, seven or eight or more zinc finger motifs, and an endonuclease domain or half-domain from at least one Type-IIS restriction enzyme and one or more TALE DNA-binding domains contemplated elsewhere herein.
(175) In particular embodiments, ZNFs contemplated herein comprise, a zinc finger DNA binding domain comprising three, four, five, six, seven or eight or more zinc finger motifs, and an endonuclease domain or half-domain from at least one Type-IIS restriction enzyme selected from the group consisting of: Aar I, Ace III, Aci I, Alo I, Alw26 I, Bae I, Bbr7 I, Bbv I, Bbv II, BbvC I, Bcc I, Bce83 I, BceA I, Bcef I, Bcg I, BciV I, Bfi I, Bin I, Bmg I, Bpu10 I, BsaX I, Bsb I, BscA I, BscG I, BseR I, BseY I, Bsi I, Bsm I, BsmA I, BsmF I, Bsp24 I, BspG I, BspM I, BspNC I, Bsr I,

BsrB I, BsrD I, BstF5 I, Btr I, Bts I, Cdi I, CjeP I, Drd II, EarI, Eci I, Eco31 I, Eco57 I, Eco57M I, Esp3 I, Fau I, Fin I, Fok I, Gdi II, Gsu I, Hga I, Hin4 II, Hph I, Ksp632 I, Mbo II, Mly I, Mme I, Mnl I, Pfl1108, I Ple I, Ppi I Psr I, RleA I, Sap I, SfaN I, Sim I, SspD5 I, Sth132 I, Sts I, TspDT I, TspGW I, Tth111 II, UbaP I, Bsa I, and BsmB I.

(176) In particular embodiments, ZNFs contemplated herein comprise, a zinc finger DNA binding domain comprising three, four, five, six, seven or eight or more zinc finger motifs, and an endonuclease domain or half-domain from the Fok I Type-IIS restriction endonuclease.

(177) In one embodiment, a ZFN contemplated herein comprises a zinc finger DNA binding domain and an endonuclease half-domain from at least one Type-IIS restriction endonuclease to enhance cleavage specificity, optionally wherein the endonuclease half-domain comprises one or more amino acid substitutions or modifications that minimize or prevent homodimerization.

(178) 5. CRISPR/Cas Nuclease System

(179) In various embodiments, a CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas (CRISPR Associated) nuclease system is introduced into a cell and engineered to bind to, and to introduce single-stranded nicks or double-strand breaks (DSBs) at a target site in a human γ-globin gene, e.g., a polynucleotide sequence as set forth in any one of SEQ ID NOs: 1-9. The CRISPR/Cas nuclease system is an engineered nuclease system based on a bacterial system that can be used for mammalian genome engineering. See, e.g., Jinek et al. (2012) *Science* 337:816-821; Cong et al. (2013) *Science* 339:819-823; Mali et al. (2013) *Science* 339:823-826; Qi et al. (2013) *Cell* 152:1173-1183; Jinek et al. (2013), *eLife* 2:e00471; David Segal (2013) *eLife* 2:e00563; Ran et al. (2013) *Nature Protocols* 8(11):2281-2308; Zetsche et al. (2015) *Cell* 163(3):759-771, each of which is incorporated herein by reference in its entirety.

(180) In one embodiment, the CRISPR/Cas nuclease system comprises Cas nuclease and one or more RNAs that recruit the Cas nuclease to the target site, e.g., a transactivating cRNA (tracrRNA) and a CRISPR RNA (crRNA), or a single guide RNA (sgRNA). crRNA and tracrRNA can engineered into one polynucleotide sequence referred to herein as a "single guide RNA" or "sgRNA."

(181) In one embodiment, the Cas nuclease is engineered as a double-stranded DNA endonuclease or a nickase or catalytically dead Cas, and forms a target complex with a crRNA and a tracrRNA, or sgRNA, for site specific DNA recognition and site-specific cleavage of the protospacer target sequence located within one or more target sites in a human γ-globin gene, e.g., a polynucleotide sequence as set forth in any one of SEQ ID NOs: 1-9. The protospacer motif abuts a short protospacer adjacent motif (PAM), which plays a role in recruiting a Cas/RNA complex. Cas polypeptides recognize PAM motifs specific to the Cas polypeptide. Accordingly, the CRISPR/Cas system can be used to target and cleave either or both strands of a double-stranded polynucleotide sequence flanked by particular 3′ PAM sequences specific to a particular Cas polypeptide. PAMs may be identified using bioinformatics or using experimental approaches. Esvelt et al., 2013, *Nature Methods.* 10(11):1116-1121, which is hereby incorporated by reference in its entirety.

(182) In one embodiment, the Cas nuclease comprises one or more heterologous DNA binding domains, e.g., a TALE DNA binding domain or zinc finger DNA binding domain. Fusion of the Cas nuclease to TALE or zinc finger DNA binding domains increases the DNA cleavage efficiency and specificity. In a particular embodiment, a Cas nuclease optionally comprises one or more linkers and/or additional functional domains, e.g., an end-processing enzymatic domain of an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. In particular embodiments, a Cas nuclease can be introduced into a T cell with an end-processing enzyme that exhibits 5-3′ exonuclease, 5-3′ alkaline exonuclease, 3-5′exonuclease (e.g., Trex2), 5′ flap endonuclease, helicase or template-independent DNA polymerases activity. The Cas nuclease and 3′ processing enzyme may be introduced separately, e.g., in different vectors or separate mRNAs, or together, e.g., as a fusion protein, or in a polycistronic construct separated by a viral

self-cleaving peptide or an IRES element.
(183) In various embodiments, the Cas nuclease is Cas9 or Cpf1.
(184) Illustrative examples of Cas9 polypeptides suitable for use in particular embodiments contemplated in particular embodiments may be obtained from bacterial species including, but not limited to: *Enterococcus faecium, Enterococcus italicus, Listeria innocua, Listeria monocytogenes, Listeria seeligeri, Listeria ivanovii, Streptococcus agalactiae, Streptococcus anginosus, Streptococcus bovis, Streptococcus dysgalactiae, Streptococcus equinus, Streptococcus gallolyticus, Streptococcus macacae, Streptococcus mutans, Streptococcus pseudoporcinus, Streptococcus pyogenes, Streptococcus thermophilus, Streptococcus gordonii, Streptococcus infantarius, Streptococcus macedonicus, Streptococcus mitis, Streptococcus pasteurianus, Streptococcus suis, Streptococcus vestibularis, Streptococcus sanguinis, Streptococcus downei, Neisseria bacilliformis, Neisseria cinerea, Neisseria flavescens, Neisseria lactamica, Neisseria meningitidis, Neisseria subflava, Lactobacillus brevis, Lactobacillus buchneri, Lactobacillus casei, Lactobacillus paracasei, Lactobacillus fermentum, Lactobacillus gasseri, Lactobacillus jensenii, Lactobacillus johnsonii, Lactobacillus rhamnosus, Lactobacillus ruminis, Lactobacillus salivarius, Lactobacillus sanfranciscensis, Corynebacterium accolens, Corynebacterium diphtheriae, Corynebacterium matruchotii, Campylobacter jejuni, Clostridium perfringens, Treponema vincentii, Treponema phagedenis,* and *Treponema denticola.*
(185) Illustrative examples of Cpf1 polypeptides suitable for use in particular embodiments contemplated in particular embodiments may be obtained from bacterial species including, but not limited to: *Francisella* spp., *Acidaminococcus* spp., *Prevotella* spp., *Lachnospiraceae* spp., among others.
(186) Conserved regions of Cas9 orthologs include a central HNH endonuclease domain and a split RuvC/RNase H domain. Cpf1 orthologs possess a RuvC/RNase H domain but no discernable HNH domain. The HNH and RuvC-like domains are each responsible for cleaving one strand of the double-stranded DNA target sequence. The HNH domain of the Cas9 nuclease polypeptide cleaves the DNA strand complementary to the tracrRNA:crRNA or sgRNA. The RuvC-like domain of the Cas9 nuclease cleaves the DNA strand that is not-complementary to the tracrRNA:crRNA or sgRNA. Cpf1 is predicted to act as a dimer wherein each RuvC-like domain of Cpf1 cleaves either the complementary or non-complementary strand of the target site. In particular embodiments, a Cas9 nuclease variant (e.g., Cas9 nickase) is contemplated comprising one or more amino acids additions, deletions, mutations, or substitutions in the HNH or RuvC-like endonuclease domains that decreases or eliminates the nuclease activity of the variant domain.
(187) Illustrative examples of Cas9 HNH mutations that decrease or eliminate the nuclease activity in the domain include, but are not limited to: *S. pyogenes* (D10A); *S. thermophilis* (D9A); *T. denticola* (D13A); and *N. meningitidis* (D16A).
(188) Illustrative examples of Cas9 RuvC-like domain mutations that decrease or eliminate the nuclease activity in the domain include, but are not limited to: *S. pyogenes* (D839A, H840A, or N863A); *S. thermophilis* (D598A, H599A, or N622A); *T. denticola* (D878A, H879A, or N902A); and *N. meningitidis* (D587A, H588A, or N611A).
(189) 6. End-Processing Enzymes
(190) Genome editing compositions and methods contemplated in particular embodiments comprise editing cellular genomes using a nuclease variant and an end-processing enzyme. In particular embodiments, a single polynucleotide encodes a homing endonuclease variant and an end-processing enzyme, separated by a linker, a self-cleaving peptide sequence, e.g., 2A sequence, or by an IRES sequence. In particular embodiments, genome editing compositions comprise a polynucleotide encoding a nuclease variant and a separate polynucleotide encoding an end-processing enzyme.
(191) The term “end-processing enzyme” refers to an enzyme that modifies the exposed ends of a polynucleotide chain. The polynucleotide may be double-stranded DNA (dsDNA), single-stranded

DNA (ssDNA), RNA, double-stranded hybrids of DNA and RNA, and synthetic DNA (for example, containing bases other than A, C, G, and T). An end-processing enzyme may modify exposed polynucleotide chain ends by adding one or more nucleotides, removing one or more nucleotides, removing or modifying a phosphate group and/or removing or modifying a hydroxyl group. An end-processing enzyme may modify ends at endonuclease cut sites or at ends generated by other chemical or mechanical means, such as shearing (for example by passing through fine-gauge needle, heating, sonicating, mini bead tumbling, and nebulizing), ionizing radiation, ultraviolet radiation, oxygen radicals, chemical hydrolysis and chemotherapy agents.

(192) In particular embodiments, genome editing compositions and methods contemplated in particular embodiments comprise editing cellular genomes using a homing endonuclease variant or megaTAL and a DNA end-processing enzyme.

(193) The term “DNA end-processing enzyme” refers to an enzyme that modifies the exposed ends of DNA. A DNA end-processing enzyme may modify blunt ends or staggered ends (ends with 5′ or 3′ overhangs). A DNA end-processing enzyme may modify single stranded or double stranded DNA. A DNA end-processing enzyme may modify ends at endonuclease cut sites or at ends generated by other chemical or mechanical means, such as shearing (for example by passing through fine-gauge needle, heating, sonicating, mini bead tumbling, and nebulizing), ionizing radiation, ultraviolet radiation, oxygen radicals, chemical hydrolysis and chemotherapy agents. DNA end-processing enzyme may modify exposed DNA ends by adding one or more nucleotides, removing one or more nucleotides, removing or modifying a phosphate group and/or removing or modifying a hydroxyl group.

(194) Illustrative examples of DNA end-processing enzymes suitable for use in particular embodiments contemplated herein include, but are not limited to: 5′-3′ exonucleases, 5′-3′ alkaline exonucleases, 3′-5′ exonucleases, 5′ flap endonucleases, helicases, phosphatases, hydrolases and template-independent DNA polymerases.

(195) Additional illustrative examples of DNA end-processing enzymes suitable for use in particular embodiments contemplated herein include, but are not limited to, Trex2, Trex1, Trex1 without transmembrane domain, Apollo, Artemis, DNA2, Exo1, ExoT, ExoIII, Fen1, Fan1, MreII, Rad2, Rad9, TdT (terminal deoxynucleotidyl transferase), PNKP, RecE, RecJ, RecQ, Lambda exonuclease, Sox, Vaccinia DNA polymerase, exonuclease I, exonuclease III, exonuclease VII, NDK1, NDK5, NDK7, NDK8, WRN, T7-exonuclease Gene 6, avian myeloblastosis virus integration protein (IN), Bloom, Antarctic Phosphatase, Alkaline Phosphatase, Poly nucleotide Kinase (PNK), ApeI, Mung Bean nuclease, Hex1, TTRAP (TDP2), Sgs1, Sae2, CUP, Pol mu, Pol lambda, MUS81, EME1, EME2, SLX1, SLX4 and UL-12.

(196) In particular embodiments, genome editing compositions and methods for editing cellular genomes contemplated herein comprise polypeptides comprising a homing endonuclease variant or megaTAL and an exonuclease. The term “exonuclease” refers to enzymes that cleave phosphodiester bonds at the end of a polynucleotide chain via a hydrolyzing reaction that breaks phosphodiester bonds at either the 3′ or 5′ end.

(197) Illustrative examples of exonucleases suitable for use in particular embodiments contemplated herein include, but are not limited to: hExoI, Yeast ExoI, *E. coli* ExoI, hTREX2, mouse TREX2, rat TREX2, hTREX1, mouse TREX1, rat TREX1, and Rat TREX1.

