United States Patent Application Publication	20250250588
Kind Code	A1
Publication Date	August 07, 2025
Inventor(s)	AUPEPIN DE LAMOTHE-DREUZY; Edouard et al.

SYSTEMS AND METHODS FOR THE TREATMENT OF HEMOGLOBINOPATHIES

Abstract

Genome editing systems, guide RNAs, and CRISPR-mediated methods are provided for altering portions of the HBG1 and HBG2 loci in cells and increasing expression of fetal hemoglobin.

Inventors: **AUPEPIN DE LAMOTHE-DREUZY; Edouard (Boston, MA), CHANG; KaiHsin (Medfield, MA), SANCHEZ; Minerva Elaine (Brookline, MA), HEATH; Jack (Winchester, MA)**

Applicant: **EDITAS MEDICINE, INC.** (Cambridge, MA)

Family ID: **66248662**

Assignee: **EDITAS MEDICINE, INC. (Cambridge, MA)**

Appl. No.: **18/747309**

Filed: **June 18, 2024**

Related U.S. Application Data

parent US continuation 17013420 20200904 ABANDONED child US 18747309
parent US continuation PCT/US2019/021244 20190307 PENDING child US 17013420
us-provisional-application US 62773055 20181129
us-provisional-application US 62672007 20180515
us-provisional-application US 62639968 20180307

Publication Classification

Int. Cl.: **C12N15/90** (20060101); **C12N9/22** (20060101); **C12N15/11** (20060101)

U.S. Cl.:

CPC **C12N15/907** (20130101); **C12N9/22** (20130101); **C12N15/11** (20130101); C12N2310/20 (20170501); C12N2800/80 (20130101)

Background/Summary

CROSS REFERENCE TO RELATED APPLICATIONS [0001] The present application is a continuation of U.S. application Ser. No. 17/013,420, filed Sep. 4, 2020, which is a continuation of International Application No. PCT/US2019/021244, filed Mar. 7, 2019; which claims the benefit of U.S. Provisional Application No. 62/639,968, filed Mar. 7, 2018; U.S. Provisional Application No. 62/672,007, filed May 15, 2018; and U.S. Provisional Application No. 62/773,055, filed Nov. 29, 2018; the contents of each of which is hereby incorporated by reference in its entirety.

SEQUENCE LISTING
[0002] This application contains a ST.26 compliant Sequence Listing, which was submitted in XML format via Patent Center, and is hereby incorporated by reference in its entirety. The XML copy, created on Apr. 22, 2025, is named SubstituteSequenceListing.xml and is 2,150,000 bytes in size.
FIELD
[0003] This disclosure relates to genome editing systems and methods for altering a target nucleic acid sequence, or modulating expression of a target nucleic acid sequence, and applications thereof in connection with the alteration of genes encoding hemoglobin subunits and/or treatment of hemoglobinopathies.
BACKGROUND
[0004] Hemoglobin (Hb) carries oxygen in erythrocytes or red blood cells (RBCs) from the lungs to tissues. During prenatal development and until shortly after birth, hemoglobin is present in the form of fetal hemoglobin (HbF), a tetrameric protein composed of two alpha (α)-globin chains and two gamma (γ)-globin chains. HbF is largely replaced by adult hemoglobin (HbA), a tetrameric protein in which the γ-globin chains of HbF are replaced with beta (β)-globin chains, through a process known as globin switching. The average adult makes less than 1% HbF out of total hemoglobin (Thein 2009). The α-hemoglobin gene is located on chromosome 16, while the β-hemoglobin gene (HBB), A gamma (γA)-globin chain (HBG1, also known as gamma globin A), and G gamma (γG)-globin chain (HBG2, also known as gamma globin G) are located on chromosome 11 within the globin gene cluster (also referred to as the globin locus).
[0005] Mutations in HBB can cause hemoglobin disorders (i.e., hemoglobinopathies) including sickle cell disease (SCD) and beta-thalassemia (β-Thal). Approximately 93,000 people in the United States are diagnosed with a hemoglobinopathy. Worldwide, 300,000 children are born with hemoglobinopathies every year (Angastiniotis 1998). Because these conditions are associated with HBB mutations, their symptoms typically do not manifest until after globin switching from HbF to HbA.
[0006] SCD is the most common inherited hematologic disease in the United States, affecting approximately 80,000 people (Brousseau 2010). SCD is most common in people of African ancestry, for whom the prevalence of SCD is 1 in 500. In Africa, the prevalence of SCD is 15 million (Aliyu 2008). SCD is also more common in people of Indian, Saudi Arabian and Mediterranean descent. In those of Hispanic-American descent, the prevalence of sickle cell disease is 1 in 1,000 (Lewis 2014).
[0007] SCD is caused by a single homozygous mutation in the HBB gene, c.17A>T (HbS mutation). The sickle mutation is a point mutation (GAG>GTG) on HBB that results in substitution of valine for glutamic acid at amino acid position 6 in exon 1. The valine at position 6 of the β-hemoglobin chain is hydrophobic and causes a change in conformation of the β-globin protein

when it is not bound to oxygen. This change of conformation causes HbS proteins to polymerize in the absence of oxygen, leading to deformation (i.e., sickling) of RBCs. SCD is inherited in an autosomal recessive manner, so that only patients with two HbS alleles have the disease. Heterozygous subjects have sickle cell trait, and may suffer from anemia and/or painful crises if they are severely dehydrated or oxygen deprived.

[0008] Sickle shaped RBCs cause multiple symptoms, including anemia, sickle cell crises, vaso-occlusive crises, aplastic crises, and acute chest syndrome. Sickle shaped RBCs are less elastic than wild-type RBCs and therefore cannot pass as easily through capillary beds and cause occlusion and ischemia (i.e., vaso-occlusion). Vaso-occlusive crisis occurs when sickle cells obstruct blood flow in the capillary bed of an organ leading to pain, ischemia, and necrosis. These episodes typically last 5-7 days. The spleen plays a role in clearing dysfunctional RBCs, and is therefore typically enlarged during early childhood and subject to frequent vaso-occlusive crises. By the end of childhood, the spleen in SCD patients is often infarcted, which leads to autosplenectomy. Hemolysis is a constant feature of SCD and causes anemia. Sickle cells survive for 10-20 days in circulation, while healthy RBCs survive for 90-120 days. SCD subjects are transfused as necessary to maintain adequate hemoglobin levels. Frequent transfusions place subjects at risk for infection with HIV, Hepatitis B, and Hepatitis C. Subjects may also suffer from acute chest crises and infarcts of extremities, end organs, and the central nervous system.

[0009] Subjects with SCD have decreased life expectancies. The prognosis for patients with SCD is steadily improving with careful, life-long management of crises and anemia. As of 2001, the average life expectancy of subjects with sickle cell disease was the mid-to-late 50's. Current treatments for SCD involve hydration and pain management during crises, and transfusions as needed to correct anemia.

[0010] Thalassemias (e.g., β-Thal, δ-Thal, and β/δ-Thal) cause chronic anemia. β-Thal is estimated to affect approximately 1 in 100,000 people worldwide. Its prevalence is higher in certain populations, including those of European descent, where its prevalence is approximately 1 in 10,000. β-Thal major, the more severe form of the disease, is life-threatening unless treated with lifelong blood transfusions and chelation therapy. In the United States, there are approximately 3,000 subjects with β-Thal major. β-Thal intermedia does not require blood transfusions, but it may cause growth delay and significant systemic abnormalities, and it frequently requires lifelong chelation therapy. Although HbA makes up the majority of hemoglobin in adult RBCs, approximately 3% of adult hemoglobin is in the form of HbA2, an HbA variant in which the two 7-globin chains are replaced with two delta (Δ)-globin chains. δ-Thal is associated with mutations in the A hemoglobin gene (HBD) that cause a loss of HBD expression. Co-inheritance of the HBD mutation can mask a diagnosis of β-Thal (i.e., β/δ-Thal) by decreasing the level of HbA2 to the normal range (Bouva 2006). β/δ-Thal is usually caused by deletion of the HBB and HBD sequences in both alleles. In homozygous (δo/δo βo/βo) patients, HBG is expressed, leading to production of HbF alone.

[0011] Like SCD, β-Thal is caused by mutations in the HBB gene. The most common HBB mutations leading to β-Thal are: c.−136C>G, c.92+1G>A, c.92+6T>C, c.93−21G>A, c.118C>T, c.316-106C>G, c.25_26delAA, c.27_28insG, c.92+5G>C, c.118C>T, c.135delC, c.315+1G>A, c.−78A>G, c.52A>T, c.59A>G, c.92+5G>C, c.124_127delTTCT, c.316-197C>T, c.−78A>G, c.52A>T, c.124_127delTTCT, c.316-197C>T, c.−138C>T, c.−79A>G, c.92+5G>C, c.75T>A, c.316-2A>G, and c.316-2A>C. These and other mutations associated with β-Thal cause mutated or absent β-globin chains, which causes a disruption of the normal Hb α-hemoglobin to β-hemoglobin ratio. Excess α-globin chains precipitate in erythroid precursors in the bone marrow.

[0012] In β-Thal major, both alleles of HBB contain nonsense, frameshift, or splicing mutations that leads to complete absence of β-globin production (denoted β.sup.0/β.sup.0). β-Thal major results in severe reduction in β-globin chains, leading to significant precipitation of α-globin chains in RBCs and more severe anemia.

[0013] β-Thal intermedia results from mutations in the 5′ or 3′ untranslated region of HBB, mutations in the promoter region or polyadenylation signal of HBB, or splicing mutations within the HBB gene. Patient genotypes are denoted βo/β+ or β+/β+. βo represents absent expression of a β-globin chain; β+ represents a dysfunctional but present β-globin chain. Phenotypic expression varies among patients. Since there is some production of β-globin, β-Thal intermedia results in less precipitation of α-globin chains in the erythroid precursors and less severe anemia than β-Thal major. However, there are more significant consequences of erythroid lineage expansion secondary to chronic anemia.

[0014] Subjects with β-Thal major present between the ages of 6 months and 2 years, and suffer from failure to thrive, fevers, hepatosplenomegaly, and diarrhea. Adequate treatment includes regular transfusions. Therapy for β-Thal major also includes splenectomy and treatment with hydroxyurea. If patients are regularly transfused, they will develop normally until the beginning of the second decade. At that time, they require chelation therapy (in addition to continued transfusions) to prevent complications of iron overload. Iron overload may manifest as growth delay or delay of sexual maturation. In adulthood, inadequate chelation therapy may lead to cardiomyopathy, cardiac arrhythmias, hepatic fibrosis and/or cirrhosis, diabetes, thyroid and parathyroid abnormalities, thrombosis, and osteoporosis. Frequent transfusions also put subjects at risk for infection with HIV, hepatitis B and hepatitis C.

[0015] β-Thal intermedia subjects generally present between the ages of 2-6 years. They do not generally require blood transfusions. However, bone abnormalities occur due to chronic hypertrophy of the erythroid lineage to compensate for chronic anemia. Subjects may have fractures of the long bones due to osteoporosis. Extramedullary erythropoiesis is common and leads to enlargement of the spleen, liver, and lymph nodes. It may also cause spinal cord compression and neurologic problems. Subjects also suffer from lower extremity ulcers and are at increased risk for thrombotic events, including stroke, pulmonary embolism, and deep vein thrombosis. Treatment of β-Thal intermedia includes splenectomy, folic acid supplementation, hydroxyurea therapy, and radiotherapy for extramedullary masses. Chelation therapy is used in subjects who develop iron overload.

[0016] Life expectancy is often diminished in β-Thal patients. Subjects with β-Thal major who do not receive transfusion therapy generally die in their second or third decade. Subjects with β-Thal major who receive regular transfusions and adequate chelation therapy can live into their fifth decade and beyond. Cardiac failure secondary to iron toxicity is the leading cause of death in β-Thal major subjects due to iron toxicity.

[0017] A variety of new treatments are currently in development for SCD and β-Thal. Delivery of an anti-sickling HBB gene via gene therapy is currently being investigated in clinical trials. However, the long-term efficacy and safety of this approach is unknown. Transplantation with hematopoietic stem cells (HSCs) from an HLA-matched allogeneic stem cell donor has been demonstrated to cure SCD and β-Thal, but this procedure involves risks including those associated with ablation therapy, which is required to prepare the subject for transplant, increases risk of life-threatening opportunistic infections, and risk of graft vs. host disease after transplantation. In addition, matched allogeneic donors often cannot be identified. Thus, there is a need for improved methods of managing these and other hemoglobinopathies.

SUMMARY

[0018] Provided herein are genome editing systems, guide RNAs (gRNAs), and CRISPR-mediated methods for altering one or more γ-globin genes (e.g., HBG1, HBG2, or HBG1 and HBG2), the erythroid specific enhancer of the BCL11A gene (BCL11Ae), or a combination thereof, and increasing expression of fetal hemoglobin (HbF). In certain embodiments, genome editing systems, gRNAs, and CRISPR-mediated methods may alter a 13 nucleotide (nt) target region that is 5′ of the transcription site of the HBG1, HBG2, or HBG1 and HBG2 gene (“13 nt target region”), one or more regions set forth in Table 13, or a combination thereof. In certain embodiments, one or more

gRNAs comprising a targeting domain set forth in SEQ ID NOs:251-901, 940-942, or Table 12 may be used to introduce alterations in the 13 nt target region. In certain embodiments, one or more gRNAs comprising a targeting domain set forth in Table 12 may be used to introduce alterations in one or more regions set forth in Table 13. In certain embodiments, genome editing systems, gRNAs, and CRISPR-mediated methods may alter a GATA1 binding motif in BCL11Ae that is in the +58 DNase I hypersensitive site (DHS) region of intron 2 of the BCL11A gene ("GATA1 binding motif in BCL11Ae"). In certain embodiments, one or more gRNAs comprising a targeting domain set forth in SEQ ID NOs:952-955 may be used to introduce alterations in the GATA1 binding motif in BCL11Ae. In certain embodiments, one or more gRNAs may be used to introduce alterations in the GATA1 binding motif in BCL11Ae and one or more gRNAs may be used to introduce alterations in the 13 nt target region of HBG1 and/or HBG2. In certain embodiments, genome editing systems, gRNAs, and CRISPR-mediated methods may alter a region within 50, 100, 200, 300, 400, or 500, 600, 700, 800, 900 or 1000 bp of a proximal HBG1/2 promoter sequence, including without limitation the 13 nt target region ("proximal HBG1/2 promoter target sequence").

[0019] In certain embodiments, genome editing systems, gRNAs, and CRISPR-mediated methods set forth herein may alter one or more regions set forth in Table 13.

[0020] The inventors have also addressed a key unmet need in the field by identifying a strategy for increasing accessibility to the chromatin using an RNA-guided helicase and dead guide RNA to unwind the DNA within or proximal to the target region to be edited (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae). This disclosure provides new and effective means of unwinding chromatin and thereby increasing accessibility of target regions to RNA-guided nucleases. Also provided herein are genome editing systems, guide RNAs, and CRISPR-mediated methods for unwinding and altering portions of a genome. Unwinding of the genome may be achieved using an RNA-guided helicase and/or a dead guide RNA configured to target an RNA-guided enzyme to a target region in DNA but not to support a cleavage event.

[0021] In one aspect, the disclosure relates to genome editing systems that may include an RNA-guided nuclease, a first guide RNA and a second guide RNA. In certain embodiments, the first and second guide RNAs may include first and second targeting domains complimentary to first and second sequences on opposite sides of positions of a 13 nt target region of a human HBG1 or HBG2 gene. One or both of the first and second sequences may overlap the 13 nt target region of the human HBG1 or HBG2 gene. The genome editing system may also include a nucleic acid template encoding a deletion of the 13 nt region of the human HBG1 or HBG2 gene. In certain embodiments, the RNA-guided nuclease may be an *S. pyogenes* Cas9 or a nickase, which optionally lacks RuvC activity. The first and second targeting domains may be complimentary to sequences immediately adjacent to a protospacer adjacent motif recognized by *S. pyogenes* Cas9. In certain embodiments, the first targeting domain may be complimentary to a sequence within positions c. −1,114 to −114 of a human HBG1 or HBG2 gene. In certain embodiments, the first targeting domain may be complimentary to a sequence within positions c. −114 to 0 of a human HBG1 or HBG2 gene. In certain embodiments, at least one of the first and second targeting domains differ by no more than 3 nucleotides from a targeting domain listed in Table 7 or Table 12. The genome editing system may include first and second RNA-guided nucleases that, in some embodiments, are complexed with the first and second guide RNAs, respectively, forming first and second ribonucleoprotein complexes.

[0022] Continuing with this aspect of the disclosure, a genome editing system including any or all of the features described above may also include a third guide RNA, and optionally a fourth guide RNA. In certain embodiments, the third and fourth guide RNAs may include third and fourth targeting domains complimentary to third and fourth sequences on opposite sides of positions of a GATA1 binding motif in BCL11A erythroid enhancer (BCL11Ae) of a human BCL11A gene. One

or both of the third and fourth sequences may optionally overlap the GATA1 binding motif in BCL11Ae of the human BCL11A gene. The genome editing systems may also include a nucleic acid template encoding a deletion of the GATA1 binding motif in BCL11Ae. In certain embodiments, the RNA-guided nuclease may be an *S. pyogenes* Cas9. In certain embodiments, the third and fourth targeting domains may be complimentary to sequences immediately adjacent to a protospacer adjacent motif recognized by *S. pyogenes* Cas9. In certain embodiments, the RNA-guided nuclease may be a nickase, which optionally lacks RuvC activity. In certain embodiments, the third targeting domain may be complimentary to a sequence within 1000 nucleotides upstream of the GATA1 binding motif in BCL11Ae. In certain embodiments, the third targeting domain may be complimentary to a sequence within 100 nucleotides upstream of the GATA1 binding motif in BCL11Ae. In certain embodiments, one of the third and fourth targeting domains may be complimentary to a sequence within 100 nucleotides downstream of the GATA1 binding motif in BCL11Ae. In certain embodiments, the fourth targeting domain may be complimentary to a sequence within 50 nucleotides downstream of the GATA1 binding motif in BCL11Ae. In certain embodiments, at least one of the third and fourth targeting domains differ by no more than 3 nucleotides from a targeting domain listed in Table 9. In certain embodiments, the genome editing systems may further include first and second RNA-guided nucleases. In certain embodiments, the first and second RNA-guided nucleases may be complexed with the third and fourth guide RNAs, respectively, forming third and fourth ribonucleoprotein complexes.

[0023] Continuing with this aspect of the disclosure, a genome editing system including any or all of the features described above may also include an RNA-guided helicase. In certain embodiments, the RNA-guided helicase may unwind nucleic acid within or proximate to the 13 nt target region or GATA1 binding motif in BCL11Ae of the human BCL11A gene. In certain embodiments, the RNA-guided helicase may be a fifth RNA-guided nuclease configured to lack nuclease activity. In certain embodiments, the RNA-guided nuclease may be complexed to a dead guide RNA including a fifth targeting domain of 15 or fewer nucleotides in length. In certain embodiments, the RNA-guided nuclease and dead guide RNA are not configured to recruit an exogenous trans-acting factor to the target region. In certain embodiments, the fifth targeting domain may be complimentary to a fifth sequence within or proximate to the 13 nt target region or GATA1 binding motif in BCL11Ae of the human BCL11A gene. In certain embodiments, the fifth targeting domain may include a nucleotide sequence that is identical to, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from a nucleotide sequence set forth in Table 10. In certain embodiments, the fifth targeting domain may include a nucleotide sequence identical to the nucleotide sequence set forth in Table 10.

[0024] Another aspect of the disclosure relates to a method of altering a cell including contacting a cell with the genome editing systems described above and disclosed herein. In certain embodiments, the step of contacting the cell with the genome editing system may include contacting the cell with a solution including first and second ribonucleoprotein complexes. In certain embodiments, the step of contacting the cell with the solution may further include electroporating the cells, thereby introducing the first and second ribonucleoprotein complexes into the cell. In certain embodiments, the genome editing systems may further include contacting the cell with the genome editing system described above, in which the step of contacting the cell with the genome editing system may include contacting the cell with a solution including first, second, third, and optionally, fourth ribonucleoprotein complexes. In certain embodiments, the step of contacting the cell with the solution may further include electroporating the cells, thereby introducing the first, second, third, and optionally, fourth ribonucleoprotein complexes into the cell. In certain embodiments, the cell may be capable of differentiating into an erythroblast or a precursor of an erythroblast. In certain embodiments, the cell may be a CD34+ cell.

[0025] In one aspect, the disclosure relates to a CRISPR-mediated method of altering a cell including introducing a first DNA single strand break (SSB) or double strand break (DSB) within a genome of the cell between positions c. −614 to −102 of a human HBG1 or HBG2 gene and

introducing a second SSB or DSB within the genome of the cell between positions c. −114 to −1 of the human HBG1 or HBG2 gene. In certain embodiments, the first and second SSBs or DSBs may be repaired by the cell in a manner that alters a 13 nt target region of the human HBG1 or HBG2 gene. In certain embodiments, the first and second SSBs or DSBs may be repaired by the cell in a manner that results in the deletion of all or part of a 13 nt target region of the human HBG1 or HBG2 gene. In certain embodiments, the first and second SSBs or DSBs may be repaired by the cell in a manner that results in the formation of at least one of an indel, a deletion, or an insertion in the 13 nt target region of the human HBG1 or HBG2 gene. In certain embodiments, the first and second SSBs or DSBs may be repaired by the cell in an error prone manner. In certain embodiments, the CRISPR-mediated method may further include introducing a third DNA single strand break (SSB) or double strand break (DSB) within 500 nucleotides upstream of a GATA1 binding motif in BCL11Ae of a human BCL11A gene and introducing a fourth SSB or DSB within the genome of the cell within 100 nucleotides downstream of the GATA1 binding motif in BCL11Ae of the human BCL11A gene. In certain embodiments, the third and fourth SSBs or DSBs may be repaired by the cell in a manner that alters the GATA1 binding motif in BCL11Ae of the human BCL11A gene. In certain embodiments, the third and fourth SSBs or DSBs may be repaired by the cell in a manner that results in the deletion of all or part of the GATA1 binding motif in BCL11Ae. In certain embodiments, the third and fourth SSBs or DSBs may be repaired by the cell in a manner that results in the formation of at least one of an indel, a deletion, or an insertion in the GATA1 binding motif in BCL11Ae. In certain embodiments, the third and fourth SSBs or DSBs may be repaired by the cell in an error prone manner.

[0026] In another aspect, the disclosure relates to CRISPR-mediated methods of altering a cell including introducing a first DNA single strand break (SSB) or double strand break (DSB) within a region of a genome of the cell set forth in Table 13. In certain embodiments, the first SSB or DSB may be repaired by the cell in a manner that alters the regulation of an HBG1 gene or an HBG2 gene. In certain embodiments, the first SSB or DSB may be repaired by the cell in a manner that results in the formation of at least one indel, insertion or deletion in the region set forth in Table 13.

[0027] Another aspect relates to a method of modifying one or more regions of interest in a HBG gene in a population of HSCs, comprising contacting the populations of cells ex vivo, in vivo or in vitro with an RNP complex comprising: a gRNA molecule; and an RNA-guided nuclease in which the one or more RNP complexes alters the one or more regions of interest in the HBG gene, and the one or more regions of interest is selected from a sequence set forth in Table 13. In certain embodiments, the alteration may result in the formation of at least one indel (e.g., insertion or deletion) in Region 6 or Region 7 set forth in Table 13.

[0028] Another aspect of this disclosure relates to a genome editing system including an RNA-guided nuclease and at least one guide RNA comprising (a) a targeting domain differing by no more than 3 nucleotides from a sequence set forth in Table 12; or (b) a targeting domain consisting of, or consisting essentially of, positions 5-20 of a sequence set forth in Table 12. In certain embodiments, the genome editing system may be configured to provide an editing event within a region set forth in Table 13. In certain embodiments, the genome editing system may further include a nucleic acid template encoding an alteration of the region set forth in Table 13. In certain embodiments, the RNA-guided nuclease may be a nickase, and optionally may lack RuvC activity. In certain embodiments, the genome editing system may further include a second guide RNA.

[0029] In certain embodiments, the gRNAs described herein may include one or more 2-o-methyl modifications, one or more phosphorothioate modifications, or a combination thereof.

[0030] In another aspect, this disclosure relates to a genome editing system configured to alter, e.g., by forming an SSB, DSB and/or an indel within, a region set forth in Table 13. In another aspect, this disclosure relates to a genome editing system comprising an RNA-guided nuclease and at least one gRNA configured to provide an editing event within one or more regions set forth in Table 13. The region is, in certain embodiments, selected from: Chr 11 (NC_000011.10): 5,247,883-

5,248,186; 5,248,509-5,249,173; 5,249,198-5,249,362; 5,249,591-5,249,712; 5,249,904-5,249,927; 5,249,955-5,249,987, 5,250,040-5,250,075; 5,250,089-5,250,129; 5,250,141-5,250,254; 5,250,464-5,250,549; 5,250,594-5,250,735; 5,253,425-5,254,121; 5,254,122-5,254,306; 5,254,511-5,254,648; 5,254,829-5,254,866; 5,254,935-5,255,009; 5,255,025-5,255,053; 5,255,076-5,255,179; 5,255,255-5,255,292; and/or 5,255,518-5,255,641.

[0031] In certain embodiments, the genome editing systems include one or more guide RNAs comprising targeting domain sequences set forth in Table 12, and/or dead guide RNAs comprising nucleotide positions 5-20 of the targeting domain sequences set forth in Table 12. In certain embodiments, the genome editing systems include one or more guide RNAs comprising (a) a targeting domain differing by no more than 3 nucleotides from a sequence set forth in Table 12; or (b) a targeting domain consisting of, or consisting essentially of, positions 5-20 of a sequence set forth in Table 12. In certain embodiments, the genome editing systems include one or more guide RNAs comprising (a) the targeting domain differing by no more than 3 nucleotides from a Tier 1 or Tier 2 targeting domain sequence set forth in Table 12; or (b) the targeting domain consisting of, or consisting essentially of, positions 5-20 of a Tier 1 or Tier 2 targeting domain sequence set forth in Table 12. In certain embodiments, the genome editing system includes at least one guide RNA comprising (a) the targeting domain differing by no more than 3 nucleotides from a targeting domain sequence set forth in Table 17; or (b) the targeting domain consisting of, or consisting essentially of, positions 5-20 of a targeting domain sequence set forth in Table 17. In certain embodiments, the genome editing system includes at least one guide RNA comprising a gRNA sequence set forth in Table 17. In certain embodiments, the gRNA comprises one or more 2-o-methyl modifications, one or more phosphorothioate modifications, or a combination thereof. In certain embodiments, the RNA-guided nuclease may be a Cas9 molecule. In certain embodiments, the RNA-guided nuclease may be a nickase, and optionally lacks RuvC activity. In certain embodiments, the genome editing system further comprises a second guide RNA.

[0032] In one aspect, the disclosure relates to compositions including a plurality of cells generated by the method disclosed above, in which at least 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% of the cells include an alteration of a sequence of a 13 nt target region of the human HBG1 or HBG2 gene or a plurality of cells generated by the method disclosed above, wherein at least 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% of the cells include an alteration of a sequence of a 13 nt target region of the human HBG1 or HBG2 gene and at least 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% of the cells include an alteration of a sequence of the GATA1 binding motif in BCL11Ae. In one aspect, the disclosure relates to compositions including a plurality of cells generated by the genome editing systems disclosed above, in which at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the cells comprise one or more alterations of one or more sequences set forth in Table 13. In certain embodiments, the alteration may result in the formation of at least one indel (e.g., insertion or deletion) in Region 6 or Region 7 as set forth in Table 13. In certain embodiments, at least a portion of the plurality of cells may be within an erythroid lineage. In certain embodiments, the plurality of cells may be characterized by an increased level of fetal hemoglobin expression relative to an unmodified plurality of cells. In certain embodiments, the level of fetal hemoglobin may be increased by at least 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90%. In certain embodiments, the compositions may further include a pharmaceutically acceptable carrier.

[0033] In one aspect, the disclosure relates to compositions including a population of cells generated by any one of the methods disclosed herein, wherein the cells comprise a higher frequency of an alteration of a sequence of an HBG1 gene, HBG2 gene, or both set forth in the region relative to an unmodified population of cells. In certain embodiments, the higher frequency may be at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% higher. In certain embodiments, at least a portion of the population of cells are within an erythroid lineage.

[0034] In one aspect, the disclosure relates to a population of cells modified by the genome editing system disclosed above, in which the population of cells comprise a higher percentage of a

productive indel relative to a population of cells not modified by the genome editing system.
[0035] In one aspect, the disclosure relates to an isolated cell comprising a modification in an HBG gene sequence generated by the delivery of a RNP complex comprising an RNA-guided nuclease and a synthetic gRNA molecule that targets the HBG gene sequence within one or more regions selected from Chr 11 (NC_000011.10): 5,247,883-5,248,186; 5,248,509-5,249,173; 5,249,198-5,249,362; 5,249,591-5,249,712; 5,249,904-5,249,927; 5,249,955-5,249,987, 5,250,040-5,250,075; 5,250,089-5,250,129; 5,250,141-5,250,254; 5,250,464-5,250,549; 5,250,594-5,250,735; 5,253,425-5,254,121; 5,254,122-5,254,306; 5,254,511-5,254,648; 5,254,829-5,254,866; 5,254,935-5,255,009; 5,255,025-5,255,053; 5,255,076-5,255,179; 5,255,255-5,255,292; and/or 5,255,518-5,255,641.
[0036] In one aspect, the disclosure relates to a population of CD34+ cells or hematopoietic stem cells (HSCs), with one or more cells comprising a disruption in one or more regions of the HBG gene, wherein the disruption is generated using an RNP complex comprising a CRISPR/RNA-guided nuclease and a synthetic gRNA that targets one or more regions of the HBG gene selected from Chr 11 (NC_000011.10): 5,247,883-5,248,186; 5,248,509-5,249,173; 5,249,198-5,249,362; 5,249,591-5,249,712; 5,249,904-5,249,927; 5,249,955-5,249,987, 5,250,040-5,250,075; 5,250,089-5,250,129; 5,250,141-5,250,254; 5,250,464-5,250,549; 5,250,594-5,250,735; 5,253,425-5,254,121; 5,254,122-5,254,306; 5,254,511-5,254,648; 5,254,829-5,254,866; 5,254,935-5,255,009; 5,255,025-5,255,053; 5,255,076-5,255,179; 5,255,255-5,255,292; and/or 5,255,518-5,255,641.
[0037] In one aspect, the disclosure relates to a population of cells modified by the genome editing system disclosed above, in which a higher percentage of the population of cells are capable of differentiating into a population of cells of an erythroid lineage that express HbF relative to a population of cells not modified by the genome editing system. In certain embodiments, the higher percentage may be at least about 15%, at least about 20%, at least about 25%, at least about 30%, or at least about 40% higher. In certain embodiments, the cells may be hematopoietic stem cells. In certain embodiments, the cells may be capable of differentiating into an erythroblast, erythrocyte, or a precursor of an erythrocyte or erythroblast.
[0038] In one aspect, the disclosure relates to methods for treating a subject with a hemoglobinopathy, the method including the steps of: isolating a hematopoietic progenitor cell from the subject providing a patient specific HSC; contacting the cell with a genome editing system disclosed herein; and implanting the cell into the subject.
[0039] In one aspect, the disclosure relates to methods of administering a population of cells to a subject, wherein the population of cells comprises an alteration in an HBG gene sequence generated by the delivery of a complex comprising an RNA-guided nuclease and a gRNA molecule that alters one or more regions of the HBG gene selected from Chr 11 (NC_000011.10): 5,247,883-5,248,186; 5,248,509-5,249,173; 5,249,198-5,249,362; 5,249,591-5,249,712; 5,249,904-5,249,927; 5,249,955-5,249,987, 5,250,040-5,250,075; 5,250,089-5,250,129; 5,250,141-5,250,254; 5,250,464-5,250,549; 5,250,594-5,250,735; 5,253,425-5,254,121; 5,254,122-5,254,306; 5,254,511-5,254,648; 5,254,829-5,254,866; 5,254,935-5,255,009; 5,255,025-5,255,053; 5,255,076-5,255,179; 5,255,255-5,255,292; and/or 5,255,518-5,255,641.
[0040] In one aspect, the disclosure relates to methods for increasing the level of fetal hemoglobin (HbF) in a human cell, the method comprising contacting the cell with: an RNA-guided nuclease; and at least one guide RNA configured to provide an editing event within one or more regions selected from Chr 11 (NC_000011.10): 5,247,883-5,248,186; 5,248,509-5,249,173; 5,249,198-5,249,362; 5,249,591-5,249,712; 5,249,904-5,249,927; 5,249,955-5,249,987, 5,250,040-5,250,075; 5,250,089-5,250,129; 5,250,141-5,250,254; 5,250,464-5,250,549; 5,250,594-5,250,735; 5,253,425-5,254,121; 5,254,122-5,254,306; 5,254,511-5,254,648; 5,254,829-5,254,866; 5,254,935-5,255,009; 5,255,025-5,255,053; 5,255,076-5,255,179; 5,255,255-5,255,292; and/or 5,255,518-5,255,641. In certain embodiments, the guide RNAs may comprise a targeting domain sequence set forth in Table 12. In certain embodiments, the gRNA may comprise (a) a targeting domain differing by no more than 3 nucleotides from a sequence set forth in Table 12; or (b) a targeting domain consisting of, or

consisting essentially of, positions 5-20 of a sequence set forth in Table 12. In certain embodiments, the gRNA may comprise (a) the targeting domain differing by no more than 3 nucleotides from a Tier 1 or Tier 2 targeting domain sequence set forth in Table 12; or (b) the targeting domain consisting of, or consisting essentially of, positions 5-20 of a Tier 1 or Tier 2 targeting domain sequence set forth in Table 12. In certain embodiments, the gRNA may comprise (a) the targeting domain differing by no more than 3 nucleotides from a targeting domain sequence set forth in Table 17; or (b) the targeting domain consisting of, or consisting essentially of, positions 5-20 of a targeting domain sequence set forth in Table 17. In certain embodiments, the gRNA may comprise a gRNA sequence set forth in Table 17. In certain embodiments, the gRNA may comprise one or more 2-o-methyl modifications, one or more phosphorothioate modifications, or a combination thereof.

[0041] In one aspect, the disclosure relates to a synthetic guide RNA molecule comprising (a) a targeting domain differing by no more than 3 nucleotides from a sequence set forth in Table 12; or (b) a targeting domain consisting of, or consisting essentially of, positions 5-20 of a sequence set forth in Table 12. In certain embodiments, a synthetic gRNA molecule further comprises (a) the targeting domain differing by no more than 3 nucleotides from a Tier 1 or Tier 2 targeting domain sequence set forth in Table 12; or (b) the targeting domain consisting of, or consisting essentially of, positions 5-20 of a Tier 1 or Tier 2 targeting domain sequence set forth in Table 12. In certain embodiments, the targeting domain may comprise a Tier 1 or Tier 2 gRNA set forth in Table 12. In certain embodiments, the targeting domain may differ by no more than 3 nucleotides from a targeting domain sequence set forth in Table 17; or the targeting domain may consists of, or may consists essentially of, positions 5-20 of a targeting domain sequence set forth in Table 17. In certain embodiments, the targeting domain may comprise a gRNA targeting domain sequence set forth in Table 17. In certain embodiments, a synthetic gRNA may comprise one or more 2-o-methyl modifications, one or more phosphorothioate modifications, or a combination thereof.

[0042] This listing is intended to be exemplary and illustrative rather than comprehensive and limiting. Additional aspects and embodiments may be set out in, or apparent from, the remainder of this disclosure and the claims.

Description

BRIEF DESCRIPTION OF THE DRAWINGS

[0043] The accompanying drawings are intended to provide illustrative, and schematic rather than comprehensive, examples of certain aspects and embodiments of the present disclosure. The drawings are not intended to be limiting or binding to any particular theory or model, and are not necessarily to scale. Without limiting the foregoing, nucleic acids and polypeptides may be depicted as linear sequences, or as schematic two- or three-dimensional structures; these depictions are intended to be illustrative rather than limiting or binding to any particular model or theory regarding their structure.

[0044] FIG. **1** depicts, in schematic form, HBG1 and HBG2 gene(s) in the context of the γ-globin gene cluster on human chromosome 11. FIG. **1**. Each gene in the γ-globin gene cluster is transcriptionally regulated by a proximal promoter. While not wishing to be bound by any particular theory, it is generally thought that A.sub.γ and/or G.sub.γ expression is activated by engagement between the proximal promoter with the distal strong erythroid-specific enhancer, the locus control region (LCR). Long-range transactivation by the LCR is thought to be mediated by alteration of chromatin configuration/confirmation. The LCR is marked by 4 erythroid specific DNase I hypersensitive sites (HS1-4) and 2 distal enhancer elements (5′ HS and 3′ HS1). γ-like gene globin gene expression is regulated in a developmental stage-specific manner, and expression of globin genes changes coincide with changes in the main site of blood production.

[0045] FIGS. **2**A-**2**B depict HBG1 and HBG2 genes, coding sequences (CDS) and small deletions and point mutations in and upstream of the HBG1 and HBG2 proximal promoters that have been identified in patients and associated with elevation of fetal hemoglobin (HbF). Core elements within the proximal promoters (CAAT box, 13 nt sequence) that have been deleted in some patients with hereditary persistence of fetal hemoglobin (HPFH). The 'target sequence' region of each locus, which has been screened for gRNA binding target sites, is also identified.
[0046] FIGS. **3**A-C show data from gRNA screening for incorporation of the 13 nt deletion in human K562 erythroleukemia cells. FIG. **3**A Gene editing as determined by T7E1 endonuclease assay analysis (referred to interchangeably as a "T7E1 analysis") of HBG1 and HBG2 locus-specific PCR products amplified from genomic DNA extracted from K562 cells after electroporation with DNA encoding *S. pyogenes*-specific gRNAs and plasmid DNA encoding *S. pyogenes* Cas9. FIG. **3**B Gene editing as determined by DNA sequence analysis of PCR products amplified from the HBG1 locus in genomic DNA extracted from K562 cells after electroporation with DNA encoding the indicated gRNA and Cas9 plasmid. FIG. **3**C Gene editing as determined by DNA sequence analysis of PCR products amplified from the HBG2 locus in genomic DNA extracted from K562 cells after electroporation with DNA encoding the indicated gRNA and Cas9 plasmid. For FIG. **3**B-C, the types of editing events (insertions, deletions) and subtypes of deletions (13 nt target partially [12 nt HPFH] or fully [13-26 nt HPFH] deleted, other sequences deleted [other deletions]) are indicated by the differently shaded/patterned bars.
[0047] FIGS. **4**A-C depict results of gene editing in human cord blood (CB) and human adult CD34.sup.+ cells after electroporation with RNPs complexed to in vitro transcribed *S. pyogenes* gRNAs that target a specific 13 nt sequence for deletion (HBG sgRNAs Sp35 and Sp37). FIG. **4**A depicts the percentage of indels detected by T7E1 analysis of HBG1 and HBG2 specific PCR products amplified from gDNA extracted from CB CD34.sup.+ cells treated with the indicated RNPs or donor matched untreated control cells (n=3 CB CD34.sup.+ cells, 3 separate experiments). Data shown represent the mean and error bars correspond to standard deviation across the 3 separate donors/experiments. FIG. **4**B depicts the percentage of indels detected by T7E1 analysis of HBG2 specific PCR product amplified from gDNA extracted from CB CD34.sup.+ cells or adult CD34.sup.+ cells treated with the indicated RNPs or donor matched untreated control cells (n=3 CB CD34.sup.+ cells, n=3 mobilized peripheral blood (mPB) CD34.sup.+ cells, 3 separate experiments). Data shown represent the mean and error bars correspond to standard deviation across the 3 separate donors/experiments. FIG. **4**C (Top panel) depicts indels as detected by T7E1 analysis of HBG2 PCR products amplified from gDNA extracted from human CB CD34.sup.+ cells electroporated with HBG Sp35 RNP or HBG Sp37 RNP+/−ssODN (unmodified or with PhTx modified 5′ and 3′ ends). The lower left panel shows the level of gene editing as determined by Sanger DNA sequence analysis of gDNA from cells edited with HBF Sp37 RNP and ssODN. The lower right panel shows the specific types of deletions detected within total deletions.
[0048] FIGS. **5**A-B depict gene editing of HBG in adult human mobilized peripheral blood (mPB) CD34.sup.+ cells and induction of fetal hemoglobin in erythroid progeny of RNP treated cells after electroporation of mPB CD34.sup.+ cells with HBG Sp37 RNP+/−ssODN encoding the 13 nt deletion. FIG. **5**A depicts the percentage of indels detected by T7E1 analysis of HBG2 PCR product amplified from gDNA extracted from mPB CD34.sup.+ cells treated with the RNP or donor matched untreated control cells. FIG. **5**B depicts the fold change in HBG mRNA expression in day 7 erythroblasts that were differentiated from RNP treated and untreated donor matched control mPB CD34.sup.+ cells. mRNA levels are normalized to GAPDH and calibrated to the levels detected in untreated controls on the corresponding days of differentiation.
[0049] FIGS. **6**A-B depict the ex vivo differentiation potential of RNP treated and untreated mPB CD34.sup.+ cells from the same donor. FIG. **6**A shows hematopoietic myeloid/erythroid colony forming cell (CFC) potential, where the number and subtype of colonies are indicated (GEMM: granulocyte-erythroid-monocyte-macrophage colony, E: erythroid colony, GM: granulocyte-

macrophage colony, M: macrophage colony, G: granulocyte colony). FIG. **6**B depicts the percentage of Glycophorin A expressed over the time course of erythroid differentiation as determined by flow cytometry analysis at the indicated time points and for the indicated samples.
[0050] FIG. **7**A depicts indels detected by T7E1 analysis of HBG PCR product amplified from gDNA extracted from human mPB CD34.sup.+ cells treated with HBG RNPs (D10A paired nickases). For a subset of samples, cells also received ssODN encoding the 13 nt deletion plus silent SNPs to monitor for HDR (ssODN). FIG. **7**B depicts DNA sequencing analysis for select subset of samples shown in FIG. **7**A. The indels were subdivided according to the type of indel (insertion, 13 nt deletion, or other deletion).
[0051] FIG. **8**A depicts the indels at the HBG target site after electroporation of mPB CD34.sup.+ cells with the indicated pairs of gRNAs complexed in D10A nickase and WT RNP pairs. FIG. **8**B depicts the large deletion events (e.g., deletion of HBG2) after electroporation of mPB CD34.sup.+ cells with the indicated pairs of gRNAs complexed in D10A nickase and WT RNPs. FIG. **8**C depicts DNA sequencing analysis and the subtypes of events (insertions, deletions) detected in gDNA from mPB CD34.sup.+ cells treated with paired D10A nickase pairs. FIG. **8**D depicts DNA sequencing analysis and the subtypes of events (insertions, deletions) detected in gDNA from mPB CD34.sup.+ cells treated with paired WT RNP pairs.
[0052] FIG. **9** depicts the summary of HbF protein and mRNA expression in the progeny of mPB CD34.sup.+ cells treated with paired RNPs targeting HBG, for the experiments shown in FIGS. **7** and **8**. HbF protein (by HPLC analysis) and HbF mRNA expression (ddPCR analysis) were evaluated in erythroid progeny of RNP treated human mPB CD34.sup.+ cells (background levels of HbF detected in donor matched untreated controls were subtracted from the levels detected in progeny of RNP treated CD34.sup.+ cells).
[0053] FIGS. **10**A-H depict the indel frequencies and ex vivo and in vivo short-term hematopoietic potential of CD34.sup.+ cells after treatment with different concentrations (0, 2.5, 3.75 M) of paired D10A nickase RNPs (SpA+Sp85). Indels were evaluated by T7E1 analysis (FIG. **10**A) and by Illumina sequencing analysis (insertions and deletions, FIG. **10**B). FIG. **10**C depicts the % of HbF protein detected by HPLC analysis (% HbF=100%×HbF/(HbF+HbA). FIG. **10**D depicts the hematopoietic activity of the RNP treated and donor matched untreated control CD34.sup.+ cells in colony forming cell (CFC) assays. CFCs shown are per thousand CD34.sup.+ cells plated. FIG. **10**E depicts human blood CD45.sup.+ cell reconstitution of the peripheral blood in immunodeficient mice (NSG) 1 month after transplantation with donor matched human mPB CD34.sup.+ that were either untreated (O M), or treated with one of two doses (2.5 and 3.75 M) of D10A RNP and paired gRNAs. FIG. **10**F depicts human blood CD45.sup.+ cell reconstitution of the peripheral blood in immunodeficient mice (NSG) 2 months after transplantation. FIGS. **10**G and **10**H depict the lineage distributions following human CD45.sup.+ blood cell reconstitution of NSG mice at 1 month (FIG. **10**G) and 2 months (FIG. **10**H).
[0054] FIG. **11***a* correlates HbF levels as assayed by HPLC and indel frequency as assessed by T7E1 analysis for two D10A nickase RNP pairs (SP37+SPB and SP37+SPA) delivered at the indicated concentrations to mPB CD34.sup.+ cells. HbF levels were analyzed in erythroid progeny (day 18) of edited CD34.sup.+ cells. HbF protein detected in donor-matched untreated controls were subtracted from edited samples. FIG. **11***b* depicts indel rates overlaid on hematopoietic colony forming cell (CFC) activity associated with CD34.sup.+ cells treated with the indicated D10A nickase pairs or untreated controls. FIG. **11***c* depicts human CD45.sup.+ blood cell reconstitution of immunodeficient NSG mice one month after transplantation of mPB CD34.sup.+ cells treated with indicated D10 RNP nickase pairs at the concentrations given or donor matched untreated controls. FIG. **11***d* depicts the human blood lineage distribution detected in the human CD45.sup.+ fraction in mouse peripheral blood one month post-transplant.
[0055] FIG. **12** depicts a target site for derepression of HbF, the GATA1 motif of the +58 DNase I hypersensitive site (DHS) erythroid specific enhancer of BCL11A (BCL11Ae) (genomic

coordinates: chr2: 60,495,265 to 60,495,270).

[0056] FIG. **13**A depicts the percentage of indels detected by T7E1 endonuclease analysis of BCL11A PCR products amplified from gDNA extracted from CB CD34.sup.+ cells treated with the indicated RNP+/−ssODN or donor matched untreated control cells. Data shown represent the mean of three 3 separate donors/experiments. FIG. **13**B depicts indels detected by T7E1 endonuclease analysis of BCL11A PCR products amplified from gDNA extracted from CB CD34.sup.+ cells treated with the indicated WT RNP (single gRNA targeting the BCL11A erythroid enhancer complexed to WT *S. pyogenes* Cas9 having both RuvC and HNH activity) or paired nickase RNP (paired gRNAs targeting the BCL11A erythroid enhancer (BCL11Ae) complexed to *S. pyogenes* Cas9 nickases sharing the same HNH single stranded cutting activity (e.g., D10A), as well as the hematopoietic activity of cells in each condition.

[0057] FIG. **14**A depicts the editing frequency of BCL11Ae (using single gRNA approach targeting the GATA1 motif) in adult human BM CD34.sup.+ cells. FIG. **14**B depicts the monoallelic and bialleleic editing detected in hematopoietic colonies (GEMMs, clonal progeny of BCL11Ae RNP treated CD34.sup.+ cells) as determined by DNA sequencing analysis. FIG. **14**C depicts the kinetics of erythroblast maturation (enucleation as determined by DRAQ5.sup.− cells detected by flow cytometry analysis). FIG. **14**D depicts the acquisition of erythroid phenotype (Glycophorin A.sup.+ cells) in differentiated control and RNP-treated BM CD34.sup.+ cells, while FIG. **14**E shows the fold increase in HbF.sup.+ cells as determined by flow cytometry analysis relative to HbF+ cells in untreated donor matched control samples.

[0058] FIGS. **15**A-C depict gene editing of BCL11Ae in adult human mPB CD34.sup.+ cells and induction of fetal hemoglobin in erythroid progeny of RNP and ssODN treated cells after electroporation of mPB CD34.sup.+ cells with BCL11Ae RNP+nonspecific ssODN (i.e., no homology to BCL11Ae target region). FIG. **15**A depicts the percentage of indels detected by T7E1 analysis of HBG2 PCR product amplified from gDNA extracted from mPB CD34.sup.+ cells treated with the BCL11Ae RNP and nonspecific ssODN or donor matched untreated control cells. FIG. **15**B depicts the fold change in HBG mRNA expression in day 10 erythroblasts that were differentiated from BCL11Ae RNP treated and untreated donor matched control mPB CD34.sup.+ cells (mRNA levels are normalized to GAPDH and calibrated to the levels detected in untreated controls on the corresponding days of differentiation). FIG. **15**C depicts the percentage of Glycophorin A expressed over the time course of erythroid differentiation of mPB CD34.sup.+ cells+/−treatment with BCL11Ae RNP and nonspecific ssODN, as determined by flow cytometry analysis at the indicated time points and for the indicated samples.

[0059] FIG. **16** depicts a schematic of the −110 nt target region in the gamma hemoglobin gene (HBG) promoter (grey box) and the relative locations of homologous sequences to dead gRNAs (dgRNAs) and wild-type gRNAs. dgRNAs that have a truncated targeting domain sequence and do not promote Cas9 cutting are depicted (i.e., Sp181 dgRNA and truncated (t)SpA dgRNA, Table 10) as white arrows. gRNAs that have a full-length targeting domain sequence, which promote Cas9 cutting are depicted as black arrows (i.e., Sp35 and Sp37 gRNAs, Table 10).

[0060] FIG. **17** shows the percentage of edits determined by T7E1 endonuclease analysis of HBG2 PCR product amplified from genomic DNA (gDNA) extracted from mobilized peripheral blood (mPB) CD34.sup.+ cells after codelivery of a dead ribonucleoprotein (dRNP) (i.e., SpA dRNP) and a wild-type (WT) RNP (i.e., Sp37 RNP). tSpA dRNP comprises WT Cas9 protein complexed to a truncated gRNA (tSpA dgRNA, Table 10) (i.e., dead (d)RNA15-mer version of SpA) and Sp37 RNP comprises WT Cas9 protein complexed to gRNA Sp37 (Table 10).

[0061] FIG. **18** depicts the percentage of edits detected by T7E1 analysis of HBG PCR product amplified from gDNA extracted from mPB CD34.sup.+ cells after delivery of Sp35 RNP alone (i.e., Sp35 gRNA complexed with WT Cas 9 protein)) or codelivery of Sp35 RNP and dRNPs that target the same or opposite DNA strand as Sp35 RNP (i.e., Sp181 dRNP (Sp181 dgRNA complexed with WT Cas9 protein) and tSpA dRNP (tSpA dgRNA complexed with WT Cas9

protein)) (see also FIG. **16**). Black bars indicate the level of indels detected in the mPB CD34.sup.+ cells. White bars indicate the level of indels detected in the mPB CD34.sup.+ cells maintained in the day 7 erythroid progeny of edited cells.

[0062] FIG. **19** shows a schematic of the variety of insertions and deletions resulting from double strand breaks repaired through NHEJ. Each unique edit (e.g., insertion or deletion) may serve as a unique identifier (or "barcode") for an individual cell or clone of cells descended therefrom.

[0063] FIG. **20** depicts a graphical rank ordering of the most abundant edited alleles in pre-infusion human HSCs and in lineages or tissue populations derived from long-term engrafting cells from two experimental replicates at 16-weeks post-infusion. Genomic DNA from cells electroporated with a ribonucleoprotein complex targeting the HBB locus was harvested and sequencing reads were aligned to an unedited or WT reference sequence. The frequency of individual edited alleles among the total number of reads from each sample was quantified and ranked. White and grey bars represent to five most abundant unique alleles in each sample, with white bars representing the most abundant unique allele, and less frequent alleles being represented by progressively darker shades of grey. Black bars represent unique alleles outside of the top 5 in terms of frequency. These data indicate that the most frequent alleles in each sample represent a comparatively small fraction of the total reads, and that the distribution of reads varies across lineages or tissue populations derived from the same pre-infusion pool, indicating that diversity of edited alleles is preserved in long-term engrafting HSCs and their progeny.

[0064] FIG. **21** depicts a graphical rank-ordering of the abundance of edited alleles in pre-infusion human HSCs and in lineages or tissue populations derived from long-term engrafting cells in two experimental replicates at 16-weeks post-infusion. Editing and analysis were performed as described for FIG. **20**, but the white bars correspond to the edited allele observed at the highest frequency in the pre-infusion edited HSC sample, and progressively darker bars correspond to less frequently observed alleles in the pre-infusion sample. Bars of the same color represent the same edited allele in each sample. Black bars represent unique alleles outside of the top 5 in any of the samples shown. The figure indicates that the frequency of individual alleles in tissue populations or lineages derived from long-term engrafting HSCs varies from the frequency of the same alleles in pre-infusion samples, consistent with the relatively low level of representation of long-term engrafting HSCs in the bulk CD34+ cell population.