(198) In particular embodiments, the DNA end-processing enzyme is a 3′ or 5′ exonuclease, preferably Trex 1 or Trex2, more preferably Trex2, and even more preferably human or mouse Trex2.

D. Target Sites

(199) The engineered nucleases contemplated herein are designed to bind to any suitable target sequence and can have a novel binding specificity, compared to a naturally-occurring nuclease. In particular embodiments, the target site is a regulatory region of a γ-globin gene including, but not limited to transcription factor binding sites. In particular embodiments, the nuclease target site is in

a polynucleotide sequence that when deleted or disrupted is associated with HPFH, e.g, at position −120 to −102 relative to the transcriptional start site of a γ-globin gene; preferably the nuclease target site is at, or disrupts or deletes, a 13 bp polynucleotide sequence in a γ-globin gene associated with HPFH (SEQ ID NO: 3); more preferably the nuclease target site is at, or disrupts or deletes, a CCAAT polynucleotide sequence at position −115 to −111 (relative to the transcriptional start site) of the promoter of the HBG1 gene; and even more preferably the nuclease target site is at, or disrupts or deletes, a CAAT polynucleotide sequence at position −114 to −111 (relative to the transcriptional start site) of the HBG1 gene.

(200) In particular embodiments, the nuclease target site is in, or near, the region of Chr11: 5249957-5249977.

(201) In particular embodiments, the nuclease target site is in, or near, the polynucleotide sequence of any one of SEQ ID NOs: 1-6 in the region of Chr11: 5249957-5249977.

(202) In particular embodiments, the nuclease target site is in, or near, the region of Chr11: 5249959-5249971 (in the HBG1 gene).

(203) In particular embodiments, the engineered nucleases introduce a DSB in, near, or flanking the 5′ and/or 3′ sequences of a 13 bp polynucleotide sequence in a γ-globin gene associated with HPFH (SEQ ID NO: 3), so that in the event that a DNA donor repair template is not inserted into the target site, the target site will still be disrupted by NHEJ in the absence of a DNA donor repair template, and thus, either HDR or NHEJ events at the target site will lead to derepression of γ-globin gene expression and/or therapeutic globin expression (either endogenous γ-globin expression or heterologous expression of another therapeutic globin, including but not limited to, γ-globin, β-globin, or an anti-sickling form of β-globin.

(204) In a preferred embodiment, an engineered nuclease cleaves double-stranded DNA and introduces a DSB into the polynucleotide sequence set forth in any one of SEQ ID NOs: 1-9, more preferably, in the human HBG1 and/or HBG2 gene.

E. Donor Repair Templates

(205) In preferred embodiments, one or more engineered nucleases are used to introduce a DSB in a target sequence in the presence of a donor repair template. The donor repair template may comprise single-stranded or double-stranded DNA or RNA. In the presence of a donor repair template the DSB may be repaired through homology directed repair (HDR) mechanisms.

(206) In particular embodiments, a donor repair template comprises a pair of homology arms, a selection cassette and an erythroid expression control sequence. In preferred embodiments, the donor repair template is inserted into a γ-globin gene and the erythroid expression control sequence is positioned such that it is operably linked to the endogenous polynucleotide sequence encoding a γ-globin.

(207) In particular embodiments, a donor repair template comprises a pair of homology arms, a polynucleotide encoding a therapeutic globin and one or more post-transcriptional control elements; and a selection cassette. In preferred embodiments, the donor repair template is inserted into a γ-globin gene at a site that derepresses the γ-globin promoter, thereby operably linking a derepressed γ-globin promoter to the polynucleotide encoding a therapeutic globin. In other preferred embodiments, the 5′ homology arm encodes a deletion in a site that derepresses the γ-globin promoter thereby operably linking a derepressed γ-globin promoter to the polynucleotide encoding a therapeutic globin.

(208) A “pair of homology arms” refers to a group of two homology arms. In particular embodiments a pair of homology arms comprises a 5′ homology arm and a 3′ homology arm. A “5′ homology arm” refers to a polynucleotide sequence that is identical, or nearly identical, or homologous to a DNA sequence 5′ of a target site (e.g., double strand break site). A “3′ homology arm” refers to a polynucleotide sequence that is identical, or nearly identical, or homologous to a DNA sequence 3′ of the target site. In particular embodiments, a pair of homology arms comprises a homology arm comprising a polynucleotide sequence that includes a target site for a double

strand break with a mutation in the target site to minimize recleavage of the target site. In particular preferred embodiments, the 5′ homology arm comprises a polynucleotide sequence within 1 kb upstream of a γ-globin transcriptional start site (SEQ ID NO: 9). In more particular preferred embodiments, the 5′ homology arm comprises a polynucleotide sequence within 1 kb upstream of a γ-globin transcriptional start site and further comprises a deletion of, or lacks the polynucleotide sequence set forth in any one of SEQ ID NOs: 1-6.

(209) In some embodiments, where the donor repair template is designed to derepress a γ-globin promoter and operably link the derepressed promoter to a polynucleotide encoding a therapeutic globin, the 5′ homology arm comprises a polynucleotide sequence within 1 kb upstream of a γ-globin transcriptional start site and further comprises a deletion of, or lacks the polynucleotide sequence set forth in any one of SEQ ID NOs: 1-6 and the 3′ homology arm comprises a sequence downstream of the 5′ homology arm. In some embodiments, the donor repair template disrupts endogenous γ-globin gene expression, optionally through deletion of genomic sequence encoding γ-globin.

(210) In some embodiments, where the donor repair template is designed to operably link a β-globin LCR responsive expression control sequence to an endogenous genomic sequence encoding a γ-globin, the 5′ homology arm comprises a polynucleotide sequence within 1 kb upstream of a γ-globin transcriptional start site and may further comprise a deletion of, or lack the polynucleotide sequence set forth in any one of SEQ ID NOs: 1-6 and the 3′ homology arm comprises a sequence downstream of the 5′ homology arm, but upstream of a γ-globin gene transcription start site.

(211) In particular embodiments, either one of, or both, homology arms in a pair of homology arms is independently located about 100 bp, about 200 bp, about 300 bp, about 400 bp, about 500 bp, about 600 bp, about 700 bp, about 800 bp, about 900 bp, about 1000 bp, about 1100 bp, about 1200 bp, about 1300 bp, about 1400 bp, about 1500 bp, about 1600 bp, about 1700 bp, about 1800 bp, about 1900 bp, about 2000 bp, about 2100 bp, about 2200 bp, about 2300 bp, about 2400 bp, about 2500 bp, about 2600 bp, about 2700 bp, about 2800 bp, about 2900 bp, or about 3000 bp, from the target site, including all intervening distances from the target site.

(212) Illustrative examples of suitable lengths of homology arms contemplated in particular embodiments, may be independently selected, and include but are not limited to: about 100 bp, about 200 bp, about 300 bp, about 400 bp, about 500 bp, about 600 bp, about 700 bp, about 800 bp, about 900 bp, about 1000 bp, about 1100 bp, about 1200 bp, about 1300 bp, about 1400 bp, about 1500 bp, about 1600 bp, about 1700 bp, about 1800 bp, about 1900 bp, about 2000 bp, about 2100 bp, about 2200 bp, about 2300 bp, about 2400 bp, about 2500 bp, about 2600 bp, about 2700 bp, about 2800 bp, about 2900 bp, or about 3000 bp, or longer homology arms, including all intervening lengths of homology arms.

(213) Additional illustrative examples of suitable homology arm lengths include, but are not limited to: about 100 bp to about 600 bp, about 100 bp to about 500 bp, about 100 bp to about 400 bp, about 100 bp to about 300 bp, about 100 bp to about 200 bp, about 200 bp to about 600 bp, about 200 bp to about 500 bp, about 200 bp to about 400 bp, about 200 bp to about 300 bp, about 300 bp to about 600 bp, about 300 bp to about 500 bp, about 100 bp to about 3000 bp, about 200 bp to about 3000 bp, about 300 bp to about 3000 bp, about 400 bp to about 3000 bp, about 500 bp to about 3000 bp, about 500 bp to about 2500 bp, about 500 bp to about 2000 bp, about 750 bp to about 2000 bp, about 750 bp to about 1500 bp, or about 1000 bp to about 1500 bp, including all intervening lengths of homology arms.

(214) In a particular embodiment, the lengths of any 5′ and 3′ homology arms present in a DNA donor repair template are independently selected from about 100 bp, about 200 bp, about 300 bp, about 400 bp, about 500 bp, or about 600 bp. In one embodiment, a 5′homology arm is about 300 bp and a 3′ homology arm is about 300 bp.

(215) Donor repair templates contemplated herein comprise a selection cassette. As used herein, the term “selection cassette” refers to an expression cassette that comprises one or more expression

control sequences operably linked to a polynucleotide sequence encoding a selectable marker and one or more post-transcriptional elements or a ribosomal skipping polypeptide.
(216) Illustrative examples of expression control sequences suitable for use in selection cassettes contemplated herein include but are not limited to: a constitutive promoter, a conditional promoter, or hematopoietic stem cell promoter. As used herein, the term “constitutive expression control sequence” refers to a promoter, enhancer, or promoter/enhancer that continually or continuously allows for transcription of an operably linked sequence. A constitutive expression control sequence may be a “ubiquitous” promoter, enhancer, or promoter/enhancer that allows expression in a wide variety of cell and tissue types or a “cell specific,” “cell type specific,” “cell lineage specific,” or “tissue specific” promoter, enhancer, or promoter/enhancer that allows expression in a restricted variety of cell and tissue types, respectively.
(217) Illustrative ubiquitous expression control sequences suitable for use in particular embodiments include, but are not limited to, a cytomegalovirus (CMV) immediate early promoter, a viral simian virus 40 (SV40) promoter, a Moloney murine leukemia virus (MoMLV) LTR promoter, a Rous sarcoma virus (RSV) LTR promoter, a herpes simplex virus (HSV) (thymidine kinase) promoter, a H5, P7.5, or P11 vaccinia virus promoter, a short elongation factor 1-alpha (EF1a-short) promoter, a long elongation factor 1-alpha (EF1a-long) promoter, an early growth response 1 (EGR1) promoter, a ferritin H (FerH) promoter, a ferritin L (FerL) promoter, a Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) promoter, a eukaryotic translation initiation factor 4A1 (EIF4A1) promoter, a heat shock 70 kDa protein 5 (HSPA5) promoter, a heat shock protein 90 kDa beta, member 1 (HSP90BT) promoter, a heat shock protein 70 kDa (HSP70) promoter, a β-kinesin (β-KIN) promoter, a human ROSA 26 promoter, a Ubiquitin C promoter (UBC) promoter, a phosphoglycerate kinase-1 (PGK) promoter, a cytomegalovirus enhancer/chicken β-actin (CAG) promoter, a β-actin promoter and a myeloproliferative sarcoma virus enhancer, negative control region deleted, dl587rev primer-binding site substituted (MND) promoter (Challita et al., *J Virol.* 69(2):748-55 (1995)).
(218) In a particular embodiment, it may be desirable to use a cell, cell type, cell lineage or tissue specific expression control sequence to achieve cell type specific, lineage specific, or tissue specific expression of a desired polynucleotide sequence (e.g., to express a particular nucleic acid encoding a polypeptide in only a subset of cell types, cell lineages, or tissues or during specific stages of development).
(219) Illustrative examples of tissue specific promoters include, but are not limited to: an B29 promoter (B cell expression), a runt transcription factor (CBFa2) promoter (stem cell specific expression), an CD14 promoter (monocytic cell expression), an CD43 promoter (leukocyte and platelet expression), an CD45 promoter (hematopoietic cell expression), an CD68 promoter (macrophage expression), a CYP450 3A4 promoter (hepatocyte expression), an desmin promoter (muscle expression), an elastase 1 promoter (pancreatic acinar cell expression, an endoglin promoter (endothelial cell expression), a fibroblast specific protein 1 promoter (FSP1) promoter (fibroblast cell expression), a fibronectin promoter (fibroblast cell expression), a fins-related tyrosine kinase 1 (FLT1) promoter (endothelial cell expression), a glial fibrillary acidic protein (GFAP) promoter (astrocyte expression), an insulin promoter (pancreatic beta cell expression), an integrin, alpha 2b (ITGA2B) promoter (megakaryocytes), an intracellular adhesion molecule 2 (ICAM-2) promoter (endothelial cells), an interferon beta (IFN-β) promoter (hematopoietic cells), a keratin 5 promoter (keratinocyte expression), a myoglobin (MB) promoter (muscle expression), a myogenic differentiation 1 (MYOD1) promoter (muscle expression), a nephrin promoter (podocyte expression), a bone gamma-carboxyglutamate protein 2 (OG-2) promoter (osteoblast expression), an 3-oxoacid CoA transferase 2B (Oxct2B) promoter, (haploid-spermatid expression), a surfactant protein B (SP-B) promoter (lung expression), a synapsin promoter (neuron expression), an interferon beta (IFN-β) promoter (hematopoietic cell expression), an α-spectrin promoter (erythroid cell expression), a β-spectrin promoter (erythroid cell expression), a β-globin LCR (erythroid cell

expression), a γ-globin promoter (erythroid cell expression), a β-globin promoter (erythroid cell expression), an α-globin HS40 enhancer (erythroid cell expression), an ankyrin-1 promoter (erythroid cell expression), and a Wiskott-Aldrich syndrome protein (WASP) promoter (hematopoietic cell expression).

(220) As used herein, “conditional expression” may refer to any type of conditional expression including, but not limited to, inducible expression; repressible expression; expression in cells or tissues having a particular physiological, biological, or disease state, etc. This definition is not intended to exclude cell type or tissue specific expression. Certain embodiments provide conditional expression of a polynucleotide-of-interest, e.g., expression is controlled by subjecting a cell, tissue, organism, etc., to a treatment or condition that causes the polynucleotide to be expressed or that causes an increase or decrease in expression of the polynucleotide encoded by the polynucleotide-of-interest.

(221) Illustrative examples of inducible promoters/systems include, but are not limited to, steroid-inducible promoters such as promoters for genes encoding glucocorticoid or estrogen receptors (inducible by treatment with the corresponding hormone), metallothionine promoter (inducible by treatment with various heavy metals), MX-1 promoter (inducible by interferon), the “GeneSwitch” mifepristone-regulatable system (Sirin et al., 2003, *Gene,* 323:67), the cumate inducible gene switch (WO 2002/088346), tetracycline-dependent regulatory systems, etc.

(222) In particular embodiments, a selection cassette comprises an EF1α promoter or an MND promoter operably linked to a polynucleotide encoding a selectable marker.

(223) Illustrative examples of selectable markers suitable for use in selection cassettes contemplated herein include, but are not limited to: hygromycin-B phosphotransferase (HPH) which may be positively selected for with hygromycin B; amino 3′-glycosyl phosphotransferase (NEO), which may be positively selected for with G418; dihydrofolate reductase (DHFR), which may be positively selected for with methotrexate; adenosine deaminase (ADA), which may be positively selected for with 2′-deoxycoformycin; multi-drug resistance protein (MDR), which may be positively selected for by anti-cancer drugs including, but not limited to vinca alkaloids, taxanes, anthracyclines, epipodophyllotoxins, colchicine, doxorubicin, and actinomycin D; O.sup.6-methylguanine-DNA-methyltransferase (MGMT), which may be selected for by O.sup.6-benzylguanine/1,3-bis(2-chloroethyl)-1-nitrosourea (BG/BCNU); Sh ble (BLE), which may be positively selected for with bleocin or zeocin; and blasticidin-S deaminase (BSR), which may be positively selected for with blastocidin.

(224) In particular embodiments, a selection cassette comprises a ubiquitous promoter operably linked to a polynucleotide encoding MGMT.

(225) Illustrative examples of post-transcriptional control sequences for use in selection cassettes include, but are not limited to: woodchuck hepatitis virus post-transcriptional response element (WPRE) or variant thereof, a hepatitis B virus post-transcriptional response element (HPRE) or variant thereof, and a polyadenylation sequence.

(226) As used herein, the terms “post-transcriptional control sequences,” “posttranscriptional regulatory element” or “PRE” refer to a cis-acting element that regulates expression at the mRNA level by, for example, regulating capping, splicing, poly(A) tail addition, and mRNA stability. Illustrative examples of PTE include, but are not limited to, woodchuck hepatitis virus posttranscriptional regulatory element (WPRE; Zufferey et al., 1999, *J. Virol.,* 73:2886); the posttranscriptional regulatory element present in hepatitis B virus (HPRE) (Huang and Yen, 1995, *Mol. Cell. Biol.,* 5:3864); and the like (Liu et al., 1995, *Genes Dev.,* 9:1766).

(227) The term “poly(A) site” or “poly(A) sequence” as used herein denotes a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript by RNA polymerase II. Polyadenylation sequences can promote mRNA stability by addition of a poly(A) tail to the 3′ end of the coding sequence and thus, contribute to increased translational efficiency. Cleavage and polyadenylation is directed by a poly(A) sequence in the RNA. The core poly(A)

sequence for mammalian pre-mRNAs has two recognition elements flanking a cleavage-polyadenylation site. Typically, an almost invariant AAUAAA hexamer lies 20-50 nucleotides upstream of a more variable element rich in U or GU residues. Cleavage of the nascent transcript occurs between these two elements and is coupled to the addition of up to 250 adenosines to the 5′ cleavage product. In particular embodiments, the core poly(A) sequence is a synthetic poly(A) sequence (e.g., AATAAA, ATTAAA, AGTAAA). Illustrative examples of poly(A) sequences include, but are not limited to an SV40 poly(A) sequence, a bovine growth hormone poly(A) sequence (BGHpA), a rabbit β-globin poly(A) sequence (rβgpA), or another suitable heterologous or endogenous poly(A) sequence known in the art.

(228) In particular embodiments, the polynucleotide encoding the selectable marker is fused to a polynucleotide encoding a viral self-cleaving peptide or ribosomal skipping sequence.

(229) Illustrative examples of ribosomal skipping sequences include, but are not limited to: a 2A or 2A-like site, sequence or domain (Donnelly et al., 2001. *J. Gen. Virol.* 82:1027-1041). In a particular embodiment, the viral 2A peptide is an aphthovirus 2A peptide, a potyvirus 2A peptide, or a cardiovirus 2A peptide.

(230) In one embodiment, the viral 2A peptide is selected from the group consisting of: a foot-and-mouth disease virus (FMDV) 2A peptide, an equine rhinitis A virus (ERAV) 2A peptide, a *Thosea asigna* virus (TaV) 2A peptide, a porcine teschovirus-1 (PTV-1) 2A peptide, a Theilovirus 2A peptide, and an encephalomyocarditis virus 2A peptide.

(231) Illustrative examples of 2A sites are provided in Table 1.

(232) TABLE-US-00001 TABLE 1 Exemplary 2A sites include the following sequences: SEQ ID NO: 60 LLNFDLLKLAGDVESNPGP SEQ ID NO: 61 TLNFDLLKLAGDVESNPGP SEQ ID NO: 62 LLKLAGDVESNPGP SEQ ID NO: 63 NFDLLKLAGDVESNPGP SEQ ID NO: 64 QLLNFDLLKLAGDVESNPGP SEQ ID NO: 65 APVKQTLNFDLLKLAGDVESNPGP SEQ ID NO: 66 VTELLYRMKRAETYCPRPLLAIHPTEARHKQKIVA PVKQT SEQ ID NO: 67 LNFDLLKLAGDVESNPGP SEQ ID NO: 68 LLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVE SNPGP SEQ ID NO: 69 EARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP

(233) In some embodiments, where the donor repair template is designed to derepress a γ-globin promoter and operably link the derepressed promoter to a polynucleotide encoding a therapeutic globin.

(234) The term “globin” as used herein refers to proteins or protein subunits that are capable of covalently or noncovalently binding a heme moiety, and can therefore transport or store oxygen. Subunits of vertebrate and invertebrate hemoglobins, vertebrate and invertebrate myoglobins or mutants thereof are included by the term globin. The term excludes hemocyanins. Examples of globins include α-globin or variant thereof, β-globin or variant thereof, a γ-globin or a variant thereof, and δ-globin or a variant thereof.

(235) In particular embodiments, the therapeutic globin or anti-sickling variant thereof includes, but is not limited to is β-globin, δ-globin, γ-globin, β-globinA-T87Q, β-globinA-T87Q/K120E/K95E, or β-globinA-T87Q/G16D/E22A.

(236) In some embodiments, the expression of the therapeutic globin is enhanced or improved by including one or more post-transcriptional elements and/or erythroid enhancers.

(237) Illustrative examples of post-transcriptional elements and/or erythroid enhancers suitable for use in particular embodiments of donor repair templates contemplated herein include, but are not limited to: post-transcription regulatory elements are selected from the group consisting of: a woodchuck hepatitis virus post-transcriptional response element (WPRE) or variant thereof, a hepatitis B virus post-transcriptional response element (HPRE) or variant thereof, and a polyadenylation sequence and an erythroid enhancer selected from the group consisting of: an HPFH-2 enhancer, an HS40 enhancer, and a β-globin gene 3′ enhancer.

(238) In particular embodiments, a donor repair template is designed to operably link an erythroid expression control sequence to an endogenous genomic sequence encoding a γ-globin.
(239) Illustrative examples of erythroid expression control sequences include, but are not limited to: a human β-globin LCR responsive promoter, an ankyrin gene promoter, an α-spectrin gene promoter, a β-spectrin gene promoter, or a β-globin gene promoter, optionally in combination with an HPFH-2 enhancer, an HS40 enhancer, and a β-globin gene 3′ enhancer.
(240) In particular embodiments, a donor repair template is inserted into a γ-globin locus to both derepress a γ-globin promoter to enable erythroid expression of a therapeutic globin and to select for genetically modified cells comprising the donor repair template. Derepression of the γ-globin promoter may occur through selection of the nuclease target sites or through engineering a deletion into one of the homology arms of the donor repair template. In particular embodiments, it is advantageous to engineer the nuclease to cleave at a transcription factor binding site associated with repression of a γ-globin gene, that way, derepression can occur by HDR; or in absence of HDR, the repressive site can still be disrupted by NHEJ, which also leads to derepression of the γ-globin gene. In preferred embodiments, the nuclease target site is designed to delete or disrupt a polynucleotide sequence associated with HPFH, e.g., at position −120 to −102 relative to the transcriptional start site of a γ-globin gene (see SEQ ID NOs: 1-6); preferably, a 13 bp polynucleotide sequence in a γ-globin gene associated with HPFH (SEQ ID NO: 3); more preferably a CCAAT polynucleotide sequence at position −115 to −111 (relative to the transcriptional start site of a γ-globin gene); and even more preferably the nuclease target site is at, or disrupts or deletes, a CAAT polynucleotide sequence at position −114 to −111 (relative to the transcriptional start site) of the HBG1 gene. In some embodiments, the donor repair template is also designed to disrupt endogenous γ-globin expression in favor of expressing the therapeutic globin of the donor repair template.
(241) In other particular embodiments, a donor repair template is inserted into a γ-globin locus to derepress a γ-globin promoter, to enable selection of genetically modified cells comprising the template, and to enable erythroid expression of endogenous γ-globin. Derepression of the γ-globin promoter may occur through selection of the nuclease target sites or through engineering a deletion into one of the homology arms of the donor repair template. In particular embodiments, it is advantageous to engineer the nuclease to cleave at a transcription factor binding site associated with repression of a γ-globin gene, that way, γ-globin expression can occur by through HDR; or through NHEJ and derepression of the γ-globin gene. In preferred embodiments, the nuclease target site is designed to delete or disrupt a polynucleotide sequence associated with HPFH, e.g., at position −120 to −102 relative to the transcriptional start site of a γ-globin gene (see SEQ ID NOs: 1-6); preferably, a 13 bp polynucleotide sequence in a γ-globin gene associated with HPFH (SEQ ID NO: 3); more preferably a CCAAT polynucleotide sequence at position −115 to −111 (relative to the transcriptional start site of a γ-globin gene); and even more preferably the nuclease target site is at, or disrupts or deletes, a CAAT polynucleotide sequence at position −114 to −111 (relative to the transcriptional start site) of the HBG1 gene.
(242) In various embodiments, an engineered nuclease is introduced into a hematopoietic cell, e.g., a hematopoietic stem or progenitor cell, or CD34+ cell, using non-viral or viral based methods and a donor repair template is introduced into a hematopoietic cell using viral methods by transducing the cell with an adeno-associated virus (AAV), retrovirus, e.g., lentivirus, IDLV, etc., herpes simplex virus, adenovirus, or vaccinia virus vector comprising the donor repair template.
(243) In particular embodiments, delivery of one or more polynucleotides encoding nucleases and/or donor repair templates may be provided by the same method or by different methods, and/or by the same vector or by different vectors.
(244) Illustrative methods of non-viral delivery of polynucleotides contemplated in particular embodiments include, but are not limited to: electroporation, sonoporation, lipofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, nanoparticles, polycation or