[0065] FIG. **22** depicts, in schematic form, the genomic region encompassing the beta globin locus on human chromosome 11 that was screened to identify cis-regulatory elements involved in the regulation of fetal globin expression. The bottom panel depicts the coverage of gRNA library where each black vertical line represents one gRNA.

[0066] FIG. **23** depicts the average enrichment in the pool of high-HbF expressing cells (over low HbF-expressing cells) of the lentiviral sequence encoding gRNAs classified as Tier 1, Tier 2, Tier 3, Tier 4, and Friend of Tier 1 (as determined by sequencing analysis of the lentiviral PCR amplicon from the cell pool gDNA extracts) (Table 12). The Y axis shows the average Log 2 enrichment value from four bioreplicates. The X axis shows the standard deviation of average enrichment. Each dot represents one gRNA.

[0067] FIG. **24** depicts the average enrichment in the pool of high-HbF expressing cells (over low HbF-expressing cells) of the lentiviral sequence encoding the gRNAs included in the screen. Each dot represents one gRNA, positioned on the X axis according to the genomic coordinate (Hg38) of its cut site (Hg38).

[0068] FIGS. **25**A-B depicts the positions, lengths, frequencies, and HbF enrichment scores of individual indels generated by gRNAs from Table 12. The HbF enrichment score represents the ratio of the frequency of an indel in the pool of high-HbF expressing cells over its frequency in the pool of low HbF-expressing cells (as determined by sequencing analysis of the HBG1, HBG2, or HBG1-2 PCR amplicon from the pool gDNA extracts). The genomic coordinates (Hg38) of the center of the indels are indicated by the value on the X-axis. FIG. **25**A depicts the indels mapped to

HBG1 (or HBG1 and HBG2 where the sequence is perfectly homologous). FIG. **25**B similarly depicts the indels mapped to HBG2 (or HBG1 and HBG2 where the sequence is perfectly homologous). For FIGS. **25**A-B, the Y-axis of the top panels depict the length of the indels where a positive number indicates an insertion and a negative number indicates a deletion. The Y-axis on the bottom panel depicts the average Log 2 HbF enrichment score (average of two biological replicates). Each dot represents one unique indel. The size of the dot represents the average frequency of the indel (average of two biological replicates). Indels enriched in the high-HbF expressing fraction are represented as black dots, other indels are represented as light grey dots.

[0069] FIG. **26** depicts the coverage by high HbF-enriched indels and non-enriched indel at each genomic position at Hg38 Chr11: 5,249,805-5,250,352. The coverage of genomic positions by high HbF-enriched indels is shown as dark grey and the non-enriched indels are shown as light grey (see Example 10). Briefly, gRNA were complexed as RNP and delivered to HUDEP-2 cells by electroporation. Following erythroid differentiation, High HbF expressing cells and low HbF expressing cells were separated by FACS (Fluorescence activated cell sorting). Following genomic DNA extraction, PCR amplification of the target regions and NGS sequencing, the count of indels enriched or not-enriched in the high-HbF fraction (over the low HbF fraction) was calculated for each genomic position. FIG. **26** shows an example of one region analyzed. Only indels with spanning 10 or less nucleotides were included in the analysis. Several genomic position covered by high frequency of HbF-enriched indels were identified.

[0070] FIG. **27** depicts the average relative fold change in gamma globin mRNA expression as measured by qRT-PCR (Y-axis) following RNP transfection of HUDEP2 cells, plotted against the average Log 2 HbF enrichment score following lentiviral transduction of HUDEP2 cells. Each dot represents one gRNA.

[0071] FIGS. **28**A-C depict the engraftment outcomes of mock-transfected (no gRNA) or RNP #3 (comprising gRNA #3 targeting domain (SEQ ID NO:295, Table 18), complexed with *S. Pyogenes* wild-type Cas9)-transfected mPB CD34+ cells. FIG. **28**A depicts the frequency of individual populations of CD19+, CD15+, GlyA+, and Lin-CD34+ cells (lineage cocktail includes antibodies against CD3, CD14, CD16, CD19, CD20, CD56 markers, Lin-CD34+ cells are defined as CD34+ cells that are negative for CD3, CD14, CD16, CD20, or CD56 marker expression) from bone marrow (BM) of nonirradiated NOD,B6.SCID Il2rγ−/− Kit(W41/W41) ("NBSGW") mice infused with mock (no gRNA) or RNP #3 transfected mPB CD34+ cells. Chimerism is defined as human CD45/(human CD45+mCD45). The frequency of GlyA+ cells was calculated as GlyA+ cells/total cells. All other markers were calculated as marker+ cells/human CD45+ cells. FIG. **28**B depicts the indels of unfractionated BM or flow-sorted individual populations of CD15+, CD19+, GlyA+, and Lin-CD34+ cells in mock-transfected (no gRNA added) or RNP #3 transfected cells. FIG. **28**C depicts the HbF expression, calculated as gamma/beta-like (%) by erythroid cells following an 18-day erythroid differentiation culture from total BM.

DETAILED DESCRIPTION

Definitions and Abbreviations

[0072] Unless otherwise specified, each of the following terms has the meaning associated with it in this section.

[0073] The indefinite articles "a" and "an" refer to at least one of the associated noun, and are used interchangeably with the terms "at least one" and "one or more." For example, "a module" means at least one module, or one or more modules.

[0074] The conjunctions "or" and "and/or" are used interchangeably as non-exclusive disjunctions.

[0075] "Domain" is used to describe a segment of a protein or nucleic acid. Unless otherwise indicated, a domain is not required to have any specific functional property.

[0076] The term "exogenous trans-acting factor" refers to any peptide or nucleotide component of a genome editing system that both (a) interacts with an RNA-guided nuclease or gRNA by means of a modification, such as a peptide or nucleotide insertion or fusion, to the RNA-guided nuclease or

gRNA, and (b) interacts with a target DNA to alter a helical structure thereof. Peptide or nucleotide insertions or fusions may include, without limitation, direct covalent linkages between the RNA-guided nuclease or gRNA and the exogenous trans-acting factor, and/or non-covalent linkages mediated by the insertion or fusion of RNA/protein interaction domains such as MS2 loops and protein/protein interaction domains such as a PDZ, Lim or SH1, 2 or 3 domains. Other specific RNA and amino acid interaction motifs will be familiar to those of skill in the art. Trans-acting factors may include, generally, transcriptional activators.

[0077] An “indel” is an insertion and/or deletion in a nucleic acid sequence. An indel may be the product of the repair of a DNA double strand break, such as a double strand break formed by a genome editing system of the present disclosure. An indel is most commonly formed when a break is repaired by an “error prone” repair pathway such as the NHEJ pathway described below.

[0078] “Gene conversion” refers to the alteration of a DNA sequence by incorporation of an endogenous homologous sequence (e.g., a homologous sequence within a gene array). “Gene correction” refers to the alteration of a DNA sequence by incorporation of an exogenous homologous sequence, such as an exogenous single-or double stranded donor template DNA. Gene conversion and gene correction are products of the repair of DNA double-strand breaks by HDR pathways such as those described below.

[0079] Indels, gene conversion, gene correction, and other genome editing outcomes are typically assessed by sequencing (most commonly by “next-gen” or “sequencing-by-synthesis” methods, though Sanger sequencing may still be used) and are quantified by the relative frequency of numerical changes (e.g., ±1, ±2 or more bases) at a site of interest among all sequencing reads. DNA samples for sequencing may be prepared by a variety of methods known in the art, and may involve the amplification of sites of interest by polymerase chain reaction (PCR), the capture of DNA ends generated by double strand breaks, as in the GUIDEseq process described in Tsai 2016 (incorporated by reference herein) or by other means well known in the art. Genome editing outcomes may also be assessed by in situ hybridization methods such as the FiberComb™ system commercialized by Genomic Vision (Bagneux, France), and by any other suitable methods known in the art.

[0080] “Alt-HDR,” “alternative homology-directed repair,” or “alternative HDR” are used interchangeably to refer to the process of repairing DNA damage using a homologous nucleic acid (e.g., an endogenous homologous sequence, e.g., a sister chromatid, or an exogenous nucleic acid, e.g., a template nucleic acid). Alt-HDR is distinct from canonical HDR in that the process utilizes different pathways from canonical HDR, and can be inhibited by the canonical HDR mediators, RAD51 and BRCA2. Alt-HDR is also distinguished by the involvement of a single-stranded or nicked homologous nucleic acid template, whereas canonical HDR generally involves a double-stranded homologous template.

[0081] “Canonical HDR,” “canonical homology-directed repair” or “cHDR” refer to the process of repairing DNA damage using a homologous nucleic acid (e.g., an endogenous homologous sequence, e.g., a sister chromatid, or an exogenous nucleic acid, e.g., a template nucleic acid). Canonical HDR typically acts when there has been significant resection at the double strand break, forming at least one single stranded portion of DNA. In a normal cell, cHDR typically involves a series of steps such as recognition of the break, stabilization of the break, resection, stabilization of single stranded DNA, formation of a DNA crossover intermediate, resolution of the crossover intermediate, and ligation. The process requires RAD51 and BRCA2, and the homologous nucleic acid is typically double-stranded.

[0082] Unless indicated otherwise, the term “HDR” as used herein encompasses both canonical HDR and alt-HDR.

[0083] “Non-homologous end joining” or “NHEJ” refers to ligation mediated repair and/or non-template mediated repair including canonical NHEJ (cNHEJ) and alternative NHEJ (altNHEJ), which in turn includes microhomology-mediated end joining (MMEJ), single-strand annealing

(SSA), and synthesis-dependent microhomology-mediated end joining (SD-MMEJ).
[0084] “Replacement” or “replaced,” when used with reference to a modification of a molecule (e.g., a nucleic acid or protein), does not require a process limitation but merely indicates that the replacement entity is present.
[0085] “Subject” means a human, mouse, or non-human primate. A human subject can be any age (e.g., an infant, child, young adult, or adult), and may suffer from a disease, or may be in need of alteration of a gene.
[0086] “Treat,” “treating,” and “treatment” mean the treatment of a disease in a subject (e.g., a human subject), including one or more of inhibiting the disease, i.e., arresting or preventing its development or progression; relieving the disease, i.e., causing regression of the disease state; relieving one or more symptoms of the disease; and curing the disease.
[0087] “Prevent,” “preventing,” and “prevention” refer to the prevention of a disease in a subject, including (a) avoiding or precluding the disease; (b) affecting the predisposition toward the disease; or (c) preventing or delaying the onset of at least one symptom of the disease.
[0088] A “kit” refers to any collection of two or more components that together constitute a functional unit that can be employed for a specific purpose. By way of illustration (and not limitation), one kit according to this disclosure can include a gRNA complexed or able to complex with an RNA-guided nuclease, and accompanied by (e.g., suspended in, or suspendable in) a pharmaceutically acceptable carrier. The kit can be used to introduce the complex into, for example, a cell or a subject, for the purpose of causing a desired genomic alteration in such cell or subject. The components of a kit can be packaged together, or they may be separately packaged. Kits according to this disclosure also optionally include directions for use (DFU) that describe the use of the kit e.g., according to a method of this disclosure. The DFU can be physically packaged with the kit, or it can be made available to a user of the kit, for instance by electronic means.
[0089] The terms “polynucleotide”, “nucleotide sequence”, “nucleic acid”, “nucleic acid molecule”, “nucleic acid sequence”, and “oligonucleotide” refer to a series of nucleotide bases (also called “nucleotides”) in DNA and RNA, and mean any chain of two or more nucleotides. The polynucleotides, nucleotide sequences, nucleic acids etc. can be chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. They can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, its hybridization parameters, etc. A nucleotide sequence typically carries genetic information, including, but not limited to, the information used by cellular machinery to make proteins and enzymes. These terms include double- or single-stranded genomic DNA, RNA, any synthetic and genetically manipulated polynucleotide, and both sense and antisense polynucleotides. These terms also include nucleic acids containing modified bases.
[0090] Conventional IUPAC notation is used in nucleotide sequences presented herein, as shown in Table 1, below (see also Cornish-Bowden 1985, incorporated by reference herein). It should be noted, however, that “T” denotes “Thymine or Uracil” in those instances where a sequence may be encoded by either DNA or RNA, for example in gRNA targeting domains.
TABLE-US-00001 TABLE 1 IUPAC nucleic acid notation Character Base A Adenine T Thymine G Guanine C Cytosine U Uracil K G or T/U M A or C R A or G Y C or T/U S C or G W A or T/U B C, G or T/U V A, C or G H A, C or T/U D A, G or T/U N A, C, G or T/U
[0091] The terms “protein,” “peptide” and “polypeptide” are used interchangeably to refer to a sequential chain of amino acids linked together via peptide bonds. The terms include individual proteins, groups or complexes of proteins that associate together, as well as fragments or portions, variants, derivatives and analogs of such proteins. Peptide sequences are presented herein using conventional notation, beginning with the amino or N-terminus on the left, and proceeding to the carboxyl or C-terminus on the right. Standard one-letter or three-letter abbreviations can be used.
[0092] In some embodiments, a guide RNA (gRNA) sequence that comprises a targeting domain hybridizes (or is complementary) to the target sequence within the target region, e.g., either the

“+”, “−” strand of the target region. In some embodiments, a genome editing system that comprises an RNA-guided nuclease and a gRNA is configured to bind to the target sequence to affect cleavage or an editing event within a target region, e.g., one or both strands of the target region.
[0093] The notations “c.−114 to −102 region,” “13 nt target region” and the like refer to a sequence that is 5′ of the transcription start site (TSS) of the HBG1 and/or HBG2 gene. The c.−114 to −102 region is exemplified in SEQ ID NO:902 at positions 2824-2836, and SEQ ID NO:903 at positions 2748-2760. The term “13 nt deletion” and the like refer to deletions of the 13 nt target region.
[0094] The term “proximal HBG1/2 promoter target sequence” denotes the region within 50, 100, 200, 300, 400, or 500 bp of a proximal HBG1/2 promoter sequence including the 13 nt target region. Alterations by genome editing systems according to this disclosure facilitate (e.g., cause, promote or tend to increase the likelihood of) upregulation of HbF production in erythroid progeny.
[0095] The term “GATA1 binding motif in BCL11Ae” refers to the sequence that is the GATA1 binding motif in the erythroid specific enhancer of BCL11A (BCL11Ae) that is in the +58 DNase I hypersensitive site (DHS) region of intron 2 of the BCL11A gene. The genomic coordinates for the GATA1 binding motif in BCL11Ae are chr2: 60,495,265 to 60,495,270. The +58 DHS site comprises a 115 base pair (bp) sequence as set forth in SEQ ID NO:968. The +58 DHS site sequence, including ˜500 bp upstream and ˜200 bp downstream is set forth in SEQ ID NO:969.
[0096] Where ranges are provided herein, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and/or the understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value within the stated ranges in different embodiments of the invention, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise. It is also to be understood that unless otherwise indicated or otherwise evident from the context and/or the understanding of one of ordinary skill in the art, values expressed as ranges can assume any subrange within the given range, wherein the endpoints of the subrange are expressed to the same degree of accuracy as the tenth of the unit of the lower limit of the range.

Overview

[0097] The various embodiments of this disclosure generally relate to genome editing systems configured to introduce alterations (e.g., a deletion or insertion, or other mutation) into chromosomal DNA that enhance transcription of the HBG1 and/or HBG2 genes, which encode the γA and γG subunits of hemoglobin, respectively. Exemplary mutations are made in or around one or more regions set forth in Table 13, the 13 nt target region, and/or into the GATA1 binding motif in BCL11Ae of HBG1 and/or HBG2. In some embodiments, a gRNA sequence that comprises a targeting domain hybridizes (or is complementary) to the target sequence within the target region, e.g., either the “+”, “−” strand of the target region. In some embodiments, a genome editing system that comprises an RNA-guided nuclease and a gRNA is configured to bind to the target sequence to affect cleavage or an editing event within a target region, e.g., one or both strands of the target region (e.g., in or around one or more regions set forth in Table 13, the 13 nt target region, and/or into the GATA1 binding motif in BCL11Ae of HBG1 and/or HBG2).

Targeted Genome Editing for Fetal Hemoglobin Induction

[0098] Fetal hemoglobin (HbF) expression can be induced using various genome strategies. For example, HbF expression can be induced through targeted disruption of the 13 nt target region, proximal HBG1/2 promoter target sequence, and or the erythroid cell specific expression of a transcriptional repressor, BCL11A (BCL11Ae) (also discussed in commonly-assigned International Patent Publication No. WO 2015/148860 by Friedland et al. (“Friedland”), published Oct. 1, 2015, which is incorporated by reference in its entirety herein), which encodes a repressor that silences HBG1 and HBG2 (Canvers 2015). In certain embodiments, the region of BCL11Ae targeted for disruption may be the GATA1 binding motif in BCL11Ae. In certain embodiments, genome editing systems disclosed herein may be used to introduce alterations into the GATA1 binding motif in BCL11Ae and the 13 nt target region of HBG1 and/or HBG2. In certain embodiments, genome

editing systems disclosed herein may be used to introduce alterations into one or more regions disclosed in Table 13.

[0099] The genome editing systems of this disclosure can include an RNA-guided nuclease such as Cas9 or Cpf1 and one or more gRNAs having a targeting domain that is complementary to a sequence in or near the target region, and optionally one or more of a DNA donor template that encodes a specific mutation (such as a deletion or insertion) in or near the target region, and/or an agent that enhances the efficiency with which such mutations are generated including, without limitation, a random oligonucleotide, a small molecule agonist or antagonist of a gene product involved in DNA repair or a DNA damage response, or a peptide agent.

[0100] A variety of approaches to the introduction of mutations into one or more regions set forth in Table 13, the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae may be employed in the embodiments of the present disclosure. In one approach, a single alteration, such as a double-strand break, is made within one or more regions set forth in Table 13, the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae, and is repaired in a way that disrupts the function of the region, for example by the formation of an indel or by the incorporation of a donor template sequence that encodes the deletion of the region. In a second approach, two or more alterations are made on either side of the region, resulting in the deletion of the intervening sequence, including the 13 nt target region and/or the GATA1 binding motif in BCL11Ae.

[0101] The treatment of hemoglobinopathies by gene therapy and/or genome editing is complicated by the fact that the cells that are phenotypically affected by the disease, erythrocytes or RBCs, are enucleated, and do not contain genetic material encoding either the aberrant hemoglobin protein (Hb) subunits nor the γA or γG subunits targeted in the exemplary genome editing approaches described above. This complication is addressed, in certain embodiments of this disclosure, by the alteration of cells that are competent to differentiate into, or otherwise give rise to, erythrocytes. Cells within the erythroid lineage that are altered according to various embodiments of this disclosure include, without limitation, hematopoietic stem and progenitor cells (HSCs), erythroblasts (including basophilic, polychromatic and/or orthochromatic erythroblasts), proerythroblasts, polychromatic erythrocytes or reticulocytes, embryonic stem (ES) cells, and/or induced pluripotent stem (iPSC) cells. These cells may be altered in situ (e.g., within a tissue of a subject) or ex vivo. Implementations of genome editing systems for in situ and ex vivo alteration of cells is described under the heading “Implementation of genome editing systems: delivery, formulations, and routes of administration” below.

[0102] In certain embodiments, alterations that result in induction of γA and/or γG expression or induction of HbF expression are obtained through the use of a genome editing system comprising an RNA-guided nuclease and at least one gRNA having a targeting domain that alters a sequence in one or more regions set forth in Table 13 or proximate thereto (e.g., within 10, 20, 30, 40, or 50, 100, 200, 300, 400 or 500 bases of the one or more regions). As is discussed in greater detail below, the RNA-guided nuclease and gRNA form a complex that is capable of associating with and altering one or more regions set forth in Table 13 or a region proximate thereto. Examples of suitable targeting domains directed to one or more regions set forth in Table 13 or proximate thereto for use in the embodiments disclosed herein include, without limitation, those set forth in Table 12. In some embodiments, a gRNA sequence that comprises a targeting domain hybridizes (or is complementary) to the target sequence within the target region, e.g., either the “+”, “−” strand of the target region. In some embodiments, a genome editing system that comprises an RNA-guided nuclease and a gRNA is configured to bind to the target sequence to effect cleavage or an editing event within a target region, e.g., one or both strands of the target region (e.g., in or around one or more regions set forth in Table 13).

[0103] In certain embodiments, alterations that result in induction of γA and/or γG expression are obtained through the use of a genome editing system comprising an RNA-guided nuclease and at

least one gRNA having a targeting domain complementary to a sequence within the 13 nt target region of HBG1 and/or HBG2 or proximate thereto (e.g., within 10, 20, 30, 40, or 50, 100, 200, 300, 400 or 500 bases of the 13 nt target region). As is discussed in greater detail below, the RNA-guided nuclease and gRNA form a complex that is capable of associating with and altering the 13 nt target region or a region proximate thereto. Examples of suitable targeting domains directed to the 13 nt target region of HBG1 and/or HBG2 or proximate thereto for use in the embodiments disclosed herein include, without limitation, those set forth in SEQ ID NOs:251-901, 940-942.

[0104] In certain embodiments, alterations that result in induction of HbF expression are obtained through the use of a genome editing system comprising an RNA-guided nuclease and at least one gRNA having a targeting domain complementary to a sequence within the GATA1 binding motif in BCL11Ae or proximate thereto (e.g., within 10, 20, 30, 40, or 50, 100, 200, 300, 400 or 500 bases of the GATA1 binding motif in BCL11Ae). In certain embodiments, the RNA-guided nuclease and gRNA form a complex that is capable of associating with and altering the GATA1 binding motif in BCL11Ae. Examples of suitable targeting domains directed to the GATA1 binding motif in BCL11Ae for use in the embodiments disclosed herein include, without limitation, those set forth in SEQ ID NOs:952-955.

[0105] The genome editing system can be implemented in a variety of ways, as is discussed below in detail. As an example, a genome editing system of this disclosure can be implemented as a ribonucleoprotein complex or a plurality of complexes in which multiple gRNAs are used. This ribonucleoprotein complex can be introduced into a target cell using art-known methods, including electroporation, as described in commonly-assigned International Patent Publication No. WO 2016/182959 by Jennifer Gori (“Gori”), published Nov. 17, 2016, which is incorporated by reference in its entirety herein.

[0106] The ribonucleoprotein complexes within these compositions are introduced into target cells by art-known methods, including without limitation electroporation (e.g., using the Nucleofection™ technology commercialized by Lonza, Basel, Switzerland or similar technologies commercialized by, for example, Maxcyte Inc. Gaithersburg, Maryland) and lipofection (e.g., using Lipofectamine™ reagent commercialized by Thermo Fisher Scientific, Waltham Massachusetts). Alternatively, or additionally, ribonucleoprotein complexes are formed within the target cells themselves following introduction of nucleic acids encoding the RNA-guided nuclease and/or gRNA. These and other delivery modalities are described in general terms below and in Gori.

[0107] Cells that have been altered ex vivo according to this disclosure can be manipulated (e.g., expanded, passaged, frozen, differentiated, de-differentiated, transduced with a transgene, etc.) prior to their delivery to a subject. The cells are, variously, delivered to a subject from which they are obtained (in an “autologous” transplant), or to a recipient who is immunologically distinct from a donor of the cells (in an “allogeneic” transplant).

[0108] In some cases, an autologous transplant includes the steps of obtaining, from the subject, a plurality of cells, either circulating in peripheral blood, or within the marrow or other tissue (e.g., spleen, skin, etc.), and manipulating those cells to enrich for cells in the erythroid lineage (e.g., by induction to generate iPSCs, purification of cells expressing certain cell surface markers such as CD34, CD90, CD49f and/or not expressing surface markers characteristic of non-erythroid lineages such as CD10, CD14, CD38, etc.). The cells are, optionally or additionally, expanded, transduced with a transgene, exposed to a cytokine or other peptide or small molecule agent, and/or frozen/thawed prior to transduction with a genome editing system targeting one or more regions set forth in Table 13, the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae. The genome editing system can be implemented or delivered to the cells in any suitable format, including as a ribonucleoprotein complex, as separated protein and nucleic acid components, and/or as nucleic acids encoding the components of the genome editing system.

[0109] In certain embodiments, CD34+ hematopoietic stem and progenitor cells (HSPCs) that have

been edited using the genome editing methods disclosed herein may be used for the treatment of a hemoglobinopathy in a subject in need thereof. In certain embodiments, the hemoglobinopathy may be severe sickle cell disease (SCD) or thalassemia, such as β-thalassemia, δ-thalassemia, or β/δ-thalassemia. In certain embodiments, an exemplary protocol for treatment of a hemoglobinopathy may include harvesting CD34+ HSPCs from a subject in need thereof, ex vivo editing of the autologous CD34+ HSPCs using the genome editing methods disclosed herein, followed by reinfusion of the edited autologous CD34+ HSPCs into the subject. In certain embodiments, treatment with edited autologous CD34+ HSPCs may result in increased HbF induction.

[0110] Prior to harvesting CD34+ HSPCs, in certain embodiments, a subject may discontinue treatment with hydroxyurea, if applicable, and receive blood transfusions to maintain sufficient hemoglobin (Hb) levels. In certain embodiments, a subject may be administered intravenous plerixafor (e.g., 0.24 mg/kg) to mobilize CD34+ HSPCs from bone marrow into peripheral blood. In certain embodiments, a subject may undergo one or more leukapheresis cycles (e.g., approximately one month between cycles, with one cycle defined as two plerixafor-mobilized leukapheresis collections performed on consecutive days). In certain embodiments, the number of leukapheresis cycles performed for a subject may be the number required to achieve a dose of edited autologous CD34+ HSPCs (e.g., ≥2×10.sup.6 cells/kg, ≥3×10.sup.6 cells/kg, ≥4×10.sup.6 cells/kg, ≥5×10.sup.6 cells/kg, 2×10.sup.6 cells/kg to 3×10.sup.6 cells/kg, 3×10.sup.6 cells/kg to 4×10.sup.6 cells/kg, 4×10.sup.6 cells/kg to 5×10.sup.6 cells/kg) to be reinfused back into the subject, along with a dose of unedited autologous CD34+ HSPCs/kg for backup storage (e.g., >1.5×10.sup.6 cells/kg). In certain embodiments, the CD34+ HSPCs harvested from the subject may be edited using any of the genome editing methods discussed herein. In certain embodiments, any one or more of the gRNAs and one or more of the RNA-guided nucleases disclosed herein may be used in the genome editing methods.

[0111] In certain embodiments, the treatment may include an autologous stem cell transplant. In certain embodiments, a subject may undergo myeloablative conditioning with busulfan conditioning (e.g., dose-adjusted based on first-dose pharmacokinetic analysis, with a test dose of 1 mg/kg). In certain embodiments, conditioning may occur for four consecutive days. In certain embodiments, following a three-day busulfan washout period, edited autologous CD34+ HSPCs (e.g., ≥2×10.sup.6 cells/kg, ≥3×10.sup.6 cells/kg, ≥4×10.sup.6 cells/kg, ≥5×10.sup.6 cells/kg, 2×10.sup.6 cells/kg to 3×10.sup.6 cells/kg, 3×10.sup.6 cells/kg to 4×10.sup.6 cells/kg, 4×10.sup.6 cells/kg to 5×10.sup.6 cells/kg) may be reinfused into the subject (e.g., into peripheral blood). In certain embodiments, the edited autologous CD34+ HSPCs may be manufactured and cryopreserved for a particular subject. In certain embodiments, a subject may attain neutrophil engraftment following a sequential myeloablative conditioning regimen and infusion of edited autologous CD34.sup.+ cells. Neutrophil engraftment may be defined as three consecutive measurements of ANC≥0.5×10.sup.9/L.

[0112] However it is implemented, a genome editing system may include, or may be co-delivered with, one or more factors that improve the viability of the cells during and after editing, including without limitation an aryl hydrocarbon receptor antagonist such as StemRegenin-1 (SR1), UM171, LGC0006, alpha-napthoflavone, and CH-223191, and/or an innate immune response antagonist such as cyclosporin A, dexamethasone, reservatrol, a MyD88 inhibitory peptide, an RNAi agent targeting Myd88, a B18R recombinant protein, a glucocorticoid, OxPAPC, a TLR antagonist, rapamycin, BX795, and a RLR shRNA. These and other factors that improve the viability of the cells during and after editing are described in Gori, under the heading “I. Optimization of Stem Cells” from page 36 through page 61, which is incorporated by reference herein.

[0113] The cells, following delivery of the genome editing system, are optionally manipulated e.g., to enrich for HSCs and/or cells in the erythroid lineage and/or for edited cells, to expand them, freeze/thaw, or otherwise prepare the cells for return to the subject. The edited cells are then returned to the subject, for instance in the circulatory system by means of intravenous delivery or

delivery or into a solid tissue such as bone marrow.
[0114] Functionally, alteration of one or more regions set forth in Table 13, the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae using the compositions, methods and genome editing systems of this disclosure results in significant induction, among hemoglobin-expressing cells, of TA and/or **7**G subunits (referred to interchangeably as HbF expression), e.g., at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50% or greater induction of γA and/or γG subunit expression relative to unmodified controls. This induction of protein expression is generally the result of alteration of one or more regions set forth in Table 13, the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae (expressed, e.g., in terms of the percentage of total genomes comprising indel mutations within the plurality of cells) in some or all of the plurality of cells that are treated, e.g., at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50% of the plurality of cells comprise at least one allele comprising a sequence alteration, including, without limitation, an indel, insertion, or deletion in or near one or more regions set forth in Table 13, the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae.
[0115] The functional effects of alterations caused or facilitated by the genome editing systems and methods of the present disclosure can be assessed in any number of suitable ways. For example, the effects of alterations on expression of fetal hemoglobin can be assessed at the protein or mRNA level. Expression of HBG1 and HBG2 mRNA can be assessed by digital droplet PCR (ddPCR), which is performed on cDNA samples obtained by reverse transcription of mRNA harvested from treated or untreated samples. Primers for HBG1, HBG2, HBB, and/or HBA may be used individually or multiplexed using methods known in the art. For example, ddPCR analysis of samples may be conducted using the QX200™ ddPCR system commercialized by Bio Rad (Hercules, CA), and associated protocols published by BioRad. Fetal hemoglobin protein may be assessed by high pressure liquid chromatography (HPLC), for example, according to the methods discussed on pp. 143-44 of Chang 2017, incorporated by reference herein, or fast protein liquid chromatography (FPLC) using ion-exchange and/or reverse phase columns to resolve HbF, HbB and HbA and/or γA and γG globin chains as is known in the art.
[0116] It should be noted that the rate at which one or more regions set forth in Table 13, the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae is altered in the target cells can be modified by the use of optional genome editing system components such as oligonucleotide donor templates. Donor template design is described in general terms below under the heading “Donor template design.” Donor templates for use in targeting the 13 nt target region may include, without limitation, donor templates encoding alterations (e.g., deletions) of HBG1 c.−114 to −102 (corresponding to nucleotides 2824-2836 of SEQ ID NO: 902), HBG1 c.−225 to −222 (corresponding to nucleotides 2716-2719 of SEQ ID NO:902)), and/or HBG2 c.−114 to −102 (corresponding to nucleotides 2748-2760 of SEQ ID NO:903). Exemplary 5′ and 3′ homology arms, and exemplary full-length donor templates encoding deletions such as c. −114 to −102 are also presented below (SEQ ID NOS: 904-909). Donor templates used herein may be non-specific templates that are non-homologous to regions of DNA within or near the target sequence. In certain embodiments, donor templates for use in targeting the 13 nt target region may include, without limitation, non-target specific templates that are nonhomologous to regions of DNA within or near the 13 nt target region. For example, a non-specific donor template for use in targeting the 13 nt target region may be non-homologous to the regions of DNA within or near the 13 nt target region and may comprise a donor template encoding the deletion of HBG1 c.−225 to −222 (corresponding to nucleotides 2716-2719 of SEQ ID NO:902). In certain embodiments, donor templates for use in targeting the GATA1 binding motif in BCL11Ae may include, without limitation, non-target specific templates that are nonhomologous to regions of DNA within or near GATA1 binding motif in BCL11Ae target sequence. Other donor templates for use in targeting BCL11Ae may include, without limitation, donor templates including

alternations (e.g., deletions) of BCL11Ae, including, without limitation, the GATA1 motif in BCL11Ae.

RNA-Guided Helicases and Dead Guide RNAs to Increase Accessibility to Edit Target Region

[0117] Various embodiments of the present disclosure also generally relate to genome editing systems configured to alter the helical structure of a nucleic acid to enhance genome editing of a target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae) in the nucleic acid, and methods and compositions thereof. Many embodiments relate to the observation that positioning an event that alters the helical structure of DNA within or adjacent to target regions in nucleic acid may improve the activity of genome editing systems directed to such target regions. Without wishing to be bound by any theory, it is thought that alterations of helical structure (e.g., by unwinding) within or proximal to DNA target regions may induce or increase accessibility of a genome editing system to the target region, resulting in increased editing of the target regions by the genome editing system.

[0118] CRISPR nucleases evolved primarily to defend bacteria against viral pathogens, whose genomes are not naturally organized into chromatin. By contrast, when eukaryotic genomes are organized into nucleosomal units comprising genomic DNA segments coiled around histones. CRISPR nucleases from several bacterial families have been found to be inactive for editing eukaryotic DNA, suggesting the ability to edit nucleosome-bound DNA might differ across enzymes (Ran 2015). Biochemical evidence shows that *S. pyogenes* Cas9 can cleave DNA efficiently at nucleosome edges, but has reduced activity when the target site is positioned near the center of nucleosome dyad (Hinz 2016).

[0119] In many cell types, target sites of interest may be strongly bound by nucleosomes, or may only possess adjacent PAMs for enzymes that do not edit efficiently in the presence of nucleosomes. In this case, the problematic nucleosomes could be displaced first by using adjacent target sites that are closer to the nucleosome edge or are bound by an enzyme that is more effective at binding nucleosomal DNA. However, cleavage at these adjacent sites could be detrimental to the therapeutic strategy. Therefore, having a programmable enzyme that binds these adjacent sites but does not cleave can enable more efficient functional editing.

[0120] It will be evident to the skilled artisan that the simplified systems and methods described herein offer several advantages over competing approaches. For example, a related strategy using catalytically inactive (dead) enzymes targeting sites adjacent to the site where editing is desired has been described in the literature (Chen 2017). However, this strategy entails a potential safety issue: if the full-length gRNAs complexed with a catalytically inactive nuclease molecule dissociates and later reassociates with a catalytically active nuclease enzyme, the gRNAs could introduce undesirable off-target edits. In contrast, the systems and methods of the present disclosure eliminates this risk because it relies on the observation that a dead gRNA (dgRNA) (gRNAs with a targeting domain of 15 nucleotides or less) allow an RNA-guided nuclease to bind, but not cleave, its target cite. Thus, the dgRNAs provided herein will not support nuclease activity irrespective of their association with any particular RNA-guided nuclease molecule. By using these dead gRNAs, adjacent target sites can be used to aid in nucleosome displacement without the risk of guide RNA swapping between active and inactive enzyme.

[0121] Another related strategy utilizes recruitment of exogenous trans-acting factors to facilitate nucleosome displacement. However, the systems and methods of this disclosure are advantageous over this strategy because they do not require gRNA modifications beyond truncation of the targeting domain, do not require the recruitment of exogenous trans-acting factors, and do not require transcriptional activation to achieve increased rates of editing.

[0122] Additionally, the use of dead gRNAs in the genome editing systems of the present disclosure are advantageous because they are not expected to result in any new delivery/solubility or folding/manufacturing considerations relative to genome editing systems utilizing full-length gRNAs. However, a skilled artisan might expect to encounter such problems in genome editing

systems that utilize a exogenous trans-acting factors, which may entail large fusion proteins and/or RNA insertions or fusions. Further, dead gRNA strategies are likely to be capable of implementation using existing manufacturing, delivery, and other commercial processes that have been designed for wild-type nuclease products with relatively few substantial changes.

[0123] A variety of approaches to the unwinding and alteration of nucleic acid are employed in the various embodiments of this disclosure. One approach comprises unwinding (or opening of) a chromatin segment within or proximal to a target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae) of a nucleic acid in a cell and generating a double stranded break (DSB) within the target region of the nucleic acid, wherein the DSB is repaired in a manner that alters the target region. Unwinding the chromatin segment using the methods provided herein may facilitate increased access of catalytically active RNPs (e.g., catalytically active RNA-guided nucleases and gRNAs) to the chromatin to allow for more efficient editing of the DNA. For example, these methods may be used to edit target regions in chromatin that are difficult for a ribonucleoprotein (e.g., RNA-guided nuclease complexed to gRNA) to access because the chromatin is occupied by nucleosomes, such as closed chromatin. In certain embodiments, the unwinding of the chromatin segment occurs via RNA-guided helicase activity. In certain embodiments, the unwinding step does not require recruiting an exogenous trans-acting factor to the chromatin segment. In certain embodiments, the step of unwinding the chromatin segment does not comprise forming a single or double-stranded break in the nucleic acid within the chromatin segment.

[0124] In certain embodiments of the approaches and methods described above, the alteration of DNA helical structure is achieved through the action of an “RNA-guided helicase,” which term is generally used to refer to a molecule, typically a peptide, that (a) interacts (e.g., complexes) with a gRNA, and (b) together with the gRNA, associates with and unwinds, but does not cleave, a target site. RNA-guided helicases may, in certain embodiments, comprise RNA-guided nucleases configured to lack nuclease activity. However, the inventors have observed that even a cleavage-competent RNA-guided nuclease may be adapted for use as an RNA-guided helicase by complexing it to a dead gRNA having a truncated targeting domain of 15 or fewer nucleotides in length. Complexes of wild-type RNA-guided nucleases with dead gRNAs (dgRNAs) exhibit reduced or eliminated RNA-cleavage activity, but appear to retain helicase activity. RNA-guided helicases and dead gRNAs are described in greater detail below.

[0125] Regarding RNA-guided helicases, according to the present disclosure an RNA-guided helicase may comprise any of the RNA-guided nucleases disclosed herein and infra under the heading entitled “RNA-guided nucleases,” including, without limitation, a Cas9 or Cpf1 RNA-guided nuclease. The helicase activity of these RNA-guided nucleases allow for unwinding of DNA, providing increased access of genome editing system components (e.g., without limitation, catalytically active RNA-guided nuclease and gRNAs) to the desired target region to be edited (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae). In certain embodiments, the RNA-guided nuclease may be a catalytically active RNA-guided nuclease with nuclease activity. In certain embodiments, the RNA-guided helicase may be configured to lack nuclease activity. For example, in certain embodiments, the RNA-guided helicase may be a catalytically inactive RNA-guided nuclease that lacks nuclease activity, such as a catalytically dead Cas9 molecule, which still provides helicase activity. In certain embodiments, an RNA-guided helicase may form a complex with a dead gRNA, forming a dead RNP that cannot cleave nucleic acid. In other embodiments, the RNA-guided helicase may be a catalytically active RNA-guided nuclease complexed to a dead gRNA, forming a dead RNP that cannot cleave nucleic acid.

[0126] Turning to dead gRNAs, these include any of the dead gRNAs discussed herein and infra under the heading entitled “Dead gRNA molecules.” Dead gRNAs may be generated by truncating the 5′ end of a gRNA targeting domain sequence, resulting in a targeting domain sequence of 15

nucleotides or fewer in length. Dead gRNA molecules may comprise targeting domains complementary to regions proximal to or within a target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae) in a target nucleic acid. In certain embodiments, “proximal to” may denote the region within 10, 25, 50, 100, or 200 nucleotides of a target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae). In certain embodiments, dead gRNAs comprise targeting domains complementary to the transcription strand or non-transcription strand of DNA. In certain embodiments, the dead guide RNA is not configured to recruit an exogenous trans-acting factor to a target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae).

[0127] Also provided herein are methods of increasing a rate of indel formation in a target nucleic acid by unwinding DNA within or proximal to the target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae) using an RNA-guided helicase, generating a DSB within the target region, and forming an indel in the target region through repair of the DSB. The step of unwinding the DNA using an RNA-guided helicase provides for increased indel formation compared to a method of forming indels that does not use a helicase.

[0128] This disclosure further encompasses methods of deleting a segment of a target nucleic acid in a cell, comprising contacting the cell with an RNA-guided helicase and generating a double strand break (DSB) within the target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae). In certain embodiments, the RNA-guided helicase is configured to associate within or proximal to a target region of the target nucleic acid and unwind double stranded DNA (dsDNA) within or proximal to the target region. In certain embodiments, the target nucleic acid is a promoter region of a gene, a coding region of a gene, a non-coding region of a gene, an intron of a gene, or an exon of a gene. In certain embodiments, the segment of the target nucleic acid to be deleted may is at least about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, or 100 base pairs in length. In certain embodiments, the DSB is repaired in a manner that deletes the segment of the target nucleic acid.

[0129] Genome editing systems configured to introduce alterations of helical structure may be implemented in a variety of ways, as is discussed below in detail. As an example, a genome editing system of this disclosure can be implemented as a ribonucleoprotein complex or a plurality of complexes in which multiple gRNAs are used. In certain embodiments, a ribonucleoprotein complex of the genome editing system may be an RNA-guided helicase complexed to a dead guide RNA. Ribonucleoprotein complexes can be introduced into a target cell using art-known methods, including electroporation, as described in Gori. Genome editing systems incorporating RNA-guided helicases may also be modified in any suitable manner, including without limitation by the inclusion of one or more of a DNA donor template that encodes a specific mutation (such as a deletion or insertion) in or near the target region, and/or an agent that enhances the efficiency with which such mutations are generated including, without limitation, a random oligonucleotide, a small molecule agonist or antagonist of a gene product involved in DNA repair or a DNA damage response, or a peptide agent. These modifications are described in greater detail below, under the heading “Genome Editing Strategies.” For clarity, this disclosure includes compositions comprising one or more gRNAs, dead gRNAs, RNA-guided helicases, RNA-guided nucleases, or a combination thereof.

[0130] While several of the exemplary embodiments above have focused on DNA unwinding, it should be noted that other helical alterations are within the scope of the present disclosure. These include, without limitation, overwinding, underwinding, increase or decrease of torsional strain on DNA strands within or proximate to a target region (e.g., through topoisomerase activity), denaturation or strand separation, and/or other suitable alterations resulting in modifications of chromatin structure. Each of these alterations may be catalyzed by an RNA-guided activity, or by

the recruitment of an endogenous factor to a target region.
[0131] This overview has focused on a handful of exemplary embodiments that illustrate the principles of genome editing systems and CRISPR-mediated methods of altering cells. For clarity, however, this disclosure encompasses modifications and variations that have not been expressly addressed above, but will be evident to those of skill in the art. With that in mind, the following disclosure is intended to illustrate the operating principles of genome editing systems more generally. What follows should not be understood as limiting, but rather illustrative of certain principles of genome editing systems and CRISPR-mediated methods utilizing these systems, which, in combination with the instant disclosure, will inform those of skill in the art about additional implementations and modifications that are within its scope.

Genome Editing Systems

[0132] The term “genome editing system” refers to any system having RNA-guided DNA editing activity. Genome editing systems of the present disclosure include at least two components adapted from naturally occurring CRISPR systems: a guide RNA (gRNA) and an RNA-guided nuclease. These two components form a complex that is capable of associating with a specific nucleic acid sequence and editing the DNA in or around that nucleic acid sequence, for instance by making one or more of a single-strand break (an SSB or nick), a double-strand break (a DSB) and/or a point mutation.
[0133] In certain embodiments, the genome editing systems in this disclosure may include a helicase for unwinding DNA. In certain embodiments, the helicase may be an RNA-guided helicase. In certain embodiments, the RNA-guided helicase may be an RNA-guided nuclease as described herein, such as a Cas9 or Cpf1 molecule. In certain embodiments, the RNA-guided nuclease is not configured to recruit an exogenous trans-acting factor to a target region. In certain embodiments, the RNA-guided nuclease may be configured to lack nuclease activity. In certain embodiments, the RNA-guided helicase may be complexed with a dead guide RNA as disclosed herein. For example, the dead guide RNA may comprise a targeting domain sequence less than 15 nucleotides in length. In certain embodiments, the dead guide RNA is not configured to recruit an exogenous trans-acting factor to a target region.
[0134] Naturally occurring CRISPR systems are organized evolutionarily into two classes and five types (Makarova 2011, incorporated by reference herein), and while genome editing systems of the present disclosure may adapt components of any type or class of naturally occurring CRISPR system, the embodiments presented herein are generally adapted from Class 2, and type II or V CRISPR systems. Class 2 systems, which encompass types II and V, are characterized by relatively large, multidomain RNA-guided nuclease proteins (e.g., Cas9 or Cpf1) and one or more guide RNAs (e.g., a crRNA and, optionally, a tracrRNA) that form ribonucleoprotein (RNP) complexes that associate with (i.e., target) and cleave specific loci complementary to a targeting (or spacer) sequence of the crRNA. Genome editing systems according to the present disclosure similarly target and edit cellular DNA sequences, but differ significantly from CRISPR systems occurring in nature. For example, the unimolecular guide RNAs described herein do not occur in nature, and both guide RNAs and RNA-guided nucleases according to this disclosure may incorporate any number of non-naturally occurring modifications.
[0135] Genome editing systems can be implemented (e.g., administered or delivered to a cell or a subject) in a variety of ways, and different implementations may be suitable for distinct applications. For instance, a genome editing system is implemented, in certain embodiments, as a protein/RNA complex (a ribonucleoprotein, or RNP), which can be included in a pharmaceutical composition that optionally includes a pharmaceutically acceptable carrier and/or an encapsulating agent, such as, without limitation, a lipid or polymer micro- or nano-particle, micelle, or liposome. In certain embodiments, a genome editing system is implemented as one or more nucleic acids encoding the RNA-guided nuclease and guide RNA components described above (optionally with one or more additional components); in certain embodiments, the genome editing system is

implemented as one or more vectors comprising such nucleic acids, for instance a viral vector such as an adeno-associated virus (see section below under the heading “Implementation of genome editing systems: delivery, formulations, and routes of administration”); and in certain embodiments, the genome editing system is implemented as a combination of any of the foregoing. Additional or modified implementations that operate according to the principles set forth herein will be apparent to the skilled artisan and are within the scope of this disclosure.

[0136] It should be noted that the genome editing systems of the present disclosure can be targeted to a single specific nucleotide sequence, or may be targeted to and capable of editing in parallel two or more specific nucleotide sequences through the use of two or more guide RNAs. The use of multiple gRNAs is referred to as “multiplexing” throughout this disclosure, and can be employed to target multiple, unrelated target sequences of interest, or to form multiple SSBs or DSBs within a single target domain and, in some cases, to generate specific edits within such target domain. For example, International Patent Publication No. WO 2015/138510 by Maeder et al. (“Maeder”), which is incorporated by reference herein, describes a genome editing system for correcting a point mutation (C.2991+1655A to G) in the human CEP290 gene that results in the creation of a cryptic splice site, which in turn reduces or eliminates the function of the gene. The genome editing system of Maeder utilizes two guide RNAs targeted to sequences on either side of (i.e., flanking) the point mutation, and forms DSBs that flank the mutation. This, in turn, promotes deletion of the intervening sequence, including the mutation, thereby eliminating the cryptic splice site and restoring normal gene function.

[0137] As another example, WO 2016/073990 by Cotta-Ramusino et al. (“Cotta-Ramusino”), which is incorporated by reference herein, describes a genome editing system that utilizes two gRNAs in combination with a Cas9 nickase (a Cas9 that makes a single strand nick such as *S. pyogenes* D10A), an arrangement termed a “dual-nickase system.” The dual-nickase system of Cotta-Ramusino is configured to make two nicks on opposite strands of a sequence of interest that are offset by one or more nucleotides, which nicks combine to create a double strand break having an overhang (5′ in the case of Cotta-Ramusino, though 3′ overhangs are also possible). The overhang, in turn, can facilitate homology directed repair events in some circumstances. And, as another example, International Patent Publication No. WO 2015/070083 by Palestrant et al. (incorporated by reference herein) describes a gRNA targeted to a nucleotide sequence encoding Cas9 (referred to as a “governing RNA”), which can be included in a genome editing system comprising one or more additional gRNAs to permit transient expression of a Cas9 that might otherwise be constitutively expressed, for example in some virally transduced cells. These multiplexing applications are intended to be exemplary, rather than limiting, and the skilled artisan will appreciate that other applications of multiplexing are generally compatible with the genome editing systems described here.

[0138] As disclosed herein, in certain embodiments, genome editing systems may comprise multiple gRNAs that may be used to introduce mutations into the GATA1 binding motif in BCL11Ae or the 13 nt target region of HBG1 and/or HBG2. In certain embodiments, genome editing systems disclosed herein may comprise multiple gRNAs used to introduce mutations into the GATA1 binding motif in BCL11Ae and the 13 nt target region of HBG1 and/or HBG2.

[0139] Genome editing systems can, in some instances, form double strand breaks that are repaired by cellular DNA double-strand break mechanisms such as NHEJ or HDR. These mechanisms are described throughout the literature (see, e.g., Davis 2014 (describing Alt-HDR), Frit 2014 (describing Alt-NHEJ), and Iyama 2013 (describing canonical HDR and NHEJ pathways generally), all of which are incorporated by reference herein).

[0140] Where genome editing systems operate by forming DSBs, such systems optionally include one or more components that promote or facilitate a particular mode of double-strand break repair or a particular repair outcome. For instance, Cotta-Ramusino also describes genome editing systems in which a single stranded oligonucleotide “donor template” is added; the donor template

is incorporated into a target region of cellular DNA that is cleaved by the genome editing system, and can result in a change in the target sequence.

[0141] In certain embodiments, genome editing systems modify a target sequence, or modify expression of a gene in or near the target sequence, without causing single- or double-strand breaks. For example, a genome editing system may include an RNA-guided nuclease fused to a functional domain that acts on DNA, thereby modifying the target sequence or its expression. As one example, an RNA-guided nuclease can be connected to (e.g., fused to) a cytidine deaminase functional domain, and may operate by generating targeted C-to-A substitutions. Exemplary nuclease/deaminase fusions are described in Komor 2016, which is incorporated by reference herein. Alternatively, a genome editing system may utilize a cleavage-inactivated (i.e., a "dead") nuclease, such as a dead Cas9 (dCas9), and may operate by forming stable complexes on one or more targeted regions of cellular DNA, thereby interfering with functions involving the targeted region(s) including, without limitation, mRNA transcription, chromatin remodeling, etc.

Guide RNA (gRNA) Molecules

[0142] The terms "guide RNA" and "gRNA" refer to any nucleic acid that promotes the specific association (or "targeting") of an RNA-guided nuclease such as a Cas9 or a Cpf1 to a target sequence such as a genomic or episomal sequence in a cell. gRNAs can be unimolecular (comprising a single RNA molecule, and referred to alternatively as chimeric), or modular (comprising more than one, and typically two, separate RNA molecules, such as a crRNA and a tracrRNA, which are usually associated with one another, for instance by duplexing). gRNAs and their component parts are described throughout the literature, for instance in Briner 2014, which is incorporated by reference), and in Cotta-Ramusino. Examples of modular and unimolecular gRNAs that may be used according to the embodiments herein include, without limitation, the sequences set forth in SEQ ID NOs:29-31 and 38-51. Examples of gRNA proximal and tail domains that may be used according to the embodiments herein include, without limitation, the sequences set forth in SEQ ID NOs:32-37.

[0143] In bacteria and archea, type II CRISPR systems generally comprise an RNA-guided nuclease protein such as Cas9, a CRISPR RNA (crRNA) that includes a 5′ region that is complementary to a foreign sequence, and a trans-activating crRNA (tracrRNA) that includes a 5′ region that is complementary to, and forms a duplex with, a 3′ region of the crRNA. While not intending to be bound by any theory, it is thought that this duplex facilitates the formation of—and is necessary for the activity of—the Cas9/gRNA complex. As type II CRISPR systems were adapted for use in gene editing, it was discovered that the crRNA and tracrRNA could be joined into a single unimolecular or chimeric guide RNA, in one non-limiting example, by means of a four nucleotide (e.g., GAAA) "tetraloop" or "linker" sequence bridging complementary regions of the crRNA (at its 3′ end) and the tracrRNA (at its 5′ end) (Mali 2013; Jiang 2013; Jinek 2012; all incorporated by reference herein).

[0144] Guide RNAs, whether unimolecular or modular, include a "targeting domain" that is fully or partially complementary to a target domain within a target sequence, such as a DNA sequence in the genome of a cell where editing is desired. Targeting domains are referred to by various names in the literature, including without limitation "guide sequences" (Hsu et al., Nat Biotechnol. 2013 September; 31(9): 827-832, ("Hsu"), incorporated by reference herein), "complementarity regions" (Cotta-Ramusino), "spacers" (Briner 2014) and generically as "crRNAs" (Jiang). Irrespective of the names they are given, targeting domains are typically 10-30 nucleotides in length, and in certain embodiments are 16-24 nucleotides in length (for instance, 16, 17, 18, 19, 20, 21, 22, 23 or 24 nucleotides in length), and are at or near the 5′ terminus of in the case of a Cas9 gRNA, and at or near the 3′ terminus in the case of a Cpf1 gRNA.