lipid:nucleic acid conjugates, naked DNA, artificial virions, DEAE-dextran-mediated transfer, gene gun, and heat-shock.
(245) Illustrative examples of polynucleotide delivery systems suitable for use in particular embodiments contemplated in particular embodiments include, but are not limited to those provided by Amaxa Biosystems, Maxcyte, Inc., BTX Molecular Delivery Systems, and Copernicus Therapeutics Inc. Lipofection reagents are sold commercially (e.g., Transfectam™ and Lipofectin™). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides have been described in the literature. See e.g., Liu et al. (2003) Gene Therapy. 10:180-187; and Balazs et al. (2011) Journal of Drug Delivery. 2011:1-12. Antibody-targeted, bacterially derived, non-living nanocell-based delivery is also contemplated in particular embodiments.
(246) Illustrative examples of viral vector systems suitable for use in particular embodiments contemplated herein include, but are not limited to adeno-associated virus (AAV), retrovirus, herpes simplex virus, adenovirus, and vaccinia virus vectors.
(247) In various embodiments, one or more polynucleotides encoding a nuclease variant and/or donor repair template are introduced into a hematopoietic cell, e.g., a hematopoietic stem or progenitor cell, or CD34+ cell, by transducing the cell with a recombinant adeno-associated virus (rAAV), comprising the one or more polynucleotides.
(248) AAV is a small (˜26 nm) replication-defective, primarily episomal, non-enveloped virus. AAV can infect both dividing and non-dividing cells and may incorporate its genome into that of the host cell. Recombinant AAV (rAAV) are typically composed of, at a minimum, a transgene and its regulatory sequences, and 5′ and 3′ AAV inverted terminal repeats (ITRs). The ITR sequences are about 145 bp in length. In particular embodiments, the rAAV comprises ITRs and capsid sequences isolated from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, or AAV10.
(249) In some embodiments, a chimeric rAAV is used the ITR sequences are isolated from one AAV serotype and the capsid sequences are isolated from a different AAV serotype. For example, a rAAV with ITR sequences derived from AAV2 and capsid sequences derived from AAV6 is referred to as AAV2/AAV6. In particular embodiments, the rAAV vector may comprise ITRs from AAV2, and capsid proteins from any one of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, or AAV10. In a preferred embodiment, the rAAV comprises ITR sequences derived from AAV2 and capsid sequences derived from AAV6. In a preferred embodiment, the rAAV comprises ITR sequences derived from AAV2 and capsid sequences derived from AAV2.
(250) In some embodiments, engineering and selection methods can be applied to AAV capsids to make them more likely to transduce cells of interest.
(251) Construction of rAAV vectors, production, and purification thereof have been disclosed, e.g., in U.S. Pat. Nos. 9,169,494; 9,169,492; 9,012,224; 8,889,641; 8,809,058; and 8,784,799, each of which is incorporated by reference herein, in its entirety.
(252) In various embodiments, one or more polynucleotides encoding a nuclease variant and/or donor repair template are introduced into a hematopoietic cell, e.g., a hematopoietic stem or progenitor cell, or CD34+ cell, by transducing the cell with a retrovirus, e.g., lentivirus, comprising the one or more polynucleotides. In one embodiment, a nuclease variant and/or donor repair template are introduced into a hematopoietic cell, e.g., a hematopoietic stem or progenitor cell, or CD34+ cell, by transducing the cell with an integrase deficient lentivirus.
(253) As used herein, the term “integration defective lentivirus” or “IDLV” refers to a lentivirus having an integrase that lacks the capacity to integrate the viral genome into the genome of the host cells. Integration-incompetent viral vectors have been described in patent application WO 2006/010834, which is herein incorporated by reference in its entirety.
(254) Illustrative mutations in the HIV-1 pol gene suitable to reduce integrase activity include, but are not limited to: H12N, H12C, H16C, H16V, S81 R, D41A, K42A, H51A, Q53C, D55V, D64E,

D64V, E69A, K71A, E85A, E87A, D116N, D1161, D116A, N120G, N1201, N120E, E152G, E152A, D35E, K156E, K156A, E157A, K159E, K159A, K160A, R166A, D167A, E170A, H171A, K173A, K186Q, K186T, K188T, E198A, R199c, R199T, R199A, D202A, K211A, Q214L, Q216L, Q221 L, W235F, W235E, K236S, K236A, K246A, G247W, D253A, R262A, R263A and K264H.

(255) In one embodiment, the HIV-1 integrase deficient pol gene comprises a D64V, D116I, D116A, E152G, or E152A mutation; D64V, D116I, and E152G mutations; or D64V, D116A, and E152A mutations.

(256) In one embodiment, the HIV-1 integrase deficient pol gene comprises a D64V mutation.

F. Genome Edited Cells

(257) The genome edited cells manufactured by the methods contemplated in particular embodiments provide improved cell-based therapeutics for the treatment of hemoglobinopathies. Without wishing to be bound to any particular theory, it is believed that the compositions and methods contemplated herein enable therapeutic globin expression and more robust selection of genome edited cells that may be used to treat, and in some embodiments potentially cure, hemoglobinopathies.

(258) Genome edited cells contemplated in particular embodiments may be autologous/autogeneic ("self") or non-autologous ("non-self," e.g., allogeneic, syngeneic or xenogeneic). "Autologous," as used herein, refers to cells from the same subject. "Allogeneic," as used herein, refers to cells of the same species that differ genetically to the cell in comparison. "Syngeneic," as used herein, refers to cells of a different subject that are genetically identical to the cell in comparison. "Xenogeneic," as used herein, refers to cells of a different species to the cell in comparison. In preferred embodiments, the cells are obtained from a mammalian subject. In a more preferred embodiment, the cells are obtained from a primate subject, optionally a non-human primate. In the most preferred embodiment, the cells are obtained from a human subject.

(259) An "isolated cell" refers to a non-naturally occurring cell, e.g., a cell that does not exist in nature, a modified cell, an engineered cell, etc., that has been obtained from an in vivo tissue or organ and is substantially free of extracellular matrix.

(260) Illustrative examples of cell types whose genome can be edited using the compositions and methods contemplated herein include, but are not limited to, cell lines, primary cells, stem cells, progenitor cells, and differentiated cells.

(261) The term "stem cell" refers to a cell which is an undifferentiated cell capable of (1) long term self-renewal, or the ability to generate at least one identical copy of the original cell, (2) differentiation at the single cell level into multiple, and in some instance only one, specialized cell type and (3) of in vivo functional regeneration of tissues. Stem cells are subclassified according to their developmental potential as totipotent, pluripotent, multipotent and oligo/unipotent. "Self-renewal" refers a cell with a unique capacity to produce unaltered daughter cells and to generate specialized cell types (potency). Self-renewal can be achieved in two ways. Asymmetric cell division produces one daughter cell that is identical to the parental cell and one daughter cell that is different from the parental cell and is a progenitor or differentiated cell. Symmetric cell division produces two identical daughter cells. "Proliferation" or "expansion" of cells refers to symmetrically dividing cells.

(262) As used herein, the term "progenitor" or "progenitor cells" refers to cells have the capacity to self-renew and to differentiate into more mature cells. Many progenitor cells differentiate along a single lineage, but may have quite extensive proliferative capacity.

(263) In particular embodiments, the cell is a primary cell. The term "primary cell" as used herein is known in the art to refer to a cell that has been isolated from a tissue and has been established for growth in vitro or ex vivo. Corresponding cells have undergone very few, if any, population doublings and are therefore more representative of the main functional component of the tissue from which they are derived in comparison to continuous cell lines, thus representing a more representative model to the in vivo state. Methods to obtain samples from various tissues and

methods to establish primary cell lines are well-known in the art (see, e.g., Jones and Wise, Methods Mol Biol. 1997). Primary cells for use in the methods contemplated herein are derived from umbilical cord blood, placental blood, mobilized peripheral blood and bone marrow. In one embodiment, the primary cell is a hematopoietic stem or progenitor cell.

(264) In one embodiment, the genome edited cell is an embryonic stem cell.

(265) In one embodiment, the genome edited cell is an adult stem or progenitor cell.

(266) In one embodiment, the genome edited cell is primary cell.

(267) In a preferred embodiment, the genome edited cell is a hematopoietic cell, e.g., hematopoietic stem cell, hematopoietic progenitor cell, an erythroid cell, or cell population comprising hematopoietic cells.

(268) As used herein, the term “population of cells” refers to a plurality of cells that may be made up of any number and/or combination of homogenous or heterogeneous cell types, as described elsewhere herein. For example, for transduction of hematopoietic stem or progenitor cells, a population of cells may be isolated or obtained from umbilical cord blood, placental blood, bone marrow, or mobilized peripheral blood. A population of cells may comprise about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, or about 100% of the target cell type to be edited. In certain embodiments, hematopoietic stem or progenitor cells may be isolated or purified from a population of heterogeneous cells using methods known in the art.

(269) Illustrative sources to obtain hematopoietic cells include, but are not limited to: cord blood, bone marrow or mobilized peripheral blood.

(270) Hematopoietic stem cells (HSCs) give rise to committed hematopoietic progenitor cells (HPCs) that are capable of generating the entire repertoire of mature blood cells over the lifetime of an organism. The term “hematopoietic stem cell” or “HSC” refers to multipotent stem cells that give rise to the all the blood cell types of an organism, including myeloid (e.g., monocytes and macrophages, neutrophils, basophils, eosinophils, erythrocytes, megakaryocytes/platelets, dendritic cells), and lymphoid lineages (e.g., T-cells, B-cells, NK-cells), and others known in the art (See Fei, R., et al., U.S. Pat. No. 5,635,387; McGlave, et al., U.S. Pat. No. 5,460,964; Simmons, P., et al., U.S. Pat. No. 5,677,136; Tsukamoto, et al., U.S. Pat. No. 5,750,397; Schwartz, et al., U.S. Pat. No. 5,759,793; DiGuisto, et al., U.S. Pat. No. 5,681,599; Tsukamoto, et al., U.S. Pat. No. 5,716,827). When transplanted into lethally irradiated animals or humans, hematopoietic stem and progenitor cells can repopulate the erythroid, neutrophil-macrophage, megakaryocyte and lymphoid hematopoietic cell pool.

(271) Additional illustrative examples of hematopoietic stem or progenitor cells suitable for use with the methods and compositions contemplated herein include hematopoietic cells that are CD34.sup.+CD38.sup.LoCD90.sup.+CD45.sup.RA−, hematopoietic cells that are CD34.sup.+, CD59.sup.+, Thy1/CD90.sup.+, CD38.sup.Lo/−, C-kit/CD117.sup.+, and Lin.sup.(−), and hematopoietic cells that are CD133.sup.+.

(272) Various methods exist to characterize hematopoietic hierarchy. One method of characterization is the SLAM code. The SLAM (Signaling lymphocyte activation molecule) family is a group of >10 molecules whose genes are located mostly tandemly in a single locus on chromosome 1 (mouse), all belonging to a subset of immunoglobulin gene superfamily, and originally thought to be involved in T-cell stimulation. This family includes CD48, CD150, CD244, etc., CD150 being the founding member, and, thus, also called slamF1, i.e., SLAM family member 1. The signature SLAM code for the hematopoietic hierarchy is hematopoietic stem cells (HSC)—CD150.sup.+CD48.sup.−CD244.sup.−; multipotent progenitor cells (MPPs)—CD150.sup.−CD48.sup.−CD244.sup.+; lineage-restricted progenitor cells (LRPs)—CD150.sup.−CD48.sup.+CD244.sup.+; common myeloid progenitor (CMP)—lin.sup.−SCA-1-c-kit.sup.+CD34.sup.+CD16/32.sup.mid; granulocyte-macrophage progenitor (GMP)—lin.sup.−SCA-1-c-kit.sup.+CD34.sup.+CD16/32.sup.hi; and megakaryocyte-erythroid progenitor (MEP)—

lin.sup.−SCA-1-c-kit.sup.+CD34.sup.−CD16/32.sup.low.
(273) Preferred target cell types edited with the compositions and methods contemplated herein include, hematopoietic cells, preferably human hematopoietic cells, more preferably human hematopoietic stem and progenitor cells, and even more preferably CD34.sup.+ human hematopoietic stem cells. The term “CD34+ cell,” as used herein refers to a cell expressing the CD34 protein on its cell surface. “CD34,” as used herein refers to a cell surface glycoprotein (e.g., sialomucin protein) that often acts as a cell-cell adhesion factor. CD34+ is a cell surface marker of both hematopoietic stem and progenitor cells.
(274) In one embodiment, the genome edited hematopoietic cells are CD150.sup.+CD48.sup.−CD244.sup.− cells.
(275) In one embodiment, the genome edited hematopoietic cells are CD34.sup.+CD133.sup.+ cells.
(276) In one embodiment, the genome edited hematopoietic cells are CD133.sup.+ cells.
(277) In one embodiment, the genome edited hematopoietic cells are CD34.sup.+ cells.
(278) In particular embodiments, a population of hematopoietic cells comprising hematopoietic stem and progenitor cells (HSPCs) comprises an edited γ-globin gene, wherein the edit is a DSB preferably repaired by HDR in the presence of a donor repair template that derepresses the γ-globin promoter, and enables therapeutic globin expression and cell selection, but where a DSB repaired by NHEJ may also be advantageous in derepressing the γ-globin promoter in certain embodiments.
(279) In particular embodiments, the genome edited cells comprise erythroid cells.
(280) In particular embodiments, the genome edited cells comprise one or more mutations in a β-globin gene. In one embodiment, the β-globin alleles of the subject are selected from the group consisting of: β.sup.E/β.sup.0, β.sup.C/β.sup.0, β.sup.0/β.sup.0, β.sup.E/β.sup.E, β.sup.C/β.sup.+, β.sup.E/β.sup.+, β.sup.0/β.sup.+, β.sup.+/β.sup.+, β.sup.C/β.sup.C, β.sup.E/β.sup.S, β.sup.0/β.sup.S, β.sup.C/β.sup.S, β.sup.+/β.sup.S or β.sup.S/β.sup.S.
(281) In particular embodiments, the genome edited cells comprise one or more one or more mutations in a β-globin gene that result in a thalassemia. In one embodiment, the thalassemia is an α-thalassemia. In one embodiment, the thalassemia is a β-thalassemia. In one embodiment, the β-globin alleles of the subject are selected from the group consisting of: β.sup.E/β.sup.0, β.sup.C/β.sup.0, β.sup.0/β.sup.0, β.sup.C/β.sup.C, β.sup.E/β.sup.E, β.sup.E/β.sup.+, β.sup.C/β.sup.E, β.sup.C/β.sup.+, β.sup.0/β.sup.+, or β.sup.+/β.sup.+.
(282) In particular embodiments, the genome edited cells comprise one or more one or more mutations in a β-globin gene that result in sickle cell disease. In one embodiment, the β-globin alleles of the subject are selected from the group consisting of: β.sup.E/β.sup.S, β.sup.0/β.sup.S, β.sup.C/β.sup.S, β.sup.+/β.sup.S or β.sup.S/β.sup.S.
G. Compositions and Formulations
(283) The compositions contemplated in particular embodiments may comprise one or more polypeptides, polynucleotides, vectors comprising same, and genome editing compositions and genome edited cell compositions, as contemplated herein. The genome editing compositions and methods contemplated in particular embodiments are useful for editing a target site in a human γ-globin gene in a cell or a population of cells. In preferred embodiments, the cell is a hematopoietic cell, e.g., a hematopoietic stem or progenitor cell, or a CD34.sup.+ cell.
(284) In particular embodiments, the compositions contemplated herein comprise a population of cells, an engineered nuclease, and a donor repair template. In particular embodiments, the compositions contemplated herein comprise a population of cells, an engineered nuclease, an end-processing enzyme, and a donor repair template.
(285) Compositions include, but are not limited to pharmaceutical compositions. A “pharmaceutical composition” refers to a composition formulated in pharmaceutically-acceptable or physiologically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy. It will also be understood that, if

desired, the compositions may be administered in combination with other agents as well, such as, e.g., cytokines, growth factors, hormones, small molecules, chemotherapeutics, pro-drugs, drugs, antibodies, or other various pharmaceutically-active agents. There is virtually no limit to other components that may also be included in the compositions, provided that the additional agents do not adversely affect the composition.