[0145] In addition to the targeting domains, gRNAs typically (but not necessarily, as discussed below) include a plurality of domains that may influence the formation or activity of gRNA/Cas9 complexes. For instance, as mentioned above, the duplexed structure formed by first and secondary

complementarity domains of a gRNA (also referred to as a repeat:anti-repeat duplex) interacts with the recognition (REC) lobe of Cas9 and can mediate the formation of Cas9/gRNA complexes (Nishimasu et al., Cell 156, 935-949, Feb. 27, 2014 ("Nishimasu 2014") and Nishimasu et al., Cell 162, 1113-1126, Aug. 27, 2015 ("Nishimasu 2015"), both incorporated by reference herein. It should be noted that the first and/or second complementarity domains may contain one or more poly-A tracts, which can be recognized by RNA polymerases as a termination signal. The sequence of the first and second complementarity domains are, therefore, optionally modified to eliminate these tracts and promote the complete in vitro transcription of gRNAs, for instance through the use of A-G swaps as described in Briner 2014, or A-U swaps. These and other similar modifications to the first and second complementarity domains are within the scope of the present disclosure.

[0146] Along with the first and second complementarity domains, Cas9 gRNAs typically include two or more additional duplexed regions that are involved in nuclease activity in vivo but not necessarily in vitro. (Nishimasu 2015). A first stem-loop one near the 3′ portion of the second complementarity domain is referred to variously as the "proximal domain," (Cotta-Ramusino) "stem loop 1" (Nishimasu 2014 and 2015) and the "nexus" (Briner 2014). One or more additional stem loop structures are generally present near the 3′ end of the gRNA, with the number varying by species: *S. pyogenes* gRNAs typically include two 3′ stem loops (for a total of four stem loop structures including the repeat:anti-repeat duplex), while *S. aureus* and other species have only one (for a total of three stem loop structures). A description of conserved stem loop structures (and gRNA structures more generally) organized by species is provided in Briner 2014.

[0147] While the foregoing description has focused on gRNAs for use with Cas9, it should be appreciated that other RNA-guided nucleases exist which utilize gRNAs that differ in some ways from those described to this point. For instance, Cpf1 ("CRISPR from Prevotella and Franciscella 1") is a recently discovered RNA-guided nuclease that does not require a tracrRNA to function (Zetsche 2015b, incorporated by reference herein). A gRNA for use in a Cpf1 genome editing system generally includes a targeting domain and a complementarity domain (alternately referred to as a "handle"). It should also be noted that, in gRNAs for use with Cpf1, the targeting domain is usually present at or near the 3′ end, rather than the 5′ end as described above in connection with Cas9 gRNAs (the handle is at or near the 5′ end of a Cpf1 gRNA).

[0148] Those of skill in the art will appreciate, however, that although structural differences may exist between gRNAs from different prokaryotic species, or between Cpf1 and Cas9 gRNAs, the principles by which gRNAs operate are generally consistent. Because of this consistency of operation, gRNAs can be defined, in broad terms, by their targeting domain sequences, and skilled artisans will appreciate that a given targeting domain sequence can be incorporated in any suitable gRNA, including a unimolecular or chimeric gRNA, or a gRNA that includes one or more chemical modifications and/or sequential modifications (substitutions, additional nucleotides, truncations, etc.). Thus, for economy of presentation in this disclosure, gRNAs may be described solely in terms of their targeting domain sequences.

[0149] More generally, skilled artisans will appreciate that some aspects of the present disclosure relate to systems, methods and compositions that can be implemented using multiple RNA-guided nucleases. For this reason, unless otherwise specified, the term gRNA should be understood to encompass any suitable gRNA that can be used with any RNA-guided nuclease, and not only those gRNAs that are compatible with a particular species of Cas9 or Cpf1. By way of illustration, the term gRNA can, in certain embodiments, include a gRNA for use with any RNA-guided nuclease occurring in a Class 2 CRISPR system, such as a type II or type V or CRISPR system, or an RNA-guided nuclease derived or adapted therefrom.

gRNA Design

[0150] Methods for selection and validation of target sequences as well as off-target analyses have been described previously (see, e.g., Mali 2013; Hsu 2013; Fu 2014; Heigwer 2014; Bae 2014; Xiao 2014; all incorporated by reference herein). As a non-limiting example, gRNA design may

involve the use of a software tool to optimize the choice of potential target sequences corresponding to a user's target sequence, e.g., to minimize total off-target activity across the genome. While off-target activity is not limited to cleavage, the cleavage efficiency at each off-target sequence can be predicted, e.g., using an experimentally-derived weighting scheme. These and other guide selection methods are described in detail in Maeder and Cotta-Ramusino.

[0151] With respect to selection of gRNA targeting domain sequences directed to HBG1/2 target sites (e.g., the 13 nt target region), an in-silico gRNA target domain identification tool was utilized, and the hits were stratified into four tiers. For *S. pyogenes*, tier 1 targeting domains were selected based on (1) distance upstream or downstream from either end of the target site (i.e., HBG1/2 13 nt target region), specifically within 400 bp of either end of the target site, (2) a high level of orthogonality, and (3) the presence of 5′ G. Tier 2 targeting domains were selected based on (1) distance upstream or downstream from either end of the target site (i.e., HBG1/2 13 nt target region), specifically within 400 bp of either end of the target site, and (2) a high level of orthogonality. Tier 3 targeting domains were selected based on (1) distance upstream or downstream from either end of the target site (i.e., HBG1/2 13 nt target region), specifically within 400 bp of either end of the target site and (2) the presence of 5′ G. Tier 4 targeting domains were selected based on distance upstream or downstream from either end of the target site (i.e., HBG1/2 13 nt target region), specifically within 400 bp of either end of the target site.

[0152] For *S. aureus*, tier 1 targeting domains were selected based on (1) distance upstream or downstream from either end of the target site (i.e., HBG1/2 13 nt target region), specifically within 400 bp of either end of the target site, (2) a high level of orthogonality, (3) the presence of 5′ G, and (4) PAM having the sequence NNGRRT (SEQ ID NO:204). Tier 2 targeting domains were selected based on (1) distance upstream or downstream from either end of the target site (i.e., HBG1/2 13 nt target), specifically within 400 bp of either end of the target site, (2) a high level of orthogonality, and (3) PAM having the sequence NNGRRT (SEQ ID NO:204). Tier 3 targeting domains were selected based on (1) distance upstream or downstream from either end of the target site (i.e., HBG1/2 13 nt target region), specifically within 400 bp of either end of the target site, and (2) PAM having the sequence NNGRRT (SEQ ID NO:204). Tier 4 targeting domains were selected based on (1) distance upstream or downstream from either end of the target site (i.e., HBG1/2 13 nt target), specifically within 400 bp of either end of the target site, and (2) PAM having the sequence NNGRRV (SEQ ID NO:205).

[0153] Table 2, below, presents targeting domains for *S. pyogenes* and *S. aureus* gRNAs, broken out by (a) tier (1, 2, 3 or 4) and (b) HBG1 or HBG2.

TABLE-US-00002 TABLE 2 gRNA targeting domain sequences for HBG1/2 target sites HBG1 HBG2 *S.* Tier 1 251-256 760-764 *pyogenes* Tier 2 257-274 765-781 Tier 3 275-300 275-281, 283-300 Tier 4 301-366 301-311, 313-342, 344- 348, 350-366, 782, 783 *S.* Tier 1 367-376 784-791 *aureus* Tier 2 343, 377-393 778, 792-803 Tier 3 357, 365, 394-461 357, 365, 394-461 Tier 4 252-254, 256, 268, 292, 295, 347, 348, 353, 272-274, 292, 295, 360-362, 366, 462-468 476- 347, 348, 353, 360- 481, 489-587, 601-607, 614- 362, 366, 598-759 620, 640-666, 674-679, 687- 693, 708-714, 733-753, 762- 764, 775, 779-781, 804-901

[0154] gRNAs may be designed to target the erythroid specific enhancer of BCL11A (BCL11Ae) to disrupt expression of a transcriptional repressor, BCL11A (Friedland). gRNAs were designed to target the GATA1 binding motif that is in the erythroid specific enhancer of BCL11A that is in the +58 DHS region of intron 2 (i.e., the GATA1 binding motif in BCL11Ae), where the +58 DHS enhancer region comprises the sequence set forth in SEQ ID NO:968. Targeting domain sequences of gRNAs that were designed to target disruption of the GATA1 binding motif in BCL11Ae, include, but are not limited to, the sequences set forth in SEQ ID NOs:952-955.

gRNA Modifications

[0155] The activity, stability, or other characteristics of gRNAs can be altered through the incorporation of certain modifications. As one example, transiently expressed or delivered nucleic

acids can be prone to degradation by, e.g., cellular nucleases. Accordingly, the gRNAs described herein can contain one or more modified nucleosides or nucleotides which introduce stability toward nucleases. While not wishing to be bound by theory it is also believed that certain modified gRNAs described herein can exhibit a reduced innate immune response when introduced into cells. Those of skill in the art will be aware of certain cellular responses commonly observed in cells, e.g., mammalian cells, in response to exogenous nucleic acids, particularly those of viral or bacterial origin. Such responses, which can include induction of cytokine expression and release and cell death, may be reduced or eliminated altogether by the modifications presented herein.
[0156] Certain exemplary modifications discussed in this section can be included at any position within a gRNA sequence including, without limitation at or near the 5′ end (e.g., within 1-10, 1-5, or 1-2 nucleotides of the 5′ end) and/or at or near the 3′ end (e.g., within 1-10, 1-5, or 1-2 nucleotides of the 3′ end). In some cases, modifications are positioned within functional motifs, such as the repeat-anti-repeat duplex of a Cas9 gRNA, a stem loop structure of a Cas9 or Cpf1 gRNA, and/or a targeting domain of a gRNA.
[0157] As one example, the 5′ end of a gRNA can include a eukaryotic mRNA cap structure or cap analog (e.g., a G(5′)ppp(5′)G cap analog, a m7G(5′)ppp(5′)G cap analog, or a 3′-O-Me-m7G(5′)ppp(5′)G anti reverse cap analog (ARCA)), as shown below:
##STR00001##
The cap or cap analog can be included during either chemical synthesis or in vitro transcription of the gRNA.
[0158] Along similar lines, the 5′ end of the gRNA can lack a 5′ triphosphate group. For instance, in vitro transcribed gRNAs can be phosphatase-treated (e.g., using calf intestinal alkaline phosphatase) to remove a 5′ triphosphate group.
[0159] Another common modification involves the addition, at the 3′ end of a gRNA, of a plurality (e.g., 1-10, 10-20, or 25-200) of adenine (A) residues referred to as a polyA tract. The polyA tract can be added to a gRNA during chemical synthesis, following in vitro transcription using a polyadenosine polymerase (e.g., *E. coli* Poly(A)Polymerase), or in vivo by means of a polyadenylation sequence, as described in Maeder.
[0160] It should be noted that the modifications described herein can be combined in any suitable manner, e.g., a gRNA, whether transcribed in vivo from a DNA vector, or in vitro transcribed gRNA, can include either or both of a 5′ cap structure or cap analog and a 3′ polyA tract.
[0161] Guide RNAs can be modified at a 3′ terminal U ribose. For example, the two terminal hydroxyl groups of the U ribose can be oxidized to aldehyde groups and a concomitant opening of the ribose ring to afford a modified nucleoside as shown below:
##STR00002##
wherein “U” can be an unmodified or modified uridine.
[0162] The 3′ terminal U ribose can be modified with a 2′3′ cyclic phosphate as shown below:
##STR00003##
wherein “U” can be an unmodified or modified uridine.
[0163] Guide RNAs can contain 3′ nucleotides which can be stabilized against degradation, e.g., by incorporating one or more of the modified nucleotides described herein. In certain embodiments, uridines can be replaced with modified uridines, e.g., 5-(2-amino)propyl uridine, and 5-bromo uridine, or with any of the modified uridines described herein; adenosines and guanosines can be replaced with modified adenosines and guanosines, e.g., with modifications at the 8-position, e.g., 8-bromo guanosine, or with any of the modified adenosines or guanosines described herein.
[0164] In certain embodiments, sugar-modified ribonucleotides can be incorporated into the gRNA, e.g., wherein the 2′ OH-group is replaced by a group selected from H, —OR, —R (wherein R can be, e.g., alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar), halo, —SH, —SR (wherein R can be, e.g., alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar), amino (wherein amino can be, e.g., NH.sub.2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino,

diheteroarylamino, or amino acid); or cyano (—CN). In certain embodiments, the phosphate backbone can be modified as described herein, e.g., with a phosphorothioate (PhTx) group. In certain embodiments, one or more of the nucleotides of the gRNA can each independently be a modified or unmodified nucleotide including, but not limited to 2′-sugar modified, such as, 2′-O-methyl, 2′-O-methoxyethyl, or 2′-Fluoro modified including, e.g., 2′-F or 2′-O-methyl, adenosine (A), 2′-F or 2′-O-methyl, cytidine (C), 2′-F or 2′-O-methyl, uridine (U), 2′-F or 2′-O-methyl, thymidine (T), 2′-F or 2′-O-methyl, guanosine (G), 2′-O-methoxyethyl-5-methyluridine (Teo), 2′-O-methoxyethyladenosine (Aeo), 2′-O-methoxyethyl-5-methylcytidine (m5Ceo), and any combinations thereof.

[0165] In certain embodiments, gRNAs as used herein may be modified or unmodified gRNAs. In certain embodiments, a gRNA may include one or more phosphorothioate modifications. In certain embodiments, the one or more phosphorothioate modifications may be at the 5′ end, 3′ end, or a combination thereof. In certain embodiments, a gRNA may include one or more 2-o-methyl modifications. In certain embodiments, the one or more 2-o-methyl modifications may be at the 5′ end, 3′ end, or a combination thereof. In certain embodiments, a gRNA may include one or more 2-o-methyl modifications, one or more phosphorothioate modifications, or a combination thereof. In certain embodiments, a gRNA comprising a targeting domain set forth in Table 12 may comprise one or more 2-o-methyl modifications, one or more phosphorothioate modifications, or a combination thereof. In certain embodiments, a gRNA modification may comprise one or more phosphorodithioate (PS2) linkage modifications.

[0166] Guide RNAs can also include “locked” nucleic acids (LNA) in which the 2′ OH-group can be connected, e.g., by a C1-6 alkylene or C1-6 heteroalkylene bridge, to the 4′ carbon of the same ribose sugar. Any suitable moiety can be used to provide such bridges, include without limitation methylene, propylene, ether, or amino bridges; O-amino (wherein amino can be, e.g., NH.sub.2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroarylamino, ethylenediamine, or polyamino) and aminoalkoxy or O(CH.sub.2).sub.n-amino (wherein amino can be, e.g., NH.sub.2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroarylamino, ethylenediamine, or polyamino).

[0167] In certain embodiments, a gRNA can include a modified nucleotide which is multicyclic (e.g., tricyclo; and “unlocked” forms, such as glycol nucleic acid (GNA) (e.g., R-GNA or S-GNA, where ribose is replaced by glycol units attached to phosphodiester bonds), or threose nucleic acid (TNA, where ribose is replaced with α-L-threofuranosyl-(3′.fwdarw.2′)).

[0168] Generally, gRNAs include the sugar group ribose, which is a 5-membered ring having an oxygen. Exemplary modified gRNAs can include, without limitation, replacement of the oxygen in ribose (e.g., with sulfur (S), selenium (Se), or alkylene, such as, e.g., methylene or ethylene); addition of a double bond (e.g., to replace ribose with cyclopentenyl or cyclohexenyl); ring contraction of ribose (e.g., to form a 4-membered ring of cyclobutane or oxetane); ring expansion of ribose (e.g., to form a 6- or 7-membered ring having an additional carbon or heteroatom, such as for example, anhydrohexitol, altritol, mannitol, cyclohexanyl, cyclohexenyl, and morpholino that also has a phosphoramidate backbone). Although the majority of sugar analog alterations are localized to the 2′ position, other sites are amenable to modification, including the 4′ position. In certain embodiments, a gRNA comprises a 4′-S, 4′-Se or a 4′-C-aminomethyl-2′-O-Me modification.

[0169] In certain embodiments, deaza nucleotides, e.g., 7-deaza-adenosine, can be incorporated into the gRNA. In certain embodiments, O- and N-alkylated nucleotides, e.g., N6-methyl adenosine, can be incorporated into the gRNA. In certain embodiments, one or more or all of the nucleotides in a gRNA are deoxynucleotides.

Dead gRNA Molecules

[0170] Dead guide RNA (dgRNA) molecules according to the present disclosure include, but are not limited to, dead guide RNA molecules that are configured such that they do not provide an

RNA guided-nuclease cleavage event. For example, dead guide RNA molecules may comprise a targeting domain comprising 15 nucleotides or fewer in length. Dead guide RNAs may be generated by removing the 5′ end of a gRNA sequence, which results in a truncated targeting domain sequence. For example, if a gRNA sequence, configured to provide a cleavage event, has a targeting domain sequence that is 20 nucleotides in length, a dead guide RNA may be created by removing 5 nucleotides from the 5′ end of the gRNA sequence. In certain embodiments, the dgRNA may be configured to bind (or associate with) a nucleic acid sequence within or proximal to a target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae) to be edited. For example, any of the dgRNAs set forth in Table 10 may be employed to bind a nucleic acid sequence proximal to the 13 nt target region. In certain embodiments, proximal to may denote the region within 10, 25, 50, 100, or 200 nucleotides of a target region (e.g., the 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae). In certain embodiments, the dead guide RNA is not configured to recruit an exogenous trans-acting factor to a target region. In certain embodiments, the dgRNA is configured such that it does not provide a DNA cleavage event when complexed with an RNA-guided nuclease. Skilled artisans will appreciate that dead guide RNA molecules may be designed to comprise targeting domains complementary to regions proximal to or within a target region in a target nucleic acid. In certain embodiments, dead guide RNAs comprise targeting domain sequences that are complementary to the transcription strand or non-transcription strand of double stranded DNA. The dgRNAs herein may include modifications at the 5′ and 3′ end of the dgRNA as described for guide RNAs in the section “gRNA modifications” herein. For example, in certain embodiments, dead guide RNAs may include an anti-reverse cap analog (ARCA) at the 5′ end of the RNA. In certain embodiments, dgRNAs may include a polyA tail at the 3′ end.

RNA-Guided Nucleases

[0171] RNA-guided nucleases according to the present disclosure include, but are not limited to, naturally-occurring Class 2 CRISPR nucleases such as Cas9, and Cpf1, as well as other nucleases derived or obtained therefrom. In functional terms, RNA-guided nucleases are defined as those nucleases that: (a) interact with (e.g., complex with) a gRNA; and (b) together with the gRNA, associate with, and optionally cleave or modify, a target region of a DNA that includes (i) a sequence complementary to the targeting domain of the gRNA and, optionally, (ii) an additional sequence referred to as a “protospacer adjacent motif,” or “PAM,” which is described in greater detail below. As the following examples will illustrate, RNA-guided nucleases can be defined, in broad terms, by their PAM specificity and cleavage activity, even though variations may exist between individual RNA-guided nucleases that share the same PAM specificity or cleavage activity. Skilled artisans will appreciate that some aspects of the present disclosure relate to systems, methods and compositions that can be implemented using any suitable RNA-guided nuclease having a certain PAM specificity and/or cleavage activity. For this reason, unless otherwise specified, the term RNA-guided nuclease should be understood as a generic term, and not limited to any particular type (e.g., Cas9 vs. Cpf1), species (e.g., *S. pyogenes* vs. *S. aureus*) or variation (e.g., full-length vs. truncated or split; naturally-occurring PAM specificity vs. engineered PAM specificity, etc.) of RNA-guided nuclease.

[0172] Various RNA-guided nucleases may require different sequential relationships between PAMs and protospacers. In general, Cas9s recognize PAM sequences that are 3′ of the protospacer.

[0173] Cpf1, on the other hand, generally recognizes PAM sequences that are 5′ of the protospacer.

[0174] In addition to recognizing specific sequential orientations of PAMs and protospacers, RNA-guided nucleases can also recognize specific PAM sequences. *S. aureus* Cas9, for instance, recognizes a PAM sequence of NNGRRT or NNGRRV. *S. pyogenes* Cas9 recognizes NGG PAM sequences. And *F. novicida* Cpf1 recognizes a TTN PAM sequence. PAM sequences have been identified for a variety of RNA-guided nucleases, and a strategy for identifying novel PAM sequences has been described by Shmakov 2015. It should also be noted that engineered RNA-

guided nucleases can have PAM specificities that differ from the PAM specificities of reference molecules (for instance, in the case of an engineered RNA-guided nuclease, the reference molecule may be the naturally occurring variant from which the RNA-guided nuclease is derived, or the naturally occurring variant having the greatest amino acid sequence homology to the engineered RNA-guided nuclease). Examples of PAMs that may be used according to the embodiments herein include, without limitation, the sequences set forth in SEQ ID NOs:199-205.

[0175] In addition to their PAM specificity, RNA-guided nucleases can be characterized by their DNA cleavage activity: naturally-occurring RNA-guided nucleases typically form DSBs in target nucleic acids, but engineered variants have been produced that generate only SSBs (discussed above; see also Ran 2013, incorporated by reference herein), or that do not cut at all.

Cas9

[0176] Crystal structures have been determined for *S. pyogenes* Cas9 (Jinek 2014), and for *S. aureus* Cas9 in complex with a unimolecular guide RNA and a target DNA (Nishimasu 2014; Anders 2014; and Nishimasu 2015).

[0177] A naturally occurring Cas9 protein comprises two lobes: a recognition (REC) lobe and a nuclease (NUC) lobe; each of which comprise particular structural and/or functional domains. The REC lobe comprises an arginine-rich bridge helix (BH) domain, and at least one REC domain (e.g., a REC1 domain and, optionally, a REC2 domain). The REC lobe does not share structural similarity with other known proteins, indicating that it is a unique functional domain. While not wishing to be bound by any theory, mutational analyses suggest specific functional roles for the BH and REC domains: the BH domain appears to play a role in gRNA:DNA recognition, while the REC domain is thought to interact with the repeat:anti-repeat duplex of the gRNA and to mediate the formation of the Cas9/gRNA complex.

[0178] The NUC lobe comprises a RuvC domain, an HNH domain, and a PAM-interacting (PI) domain. The RuvC domain shares structural similarity to retroviral integrase superfamily members and cleaves the non-complementary (i.e., bottom) strand of the target nucleic acid. It may be formed from two or more split RuvC motifs (such as RuvC I, RuvCII, and RuvCIII in *S. pyogenes* and *S. aureus*). The HNH domain, meanwhile, is structurally similar to HNN endonuclease motifs, and cleaves the complementary (i.e., top) strand of the target nucleic acid. The PI domain, as its name suggests, contributes to PAM specificity. Examples of polypeptide sequences encoding Cas9 RuvC-like and Cas9 HNH-like domains that may be used according to the embodiments herein are set forth in SEQ ID NOs:15-23, 52-123 (RuvC-like domains) and SEQ ID NOs:24-28, 124-198 (HNH-like domains).

[0179] While certain functions of Cas9 are linked to (but not necessarily fully determined by) the specific domains set forth above, these and other functions may be mediated or influenced by other Cas9 domains, or by multiple domains on either lobe. For instance, in *S. pyogenes* Cas9, as described in Nishimasu 2014, the repeat:antirepeat duplex of the gRNA falls into a groove between the REC and NUC lobes, and nucleotides in the duplex interact with amino acids in the BH, PI, and REC domains. Some nucleotides in the first stem loop structure also interact with amino acids in multiple domains (PI, BH and REC1), as do some nucleotides in the second and third stem loops (RuvC and PI domains). Examples of polypeptide sequences encoding Cas9 molecules that may be used according to the embodiments herein are set forth in SEQ ID NOs:1-2, 4-6, 12, 14.

Cpf1

[0180] The crystal structure of Acidaminococcus sp. Cpf1 in complex with crRNA and a double-stranded (ds) DNA target including a TTTN PAM sequence has been solved (Yamano 2016, incorporated by reference herein). Cpf1, like Cas9, has two lobes: a REC (recognition) lobe, and a NUC (nuclease) lobe. The REC lobe includes REC1 and REC2 domains, which lack similarity to any known protein structures. The NUC lobe, meanwhile, includes three RuvC domains (RuvC-I, -II and -III) and a BH domain. However, in contrast to Cas9, the Cpf1 REC lobe lacks an HNH domain, and includes other domains that also lack similarity to known protein structures: a

structurally unique PI domain, three Wedge (WED) domains (WED-I, -II and -III), and a nuclease (Nuc) domain.

[0181] While Cas9 and Cpf1 share similarities in structure and function, it should be appreciated that certain Cpf1 activities are mediated by structural domains that are not analogous to any Cas9 domains. For instance, cleavage of the complementary strand of the target DNA appears to be mediated by the Nuc domain, which differs sequentially and spatially from the HNH domain of Cas9. Additionally, the non-targeting portion of Cpf1 gRNA (the handle) adopts a pseudoknot structure, rather than a stem loop structure formed by the repeat:antirepeat duplex in Cas9 gRNAs.

Modifications of RNA-Guided Nucleases

[0182] The RNA-guided nucleases described above have activities and properties that can be useful in a variety of applications, but the skilled artisan will appreciate that RNA-guided nucleases can also be modified in certain instances, to alter cleavage activity, PAM specificity, or other structural or functional features.

[0183] Turning first to modifications that alter cleavage activity, mutations that reduce or eliminate the activity of domains within the NUC lobe have been described above. Exemplary mutations that may be made in the RuvC domains, in the Cas9 HNH domain, or in the Cpf1 Nuc domain are described in Ran 2013 and Yamano 2016, as well as in Cotta-Ramusino. In general, mutations that reduce or eliminate activity in one of the two nuclease domains result in RNA-guided nucleases with nickase activity, but it should be noted that the type of nickase activity varies depending on which domain is inactivated. As one example, inactivation of a RuvC domain of a Cas9 will result in a nickase that cleaves the complementary or top strand as shown below (where C denotes the site of cleavage).

[0184] On the other hand, inactivation of a Cas9 HNH domain results in a nickase that cleaves the bottom or non-complementary strand.

[0185] Modifications of PAM specificity relative to naturally occurring Cas9 reference molecules has been described for both *S. pyogenes* (Kleinstiver 2015a) and *S. aureus* (Kleinstiver 2015b). Modifications that improve the targeting fidelity of Cas9 have also been described (Kleinstiver 2016). Each of these references is incorporated by reference herein.

[0186] RNA-guided nucleases have been split into two or more parts (see, e.g., Zetsche 2015a; Fine 2015; both incorporated by reference).