(286) The phrase “pharmaceutically acceptable” is employed herein to refer 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 of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

(287) The term “pharmaceutically acceptable carrier” refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic cells are administered. Illustrative examples of pharmaceutical carriers can be sterile liquids, such as cell culture media, water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients in particular embodiments, include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

(288) In one embodiment, a composition comprising a pharmaceutically acceptable carrier is suitable for administration to a subject. In particular embodiments, a composition comprising a carrier is suitable for parenteral administration, e.g., intravascular (intravenous or intraarterial), intraperitoneal or intramuscular administration. In particular embodiments, a composition comprising a pharmaceutically acceptable carrier is suitable for intraventricular, intraspinal, or intrathecal administration. Pharmaceutically acceptable carriers include sterile aqueous solutions, cell culture media, or dispersions. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the transduced cells, use thereof in the pharmaceutical compositions is contemplated.

(289) In particular embodiments, compositions contemplated herein comprise genetically modified hematopoietic stem and/or progenitor cells and a pharmaceutically acceptable carrier. A composition comprising a cell-based composition contemplated herein can be administered separately by enteral or parenteral administration methods or in combination with other suitable compounds to effect the desired treatment goals.

(290) The pharmaceutically acceptable carrier must be of sufficiently high purity and of sufficiently low toxicity to render it suitable for administration to the human subject being treated. It further should maintain or increase the stability of the composition. The pharmaceutically acceptable carrier can be liquid or solid and is selected, with the planned manner of administration in mind, to provide for the desired bulk, consistency, etc., when combined with other components of the composition. For example, the pharmaceutically acceptable carrier can be, without limitation, a binding agent (e.g. pregelatinized maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose, etc.), a filler (e.g., lactose and other sugars, microcrystalline cellulose, pectin, gelatin, calcium sulfate, ethyl cellulose, polyacrylates, calcium hydrogen phosphate, etc.), a lubricant (e.g., magnesium stearate, talc, silica, colloidal silicon dioxide, stearic acid, metallic stearates, hydrogenated vegetable oils, corn starch, polyethylene glycols, sodium benzoate, sodium acetate, etc.), a disintegrant (e.g., starch, sodium starch glycolate, etc.), or a wetting agent (e.g., sodium lauryl sulfate, etc.). Other suitable pharmaceutically acceptable carriers for the compositions contemplated herein include, but are not limited to, water, salt solutions, alcohols,

polyethylene glycols, gelatins, amyloses, magnesium stearates, talcs, silicic acids, viscous paraffins, hydroxymethylcelluloses, polyvinylpyrrolidones and the like.
(291) Such carrier solutions also can contain buffers, diluents and other suitable additives. The term “buffer” as used herein refers to a solution or liquid whose chemical makeup neutralizes acids or bases without a significant change in pH. Examples of buffers contemplated herein include, but are not limited to, Dulbecco's phosphate buffered saline (PBS), Ringer's solution, 5% dextrose in water (D5W), normal/physiologic saline (0.9% NaCl).
(292) The pharmaceutically acceptable carriers may be present in amounts sufficient to maintain a pH of the composition of about 7. Alternatively, the composition has a pH in a range from about 6.8 to about 7.4, e.g., 6.8, 6.9, 7.0, 7.1, 7.2, 7.3, and 7.4. In still another embodiment, the composition has a pH of about 7.4.
(293) Compositions contemplated herein may comprise a nontoxic pharmaceutically acceptable medium. The compositions may be a suspension. The term “suspension” as used herein refers to non-adherent conditions in which cells are not attached to a solid support. For example, cells maintained as a suspension may be stirred or agitated and are not adhered to a support, such as a culture dish.
(294) In particular embodiments, compositions contemplated herein are formulated in a suspension, where the genome edited hematopoietic stem and/or progenitor cells are dispersed within an acceptable liquid medium or solution, e.g., saline or serum-free medium, in an intravenous (IV) bag or the like. Acceptable diluents include, but are not limited to water, PlasmaLyte, Ringer's solution, isotonic sodium chloride (saline) solution, serum-free cell culture medium, and medium suitable for cryogenic storage, e.g., Cryostor® medium.
(295) In certain embodiments, a pharmaceutically acceptable carrier is substantially free of natural proteins of human or animal origin, and suitable for storing a composition comprising a population of genome edited cells, e.g., hematopoietic stem and progenitor cells. The therapeutic composition is intended to be administered into a human patient, and thus is substantially free of cell culture components such as bovine serum albumin, horse serum, and fetal bovine serum.
(296) In some embodiments, compositions are formulated in a pharmaceutically acceptable cell culture medium. Such compositions are suitable for administration to human subjects. In particular embodiments, the pharmaceutically acceptable cell culture medium is a serum free medium.
(297) Serum-free medium has several advantages over serum containing medium, including a simplified and better defined composition, a reduced degree of contaminants, elimination of a potential source of infectious agents, and lower cost. In various embodiments, the serum-free medium is animal-free, and may optionally be protein-free. Optionally, the medium may contain biopharmaceutically acceptable recombinant proteins. “Animal-free” medium refers to medium wherein the components are derived from non-animal sources. Recombinant proteins replace native animal proteins in animal-free medium and the nutrients are obtained from synthetic, plant or microbial sources. “Protein-free” medium, in contrast, is defined as substantially free of protein.
(298) Illustrative examples of serum-free media used in particular compositions include, but are not limited to QBSF-60 (Quality Biological, Inc.), StemPro-34 (Life Technologies), and X-VIVO 10.
(299) In a preferred embodiment, the compositions comprising genome edited hematopoietic stem and/or progenitor cells are formulated in PlasmaLyte.
(300) In various embodiments, compositions comprising hematopoietic stem and/or progenitor cells are formulated in a cryopreservation medium. For example, cryopreservation media with cryopreservation agents may be used to maintain a high cell viability outcome post-thaw. Illustrative examples of cryopreservation media used in particular compositions include, but are not limited to, CryoStor CS10, CryoStor CS5, and CryoStor CS2.
(301) In one embodiment, the compositions are formulated in a solution comprising 50:50 PlasmaLyte A to CryoStor CS10.
(302) In particular embodiments, the composition is substantially free of mycoplasma, endotoxin,

and microbial contamination. By “substantially free” with respect to endotoxin is meant that there is less endotoxin per dose of cells than is allowed by the FDA for a biologic, which is a total endotoxin of 5 EU/kg body weight per day, which for an average 70 kg person is 350 EU per total dose of cells. In particular embodiments, compositions comprising hematopoietic stem or progenitor cells transduced with a retroviral vector contemplated herein contains about 0.5 EU/mL to about 5.0 EU/mL, or about 0.5 EU/mL, 1.0 EU/mL, 1.5 EU/mL, 2.0 EU/mL, 2.5 EU/mL, 3.0 EU/mL, 3.5 EU/mL, 4.0 EU/mL, 4.5 EU/mL, or 5.0 EU/mL.

(303) In certain embodiments, compositions and formulations suitable for the delivery of polynucleotides are contemplated including, but not limited to, one or more mRNAs encoding one or more reprogrammed nucleases, and optionally end-processing enzymes.

(304) Exemplary formulations for ex vivo delivery may also include the use of various transfection agents known in the art, such as calcium phosphate, electroporation, heat shock and various liposome formulations (i.e., lipid-mediated transfection). Liposomes, as described in greater detail below, are lipid bilayers entrapping a fraction of aqueous fluid. DNA spontaneously associates to the external surface of cationic liposomes (by virtue of its charge) and these liposomes will interact with the cell membrane.

(305) In particular embodiments, formulation of pharmaceutically-acceptable carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including e.g., enteral and parenteral, e.g., intravascular, intravenous, intraarterial, intraosseously, intraventricular, intracerebral, intracranial, intraspinal, intrathecal, and intramedullary administration and formulation. It would be understood by the skilled artisan that particular embodiments contemplated herein may comprise other formulations, such as those that are well known in the pharmaceutical art, and are described, for example, in *Remington: The Science and Practice of Pharmacy*, volume I and volume II. 22.sup.nd Edition. Edited by Loyd V. Allen Jr. Philadelphia, PA: Pharmaceutical Press; 2012, which is incorporated by reference herein, in its entirety.

H. Genome Edited Cell Therapies

(306) The genome edited cells manufactured by the methods contemplated in particular embodiments provide improved drug products for use in the prevention, treatment, and amelioration of a hemoglobinopathy or for preventing, treating, or ameliorating at least one symptom associated with a hemoglobinopathy or a subject having a hemoglobinopathic mutation in a β-globin gene. As used herein, the term “drug product” refers to genetically modified cells produced using the compositions and methods contemplated herein. In particular embodiments, the drug product comprises genetically modified hematopoietic stem or progenitor cells, e.g., CD34.sup.− cells. The genetically modified hematopoietic stem or progenitor cells can be selected through positive selection in vitro, ex vivo or in vivo and give rise to adult erythroid cells with increased γ-globin gene expression and allow treatment of subjects having no or minimal expression of the γ-globin gene in vivo, thereby significantly expanding the opportunity to bring genome edited cell therapies to subjects for which this type of treatment was not previously a viable treatment option.

(307) In particular embodiments, genome edited hematopoietic stem or progenitor cells comprise a selection cassette and a mechanism to express a therapeutic globin through derepression of a γ-globin promoter or through operably linking an erythroid expression control sequence to an endogenous γ-globin coding sequence. The genetically modified cells may be positively selected for a selectable marker in a donor repair template through in vitro or ex vivo culture with the appropriate drug, or though in vivo selection by administration of the drug to a subject that has been administered a population of cells comprising the genetically modified cells that comprise the donor repair template. Drug-based selection, including dose and dosing schedule may be determined using methods known in the art.

(308) In particular embodiments, genome edited hematopoietic stem or progenitor cells provide a curative, preventative, or ameliorative therapy to a subject diagnosed with or that is suspected of having a hemoglobinopathy.
(309) As used herein, "hematopoiesis," refers to the formation and development of blood cells from progenitor cells, as well as formation of progenitor cells from stem cells. Blood cells include but are not limited to erythrocytes or red blood cells (RBCs), reticulocytes, monocytes, neutrophils, megakaryocytes, eosinophils, basophils, B-cells, macrophages, granulocytes, mast cells, thrombocytes, and leukocytes.
(310) As used herein, the term "hemoglobinopathy" or "hemoglobinopathic condition" refers to a diverse group of inherited blood disorders that involve the presence of abnormal hemoglobin molecules resulting from alterations in the structure and/or synthesis of hemoglobin. Normally, hemoglobin consists of four protein subunits: two subunits of β-globin and two subunits of α-globin. Each of these protein subunits is attached (bound) to an iron-containing molecule called heme; each heme contains an iron molecule in its center that can bind to one oxygen molecule. Hemoglobin within red blood cells binds to oxygen molecules in the lungs. These cells then travel through the bloodstream and deliver oxygen to tissues throughout the body.
(311) Hemoglobin A (HbA) is the designation for the normal hemoglobin that exists after birth. Hemoglobin A is a tetramer with two alpha chains and two beta chains (α.sub.2β.sub.2). Hemoglobin A2 is a minor component of the hemoglobin found in red cells after birth and consists of two alpha chains and two delta chains (α.sub.2δ.sub.2). Hemoglobin A2 generally comprises less than 3% of the total red cell hemoglobin. Hemoglobin F (HbF) is the predominant hemoglobin during fetal development. The molecule is a tetramer of two alpha chains and two gamma chains (α.sub.2γ.sub.2). In preferred embodiments, subjects are administered genome edited hematopoietic stem or progenitor cells that give rise to erythroid cells that have increased γ-globin gene expression and/or decreased hemoglobinopathic β-globin gene expression, thereby increasing the amount of HbF in the subject.
(312) The most common hemoglobinopathies include sickle cell disease, β-thalassemia, and α-thalassemia.
(313) In particular embodiments, the compositions and methods contemplated herein provide genome edited cell therapies for subjects having a sickle cell disease. The term "sickle cell anemia" or "sickle cell disease" is defined herein to include any symptomatic anemic condition which results from sickling of red blood cells. Sickle cell anemia β.sup.S/β.sup.S, a common form of sickle cell disease (SCD), is caused by Hemoglobin S (HbS). HbS is generated by replacement of glutamic acid (E) with valine (V) at position 6 in β-globin, noted as Glu6Val or E6V. Replacing glutamic acid with valine causes the abnormal HbS subunits to stick together and form long, rigid molecules that bend red blood cells into a sickle (crescent) shape. The sickle-shaped cells die prematurely, which can lead to a shortage of red blood cells (anemia). In addition, the sickle-shaped cells are rigid and can block small blood vessels, causing severe pain and organ damage.
(314) Additional mutations in the β-globin gene can also cause other abnormalities in β-globin, leading to other types of sickle cell disease. These abnormal forms of β-globin are often designated by letters of the alphabet or sometimes by a name. In these other types of sickle cell disease, one β-globin subunit is replaced with HbS and the other β-globin subunit is replaced with a different abnormal variant, such as hemoglobin C (HbC; β-globin allele noted as β.sup.C) or hemoglobin E (HbE; β-globin allele noted as β.sup.E).
(315) In hemoglobin SC (HbSC) disease, the β-globin subunits are replaced by HbS and HbC. HbC results from a mutation in the β-globin gene and is the predominant hemoglobin found in people with HbC disease (α.sub.2β.sup.C.sub.2). HbC results when the amino acid lysine replaces the amino acid glutamic acid at position 6 in β-globin, noted as Glu6Lys or E6K. HbC disease is relatively benign, producing a mild hemolytic anemia and splenomegaly. The severity of HbSC disease is variable, but it can be as severe as sickle cell anemia.

(316) HbE is caused when the amino acid glutamic acid is replaced with the amino acid lysine at position 26 in β-globin, noted as Glu26Lys or E26K. People with HbE disease have a mild hemolytic anemia and mild splenomegaly. HbE is extremely common in Southeast Asia and in some areas equals hemoglobin A in frequency. In some cases, the HbE mutation is present with HbS. In these cases, a person may have more severe signs and symptoms associated with sickle cell anemia, such as episodes of pain, anemia, and abnormal spleen function.
(317) Other conditions, known as hemoglobin sickle-β-thalassemias (HbSBetaThal), are caused when mutations that produce hemoglobin S and β-thalassemia occur together. Mutations that combine sickle cell disease with beta-zero (β.sup.0; gene mutations that prevent β-globin production) thalassemia lead to severe disease, while sickle cell disease combined with beta-plus (β.sup.+; gene mutations that decrease β-globin production) thalassemia is milder.
(318) As used herein, “thalassemia” refers to a hereditary disorder characterized by defective production of hemoglobin. Examples of thalassemias include α- and β-thalassemia.
(319) In particular embodiments, the compositions and methods contemplated herein provide genome edited cell therapies for subjects having a β-thalassemia. β-thalassemias are caused by a mutation in the β-globin chain, and can occur in a major or minor form. Nearly 400 mutations in the β-globin gene have been found to cause β-thalassemia. Most of the mutations involve a change in a single DNA building block (nucleotide) within or near the β-globin gene. Other mutations insert or delete a small number of nucleotides in the β-globin gene. As noted above, β-globin gene mutations that decrease β-globin production result in a type of the condition called beta-plus (β.sup.+) thalassemia. Mutations that prevent cells from producing any beta-globin result in beta-zero (β.sup.0) thalassemia. In the major form of β-thalassemia, children are normal at birth, but develop anemia during the first year of life. The minor form of β-thalassemia produces small red blood cells. Thalassemia minor occurs if you receive the defective gene from only one parent. Persons with this form of the disorder are carriers of the disease and usually do not have symptoms.
(320) HbE/β-thalassemia results from combination of HbE and β-thalassemia (β.sup.E/β.sup.0, β.sup.E/β.sup.+) and produces a condition more severe than is seen with either HbE trait or β-thalassemia trait. The disorder manifests as a moderately severe thalassemia that falls into the category of thalassemia intermedia. HbE/β-thalassemia is most common in people of Southeast Asian background.
(321) In a preferred embodiment, genome edited cell therapies contemplated herein are used to treat, prevent, or ameliorate a hemoglobinopathy is selected from the group consisting of hemoglobin C disease, hemoglobin E disease, sickle cell anemia, sickle cell disease (SCD), thalassemia, β-thalassemia, thalassemia major, thalassemia intermedia, hemoglobin Bart syndrome and hemoglobin H disease.
(322) In various embodiments, the genome editing compositions are administered by direct injection to a cell, tissue, or organ of a subject in need of gene therapy, in vivo, e.g., bone marrow. In various other embodiments, cells are edited in vitro or ex vivo, and optionally selected and expanded ex vivo. The genome edited cells are then administered to a subject in need of therapy. In certain embodiments, the cells are edited in vitro or ex vivo and selected in vivo after administration to a subject in need of therapy.
(323) Preferred cells for use in the genome editing methods contemplated herein include autologous/autogeneic (“self”) cells, preferably hematopoietic cells, more preferably hematopoietic stem or progenitor cell, and even more preferably CD34.sup.+ cells.
(324) As used herein, the terms “individual” and “subject” are often used interchangeably and refer to any animal that exhibits a symptom of a hemoglobinopathy that can be treated with the reprogrammed nucleases, genome editing compositions, gene therapy vectors, genome editing vectors, genome edited cells, and methods contemplated elsewhere herein. Suitable subjects (e.g., patients) include laboratory animals (such as mouse, rat, rabbit, or guinea pig), farm animals, and domestic animals or pets (such as a cat or dog). Non-human primates and, preferably, human

subjects, are included. Typical subjects include human patients that have, have been diagnosed with, or are at risk of having a hemoglobinopathy.
(325) As used herein, the term “patient” refers to a subject that has been diagnosed with hemoglobinopathy that can be treated with the reprogrammed nucleases, genome editing compositions, gene therapy vectors, genome editing vectors, genome edited cells, and methods contemplated elsewhere herein.
(326) As used herein “treatment” or “treating,” includes any beneficial or desirable effect on the symptoms or pathology of a hemoglobinopathy or hemoglobinopathic condition, and may include even minimal reductions in one or more measurable markers of the hemoglobinopathy or hemoglobinopathic condition. Treatment can optionally involve delaying of the progression of the hemoglobinopathy or hemoglobinopathic condition. “Treatment” does not necessarily indicate complete eradication or cure of the hemoglobinopathy or hemoglobinopathic condition, or associated symptoms thereof.
(327) As used herein, “prevent,” and similar words such as “prevention,” “prevented,” “preventing” etc., indicate an approach for preventing, inhibiting, or reducing the likelihood of the occurrence or recurrence of, hemoglobinopathy or hemoglobinopathic condition. It also refers to delaying the onset or recurrence of a hemoglobinopathy or hemoglobinopathic condition or delaying the occurrence or recurrence of the symptoms of hemoglobinopathy or hemoglobinopathic condition. As used herein, “prevention” and similar words also includes reducing the intensity, effect, symptoms and/or burden of a hemoglobinopathy or hemoglobinopathic condition prior to its onset or recurrence.
(328) As used herein, the phrase “ameliorating at least one symptom of” refers to decreasing one or more symptoms of the hemoglobinopathy or hemoglobinopathic condition for which the subject is being treated, e.g., thalassemia, sickle cell disease, etc. In particular embodiments, the hemoglobinopathy or hemoglobinopathic condition being treated is β-thalassemia, wherein the one or more symptoms ameliorated include, but are not limited to, weakness, fatigue, pale appearance, jaundice, facial bone deformities, slow growth, abdominal swelling, dark urine, iron deficiency (in the absence of transfusion), requirement for frequent transfusions. In particular embodiments, the hemoglobinopathy or hemoglobinopathic condition being treated is sickle cell disease (SCD) wherein the one or more symptoms ameliorated include, but are not limited to, anemia; unexplained episodes of pain, such as pain in the abdomen, chest, bones or joints; swelling in the hands or feet; abdominal swelling; fever; frequent infections; pale skin or nail beds; jaundice; delayed growth; vision problems; signs or symptoms of stroke; iron deficiency (in the absence of transfusion), requirement for frequent transfusions.
(329) As used herein, the term “amount” refers to “an amount effective” or “an effective amount” of a nuclease variant, genome editing composition, or genome edited cell sufficient to achieve a beneficial or desired prophylactic or therapeutic result, including clinical results.
(330) A “prophylactically effective amount” refers to an amount of a nuclease variant, genome editing composition, or genome edited cell sufficient to achieve the desired prophylactic result. Typically but not necessarily, since a prophylactic dose is used in subjects prior to or at an earlier stage of disease, the prophylactically effective amount is less than the therapeutically effective amount.
(331) A “therapeutically effective amount” of a nuclease variant, genome editing composition, or genome edited cell may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability to elicit a desired response in the individual. A therapeutically effective amount is also one in which any toxic or detrimental effects are outweighed by the therapeutically beneficial effects. The term “therapeutically effective amount” includes an amount that is effective to “treat” a subject (e.g., a patient). When a therapeutic amount is indicated, the precise amount of the compositions contemplated in particular embodiments, to be administered, can be determined by a physician in view of the specification and with consideration of individual

differences in age, weight, tumor size, extent of infection or metastasis, and condition of the patient (subject).

(332) The genome edited cells may be administered as part of a bone marrow or cord blood transplant in an individual that has or has not undergone bone marrow ablative therapy. In one embodiment, genome edited cells contemplated herein are administered in a bone marrow transplant to an individual that has undergone chemoablative or radioablative bone marrow therapy.

(333) In one embodiment, a dose of genome edited cells is delivered to a subject intravenously. In preferred embodiments, genome edited hematopoietic stem cells are intravenously administered to a subject.

(334) In one illustrative embodiment, the effective amount of genome edited cells provided to a subject is at least 2×10.sup.6 cells/kg, at least 3×10.sup.6 cells/kg, at least 4×10.sup.6 cells/kg, at least 5×10.sup.6 cells/kg, at least 6×10.sup.6 cells/kg, at least 7×10.sup.6 cells/kg, at least 8×10.sup.6 cells/kg, at least 9×10.sup.6 cells/kg, or at least 10×10.sup.6 cells/kg, or more cells/kg, including all intervening doses of cells.

(335) In another illustrative embodiment, the effective amount of genome edited cells provided to a subject is about 2×10.sup.6 cells/kg, about 3×10.sup.6 cells/kg, about 4×10.sup.6 cells/kg, about 5×10.sup.6 cells/kg, about 6×10.sup.6 cells/kg, about 7×10.sup.6 cells/kg, about 8×10.sup.6 cells/kg, about 9×10.sup.6 cells/kg, or about 10×10.sup.6 cells/kg, or more cells/kg, including all intervening doses of cells.

(336) In another illustrative embodiment, the effective amount of genome edited cells provided to a subject is from about 2×10.sup.6 cells/kg to about 10×10.sup.6 cells/kg, about 3×10.sup.6 cells/kg to about 10×10.sup.6 cells/kg, about 4×10.sup.6 cells/kg to about 10×10.sup.6 cells/kg, about 5×10.sup.6 cells/kg to about 10×10.sup.6 cells/kg, 2×10.sup.6 cells/kg to about 6×10.sup.6 cells/kg, 2×10.sup.6 cells/kg to about 7×10.sup.6 cells/kg, 2×10.sup.6 cells/kg to about 8×10.sup.6 cells/kg, 3×10.sup.6 cells/kg to about 6×10.sup.6 cells/kg, 3×10.sup.6 cells/kg to about 7×10.sup.6 cells/kg, 3×10.sup.6 cells/kg to about 8×10.sup.6 cells/kg, 4×10.sup.6 cells/kg to about 6×10.sup.6 cells/kg, 4×10.sup.6 cells/kg to about 7×10.sup.6 cells/kg, 4×10.sup.6 cells/kg to about 8×10.sup.6 cells/kg, 5×10.sup.6 cells/kg to about 6×10.sup.6 cells/kg, 5×10.sup.6 cells/kg to about 7×10.sup.6 cells/kg, 5×10.sup.6 cells/kg to about 8×10.sup.6 cells/kg, or 6×10.sup.6 cells/kg to about 8×10.sup.6 cells/kg, including all intervening doses of cells.

(337) Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject.

(338) In particular embodiments, a genome edited cell therapy is used to treat, prevent, or ameliorate a hemoglobinopathy, or condition associated therewith, comprising administering to subject having a β-globin genotype selected from the group consisting of: β.sup.E/β.sup.0, β.sup.C/β.sup.0, β.sup.0/β.sup.0, β.sup.E/β.sup.E, β.sup.C/β.sup.+, β.sup.E/β.sup.+, β.sup.0/β.sup.+, β.sup.+/β.sup.+, β.sup.C/β.sup.C, β.sup.E/β.sup.S, β.sup.0/β.sup.S, β.sup.C/β.sup.S, β.sup.+/β.sup.S or β.sup.S/β.sup.S, a therapeutically effective amount of the genome edited cells contemplated herein.