[0187] RNA-guided nucleases can be, in certain embodiments, size-optimized or truncated, for instance via one or more deletions that reduce the size of the nuclease while still retaining gRNA association, target and PAM recognition, and cleavage activities. In certain embodiments, RNA guided nucleases are bound, covalently or non-covalently, to another polypeptide, nucleotide, or other structure, optionally by means of a linker. Exemplary bound nucleases and linkers are described by Guilinger 2014, which is incorporated by reference herein.

[0188] RNA-guided nucleases also optionally include a tag, such as, but not limited to, a nuclear localization signal to facilitate movement of RNA-guided nuclease protein into the nucleus. In certain embodiments, the RNA-guided nuclease can incorporate C- and/or N-terminal nuclear localization signals. Nuclear localization sequences are known in the art and are described in Maeder and elsewhere.

[0189] The foregoing list of modifications is intended to be exemplary in nature, and the skilled artisan will appreciate, in view of the instant disclosure, that other modifications may be possible or desirable in certain applications. For brevity, therefore, exemplary systems, methods and compositions of the present disclosure are presented with reference to particular RNA-guided nucleases, but it should be understood that the RNA-guided nucleases used may be modified in ways that do not alter their operating principles. Such modifications are within the scope of the present disclosure.

RNA-Guided Helicases

[0190] RNA-guided helicases according to the present disclosure include, but are not limited to,

naturally-occurring RNA-guided helicases that are capable of unwinding nucleic acid. As discussed supra, catalytically active RNA-guided nucleases cleave or modify a target region of DNA. It has also been shown that certain RNA-guided nucleases, such as Cas9, also have helicase activity that enables them to unwind nucleic acid. In certain embodiments, the RNA-guided helicases according to the present disclosure may be any of the RNA-nucleases described herein and supra in the section entitled "RNA-guided nucleases." In certain embodiments, the RNA-guided nuclease is not configured to recruit an exogenous trans-acting factor to a target region. In certain embodiments, an RNA-guided helicase may be an RNA-guided nuclease configured to lack nuclease activity. For example, in certain embodiments, an RNA-guided helicase may be a catalytically inactive RNA-guided nuclease that lacks nuclease activity, but still retains its helicase activity. In certain embodiments, an RNA-guided nuclease may be mutated to abolish its nuclease activity (e.g., dead Cas9), creating a catalytically inactive RNA-guided nuclease that is unable to cleave nucleic acid, but which can still unwind DNA. In certain embodiments, an RNA-guided helicase may be complexed with any of the dead guide RNAs as described herein. For example, a catalytically active RNA-guided helicase (e.g., Cas9 or Cpf1) may form an RNP complex with a dead guide RNA, resulting in a catalytically inactive dead RNP (dRNP). In certain embodiments, a catalytically inactive RNA-guided helicase (e.g., dead Cas9) and a dead guide RNA may form a dRNP. These dRNPs, although incapable of providing a cleavage event, still retain their helicase activity that is important for unwinding nucleic acid.

Nucleic Acids Encoding RNA-Guided Nucleases

[0191] Nucleic acids encoding RNA-guided nucleases, e.g., Cas9, Cpf1 or functional fragments thereof, are provided herein. Examples of nucleic acid sequences encoding Cas9 molecules that may be used according to the embodiments herein are set forth in SEQ ID NOs:3, 7-11, 13. Exemplary nucleic acids encoding RNA-guided nucleases have been described previously (see, e.g., Cong 2013; Wang 2013; Mali 2013; Jinek 2012).

[0192] In some cases, a nucleic acid encoding an RNA-guided nuclease can be a synthetic nucleic acid sequence. For example, the synthetic nucleic acid molecule can be chemically modified. In certain embodiments, an mRNA encoding an RNA-guided nuclease will have one or more (e.g., all) of the following properties: it can be capped; polyadenylated; and substituted with 5-methylcytidine and/or pseudouridine.

[0193] Synthetic nucleic acid sequences can also be codon optimized, e.g., at least one non-common codon or less-common codon has been replaced by a common codon. For example, the synthetic nucleic acid can direct the synthesis of an optimized messenger mRNA, e.g., optimized for expression in a mammalian expression system, e.g., described herein. Examples of codon optimized Cas9 coding sequences are presented in Cotta-Ramusino.

[0194] In addition, or alternatively, a nucleic acid encoding an RNA-guided nuclease may comprise a nuclear localization sequence (NLS). Nuclear localization sequences are known in the art.

Functional Analysis of Candidate Molecules

[0195] Candidate RNA-guided nucleases, gRNAs, and complexes thereof, can be evaluated by standard methods known in the art (see, e.g., Cotta-Ramusino). The stability of RNP complexes may be evaluated by differential scanning fluorimetry, as described below.

Differential Scanning Fluorimetry (DSF)

[0196] The thermostability of ribonucleoprotein (RNP) complexes comprising gRNAs and RNA-guided nucleases can be measured via DSF. The DSF technique measures the thermostability of a protein, which can increase under favorable conditions such as the addition of a binding RNA molecule, e.g., a gRNA.

[0197] A DSF assay can be performed according to any suitable protocol, and can be employed in any suitable setting, including without limitation (a) testing different conditions (e.g., different stoichiometric ratios of gRNA: RNA-guided nuclease protein, different buffer solutions, etc.) to identify optimal conditions for RNP formation; and (b) testing modifications (e.g., chemical

modifications, alterations of sequence, etc.) of an RNA-guided nuclease and/or a gRNA to identify those modifications that improve RNP formation or stability. One readout of a DSF assay is a shift in melting temperature of the RNP complex; a relatively high shift suggests that the RNP complex is more stable (and may thus have greater activity or more favorable kinetics of formation, kinetics of degradation, or another functional characteristic) relative to a reference RNP complex characterized by a lower shift. When the DSF assay is deployed as a screening tool, a threshold melting temperature shift may be specified, so that the output is one or more RNPs having a melting temperature shift at or above the threshold. For instance, the threshold can be 5-10° C. (e.g., 5°, 6°, 7° 8°, 9°, 10°) or more, and the output may be one or more RNPs characterized by a melting temperature shift greater than or equal to the threshold.

[0198] Two non-limiting examples of DSF assay conditions are set forth below:

[0199] To determine the best solution to form RNP complexes, a fixed concentration (e.g., 2 M) of Cas9 in water+10× SYPRO Orange® (Life Technologies cat #S-6650) is dispensed into a 384 well plate. An equimolar amount of gRNA diluted in solutions with varied pH and salt is then added. After incubating at room temperature for 10′ and brief centrifugation to remove any bubbles, a Bio-Rad CFX384™ Real-Time System C1000 Touch™ Thermal Cycler with the Bio-Rad CFX Manager software is used to run a gradient from 20° C. to 90° C. with a 1° C. increase in temperature every 10 seconds.

[0200] The second assay consists of mixing various concentrations of gRNA with fixed concentration (e.g., 2 M) Cas9 in optimal buffer from assay 1 above and incubating (e.g., at RT for 10′) in a 384 well plate. An equal volume of optimal buffer+10× SYPRO Orange® (Life Technologies cat #S-6650) is added and the plate sealed with Microseal® B adhesive (MSB-1001). Following brief centrifugation to remove any bubbles, a Bio-Rad CFX384™ Real-Time System C1000 Touch™ Thermal Cycler with the Bio-Rad CFX Manager software is used to run a gradient from 20° C. to 90° C. with a 1° C. increase in temperature every 10 seconds.

Genome Editing Strategies

[0201] The genome editing systems described above are used, in various embodiments of the present disclosure, to generate edits in (i.e., to alter) targeted regions of DNA within or obtained from a cell. Various strategies are described herein to generate particular edits, and these strategies are generally described in terms of the desired repair outcome, the number and positioning of individual edits (e.g., SSBs or DSBs), and the target sites of such edits.

[0202] Genome editing strategies that involve the formation of SSBs or DSBs are characterized by repair outcomes including: (a) deletion of all or part of a targeted region; (b) insertion into or replacement of all or part of a targeted region; or (c) interruption of all or part of a targeted region. This grouping is not intended to be limiting, or to be binding to any particular theory or model, and is offered solely for economy of presentation. Skilled artisans will appreciate that the listed outcomes are not mutually exclusive and that some repairs may result in other outcomes. The description of a particular editing strategy or method should not be understood to require a particular repair outcome unless otherwise specified.

[0203] Replacement of a targeted region generally involves the replacement of all or part of the existing sequence within the targeted region with a homologous sequence, for instance through gene correction or gene conversion, two repair outcomes that are mediated by HDR pathways. HDR is promoted by the use of a donor template, which can be single-stranded or double stranded, as described in greater detail below. Single or double stranded templates can be exogenous, in which case they will promote gene correction, or they can be endogenous (e.g., a homologous sequence within the cellular genome), to promote gene conversion. Exogenous templates can have asymmetric overhangs (i.e., the portion of the template that is complementary to the site of the DSB may be offset in a 3′ or 5′ direction, rather than being centered within the donor template), for instance as described by Richardson 2016 (incorporated by reference herein). In instances where the template is single stranded, it can correspond to either the complementary (top) or non-

complementary (bottom) strand of the targeted region.

[0204] Gene conversion and gene correction are facilitated, in some cases, by the formation of one or more nicks in or around the targeted region, as described in Ran and Cotta-Ramusino. In some cases, a dual-nickase strategy is used to form two offset SSBs that, in turn, form a single DSB having an overhang (e.g., a 5′ overhang).

[0205] Interruption and/or deletion of all or part of a targeted sequence can be achieved by a variety of repair outcomes. As one example, a sequence can be deleted by simultaneously generating two or more DSBs that flank a targeted region, which is then excised when the DSBs are repaired, as is described in Maeder for the LCA10 mutation. As another example, a sequence can be interrupted by a deletion generated by formation of a double strand break with single-stranded overhangs, followed by exonucleolytic processing of the overhangs prior to repair.

[0206] One specific subset of target sequence interruptions is mediated by the formation of an indel within the targeted sequence, where the repair outcome is typically mediated by NHEJ pathways (including Alt-NHEJ). NHEJ is referred to as an “error prone” repair pathway because of its association with indel mutations. In some cases, however, a DSB is repaired by NHEJ without alteration of the sequence around it (a so-called “perfect” or “scarless” repair); this generally requires the two ends of the DSB to be perfectly ligated. Indels, meanwhile, are thought to arise from enzymatic processing of free DNA ends before they are ligated that adds and/or removes nucleotides from either or both strands of either or both free ends.

[0207] Because the enzymatic processing of free DSB ends may be stochastic in nature, indel mutations tend to be variable, occurring along a distribution, and can be influenced by a variety of factors, including the specific target site, the cell type used, the genome editing strategy used, etc. Even so, it is possible to draw limited generalizations about indel formation: deletions formed by repair of a single DSB are most commonly in the 1-50 bp range, but can reach greater than 100-200 bp. Insertions formed by repair of a single DSB tend to be shorter and often include short duplications of the sequence immediately surrounding the break site. However, it is possible to obtain large insertions, and in these cases, the inserted sequence has often been traced to other regions of the genome or to plasmid DNA present in the cells.

[0208] Indel mutations—and genome editing systems configured to produce indels—are useful for interrupting target sequences, for example, when the generation of a specific final sequence is not required and/or where a frameshift mutation would be tolerated. They can also be useful in settings where particular sequences are preferred, insofar as the certain sequences desired tend to occur preferentially from the repair of an SSB or DSB at a given site. Indel mutations are also a useful tool for evaluating or screening the activity of particular genome editing systems and their components. In these and other settings, indels can be characterized by (a) their relative and absolute frequencies in the genomes of cells contacted with genome editing systems and (b) the distribution of numerical differences relative to the unedited sequence, e.g., ±1, ±2, ±3, etc. As one example, in a lead-finding setting, multiple gRNAs can be screened to identify those gRNAs that most efficiently drive cutting at a target site based on an indel readout under controlled conditions. Guides that produce indels at or above a threshold frequency, or that produce a particular distribution of indels, can be selected for further study and development. Indel frequency and distribution can also be useful as a readout for evaluating different genome editing system implementations or formulations and delivery methods, for instance by keeping the gRNA constant and varying certain other reaction conditions or delivery methods.

Multiplex Strategies

[0209] Genome editing systems according to this disclosure may also be employed for multiplex gene editing to generate two or more DSBs, either in the same locus or in different loci. Any of the RNA-guided nucleases and gRNAs disclosed herein may be used in genome editing systems for multiplex gene editing. Strategies for editing that involve the formation of multiple DSBs, or SSBs, are described in, for instance, Cotta-Ramusino.

[0210] As disclosed herein, multiple gRNAs may be used in genome editing systems to introduce alterations (e.g., deletions, insertions) into the 13 nt target region of HBG1 and/or HBG2. In certain embodiments, one or more gRNAs comprising a targeting domain set forth in SEQ ID NOs:251-901, 940-942 may be used to introduce alterations in the 13 nt target region of HBG1 and/or HBG2. In other embodiments, multiple gRNAs may be used in genome editing systems to introduce alterations into the GATA1 binding motif in BCL11Ae. In certain embodiments, one or more gRNAs comprising a targeting domain set forth in SEQ ID NOs:952-955 may be used to introduce alterations in the GATA1 binding motif in BCL11Ae. Multiple gRNAs may also be used in genome editing systems to introduce alterations into the GATA1 binding motif in BCL11Ae and the 13 nt target region of HBG1 and/or HBG2. In certain embodiments, one or more gRNAs comprising a targeting domain set forth in SEQ ID NOs:952-955 may be used to introduce alterations in the GATA1 binding motif in BCL11Ae and one or more gRNAs comprising a targeting domain set forth in SEQ ID NOs:251-901, 940-942 may be used to introduce alterations in the 13 nt target region of HBG1 and/or HBG2.

Donor Template Design

[0211] Donor template design is described in detail in the literature, for instance in Cotta-Ramusino. DNA oligomer donor templates (oligodeoxynucleotides or ODNs), which can be single stranded (ssODNs) or double-stranded (dsODNs), can be used to facilitate HDR-based repair of DSBs or to boost overall editing rate, and are particularly useful for introducing alterations into a target DNA sequence, inserting a new sequence into the target sequence, or replacing the target sequence altogether.

[0212] Whether single-stranded or double stranded, donor templates generally include regions that are homologous to regions of DNA within or near (e.g., flanking or adjoining) a target sequence to be cleaved. These homologous regions are referred to here as “homology arms,” and are illustrated schematically below: [0213] [5′ homology arm]-[replacement sequence]-[3′ homology arm].

[0214] The homology arms can have any suitable length (including 0 nucleotides if only one homology arm is used), and 3′ and 5′ homology arms can have the same length, or can differ in length. The selection of appropriate homology arm lengths can be influenced by a variety of factors, such as the desire to avoid homologies or microhomologies with certain sequences such as Alu repeats or other very common elements. For example, a 5′ homology arm can be shortened to avoid a sequence repeat element. In other embodiments, a 3′ homology arm can be shortened to avoid a sequence repeat element. In some embodiments, both the 5′ and the 3′ homology arms can be shortened to avoid including certain sequence repeat elements. In addition, some homology arm designs can improve the efficiency of editing or increase the frequency of a desired repair outcome. For example, Richardson 2016, which is incorporated by reference herein, found that the relative asymmetry of 3′ and 5′ homology arms of single stranded donor templates influenced repair rates and/or outcomes.

[0215] Replacement sequences in donor templates have been described elsewhere, including in Cotta-Ramusino et al. A replacement sequence can be any suitable length (including zero nucleotides, where the desired repair outcome is a deletion), and typically includes one, two, three or more sequence modifications relative to the naturally-occurring sequence within a cell in which editing is desired. One common sequence modification involves the alteration of the naturally-occurring sequence to repair a mutation that is related to a disease or condition of which treatment is desired. Another common sequence modification involves the alteration of one or more sequences that are complementary to, or then, the PAM sequence of the RNA-guided nuclease or the targeting domain of the gRNA(s) being used to generate an SSB or DSB, to reduce or eliminate repeated cleavage of the target site after the replacement sequence has been incorporated into the target site.

[0216] Where a linear ssODN is used, it can be configured to (i) anneal to the nicked strand of the target nucleic acid, (ii) anneal to the intact strand of the target nucleic acid, (iii) anneal to the plus

strand of the target nucleic acid, and/or (iv) anneal to the minus strand of the target nucleic acid. An ssODN may have any suitable length, e.g., about, at least, or no more than 100-150 or 150-200 nucleotides (e.g., 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, or 200 nucleotides).

[0217] It should be noted that a template nucleic acid can also be a nucleic acid vector, such as a viral genome or circular double stranded DNA, e.g., a plasmid. Nucleic acid vectors comprising donor templates can include other coding or non-coding elements. For example, a template nucleic acid can be delivered as part of a viral genome (e.g., in an AAV or lentiviral genome) that includes certain genomic backbone elements (e.g., inverted terminal repeats, in the case of an AAV genome) and optionally includes additional sequences coding for a gRNA and/or an RNA-guided nuclease. In certain embodiments, the donor template can be adjacent to, or flanked by, target sites recognized by one or more gRNAs, to facilitate the formation of free DSBs on one or both ends of the donor template that can participate in repair of corresponding SSBs or DSBs formed in cellular DNA using the same gRNAs. Exemplary nucleic acid vectors suitable for use as donor templates are described in Cotta-Ramusino, which is incorporated by reference.

[0218] Whatever format is used, a template nucleic acid can be designed to avoid undesirable sequences. In certain embodiments, one or both homology arms can be shortened to avoid overlap with certain sequence repeat elements, e.g., Alu repeats, LINE elements, etc.

[0219] In certain embodiments, silent, non-pathogenic SNPs may be included in the ssODN donor template to allow for identification of a gene editing event.

[0220] In certain embodiments, a donor template may be a non-specific template that is non-homologous to regions of DNA within or near a target sequence to be cleaved. In certain embodiments, donor templates for use in targeting the GATA1 binding motif in BCL11Ae may include, without limitation, non-target specific templates that are nonhomologous to regions of DNA within or near the GATA1 binding motif in BCL11Ae. In certain embodiments, donor templates for use in targeting the 13 nt target region may include, without limitation, non-target specific templates that are nonhomologous to regions of DNA within or near the 13 nt target region.

Target Cells

[0221] Genome editing systems according to this disclosure can be used to manipulate or alter a cell, e.g., to edit or alter a target nucleic acid. The manipulating can occur, in various embodiments, in vivo or ex vivo.

[0222] A variety of cell types can be manipulated or altered according to the embodiments of this disclosure, and in some cases, such as in vivo applications, a plurality of cell types are altered or manipulated, for example by delivering genome editing systems according to this disclosure to a plurality of cell types. In other cases, however, it may be desirable to limit manipulation or alteration to a particular cell type or types. For instance, it can be desirable in some instances to edit a cell with limited differentiation potential or a terminally differentiated cell, such as a photoreceptor cell in the case of Maeder, in which modification of a genotype is expected to result in a change in cell phenotype. In other cases, however, it may be desirable to edit a less differentiated, multipotent or pluripotent, stem or progenitor cell. By way of example, the cell may be an embryonic stem cell, induced pluripotent stem cell (iPSC), hematopoietic stem/progenitor cell (HSPC), or other stem or progenitor cell type that differentiates into a cell type of relevance to a given application or indication.

[0223] As a corollary, the cell being altered or manipulated is, variously, a dividing cell or a non-dividing cell, depending on the cell type(s) being targeted and/or the desired editing outcome.

[0224] When cells are manipulated or altered ex vivo, the cells can be used (e.g., administered to a subject) immediately, or they can be maintained or stored for later use. Those of skill in the art will appreciate that cells can be maintained in culture or stored (e.g., frozen in liquid nitrogen) using any suitable method known in the art.

Implementation of Genome Editing Systems: Delivery, Formulations, and Routes of Administration

[0225] As discussed above, the genome editing systems of this disclosure can be implemented in any suitable manner, meaning that the components of such systems, including without limitation the RNA-guided nuclease, gRNA, and optional donor template nucleic acid, can be delivered, formulated, or administered in any suitable form or combination of forms that results in the transduction, expression or introduction of a genome editing system and/or causes a desired repair outcome in a cell, tissue or subject. Tables 3 and 4 set forth several, non-limiting examples of genome editing system implementations. Those of skill in the art will appreciate, however, that these listings are not comprehensive, and that other implementations are possible. With reference to Table 3 in particular, the table lists several exemplary implementations of a genome editing system comprising a single gRNA and an optional donor template. However, genome editing systems according to this disclosure can incorporate multiple gRNAs, multiple RNA-guided nucleases, and other components such as proteins, and a variety of implementations will be evident to the skilled artisan based on the principles illustrated in the table. In the table, [N/A] indicates that the genome editing system does not include the indicated component.

TABLE-US-00003 TABLE 3 Genome editing components RNA-guided Donor Nuclease gRNA Template Comments Protein RNA [N/A] An RNA-guided nuclease protein complexed with a gRNA molecule (an RNP complex) Protein RNA DNA An RNP complex as described above plus a single-stranded or double stranded donor template. Protein DNA [N/A] An RNA-guided nuclease protein plus gRNA transcribed from DNA. Protein DNA DNA An RNA-guided nuclease protein plus gRNA-encoding DNA and a separate DNA donor template. Protein DNA An RNA-guided nuclease protein and a single DNA encoding both a gRNA and a donor template. DNA A DNA or DNA vector encoding an RNA-guided nuclease, a gRNA and a donor template. DNA DNA [N/A] Two separate DNAs, or two separate DNA vectors, encoding the RNA-guided nuclease and the gRNA, respectively. DNA DNA DNA Three separate DNAs, or three separate DNA vectors, encoding the RNA-guided nuclease, the gRNA and the donor template, respectively. DNA [N/A] A DNA or DNA vector encoding an RNA-guided nuclease and a gRNA DNA DNA A first DNA or DNA vector encoding an RNA-guided nuclease and a gRNA, and a second DNA or DNA vector encoding a donor template. DNA DNA A first DNA or DNA vector encoding an RNA-guided nuclease and second DNA or DNA vector encoding a gRNA and a donor template. DNA A first DNA or DNA vector DNA encoding an RNA-guided nuclease and a donor template, and a second DNA or DNA vector encoding a gRNA DNA A DNA or DNA vector encoding RNA an RNA-guided nuclease and a donor template, and a gRNA RNA [N/A] An RNA or RNA vector encoding an RNA-guided nuclease and comprising a gRNA RNA DNA An RNA or RNA vector encoding an RNA-guided nuclease and comprising a gRNA, and a DNA or DNA vector encoding a donor template.

[0226] Table 4 summarizes various delivery methods for the components of genome editing systems, as described herein. Again, the listing is intended to be exemplary rather than limiting.

TABLE-US-00004 TABLE 4 Delivery into Non- Type of Dividing Duration of Genome Molecule Delivery Vector/Mode Cells Expression Integration Delivered Physical (e.g., electroporation, YES Transient NO Nucleic Acids particle gun, Calcium Phosphate and Proteins transfection, cell compression or squeezing) Viral Retrovirus NO Stable YES RNA Lentivirus YES Stable YES/NO with RNA modifications Adenovirus YES Transient NO DNA Adeno- YES Stable NO DNA Associated Virus (AAV) Vaccinia Virus YES Very NO DNA Transient Herpes Simplex YES Stable NO DNA Virus Non-Viral Cationic YES Transient Depends on Nucleic Acids Liposomes what is and Proteins delivered Polymeric YES Transient Depends on Nucleic Acids Nanoparticles what is and Proteins delivered Biological Attenuated YES Transient NO Nucleic Acids Non-Viral Bacteria Delivery Engineered YES Transient NO Nucleic Acids Vehicles Bacteriophages Mammalian YES Transient NO Nucleic Acids Virus-like Particles Biological YES Transient NO Nucleic Acids liposomes: Erythrocyte Ghosts and Exosomes

Nucleic Acid-Based Delivery of Genome Editing Systems

[0227] Nucleic acids encoding the various elements of a genome editing system according to the present disclosure can be administered to subjects or delivered into cells by art-known methods or as described herein. For example, RNA-guided nuclease-encoding and/or gRNA-encoding DNA, as well as donor template nucleic acids can be delivered by, e.g., vectors (e.g., viral or non-viral vectors), non-vector based methods (e.g., using naked DNA or DNA complexes), or a combination thereof.
[0228] Nucleic acids encoding genome editing systems or components thereof can be delivered directly to cells as naked DNA or RNA, for instance by means of transfection or electroporation, or can be conjugated to molecules (e.g., N-acetylgalactosamine) promoting uptake by the target cells (e.g., erythrocytes, HSCs). Nucleic acid vectors, such as the vectors summarized in Table 4, can also be used.
[0229] Nucleic acid vectors can comprise one or more sequences encoding genome editing system components, such as an RNA-guided nuclease, a gRNA and/or a donor template. A vector can also comprise a sequence encoding a signal peptide (e.g., for nuclear localization, nucleolar localization, or mitochondrial localization), associated with (e.g., inserted into or fused to) a sequence coding for a protein. As one example, a nucleic acid vectors can include a Cas9 coding sequence that includes one or more nuclear localization sequences (e.g., a nuclear localization sequence from SV40).
[0230] The nucleic acid vector can also include any suitable number of regulatory/control elements, e.g., promoters, enhancers, introns, polyadenylation signals, Kozak consensus sequences, or internal ribosome entry sites (IRES). These elements are well known in the art, and are described in Cotta-Ramusino.
[0231] Nucleic acid vectors according to this disclosure include recombinant viral vectors. Exemplary viral vectors are set forth in Table 4, and additional suitable viral vectors and their use and production are described in Cotta-Ramusino. Other viral vectors known in the art can also be used. In addition, viral particles can be used to deliver genome editing system components in nucleic acid and/or peptide form. For example, “empty” viral particles can be assembled to contain any suitable cargo. Viral vectors and viral particles can also be engineered to incorporate targeting ligands to alter target tissue specificity.
[0232] In addition to viral vectors, non-viral vectors can be used to deliver nucleic acids encoding genome editing systems according to the present disclosure. One important category of non-viral nucleic acid vectors are nanoparticles, which can be organic or inorganic. Nanoparticles are well known in the art, and are summarized in Cotta-Ramusino. Any suitable nanoparticle design can be used to deliver genome editing system components or nucleic acids encoding such components. For instance, organic (e.g., lipid and/or polymer) nanoparticles can be suitable for use as delivery vehicles in certain embodiments of this disclosure. Exemplary lipids for use in nanoparticle formulations, and/or gene transfer are shown in Table 5, and Table 6 lists exemplary polymers for use in gene transfer and/or nanoparticle formulations.
TABLE-US-00005 TABLE 5 Lipids used for gene transfer Lipid Abbreviation Feature 1,2-Dioleoyl-sn-glycero-3-phosphatidylcholine DOPC Helper 1,2-Dioleoyl-sn-glycero-3-phosphatidylethanolamine DOPE Helper Cholesterol Helper N-[1-(2,3-Dioleyloxy)propyl]N,N,N-trimethylammonium chloride DOTMA Cationic 1,2-Dioleoyloxy-3-trimethylammonium-propane DOTAP Cationic Dioctadecylamidoglycylspermine DOGS Cationic N-(3-Aminopropyl)-N,N-dimethyl-2,3-bis(dodecyloxy)-1- GAP-DLRIE Cationic propanaminium bromide Cetyltrimethylammonium bromide CTAB Cationic 6-Lauroxyhexyl ornithinate LHON Cationic 1-(2,3-Dioleoyloxypropyl)-2,4,6-trimethylpyridinium 2Oc Cationic 2,3-Dioleyloxy-N-[2(sperminecarboxamido-ethyl]-N,N-dimethyl- DOSPA Cationic 1-propanaminium trifluoroacetate 1,2-Dioleyl-3-trimethylammonium-propane DOPA Cationic N-(2-Hydroxyethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1- MDRIE Cationic propanaminium bromide Dimyristooxypropyl dimethyl hydroxyethyl ammonium bromide DMRI Cationic 3β-[N-(N′,N′-Dimethylaminoethane)-carbamoyl]cholesterol DC-Chol Cationic Bis-guanidium-tren-cholesterol BGTC Cationic 1,3-