(339) In particular embodiments, genome edited cell therapies contemplated herein are used to treat, prevent, or ameliorate a thalassemia, or condition associated therewith. Thalassemias treatable with the genome edited cell contemplated herein include, but are not limited to α-thalassemias and β-thalassemias. In particular embodiments, a genome edited cell therapy is used to treat, prevent, or ameliorate a β-thalassemia, or condition associated therewith, comprising administering to subject having a β-globin genotype selected from the group consisting of: β.sup.E/β.sup.0, β.sup.C/β.sup.0, β.sup.0/β.sup.0, β.sup.C/β.sup.C, β.sup.E/β.sup.E, β.sup.E/β.sup.+, β.sup.C/β.sup.E, β.sup.C/β.sup.+, β.sup.0/β.sup.+, or β.sup.+/β.sup.+, a therapeutically effective amount of the genome edited cells contemplated herein.

(340) In particular embodiments, genome edited cell therapies contemplated herein are used to

treat, prevent, or ameliorate a sickle cell disease or condition associated therewith. In particular embodiments, a genome edited cell therapy is used to treat, prevent, or ameliorate a sickle cell disease or condition associated therewith, comprising administering to subject having a β-globin genotype selected from the group consisting of: β.sup.E/β.sup.S, β.sup.0/β.sup.S, β.sup.C/β.sup.S, β.sup.+/β.sup.S or β.sup.S/β.sup.S, a therapeutically effective amount of the genome edited cells contemplated herein.

(341) One of ordinary skill in the art would be able to use routine methods in order to determine the appropriate route of administration and the correct dosage of an effective amount of a composition comprising genome edited cells contemplated herein. It would also be known to those having ordinary skill in the art to recognize that in certain therapies, multiple administrations of pharmaceutical compositions contemplated herein may be required to effect therapy.

(342) One of the prime methods used to treat subjects amenable to treatment with genome edited hematopoietic stem and progenitor cell therapies is blood transfusion. Thus, one of the chief goals of the compositions and methods contemplated herein is to reduce the number of, or eliminate the need for, transfusions.

(343) In particular embodiments, the drug product is administered once.

(344) In certain embodiments, the drug product is administered 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more times over a span of 1 year, 2 years, 5, years, 10 years, or more.

(345) All publications, patent applications, and issued patents cited in this specification are herein incorporated by reference as if each individual publication, patent application, or issued patent were specifically and individually indicated to be incorporated by reference.

(346) Although the foregoing embodiments have been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings contemplated herein that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims. The following examples are provided by way of illustration only and not by way of limitation. Those of skill in the art will readily recognize a variety of noncritical parameters that could be changed or modified to yield essentially similar results.

EXAMPLES

Example 1

(347) Genome Editing Strategies for Treatment of Hemoglobinopathies

(348) Human CD34.sup.+ hematopoietic stem and progenitor cells were selected and cultured for 48-72 hours in growth medium containing SCF, TPO, and FLT3. The hCD34.sup.+ cells were then centrifuged, re-suspended in electroporation media, and electroporated with a polynucleotide encoding a nuclease. In this particular example, hCD34.sup.+ cells were electroporated with an mRNA encoding TALENs nucleases that target a 13 base pair region (SEQ ID NO: 3) in the γ-globin promoter (e.g., SEQ ID NO: 9), that when deleted, has shown to be associated with HPFH. FIG. **1**. The electroporated hCD34.sup.+ cells were cultured for 4 hours and were then transduced with AAV vectors (packaged with AAV6 capsid) comprising the cassettes schematized in FIG. **2**. The transduced hCD34.sup.+ cells were then cultured in methyl-cellulose to assess CFU potential and flow cytometric analysis of fetal globin expression following erythroid differentiation. FIG. **3**.

(349) The γ-globin locus was targeted by HDR using the illustrative donor repair templates and TALENs nucleases and resulted in upregulation of fetal globin (HbF) expression (2.sup.nd, and 4.sup.th bars), and was also detectable when a GFP selection cassette was used (11.sup.th and 12.sup.th bars). Analysis of erythroid-differentiated colonies in CFU assays from cells targeted with the GFP-vector showed GFP expression, thereby demonstrating that integration of a donor repair template comprising a selection cassette into this region of the γ-globin gene did not disrupt γ-globin expression or erythroid differentiation. HbF expression was further increased when the nuclease was co-transfected with donor repair templates that comprised a 13 bp deletion in the γ-globin promoter associated with HPFH or the β-globin promoter (see bars 5/6, and 7-10

respectively).
Example 2
(350) Approach
(351) A gene editing strategy was developed with TALEN's or Crispr/Cas9 ribonucleoprotein (RNP) delivery to create clinically useful deletions, including a naturally occurring 13 bp deletion, at the HBG1 and HBG2 promoter region that drives increased fetal hemoglobin expression. The deletions created by these designer nucleases eliminate suppressive elements that function to block fetal globin expression. Upon editing this region at the HBG1 or HBG2 loci, fetal hemoglobin is induced. Re-induction of fetal hemoglobin can be therapeutic in sickle-cell and β-thalassemia patients and potentially curative to alleviate symptoms. In parallel with generation of clinically useful indels/deletions, we have developed AAV gene delivery cassettes that, following introduction by homology-directed-repair (HDR), will mediate expression of functional hemoglobin based upon a series of alternative strategies described below. This overall approach thereby effectively partners: a) deletional events that promote fetal hemoglobin induction with b) additional HDR-mediated gene expression events that drive therapeutic hemoglobin production, that together, synergize to provide increased overall therapeutic benefit.
(352) Strategies
(353) Multiple strategies have been optimized for inducing hemoglobin at the HBG1 and HBG2 loci: 1. Nucleases (TALEN's and Crispr/Cas9) that edit at HBG1 and HBG2 loci and drive indels and re-induce both G1 and G2 globin. 2. Early parental constructs that integrate at the HBG1 locus and test expression of globin. 3. Optimized homology-directed repair templates that integrate at the HBG1 locus and drive: (A) HBG1 expression or (B) β.sup.T87Q expression in human or non-human primate cells. 4. Optimized homology-directed repair templates that integrate at the HBG1 locus and drive HBG1 expression or β.sup.T87Q expression and, in parallel, allow for chemo-therapeutic selection of HDR-edited human or non-human primate cells.
Strategy 1: Nucleases (TALEN's and Crispr/Cas9) that Edit at HBG1 and HBG2 Loci in Mobilized Primary Human CD34+ HSC Cells Drive Indels and Re-Induce Both G1 and G2 Globin.
(354) 1A. 13 bp deletion drives fetal hemoglobin expression in gene edited mobilized primary human CD34+ HSC cells. (FIGS. **4**-**6**) TALEN and RNP nucleases drive non-homologus end joining (NHEJ) mediated re-creation of 13 bp deletion and other useful deletions at the HBG1 and HBG2 loci. These deletions obliterate the distal CCAAT box along with sites that bind multiple transcription factors including the BCL11A binding site TGACCA and results in the induction of both HBG1 and HBG2 fetal hemoglobin.
(355) 1B. Edited mobilized primary human CD34+ HSC cells engraft and retain deletions in the gamma hemoglobin promoter (including the 13 bp and other useful deletions) in their LT-HSC population (FIG. **7**) which allows for the de-repression of fetal hemoglobin. TALEN-edited cells were engrafted into W41 NSG mice and the data supports that the edited cells robustly drive multi-lineage engraftment. The edits remain in the erythroid population up to 24 weeks and are sustained across multiple animals in three independent transplant experiments. Edits remain detectable in LT-HSC populations that engraft and re-populate the bone marrow in primary and secondary transplant recipients. These data prove that creating deletions in the gamma hemoglobin promoter that drive fetal hemoglobin (including the 13 bp deletion) may be an effective long-term therapeutic approach. FIG. **8** shows use of RNPs to efficiently target the same region in HBG1 and 2 in mobilized primary human CD34+ HSC cells.
(356) Strategy 2: Development and Testing of Initial/Parental Constructs for HDR Based Editing of the HBG1 Locus (FIGS. **9**-**18**).
(357) The initial constructs tested various elements including alternative promoters, enhancers, polyA tails, introns, varying homology arm lengths, deletional versus non-deletional constructs to identify the best design to maximize globin expression. Various constructs were designed and tested to identify the best regulatory regions to promote maximal globin expression. The data

supports the following: HBB and HBG1-d13 promoters work equally well with respect to globin expression. Tissue-specific enhancers like HS-40, HPFH-2 produce the maximum protein expression from the donor templates. wPRE-3, wPRE-O elements work well and enhance stability of the mRNA and therefore globin expression. Longer HR arms mediated higher rates of HDR using AAV donor templates. Deletional templates (e.g. ones that had deletions near the nuclease cleavage site) yielded lower HR rates than non-deletional templates.
The early designs helped identify that HBG1 locus is amenable to HDR and with a positive control AAV delivering GFP, a 30% HR rate was observed at the HBG1 locus.
(358) The basic function of elements described in the naming conventions of the described templates are as follows: HPFH2 Enhancer—used in HBG1 cassettes to enhance promoter activity; d13HBG1 Promoter—HBG1 promoter with 13 bp HPFH deletion that promotes HBG1 expression; HBBpro—utilizes the endogenous HBB promoter to drive HBG1 expression; HS40 Enhancer—used in combination with HBBpro and HBG1pro to enhance the promoter activity; T2A—used in constructs to use the exons and polyA from the native gene; wPRE3.SV40USE.pA—"minimal" wPRE and modified SV40 polyA; MND—denotes MND-CMV1 short version promoter; T87Q—Anti-sickling hemoglobin gene; MGMT—Anti-sickling P140K mutant MGMT for chemoselection; and GFP—Green Fluorescent Protein.
(359) Data optimizing HDR templates with GFP include control templates expressing only GFP (Constructs 1263 and 1264); templates inducing G1 globin (Constructs 1324 and 1325); and templates driving β.sup.T87Q (Construct 1345).
(360) Strategy 3: Homology-Directed Repair Templates that Integrate at the HBG1 Locus and Drive: (A) HBG1 Expression or (B) β.sup.T87Q Expression in Mobilized Primary Human CD34+ HSC Cells and Non-Human Primate BM CD34+ Cells (FIGS. **19**-**27**).
(361) rAAV-6 and rAAV-5 delivery of HDR templates have been designed and optimized for delivery of donor templates into human and Rhesus cells, respectively.
(362) 3A. Homology-directed repair templates that integrate at the HBG1 locus and drive HBG1 expression. Constructs 1324 and 1325 (FIGS. **19**-**20**) are donor templates with optimized promoters (HBG1, HBB, PGK) and enhancers (HS-40, HPFH2), poly-A tail (SV-40) with varying homology arm lengths (400-1000 bp) and drive insertion at the HBG1 promoter region and recapitulates the 13 bp deletion that drives HBG1 native promoter-mediated induction of Gamma 1 globin. Data shown demonstrates results following co-delivery of these templates with AAV and RNP in mobilized primary human CD34+ HSC cells.
(363) 3B. Homology-directed repair templates that integrate at the HBG1 locus and drive β.sup.T87Q expression. Constructs 1345 (FIG. **27**) is a donor template with optimized promoters (HBG1, HBB, PGK) and enhancers (HS-40, HPFH2), poly-A tail (SV-40) with varying homology arm lengths (400-1000 bp) that drive insertion at the HBG1 promoter region and drives HBB promoter-mediated induction of β.sup.T87Q globin (anti-sickling globin) at the HBG1 locus. (Rhesus version of this construct: 1348 is shown on FIG. **28**). Data shown demonstrates results following co-delivery of these templates as AAV and RNP in mobilized primary human CD34+ HSC cells.
(364) Strategy 4: Homology-Directed Repair Templates that Integrate at the HBG1 Locus and Drive HBG1 Expression or β.sup.T87Q Expression and have a P140K MGMT Cassette that Allows for Chemo-Therapeutic Selection of Edited Human or Non-Human Primate CD34+ Cells (FIGS. **28**-**31**; **34**-**36**).
(365) 4A. Homology-directed repair templates that integrate at the HBG1 locus and drive HBG1 expression and allows for chemo-selection. Construct 1333 is a donor template with optimized promoters (HBG1, HBB, PGK) and enhancers (HS-40, HPFH2), poly-A tail (SV-40) with varying homology arm lengths (400-1000 bp) and drive insertion at the HBG1 promoter region and induces HBG1 native promoter-mediated induction of G1 globin and has the P140K MGMT cassette that allows for enrichment of edited cells ex vivo before transplant or in vivo in the patient following

infusion of edited cells.
(366) 4B. Homology-directed repair templates that integrate at the HBG1 locus and drive β.sup.T87Q expression. Constructs 1336, 1346, 1343, 1344: These are donor templates with optimized promoters (HBG1, HBB, PGK) and enhancers (HS-40, HPFH2), poly-A tail (SV-40) with varying homology arm lengths (400-1000 bp) that drive insertion at the HBG1 promoter region and drives HBB promoter-mediated induction of T87Q globin (anti-sickling globin) at the HBG1 locus. All constructs also have a P140K MGMT cassette that allows for enrichment of edited cells ex vivo before transplant or in vivo following infusion of edited cells. (Rhesus version: Construct 1348).
(367) Conclusions
(368) 1. The data proves the idea that creating deletions using TALEN or RNP nuclease delivery in the promoter of gamma hemoglobin including re-creating the 13 bp deletion at the HBG1 and/or HBG2 loci drive fetal hemoglobin expression. 2. TALEN's as well as Crispr/Cas9 nucleases can create deletions in the promoter including re-creating the 13 bp deletion as well as a range of other useful deletions. 3. Edited mobilized human CD34+ HSC engraft and retain their multi-lineage engraftment potential in primary and secondary recipient mice. 4. Edits are sustained in the LT-HSC population and are able to re-populate the bone marrow in the scenario of a primary and secondary transplant and to continue to facilitate fetal hemoglobin production. 5. The HBG1 locus is amenable to HDR. Construct 1263 drives HDR rates of ˜30% as assessed by GFP+ cells following co-delivery of TALEN and 1263 and confirms that the locus supports HDR. 6. Constructs 1324, 1325, and 1345 drive HDR following co-delivery of RNP+AAV donors. Construct 1345 produces 3.5% HDR in RNP+1345 treated cells. 7. AAV HDR-donor cassettes promote gamma globin or T87Q globin expression following HDR depending on the respective cassette. 8. MGMT donor templates are effective in permitting selection for HDR-edited cells. There is 5-fold enrichment of HDR edited cells when chemo-selection is used in vitro in CD34+ cells containing the MGMT cassettes.
(369) The data confirms that the d13 deletion and related useful deletions can be used effectively as a therapeutic approach for treating sickle cell anemia and thalassemia. Most importantly, its shown that HDR cassettes can be delivered to HBG1 locus following nuclease cleavage at this site. The approach is useful and novel as both NHEJ and HDR outcomes will drive a functional response that is desirable. This combined strategy is unique as all edited cells (including HDR and NHEJ edited outcomes) have a therapeutic benefit and these combined events are more likely to provide a curative approach in sickle or β-thalassemia patients.
Example 3
Talen Mediated Therapeutic Gene Editing Strategy for β-Hemoglobinopathies
(370) Hemoglobinopathies including sickle cell disease (SCD) and β-thalassemia are the most common single-gene disorders in the world and represent a major global public health concern. The unifying principle of this heterogeneous mix of gene mutations is the decreased production of wild type hemoglobin molecules either due to structural defects in the case of SCD or insufficient production of β-globin subunits.
(371) Patients who carry both a mutation causing a hemoglobinopathy as well as increased expression of fetal hemoglobin (HPFH) tend to exhibit a milder phenotype. These mutations range from large deletions to single nucleotide polymorphisms. The focus of this Example is on a unique naturally occurring 13 bp deletion in the γ-hemoglobin promoter that has been shown to induce high levels of fetal hemoglobin expression.
(372) The 13 bp deletion site offers a unique target for therapeutic gene editing in the treatment of hemoglobinopathies. The sequence specific introduction of double strand breaks using targeted nucleases, such as TALENs or RNPs, has the potential to generate a HPFH phenotype by NHEJ (via disruption of the distal CCAAT box) but also allows for the integration of therapeutic repair templates and selection elements via HDR.

(373) Methods:
(374) TALEN Design & Testing—Multiple TALEN pairs were designed, Golden Gate assembled into a novel expression vector with an encoded poly-A tail, mRNA was generated by CellScript IVT. mRNA was transfected into human mobilized peripheral blood CD34 cells using the Neon Transfection System.
(375) Editing Efficiency—INDEL generation detected by T7 Assay following globin specific nested PCR. Colony sequencing and next generation sequencing were carried out to determine specific sequence variants. A globin specific ddPCR assay was developed to detect INDEL generation.
(376) Erythroid Differentiation—Following TALEN editing, CD34 cells are moved to erythroid differentiation media and cultured for 7-10 days. HbF expression is assessed by flow and HPLC.
(377) Murine Transplants—1e6 CD34 cells (Control & TALEN transfected) were injected by tail vein into W41 mice following minimal radiation (150rad) 18 hours post electroporation. Marrow was harvested and sort/analyzed 24 weeks post transplant. Secondary animals were transplanted with 50% of the primary harvested cells and were subsequently harvested at 9 weeks.
(378) HR Template Testing—HR Templates were designed and synthesized, packaged into AAV6 constructs and introduced to the CD34 cells at the time of TALEN transfection. Cells were differentiated and analyzed by flow and HPLC.
(379) Results (In Vitro)
(380) An experimental timeline (shown below) provided for neon transfection of human mobilized peripheral blood CD34 cells.
(381) ##STR00001##
(382) T7 Analysis showed that del13 TALEN pair transfection induces INDELs in human CD34 cells at both the γ1(HBG1) and γ2(HBG2) locus (FIG. **38**). ddPCR showed 20-30% INDEL rate at both loci using γ1(HBG1) and γ2(HBG2) specific probes. This increased to 50% with a 30 C recovery step (FIG. **37**). Next gen sequencing then provided confirmed editing rates (50% in this example). There was an overrepresentation of the 13 bp deletion, which was likely the result of microhomology in the region (FIG. **39**). HbF induction by flow cytometry demonstrated TALEN editing of peripheral blood CD34 cells followed by erythroid differentiation, which resulted in significantly increased number of F-cells (FIG. **40**). HbF induction by HPLC demonstrated TALEN editing of peripheral blood CD34 cells followed by erythroid differentiation, which resulted in significantly increased HbF protein expression (FIG. **41**). The difference in protein expression was greater than the percent F-cells suggesting that the increased F-cells expressed higher levels of HbF than control F-cells. HR template integration yielded HbF induction and anti-sickling T87Q expression (FIGS. **16**-**17**). Combined, over 60% of the globin expressed in edited cells is potentially clinically beneficial.
(383) Results (In Vivo)
(384) An experimental timeline (shown below) provided for neon transfected human CD34 cells transplanted by tail vein injection in W41 mice.
(385) ##STR00002##
(386) Week 24 human engraftment showed no significant differences in percent human engraftment between control and edited cells (FIG. **42**). Engrafted edited CD34 cells generated all human hematopoietic lineages. Human erythroid, CD34+ Lymphoid and Myeloid, and CD19+ cells were all identified and sorted from harvested marrows following transplant (W24) (FIG. **43**). The TALEN edited CD34 cells produced more F-Cells. At sac (A) there was a significantly higher rate of human F-Cells detected in the marrow, while differentiated CD34 cells from the marrow produced more F-Cells (B) (FIG. **44**). Sorted cells from all lineages retained INDELs from TALEN editing. T7 Analysis demonstrated INDELS are present in vitro (FIG. **45**). Secondary mice engraftments and marrow cells retained edits made; human engraftment at 9 weeks was low (<2%) but present and marrow cells retained edits at both loci.

(387) Conclusions

(388) It was demonstrated that TALEN mRNA mediated disruption of repressive elements in the promoters of the γ-hemoglobin genes HBG1 and HBG2 can result in the induction of fetal hemoglobin expression. These cells retain the ability to engraft in W41 mice and differentiate into multiple lineages while retaining the TALEN induced gene edits. The successful engraftment of hematopoietic cells in secondary mice suggests that at least a portion of the edited peripheral blood CD34 cells represent hematopoietic stem cells. A homologous repair template was successfully integrated at the cut site using the same TALEN mRNA to generate the double strand break resulting in the expression of both anti-sickling T87Q hemoglobin as well as increased levels of fetal hemoglobin. Additional repair templates allow for further exploration of controlling hemoglobin expression at this locus.

Example 4

CRISPR/Cas9-Edited Hematopoietic Stem and Progenitor Cells for the Reactivation of Fetal Hemoglobin

(389) A promising therapeutic strategy for hemoglobinopathies consists in the genome engineering of patients' hematopoietic stem and progenitor cells (HSPCs) to reactivate fetal hemoglobin (HbF) production, which can serve as substitute for defective or absent adult hemoglobin molecules. Here, the nonhuman primate (NHP) large animal transplantation model was used to address existing challenges for clinical translation of this approach to ensure efficient gene editing in scale-up conditions and optimize long-term engraftment of gene-edited cells.

(390) The CRISPR/Cas9 nuclease platform was employed to recapitulate a 13-nucleotide (nt) deletion in the gamma globin gene promoter identified in individuals with hereditary persistence of fetal hemoglobin (HPFH). Two rhesus macaques were transplanted with bone marrow-derived CD34+ cells edited ex vivo by CRISPR/Cas9 ribonucleoprotein electroporation. 70% editing efficiency was detected in the infusion product, with over 25% of cells containing the 13-nt deletion. Both animals showed rapid hematopoietic recovery and peripheral blood gene editing levels stabilized at 15% and 30%, respectively, at 6 months post transplantation (FIG. **61**A). HbF production, as determined by peripheral blood F-cells staining (FIG. **61**B) and quantitative PCR, was substantially increased in both animals as compared to controls and correlated with in vivo editing levels.

(391) To circumvent challenges associated with scale up and cost of editing reagents, the transplantation protocol was further optimized by purifying a refined and more highly enriched target cell population (CD34.sup.+CD45RA.sup.−CD90.sup.+) (described in PCT Application No. WO2017/218948, which is incorporated herein by reference in its entirety), and capable of both rapid short-term and as durable multilineage hematopoietic reconstitution. Two rhesus macaques were co-infused with this CRISPR/Cas9-edited subset (comprising only 5-7% of total CD34+ cells) along with the remaining un-edited cells. In vivo gene-editing levels started at less than 5% but rapidly increased to 50% within a week, and persisted at efficiencies comparable to animals receiving edited CD34+ cells, consistent with this refined cell subset as major contributor to hematopoietic recovery (FIG. **60**A).

(392) Taken together, these data demonstrate robust engraftment of CRISPR/Cas9-edited HSPCs following targeting of the 13 nt-HPFH site in the NHP model leading to high levels of HbF production. In addition, efficient editing and engraftment of the CD90+ cell subset is shown, an approach that reduces the required amount of editing reagents by 95%, circumvents challenges associated with scale up, without compromising editing or engraftment efficiencies, and thereby facilitating clinical translation of gene editing for the treatment of hemoglobinopathies.

(393) In general, in the following claims, the terms used should not be construed to limit the claims to the specific embodiments disclosed in the specification and the claims, but should be construed to include all possible embodiments along with the full scope of equivalents to which such claims are entitled. Accordingly, the claims are not limited by the disclosure.

(394) TABLE-US-00002 Lengthy table referenced here US12385070-20250812-T00001 Please refer to the end of the specification for access instructions.
(395) TABLE-US-LTS-00001 LENGTHY TABLES The patent contains a lengthy table section. A copy of the table is available in electronic form from the USPTO web site (). An electronic copy of the table will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

Claims

1. A DNA donor repair template comprising: a 5′ homology arm, a selection cassette, an erythroid expression control sequence, and a 3′ homology arm, wherein the DNA donor repair template has a sequence with at least 95% sequence identity to SEQ ID NOs: 17, 18, 19, 20, or 21.
2. The DNA donor repair template of claim 1, wherein the DNA donor repair template comprises a sequence having at least 95% sequence identity to the sequence as set forth in SEQ ID NO: 17.
3. The DNA donor repair template of claim 1, wherein the DNA donor repair template has a sequence as set forth in SEQ ID NO: 17.
4. The DNA donor repair template of claim 1, wherein the DNA donor repair template comprises a sequence having at least 95% sequence identity to the sequence as set forth in SEQ ID NO: 18.
5. The DNA donor repair template of claim 1, wherein the DNA donor repair template has a sequence as set forth in SEQ ID NO: 18.
6. The DNA donor repair template of claim 1, wherein the DNA donor repair template comprises a sequence having at least 95% sequence identity to the sequence as set forth in SEQ ID NO: 19.
7. The DNA donor repair template of claim 1, wherein the DNA donor repair template has a sequence as set forth in SEQ ID NO: 19.
8. The DNA donor repair template of claim 1, wherein the DNA donor repair template comprises a sequence having at least 95% sequence identity to the sequence as set forth in SEQ ID NO: 20.
9. The DNA donor repair template of claim 1, wherein the DNA donor repair template has a sequence as set forth in SEQ ID NO: 20.
10. The DNA donor repair template of claim 1, wherein the DNA donor repair template comprises a sequence having at least 95% sequence identity to the sequence as set forth in SEQ ID NO: 21.
11. The DNA donor repair template of claim 1, wherein the DNA donor repair template has a sequence as set forth in SEQ ID NO: 21.
12. A viral vector comprising the DNA donor repair template of claim 1.
13. The viral vector of claim 12, wherein the viral vector is a recombinant adeno-associated viral vector (rAAV) or a retrovirus.
14. The viral vector of claim 13, wherein the rAAV has one or more ITRs from AAV2.
15. The viral vector of claim 13, wherein the rAAV has a serotype selected from the group consisting of: AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, and AAV10.
16. The viral vector of claim 13, wherein the rAAV has an AAV6 serotype.
17. The viral vector of claim 13, wherein the retrovirus is a lentivirus.
18. The viral vector of claim 17, wherein the lentivirus is an integrase deficient lentivirus.
19. A cell comprising the DNA donor repair template of claim 1.
20. The cell of claim 19, wherein the DNA donor repair template has been inserted into a human gamma globin gene target site by homology directed repair.
21. The cell of claim 20, wherein the target site is an engineered nuclease target site set forth in SEQ ID NO: 7 or SEQ ID NO: 8.
22. The cell of claim 19, wherein the cell is a hematopoietic cell.
23. The cell of claim 19, wherein the cell is CD34+ cell.
24. The cell of claim 19, wherein the cell is CD133+ cell.
25. A cell comprising the viral vector of claim 12.

26. A composition comprising the DNA donor repair template of claim 1 and a physiologically acceptable excipient.
27. A composition comprising a cell of claim 19 and a physiologically acceptable excipient.