Diodeoxy-2-(6-carboxy-spermyl)-propylamide DOSPER Cationic Dimethyloctadecylammonium bromide DDAB Cationic Dioctadecylamidoglicylspermidin DSL Cationic rac-[(2,3-Dioctadecyloxypropyl)(2-hydroxyethyl)]- CLIP-1 Cationic dimethylammonium chloride rac-[2(2,3-Dihexadecyloxypropyl- CLIP-6 Cationic oxymethyloxy)ethyl]trimethylammonium bromide Ethyldimyristoylphosphatidylcholine EDMPC Cationic 1,2-Distearyloxy-N,N-dimethyl-3-aminopropane DSDMA Cationic 1,2-Dimyristoyl-trimethylammonium propane DMTAP Cationic O,O′-Dimyristyl-N-lysyl aspartate DMKE Cationic 1,2-Distearoyl-sn-glycero-3-ethylphosphocholine DSEPC Cationic N-Palmitoyl D-erythro-sphingosyl carbamoyl-spermine CCS Cationic N-t-Butyl-N0-tetradecyl-3-tetradecylaminopropionamidine diC14-amidine Cationic Octadecenolyoxy[ethyl-2-heptadecenyl-3 hydroxyethyl] DOTIM Cationic imidazolinium chloride N1-Cholesteryloxycarbonyl-3,7-diazanonane-1,9-diamine CDAN Cationic 2-(3-[Bis(3-amino-propyl)-amino]propylamino)-N- RPR209120 Cationic ditetradecylcarbamoylme-ethyl-acetamide 1,2-dilinoleyloxy-3-dimethylaminopropane DLinDMA Cationic 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane DLin-KC2- Cationic DMA dilinoleyl-methyl-4-dimethylaminobutyrate DLin-MC3- Cationic DMA

TABLE-US-00006 TABLE 6 Polymers used for gene transfer Polymer Abbreviation Poly(ethylene)glycol PEG Polyethylenimine PEI Dithiobis(succinimidylpropionate) DSP Dimethyl-3,3′-dithiobispropionimidate DTBP Poly(ethylene imine) biscarbamate PEIC Poly(L-lysine) PLL Histidine modified PLL Poly(N-vinylpyrrolidone) PVP Poly(propylenimine) PPI Poly(amidoamine) PAMAM Poly(amido ethylenimine) SS-PAEI Triethylenetetramine TETA Poly(β-aminoester) Poly(4-hydroxy-L-proline ester) PHP Poly(allylamine) Poly(α-[4-aminobutyl]-L-glycolic acid) PAGA Poly(D,L-lactic-co-glycolic acid) PLGA Poly(N-ethyl-4-vinylpyridinium bromide) Poly(phosphazene)s PPZ Poly(phosphoester)s PPE Poly(phosphoramidate)s PPA Poly(N-2-hydroxypropylmethacrylamide) pHPMA Poly (2-(dimethylamino)ethyl methacrylate) pDMAEMA Poly(2-aminoethyl propylene phosphate) PPE-EA Chitosan Galactosylated chitosan N-Dodacylated chitosan Histone Collagen Dextran-spermine D-SPM

[0233] Non-viral vectors optionally include targeting modifications to improve uptake and/or selectively target certain cell types. These targeting modifications can include e.g., cell specific antigens, monoclonal antibodies, single chain antibodies, aptamers, polymers, sugars (e.g., N-acetylgalactosamine (GalNAc)), and cell penetrating peptides. Such vectors also optionally use fusogenic and endosome-destabilizing peptides/polymers, undergo acid-triggered conformational changes (e.g., to accelerate endosomal escape of the cargo), and/or incorporate a stimuli-cleavable polymer, e.g., for release in a cellular compartment. For example, disulfide-based cationic polymers that are cleaved in the reducing cellular environment can be used.

[0234] In certain embodiments, one or more nucleic acid molecules (e.g., DNA molecules) other than the components of a genome editing system, e.g., the RNA-guided nuclease component and/or the gRNA component described herein, are delivered. In certain embodiments, the nucleic acid molecule is delivered at the same time as one or more of the components of the Genome editing system. In certain embodiments, the nucleic acid molecule is delivered before or after (e.g., less than about 30 minutes, 1 hour, 2 hours, 3 hours, 6 hours, 9 hours, 12 hours, 1 day, 2 days, 3 days, 1 week, 2 weeks, or 4 weeks) one or more of the components of the Genome editing system are delivered. In certain embodiments, the nucleic acid molecule is delivered by a different means than one or more of the components of the genome editing system, e.g., the RNA-guided nuclease component and/or the gRNA component, are delivered. The nucleic acid molecule can be delivered by any of the delivery methods described herein. For example, the nucleic acid molecule can be delivered by a viral vector, e.g., an integration-deficient lentivirus, and the RNA-guided nuclease molecule component and/or the gRNA component can be delivered by electroporation, e.g., such that the toxicity caused by nucleic acids (e.g., DNAs) can be reduced. In certain embodiments, the nucleic acid molecule encodes a therapeutic protein, e.g., a protein described herein. In certain embodiments, the nucleic acid molecule encodes an RNA molecule, e.g., an RNA molecule

described herein.

Delivery of RNPs and/or RNA Encoding Genome Editing System Components

[0235] RNPs (complexes of gRNAs and RNA-guided nucleases) and/or RNAs encoding RNA-guided nucleases and/or gRNAs, can be delivered into cells or administered to subjects by art-known methods, some of which are described in Cotta-Ramusino. In vitro, RNA-guided nuclease-encoding and/or gRNA-encoding RNA can be delivered, e.g., by microinjection, electroporation, transient cell compression or squeezing (see, e.g., Lee **2012**). Lipid-mediated transfection, peptide-mediated delivery, GalNAc- or other conjugate-mediated delivery, and combinations thereof, can also be used for delivery in vitro and in vivo. A protective, interactive, non-condensing (PINC) system may be used for delivery.

[0236] In vitro delivery via electroporation comprises mixing the cells with the RNA encoding RNA-guided nucleases and/or gRNAs, with or without donor template nucleic acid molecules, in a cartridge, chamber or cuvette and applying one or more electrical impulses of defined duration and amplitude. Systems and protocols for electroporation are known in the art, and any suitable electroporation tool and/or protocol can be used in connection with the various embodiments of this disclosure.

Route of Administration

[0237] Genome editing systems, or cells altered or manipulated using such systems, can be administered to subjects by any suitable mode or route, whether local or systemic. Systemic modes of administration include oral and parenteral routes. Parenteral routes include, by way of example, intravenous, intramarrow, intrarterial, intramuscular, intradermal, subcutaneous, intranasal, and intraperitoneal routes. Components administered systemically can be modified or formulated to target, e.g., HSCs, hematopoietic stem/progenitor cells, or erythroid progenitors or precursor cells.

[0238] Local modes of administration include, by way of example, intramarrow injection into the trabecular bone or intrafemoral injection into the marrow space, and infusion into the portal vein. In certain embodiments, significantly smaller amounts of the components (compared with systemic approaches) can exert an effect when administered locally (for example, directly into the bone marrow) compared to when administered systemically (for example, intravenously). Local modes of administration can reduce or eliminate the incidence of potentially toxic side effects that may occur when therapeutically effective amounts of a component are administered systemically.

[0239] Administration can be provided as a periodic bolus (for example, intravenously) or as continuous infusion from an internal reservoir or from an external reservoir (for example, from an intravenous bag or implantable pump). Components can be administered locally, for example, by continuous release from a sustained release drug delivery device.

[0240] In addition, components can be formulated to permit release over a prolonged period of time. A release system can include a matrix of a biodegradable material or a material which releases the incorporated components by diffusion. The components can be homogeneously or heterogeneously distributed within the release system. A variety of release systems can be useful, however, the choice of the appropriate system will depend upon rate of release required by a particular application. Both non-degradable and degradable release systems can be used. Suitable release systems include polymers and polymeric matrices, non-polymeric matrices, or inorganic and organic excipients and diluents such as, but not limited to, calcium carbonate and sugar (for example, trehalose). Release systems may be natural or synthetic. However, synthetic release systems are preferred because generally they are more reliable, more reproducible and produce more defined release profiles. The release system material can be selected so that components having different molecular weights are released by diffusion through or degradation of the material.

[0241] Representative synthetic, biodegradable polymers include, for example: polyamides such as poly(amino acids) and poly(peptides); polyesters such as poly(lactic acid), poly(glycolic acid), poly(lactic-co-glycolic acid), and poly(caprolactone); poly(anhydrides); polyorthoesters; polycarbonates; and chemical derivatives thereof (substitutions, additions of chemical groups, for

example, alkyl, alkylene, hydroxylations, oxidations, and other modifications routinely made by those skilled in the art), copolymers and mixtures thereof. Representative synthetic, non-degradable polymers include, for example: polyethers such as poly(ethylene oxide), poly(ethylene glycol), and poly(tetramethylene oxide); vinyl polymers-polyacrylates and polymethacrylates such as methyl, ethyl, other alkyl, hydroxyethyl methacrylate, acrylic and methacrylic acids, and others such as poly(vinyl alcohol), poly(vinyl pyrolidone), and poly(vinyl acetate); poly(urethanes); cellulose and its derivatives such as alkyl, hydroxyalkyl, ethers, esters, nitrocellulose, and various cellulose acetates; polysiloxanes; and any chemical derivatives thereof (substitutions, additions of chemical groups, for example, alkyl, alkylene, hydroxylations, oxidations, and other modifications routinely made by those skilled in the art), copolymers and mixtures thereof.

[0242] Poly(lactide-co-glycolide) microsphere can also be used. Typically the microspheres are composed of a polymer of lactic acid and glycolic acid, which are structured to form hollow spheres. The spheres can be approximately 15-30 microns in diameter and can be loaded with components described herein. In some embodiments, genome editing systems, system components and/or nucleic acids encoding system components, are delivered with a block copolymer such as a poloxamer or a poloxamine.

Multi-Modal or Differential Delivery of Components

[0243] Skilled artisans will appreciate, in view of the instant disclosure, that different components of genome editing systems disclosed herein can be delivered together or separately and simultaneously or nonsimultaneously. Separate and/or asynchronous delivery of genome editing system components can be particularly desirable to provide temporal or spatial control over the function of genome editing systems and to limit certain effects caused by their activity.

[0244] Different or differential modes as used herein refer to modes of delivery that confer different pharmacodynamic or pharmacokinetic properties on the subject component molecule, e.g., a RNA-guided nuclease molecule, gRNA, template nucleic acid, or payload. For example, the modes of delivery can result in different tissue distribution, different half-life, or different temporal distribution, e.g., in a selected compartment, tissue, or organ.

[0245] Some modes of delivery, e.g., delivery by a nucleic acid vector that persists in a cell, or in progeny of a cell, e.g., by autonomous replication or insertion into cellular nucleic acid, result in more persistent expression of and presence of a component. Examples include viral, e.g., AAV or lentivirus, delivery.

[0246] By way of example, the components of a genome editing system, e.g., a RNA-guided nuclease and a gRNA, can be delivered by modes that differ in terms of resulting half-life or persistent of the delivered component the body, or in a particular compartment, tissue or organ. In certain embodiments, a gRNA can be delivered by such modes. The RNA-guided nuclease molecule component can be delivered by a mode which results in less persistence or less exposure to the body or a particular compartment or tissue or organ.

[0247] More generally, in certain embodiments, a first mode of delivery is used to deliver a first component and a second mode of delivery is used to deliver a second component. The first mode of delivery confers a first pharmacodynamic or pharmacokinetic property. The first pharmacodynamic property can be, e.g., distribution, persistence, or exposure, of the component, or of a nucleic acid that encodes the component, in the body, a compartment, tissue or organ. The second mode of delivery confers a second pharmacodynamic or pharmacokinetic property. The second pharmacodynamic property can be, e.g., distribution, persistence, or exposure, of the component, or of a nucleic acid that encodes the component, in the body, a compartment, tissue or organ.

[0248] In certain embodiments, the first pharmacodynamic or pharmacokinetic property, e.g., distribution, persistence or exposure, is more limited than the second pharmacodynamic or pharmacokinetic property.

[0249] In certain embodiments, the first mode of delivery is selected to optimize, e.g., minimize, a pharmacodynamic or pharmacokinetic property, e.g., distribution, persistence or exposure.

[0250] In certain embodiments, the second mode of delivery is selected to optimize, e.g., maximize, a pharmacodynamic or pharmacokinetic property, e.g., distribution, persistence or exposure.
[0251] In certain embodiments, the first mode of delivery comprises the use of a relatively persistent element, e.g., a nucleic acid, e.g., a plasmid or viral vector, e.g., an AAV or lentivirus. As such vectors are relatively persistent product transcribed from them would be relatively persistent.
[0252] In certain embodiments, the second mode of delivery comprises a relatively transient element, e.g., an RNA or protein.
[0253] In certain embodiments, the first component comprises gRNA, and the delivery mode is relatively persistent, e.g., the gRNA is transcribed from a plasmid or viral vector, e.g., an AAV or lentivirus. Transcription of these genes would be of little physiological consequence because the genes do not encode for a protein product, and the gRNAs are incapable of acting in isolation. The second component, a RNA-guided nuclease molecule, is delivered in a transient manner, for example as mRNA or as protein, ensuring that the full RNA-guided nuclease molecule/gRNA complex is only present and active for a short period of time.
[0254] Furthermore, the components can be delivered in different molecular form or with different delivery vectors that complement one another to enhance safety and tissue specificity.
[0255] Use of differential delivery modes can enhance performance, safety, and/or efficacy, e.g., the likelihood of an eventual off-target modification can be reduced. Delivery of immunogenic components, e.g., Cas9 molecules, by less persistent modes can reduce immunogenicity, as peptides from the bacterially-derived Cas enzyme are displayed on the surface of the cell by MHC molecules. A two-part delivery system can alleviate these drawbacks.
[0256] Differential delivery modes can be used to deliver components to different, but overlapping target regions. The formation active complex is minimized outside the overlap of the target regions. Thus, in certain embodiments, a first component, e.g., a gRNA is delivered by a first delivery mode that results in a first spatial, e.g., tissue, distribution. A second component, e.g., a RNA-guided nuclease molecule is delivered by a second delivery mode that results in a second spatial, e.g., tissue, distribution. In certain embodiments, the first mode comprises a first element selected from a liposome, nanoparticle, e.g., polymeric nanoparticle, and a nucleic acid, e.g., viral vector. The second mode comprises a second element selected from the group. In certain embodiments, the first mode of delivery comprises a first targeting element, e.g., a cell specific receptor or an antibody, and the second mode of delivery does not include that element. In certain embodiments, the second mode of delivery comprises a second targeting element, e.g., a second cell specific receptor or second antibody.
[0257] When the RNA-guided nuclease molecule is delivered in a virus delivery vector, a liposome, or polymeric nanoparticle, there is the potential for delivery to and therapeutic activity in multiple tissues, when it may be desirable to only target a single tissue. A two-part delivery system can resolve this challenge and enhance tissue specificity. If the gRNA and the RNA-guided nuclease molecule are packaged in separated delivery vehicles with distinct but overlapping tissue tropism, the fully functional complex is only be formed in the tissue that is targeted by both vectors.

EXAMPLES

[0258] The principles and embodiments described above are further illustrated by the non-limiting examples that follow:

Example 1: Screening of *S. pyogenes* gRNAs Delivered to K562 Cells as Ribonucleoprotein Complexes for Use in Causing 13 nt Deletions in HBG1 and HBG2 Regulatory Regions

[0259] gRNAs targeting a 26 nt fragment spanning and including the 13 nucleotides at the 13 nt target region of HBG1 and HBG2 were designed by standard methods. After gRNAs were designed in silico and tiered, a subset of the gRNAs were selected and screened for activity and specificity in human K562 cells. The gRNAs selected for screening are set forth in Table 7. Briefly, gRNAs were

in vitro transcribed and then complexed with *S. pyogenes* wildtype (Wt) Cas9 protein to form ribonucleoprotein complexes (RNPs). The gRNAs complexed to *S. pyogenes* Cas9 protein were modified sgRNAs ((e.g., 5′ ARCA capped and 3′ polyA (**20**A) tail; Table 7) and target the HBG1 and HBG2 regulatory regions. To allow for direct comparison of the activity of these RNPs in K562 cells and human CD34.sup.+ cells, RNPs were first delivered to K562 cells by electroporation (Amaxa Nucleofector).

[0260] Three days after RNP electroporation, gDNA was extracted from K562 cells and then the HBG1 and HBG2 loci were PCR amplified from the gDNA. Gene editing was evaluated in the PCR products by T7E1 endonuclease assay analysis. Eight out of nine RNPs supported a high percentage of NHEJ. Sp37 RNP, the only gRNA shown to be active in human CD34.sup.+ cells (<10% editing in CD34.sup.+ cells) was highly active in K562 cells, with >60% indels detected at both HBG1 and HBG2 and eight cut in both the HBG1 and HBG2 targeted regions in the promoter sequences (FIG. **3**A).

TABLE-US-00007 TABLE 7 Selected gRNAs for screening in K562 cells or CD34.sup.+ cells Targeting Targeting domain domain Targeting Targeting sequence sequence domain domain plus PAM plus PAM gRNA sequence sequence (NGG) (NGG) ID (RNA) (DNA) (RNA) (DNA) Sense Sp9 GGCUAUUGG GGCTATTGG GGCUAUUGG GGCTATTGG Anti- UCAAGGCA TCAAGGCA UCAAGGCAA TCAAGGCAA sense (SEQ ID (SEQ ID GG GG NO: 277) NO: 910) (SEQ ID (SEQ ID NO: 920) NO: 930) Sp36 CAAGGCUAU CAAGGCTAT CAAGGCUAU CAAGGCTAT Anti- UGGUCAAGG TGGTCAAGG UGGUCAAGG TGGTCAAGG sense CA CA CAAGG CAAGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 338) NO: 911) NO: 921) NO: 931) Sp40 UGCCUUGUC TGCCTTGTC UGCCUUGUC TGCCTTGTC Anti- AAGGCUAU GAAGCTAT AAGGCUAUU AAGGCTATT sense (SEQ ID (SEQ ID GG GG NO: 327) NO: 912) (SEQ ID (SEQ ID NO: 922) NO: 932) Sp42 GUUUGCCUU GTTTGCCTT GUUUGCCUU GTTTGCCTT Anti- GUCAAGGCU GTCAAGGCT GUCAAGGCU GTCAAGGCT sense AU AT AUUGG ATTGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 299) NO: 913) NO: 923) NO: 933) Sp38 GACCAAUAG GACCAATAG GACCAAUAG GACCAATAG Sense CCUUGACA CCTTGACA CCUUGACAA CCTTGACAA (SEQ ID (SEQ ID GG GG NO: 276) NO: 914) (SEQ ID (SEQ ID NO: 924) NO: 934) Sp37 CUUGACCAA CTTGACCAA CUUGACCAA CTTGACCAA Sense UAGCCUUGA TAGCCTTGA UAGCCUUGA TAGCCTTGA CA CA CAAGG CAAGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 333) NO: 915) NO: 925) NO: 935) Sp43 GUCAAGGCU GTCAAGGCT GUCAAGGCU GTCAAGGCT Anti- AUUGGUCA ATTGGTCA AUUGGUCAA ATTGGTCAA sense (SEQ ID (SEQ ID GG GG NO: 278) NO: 916) (SEQ ID (SEQ ID NO: 926) NO: 936) Sp35 CUUGUCAAG CTTGTCAAG CUUGUCAAG CTTGTCAAG Anti- GCUAUUGGU GCTATTGGT GCUAUUGGU GCTATTGGT sense CA CA CAAGG CAAGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 339) NO: 917) NO: 927) NO: 937) Sp41 UCAAGUUUG TCAAGTTTG UCAAGUUUG TCAAGTTTG Anti- CCUUGUCA CCTTGTCA CCUUGUCAA CCTTGTCAA sense (SEQ ID (SEQ ID GG GG NO: 310) NO: 918) (SEQ ID (SEQ ID NO: 928) NO: 938) Sp34 UGGUCAAGU TGGTCAAGT UGGUCAAGU TGGTCAAGT Anti- UUGCCUUGU TTGCCTTGT UUGCCUUGU TTGCCTTGT sense CA CA CAAGG CAAGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 340) NO: 919) NO: 929) NO: 939) Sp85 AGUAUCCAG AGTATCCAG AGUAUCCAG AGTATCCAG Anti- UGAGGCCA TGAGGCCA UGAGGCCAG TGAGGCCAG sense (SEQ ID (SEQ ID GG GG NO: 940) NO: 943) (SEQ ID (SEQ ID NO: 946) NO: 949) SpA GGCAAGGCU GGCAAGGCT GGCAAGGCU GGCAAGGCT Sense GGCCAACCC GGCCAACCC GGCCAACCC GGCCAACCC AU AT AUGGG ATGGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 941) NO: 944) NO: 947) NO: 950) SpB UAUUUGCAU

TATTTGCAT UAUUUGCAU TATTTGCAT Sense UGAGAUAGU TGAGATAGT UGAGAUAGU TGAGATAGT GU GT GUGGG GTGGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 942) NO: 945) NO: 948) NO: 951)

[0261] The HBG1 and HBG2 PCR products for the K562 cells that were targeted with the eight active sgRNAs were then analyzed by DNA sequencing analysis and scored for insertions and deletions detected. The deletions were subdivided into precise 13 nt deletions at the target site, 13 nt target site inclusive and proximal small deletions (18-26 nt), 12 nt deletions (i.e., partial deletion) of the 13 nt target site, >26 nt deletions that span a portion of the HPFH target site, and other deletions, e.g., deletions proximal to but outside the HPFH target site. Seven of the eight sgRNAs targeted deletion of the 13 nt (HPFH mutation induction) (FIG. **3**B) for HBG1. At least five of the eight sgRNAs also supported targeted deletion of the 13 nt in HBG2 promoter region (FIG. **3**C). Note that DNA sequence results for HBG2 in cells treated with HBG Sp34 sgRNA were not available. These data indicate that Cas9 and sgRNA support precise induction of the 13 nt deletions. FIGS. **3**B-**3**C depict examples of the types of deletions observed in target sequences in HBG1.

Example 2: Cas9 RNP Containing gRNA Targeting the 13 nt Deletion Mutation Supports Gene Editing in Human Hematopoietic Stem/Progenitor Cells

[0262] Of the RNPs containing different gRNAs tested in human cord blood (CB) CD34.sup.+ cells, only Sp37 resulted in detectable editing at the target site in the HBG1 and HBG2 promoters as determined by T7E1 analysis of indels in HBG1 and HBG2 specific PCR products amplified from gDNA extracted from electroporated CB CD34.sup.+ cells from a three cord blood donors (FIG. **4**A). The average level of editing detected in cells electroporated with Cas9 protein complexed to Sp37 was 5±2% indels at HBG1 and 3±1% indels detected at HBG2 (3 separate experiments, and CB donors).

[0263] Next, three *S. pyogenes* gRNAs whose target sites are within the HBG promoter (Sp35, Sp36, Sp37) were complexed to wild-type *S. pyogenes* Cas9 protein to form ribonucleoprotein complexes. These HBG targeted RNPS were electroporated into CB CD34.sup.+ cells (n=3 donors) and adult mobilized peripheral blood (mPB) CD34.sup.+ cell donors (n=3 donors). Then the level of insertions/deletions at the target site was analyzed by T7E1 endonuclease analysis of the HBG2 PCR products amplified from genomic DNA extracted from the samples approximately 3 days after Cas9 RNP delivery. Each of these RNPs supported only low level gene editing in both the CB and adult CD34.sup.+ cells across 3 donors and 3 separate experiments (FIG. **4**B).

[0264] To increase gene editing and the occurrence of the 13 nt deletion at the target site, single strand deoxynucleotide donor repair templates (ssODNs) that encoded 87 nt and 89 nt of homology on each side of the targeted deletion site was generated. The ssODNs, either unmodified at the ends (i.e., ssODN1, SEQ ID NO:906, Table 8) or modified to contain phosphorothioates (PhTx) at the 5′ and 3′ ends (i.e., PhTx ssODN1, SEQ ID NO:909, Table 8). The ssODN was designed to ‘encode’ the 13 nt deletion with sequence homology arms engineered flanking this absent sequence to create a perfect deletion.

TABLE-US-00008 TABLE 8 Single strand deoxynucleotide donor repair templates (ssODN) SEQ ssODN ID ID NO Sequence ssODN1 904 **GGGTGCTTCCTTTTATTCTT** 5′ **CATCCCTAGCCAGCCGCCGG** homology **CCCCTGGCCTCACTGGATAC** arm **TCTAAGACTATTGGTCAAGT TTGCCTT** ssODN1 905 GTCAAGGCAAGGCTGGCCAA 3′ CCCATGGGTGGAGTTTAGCC homology AGGGACCGTTTCAGACAGAT arm ATTTGCATTGAGATAGTGTG GGGAAGGGG ssODN1 906 **GGGTGCTTCCTTTTATTCTT CATCCCTAGCCAGCCGCCGG CCCCTGGCCTCACTGGATAC TCTAAGACTATTGGTCAAGT TTGCCTT**GTCAAGGCAAGGC TGGCCAACCCATGGGTGGAG TTTAGCCAGGGACCGTTTCA GACAGATATTTGCATTGAGA TAGTGTGGGGAAGGGG PhTx 907 *__GGGTGCTTCCTTTTATTCT__ ssODN1 **TCATCCCTAGCCAGCCGCCG** 5′

GCCCCTGGCCTCACTGGATA homology **CTCTAAGACTATTGGTCAAG** arm **TTTGCCTT** PhTx 908 GTCAAGGCAAGGCTGGCCAA ssODN1 CCCATGGGTGGAGTTTAGCC 3′ AGGGACCGTTTCAGACAGAT homology ATTTGCATTGAGATAGTGTG arm GGGAAGGGG* PhTx 909 *__GGGTGCTTCCTTTTATTCT__ ssODN1 **TCATCCCTAGCCAGCCGCCG GCCCCTGGCCTCACTGGATA CTCTAAGACTATTGGTCAAG TTTGCCTT**GTCAAGGCAAGG CTGGCCAACCCATGGGTGGA GTTTAGCCAGGGACCGTTTC AGACAGATATTTGCATTGAG ATAGTGTGGGGAAGGGG* The homology arms flanking the deletion are indicated by bold [5′ homology arm] and underline [3′ homology arm]). Note the absence of the 13 bp sequence in ssODN1 and PhTx ssODN1. *Represents modification by phosphorothioate.

[0265] ssODN1 and PhTx ssODN1 were co-delivered with RNP targeting HBG containing the Sp37 gRNA (HBG Sp37 RNP) or HBG Sp35 (HBG Sp35 RNP) to CB CD34.sup.+ cells. Co-delivery of the ssODN donor encoding the 13 nt deletion with HBG Sp35 RNP or HBG Sp37 RNP led to a 6-fold and 5-fold increase in gene editing of the target site, respectively, as determined by T7E1 analysis of the HBG2 PCR product (FIG. **4**C). DNA sequencing analysis (Sanger sequencing) of the HBG2 PCR product indicated that 20% gene editing in cells that were treated with HBG Sp37 RNP and the PhTx modified ssODN1, with 15% deletions and 5% insertions (FIG. **4**C, lower left panel). Further analysis of the specific type and size of deletions at the target site revealed that 75% of the total deletions detected contained the 13 nt deletion (which included deletion at c. −110-−115 of the CAAT box in the proximal promoter), the absence of which is associated with elevation of HbF expression (FIG. **4**C, lower right panel). The remaining ¼ of deletions were partial deletions that did not span the full 13 nt deletion. These data indicate that co-delivery of a homologous ssODN that is engineered to have a deletion supported precise gene editing (deletion) at HBG in human CD34.sup.+ cells.

Example 3: Cas9 RNP Targeting the 13 nt Deletion Mutation Supports Gene Editing in Human Adult Mobilized Peripheral Blood Hematopoietic Stem/Progenitor Cells with Increased HBG Expression in Erythroblast Progeny

[0266] To determine whether editing HBG with Cas9 RNP complexed to Sp37 gRNA or Sp35 gRNA (i.e., the gRNAs that target the 13 nt deletion that is associated with HPFH) in the promoter of HBG supports an increase in HBG expression in erythroid progeny of edited CD34.sup.+ cells, human adult CD34.sup.+ cells from mobilized peripheral blood (mPB) were electroporated with the RNPs. Briefly, mPB CD34.sup.+ cells were prestimulated for 2 days with human cytokines and PGE2 in StemSpan SFEM and then electroporated with Cas9 protein precomplexed to Sp35 and Sp37, respectively. T7E1 analysis of HBG PCR product indicated ˜3% indels detected for mPB CD34.sup.+ cells treated with RNP complexed to Sp37 while no editing was detected for cells that were treated with RNP complexed to Sp35 (FIG. **5**A).

[0267] In order to increase gene editing at the target site and to increase the occurrence of the 13 nt deletion at the target site, PhTx ssODN1 (SEQ ID NO:909) was co-delivered with the precomplexed RNP targeting HBG containing the Sp37 gRNA. Co-delivery of the ssODN donor encoding the 13 nt deletion led to a nearly 2-fold increase in gene editing of the target site (FIG. **5**A). To determine whether editing HBG increases production of fetal hemoglobin in erythroid progeny of edited adult CD34.sup.+ cells, the cells were differentiated into erythroblasts by culture for up to 18 days in the presence of human cytokines (erythropoietin, SCF, IL3), human plasma (Octoplas), and other supplements (hydrocortisone, heparin, transferrin). Over the time course of differentiation, mRNA was collected to evaluate HBG gene expression in the erythroid progeny of RNP treated mPB CD34.sup.+ cells and donor matched negative (untreated) controls. By day 7 of differentiation, erythroblast progeny of human CD34.sup.+ cells that were treated with HBG Sp37 RNP and 13 nt deletion encoding ssODN (˜5% indels detected in gDNA from the bulk cell population by T7E1 analysis) exhibited a 2-fold increase in HBG mRNA production (FIG. **5**B).

Importantly, CD34.sup.+ cells that were electroporated with HBG RNP maintained their ex vivo hematopoietic activity (i.e., no difference in the quantity or diversity of erythroid and myeloid colonies compared to untreated donor matched CD34.sup.+ cell negative control), as determined in hematopoietic colony forming cell (CFC) assays (FIG. **6**A). Furthermore, the erythroblasts differentiated from RNP treated CD34.sup.+ cells maintained the kinetics of differentiation observed for donor matched untreated control cells as determined by flow analysis for acquisition of erythroid phenotype (% Glycophorin A.sup.+ cells) (FIG. **6**B). These data indicate that targeted disruption of HBG1/HBG2 proximal promoter region supported an increase in HBG expression in erythroid progeny of RNP treated adult hematopoietic stem/progenitor cells without altering differentiation potential.

Example 4: Cas9 RNP Targeting the HPFH Mutation Supports Gene Editing in Human Adult Mobilized Peripheral Blood Hematopoietic Stem/Progenitor Cells with Increased HBG Expression in Erythroblast Progeny

[0268] To determine whether co-delivery of paired nickase RNPs targeting HBG would increase targeted disruption of the proximal HBG promoter, mPB CD34.sup.+ cells were cultured for 2 days with human cytokines and PGE2 in StemSpan SFEM and then electroporated with *S. pyogenes* D10A Cas9 protein precomplexed to two gRNAs that target sites flanking the site of the 13 nt deletion. The targeting domain sequences for gRNAs used in nickase pairs in this example (including, without limitation, SpA, Sp85 and SpB) are presented in Table 7. D10A nickase pairs were selected such that the PAMs for the targets were oriented outward and the distance between the cut sites were <100 nt. gRNAs were complexed with D10A Cas9 protein to form RNP complexes and then human CD34.sup.+ cells and paired nickase were subject to electroporation. To determine whether co-delivery of an ssODN that encoded the 13 nt deletion would increase editing and introduction of the mutation into the cells, in some experiments, ssODN1 was added to the cell RNP mixture prior to electroporation. Approximately 3 days after electroporation, gDNA was extracted from the RNP treated cells and analyzed by T7E1 endonuclease assay and/or Sanger DNA sequencing of HBG2 PCR products amplified from the extracted gDNA. Of the three D10A nickase pairs tested, indels detected by T7E1 endonuclease analysis were increased for one nickase pair (gRNAs SpA+Sp85) samples for which ssODN1 was included (FIG. **7**A). DNA sequencing analysis was performed on limited samples shown in FIG. **7**A. DNA sequencing analysis showed up to ˜27% indels at the target site, with insertions as the dominant indel detected, followed by deletions of the targeted region (area between the cut sites of the paired nickases), and the 13 nt deletion mutation was also detected at a frequency of 2-3% when ssODN1 encoding the deletion was co-delivered (FIG. **7**B). Silent, non-pathogenic SNPs were included in the ssODN1 donor template, and were detected in the sequences that contained the 13 nt deletion, indicating that creation of the HFPH mutation occurred through an HDR event.

Example 5: D10A Paired RNPs Electroporated into Adult CD34.SUP.+ Cells Supports Induction of HbF Protein in Erythroid Progeny

[0269] To further optimize editing conditions in mPB CD34.sup.+ cells at the target site and to evaluate editing in additional human cell donors, human mPB CD34.sup.+ cells were electroporated with D10A Cas9 and WT Cas9 paired RNPs targeting HBG. The most efficient guide pair for both D10A Cas9 and WT Cas9 RNPs was Sp37+SpA, which supported >30% indels as determined by T7E1 endonuclease analysis of HBG2 PCR products (FIG. **8**A). Given that editing at both HBG1 and HBG2 could result in large deletions of HBG2 and the intergenic region between HBG2 and HBG1, indels were further characterized in order to capture local indels by T7E1 endonuclease assay and sequencing and large deletion by ddPCR analysis. Large deletions were detected in all samples at variable frequencies for both D10A Cas9 and WT Cas9 RNP nickase pairs (FIG. **8**B). Illumina sequencing analysis of indels correlated with indels determined by T7E1 analysis (FIG. **8**C-**8**D).

[0270] To determine whether CD34.sup.+ cells edited with dual nickases at the HBG promoter

gave rise to erythroid progeny with elevated HbF expression, donor matched RNP treated and untreated controls were induced toward erythroid differentiation and then evaluated for maintenance of indels during differentiation and for expression of HbF mRNA and protein. The level of editing (as determined by T7E1 endonuclease assay) was evaluated over the first 2 weeks of erythroid differentiation in the progeny of RNP treated cells prior to enucleation. Indels were detected in the erythroid progeny at every time point assayed suggesting that the editing that occurred in the CD34.sup.+ cells was maintained during erythroid differentiation and that edited CD34.sup.+ cells maintain erythroid differentiation potential.

[0271] The levels of HBG mRNA (day 10 of differentiation) and HbF protein (day 20-23 of differentiation) were quantified by ddPCR and HPLC analysis (according to the HPLC method described in Chang 2017 at pp. 143-44, incorporated by reference herein), respectively (FIG. **9**). A ˜2-fold increase (+40% in in HBG transcripts vs. unedited donor matched control) was observed for HBG:HBA ratio (data not shown) and the ratio of HbF/HbF+HbA (i.e., HBG mRNA/HGB+HBB mRNA) increased to 30% above the level detected in donor matched untreated control samples.

[0272] For the D10A Cas9 nickase pairs, upregulation of HbF mRNA and protein was detected in erythroid progeny (FIG. **9**). With respect to HbF protein analysis, two pairs supported 20% HbF induction for two D10A nickase pairs. No HbF upregulation was detected in erythroid progeny of WT Cas9 RNP treated CD34.sup.+ cells (data not shown).

Example 6: Increasing the Dose of RNP Increases Total Editing Efficiency in Human Adult CD34.SUP.+ Cells at the HBG Locus

[0273] The concentration of D10A Cas9 RNP for the nickase pair SpA+Sp85 was increased (2.5 μM standard concentration and 3.7 μM) and delivered to mPB CD34.sup.+ cells by electroporation. The increased RNP concentration supported an increase in indels at the HBG target site to >30% (FIG. **10**A) as determined by T7E1 endonuclease analysis of the HBG PCR product amplified for gDNA extracted 3 days after electroporation of CD34.sup.+ cells. Sequencing analysis indicated that increasing the RNP concentration increased insertions (FIG. **10**B). Erythroid progeny of RNP treated CD34.sup.+ cells also had an increase in HbF protein production (FIG. **10**C). Importantly, the hematopoietic colony forming potential was maintained after editing (FIG. **10**D). These cells were then transplanted into immunodeficient mice and their engraftment 1 month (FIG. **10**E) and 2 months (FIG. **10**F) after transplantation was evaluated by sampling the peripheral blood and measuring the percentage of human CD45.sup.+ cells. Early engraftment data showed no difference in engraftment between recipient cohorts of donor matched untreated controls (0 μM RNP) and mice transplanted with RNP treated cells. Furthermore, there was no difference in human blood lineage distribution (myeloid, B cell, T cell) within the human CD45.sup.+ fraction among cohorts at indicated time points (FIG. **10**G-H).

[0274] Two additional D10A nickase pairs were also tested in RNP dose response studies in adult mPB CD34.sup.+ cells (Sp37+SpA, Sp37+SpB). Here, mPB CD34.sup.+ cells were electroporated with D10A paired nickases delivered at 0, 2.5, and 3.75 M of total RNP. RNP treated cells were differentiated into erythroid progeny and the HbF protein levels (% HbF/HbF+HbA) were analyzed by HPLC analysis. The indel frequency detected in CD34.sup.+ cells was plotted with the HbF levels detected in erythroid progeny in order to correlate editing and HbF induction (FIG. **11**A). RNP treated and untreated control mPB CD34.sup.+ cells were also differentiated into colonies to evaluate ex vivo hematopoietic activity. Colony forming cell (CFC) activity was maintained for the progeny of RNP treated and donor matched untreated control CD34.sup.+ cells (FIG. **11**B). There was no difference in the percentage of human CD45.sup.+ cells in the mouse peripheral blood 1 month after transplantation and no difference in blood lineage distribution (FIG. **11**C-D) for cells exposed to different D10A RNP pairs at different doses compared to untreated donor matched control CD34.sup.+ cells.

Example 7: Co-Delivery of RNP Targeting the Erythroid Specific Enhancer of BCL11A and a Non-Specific (N) Single Strand Deoxynucleotide Sequence or Paired RNPs Increases Gene Editing in

Human CD34.SUP.+ Cells and Supports Induction of Fetal Hemoglobin Expression in Erythroid Progeny

[0275] Fetal hemoglobin expression can be induced through targeted disruption of the erythroid cell specific expression of a transcriptional repressor, BCL11A (Canvers 2015). One potential strategy to increase HbF expression through a gene editing strategy is to multiplex gene editing for introduction of 13 nt deletion associated in the HBG proximal promoter and also for targeted disruption of the GATA1 binding motif in the erythroid specific enhancer of BCL11A that is in the +58 DHS region of intron 2 of the BCL11A gene (FIG. **12**). In order to accomplish this multiplex strategy to increase HbF expression through multiplex gene editing, the effect of disruption of BCL11A erythroid enhancer (BCL11Ae) must first be determined as a single editing event.

[0276] In this experiment, CB CD34.sup.+ cells were electroporated with *S. pyogenes* WT Cas9 complexed to in vitro transcribed sgRNA targeting the GATA1 motif in the +58 DHS region of intron 2 of BCL11A gene (gRNA SpK, Table 9) (FIG. **13**A). To determine whether co-delivery of a non-target specific ssODN would increase editing of the target sequence, BCL11Ae RNP was co-delivered with ssODN (which is nonhomologous to the BCL11Ae target sequence, also called a non-specific ssODN) in CB CD34.sup.+ cells. T7E1 analysis of BCL11A erythroid enhancer PCR product from gDNA extracted from CB CD34.sup.+ cells treated with BCL11Ae RNP indicated that ˜5% indels was achieved (FIG. **13**A). Co-delivery of BCL11Ae RNP with a non-target specific ssODN increase in indels by 5-fold to 20% as detected by T7E1 endonuclease analysis. Illumina sequencing analysis indicated that >90% of edits had disruption of the GATA1 motif in the +DHS 58 region enhancer in intron 2 of the BCL11A gene (data not shown). To increase editing, human CB CD34.sup.+ cells were electroporated with WT Cas9 RNP (single gRNAs complexed to WT Cas9) or with WT Cas9 paired RNPs (paired gRNAs complexed to WT Cas9), so that the cut sites in each pair flank the target site for excision of the GATA1 motif (gRNAs SpC, SpK, SpM, SpN) (Table 9). Two of the single gRNAs and two pairs had >50% indels as determined by T7E1 endonuclease analysis (FIG. **13**B).

TABLE-US-00009 TABLE 9 Select gRNA sequences targeting BCL11A erythroid enhancer for screening in CD34.sup.+ cells Targeting Targeting domain domain Targeting Targeting sequence sequence domain domain plus PAM plus PAM gRNA sequence sequence (NGG) (NGG) ID (RNA) (DNA) (RNA) (DNA) Sense SpK CUAACAGUU CTAACAGTT CUAACAGUU CTAACAGTT Anti- GCUUUUAUC GCTTTTATC GCUUUUAUC GCTTTTATC sense AC AC ACAGG ACAGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 952) NO: 956) NO: 960) NO: 964) SpM GGGCGUGGG GGGCGTGGG GGGCGUGGG GGGCGTGGG Anti- UGGGGUAGA TGGGGTAGA UGGGGUAGA TGGGGTAGA sense AG AG GAGAG AGAGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 953) NO: 957) NO: 961) NO: 965) SpN CUCUUAGAC CTCTTAGAC CUCUUAGAC CTCTTAGAC Anti- AUAACACAC ATAACACAC AUAACACAC ATAACACAC sense CA CA CAGGG CAGGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 954) NO: 958) NO: 962) NO: 966) SpC AUCAGAGGC ATCAGAGGC AUCAGAGGC ATCAGAGGC Sense CAAACCCUU CAAACCCTT CAAACCCUU CAAACCCTT CC CC CCUGG CCTGG (SEQ ID (SEQ ID (SEQ ID (SEQ ID NO: 955) NO: 959) NO: 963) NO: 967)

[0277] Next, human adult bone marrow CD34.sup.+ cells were electroporated with the BCL11Ae RNP. DNA sequencing analysis of the BCL11A PCR product amplified from gDNA extracted from marrow CD34.sup.+ cells indicated 15% gene editing comprised of insertions and deletions (FIG. **14**A). Importantly, all deletions resulted in deletion of the GATA1 motif and all insertions disrupted GATA1 motif through addition of a small number of bp in the motif. CD34.sup.+ cells were plated into colony forming assays and the mixed hematopoietic colonies (GEMMs), which correspond to CD34.sup.+ cell clones, were picked. gDNA was isolated and analyzed by Illumina sequencing to quantify monoallelic and biallelic disruption of the target site. Most GEMMs differentiated from

the CD34.sup.+ cell clones had monoallelic disruption and biallelic disruption was also detected, with the overall indel rate ˜2/3 higher compared to what was detected in the bulk CD34.sup.+ cell population (FIG. **14**B). This was likely a reflection of the percentage of common myeloid progenitors (CMPs) that give rise to GEMMs that make up a larger fraction of the heterogeneous CD34.sup.+ cells versus the other lineages present, but not captured/differentiated in the short-term CFC assays. The RNP treated marrow CD34.sup.+ cells also maintained similar kinetics of erythroid maturation (enucleation, FIG. **14**C) and differentiation (phenotype acquisition, FIG. **14**D) compared to donor matched untreated control cells. Erythroid progeny of edited marrow CD34.sup.+ cells exhibited ˜5-fold increase in HbF induction as determined by flow cytometry analysis (FIG. **14**E).

[0278] Gene editing and induction of fetal hemoglobin was also evaluated in human adult mPB CD34.sup.+ cells. Co-delivery of BCL11Ae RNP and a non-specific ssODN supported ˜20% indels at the target site (FIG. **15**A). To evaluate early induction of fetal hemoglobin in erythroid progeny of edited cells, mPB CD34.sup.+ cells were differentiated into erythroblasts and induction of fetal hemoglobin transcription (HBG mRNA) was evaluated by qRT-PCR analysis. The erythroid progeny of BCL11Ae RNP treated CD34.sup.+ cells exhibited a 2-fold induction of HBG mRNA compared to untreated controls, suggesting induction of fetal hemoglobin expression (FIG. **15**B). The RNP treated marrow CD34.sup.+ cells also maintained similar kinetics of differentiation (phenotype acquisition, FIG. **15**C) compared to donor matched untreated control cells.

Example 8: Co-Delivery of *S. pyogenes* Cas9 Protein Complexed to a Truncated (15-Mer) "Dead" gRNA Increases Editing of the HBG Promoter Region in Adult Mobilized Peripheral Blood (mPB) CD34.SUP.+ Cells

[0279] Delivery of a single wild-type (WT) ribonucleoprotein (RNP) (e.g., WT Cas9 protein complexed to Sp37 guide RNA (gRNA), see Table 10) targeting the HBG promoter supports ˜1.5% indels editing in human CD34.sup.+ cells (see International Patent Application No. PCT/US17/22377 by Gori et al., filed Mar. 14, 2017, which is incorporated by reference herein). It was hypothesized that co-delivery of a dead RNP (dRNP), comprised of a catalytically active WT Cas9 protein and a truncated dead gRNA (dgRNA) that binds proximal to the target site in the HBG promoter (−110 nt) would increase the accessibility of a WT RNP (e.g., catalytically active WT Cas9 complexed to a full-length gRNA (e.g., Sp37 gRNA, see Table 10)) to the target site. Therefore, dead guide RNAs (dgRNAs) were designed that target the regions proximal to the −110 target site in the HBG promoter and have a truncated targeting domain (see FIG. **16** and Table 10).

[0280] To increase editing at the target site in mobilized peripheral blood (mPB) CD34.sup.+ cells, WT Cas9 protein was complexed to a truncated gRNA (i.e., dead (d)RNA15-mer version of wild-type SpA, which was truncated (t) at the 5′ end of the gRNA sequence (tSpA dgRNA, see Table 10); tSpA dRNP). RNP comprised of dgRNA complexed to WT Cas9 is able to bind to sequence but does not cut genomic DNA homologous to the gRNA sequence. To determine the optimal ratio of dead RNP:WT RNP for the assay, tSpA dRNP:Sp37 WT RNP mixed at different ratios were Nucleofected™ into mPB CD34.sup.+ cells, keeping the total RNP concentration constant at 3.75 μM.

TABLE-US-00010 TABLE 10 List of selected guide RNAs and dead guide RNAs gRNA* Targeting Targeting or domain domain dgRNA sequence sequence ID (RNA) (DNA) Sense Sp35 CUUGUCAAGGCUAUUG CTTGTCAAGGCTATTG Anti- gRNA GUCA GTCA sense (SEQ ID NO: 339) (SEQ ID NO: 917) Sp37 CUUGACCAAUAGCCUU CTTGACCAATAGCCTT Sense gRNA GACA GACA (SEQ ID NO: 333) (SEQ ID NO: 915) SpA GGCAAGGCUGGCCAAC GGCAAGGCTGGCCAAC Sense gRNA CCAU CCAT (SEQ ID NO: 340) (SEQ ID NO: 919) tSpA GGCUGGCCAACCCAU GGCTGGCCAACCCAT Sense dgRNA (SEQ ID NO: 970) (SEQ ID NO: 971) Sp180 GCCGGCGGCUGGCUA GCCGGCGGCTGGCTA dgRNA (SEQ ID NO: 972) (SEQ ID NO: 973) Sp181 AGUGAGGCCAGGGGC AGTGAGGCCAGGGGC dgRNA (SEQ ID

NO: 974) (SEQ ID NO: 975) Sp182 UUAGAGUAUCCAGUG TTAGAGTATCCAGTG dgRNA (SEQ ID NO: 976) (SEQ ID NO: 977) *None of the guide RNAs or dead guide RNAs in Table 10 and used in the experiments of Example 8 are modified to recruit an exogenous trans-acting factor.

[0281] tSpA dRNP co-delivered with Sp37 WT RNP at a ratio of 1:4 (dRNP:Total RNP ratio 1:5; 0.75 μM dRNP:3.75 μM Total RNP) supported a ˜4.3-fold increase in indels (as determined by T7E1 endonuclease analysis of HBG2 PCR product amplified from gDNA extracted from CD34.sup.+ cells) compared to CD34.sup.+ cells treated with 3.75 μM live Sp37 WT RNP alone (FIG. **17**). These data show that dRNP paired with WT RNP can increase editing at a target in adult CD34.sup.+ cells.

[0282] To determine whether co-delivery of dead RNP would increase editing of HBG target site, Sp181 dRNP (comprising Sp181 dgRNA (Table 10)) and tSpA dRNP (comprising tSpA dgRNA (Table 10) targeting the same strand of Sp35) were co-delivered with Sp35 by Maxcyte electroporation into mPB CD34.sup.+ cells.

[0283] Electroporation of Sp35 WT RNP (3.75 μM) alone does not support detectable indels (by T7E1 endonuclease analysis) (FIG. **18**). However, co-delivery of Sp35 WT RNP (3 μM) with either dRNP (Sp181 dRNP and tSpA dRNP) (Sp181 dgRNA (Table 10) or tSpA dgRNA (Table 10), at 0.75 μM) increased indels from 0% to 10% compared to CD34.sup.+ cells electroporated with Sp35 WT RNP alone (by T7E1 endonuclease analysis) (FIG. **18**). Importantly, the level of indels detected in the mPB CD34.sup.+ cells was maintained in the day 7 erythroid progeny of edited cells (FIG. **18**, white bars).

[0284] Additional dead/WT pairs of RNPs were tested to determine the effect of co-delivery of dead/WT RNPs on editing the target site in the HBG promoter (−110 nt) and resulting expression of HbF. Dead/WT pairs of RNPs (as shown in Table 11) were co-delivered by electroporation into mPB CD34.sup.+ cells. dRNP was codelivered with WT RNP (i.e., Sp35 gRNA+tSpA dgRNA, Sp35 gRNA+Sp181 dgRNA, and Sp37 gRNA+tSpA dgRNA) at a ratio of 1:4 (dRNP:Total RNP ratio 1:5; 0.75 μM dRNP:3.75 μM Total RNP).

TABLE-US-00011 TABLE 11 Percentage Editing and HbF Production by Co-Delivery of RNPs % Editing gRNA Pairs RNPs Strategy (T7E1) % HbF Sp37 gRNA + Sp37 D10A RNP + D10A* 22.44 11.2 SpA gRNA SpA D10A RNP Sp85 gRNA + Sp85 D10A RNP + D10A 10.5 5 SpA gRNA SpA D10A RNP Sp36 gRNA + Sp36 D10A RNP + D10A 5 5.3 Sp85 gRNA Sp85 D10A RNP Sp35 gRNA + Sp35 WT RNP + WT 10 9.6 tSpA dgRNA tSpA dRNP Live/Dead Sp35 gRNA + Sp35 WT RNP + WT 8.27 12.71 Sp181 dgRNA Sp181 dRNP Live/Dead Sp37 gRNA + Sp37 WT RNP + WT 33.66 13.48 tSpA dgRNA tSpA dRNP Live/Dead *= D10A is a Cas9 nickase that makes a single strand nick.

[0285] tSpA dRNP co-delivered with Sp35 WT RNP, Sp181 dRNP co-delivered with Sp35 WT RNP, and tSpA dRNP co-delivered with Sp37 WT RNP supported editing of the HBG promoter (as determined by T7E1 endonuclease analysis of HBG2 PCR product amplified from gDNA extracted from CD34.sup.+ cells) and resulted in induction of HbF protein (as determined by HPLC analysis of hemoglobin expression in erythroid progeny according to the HPLC method described in Chang 2017 at pp. 143-44 and/or UPLC analysis, incorporated by reference herein) (Table 3). These data show that dRNP paired with WT RNP can support editing at a target region in adult CD34.sup.+ cells, resulting in HbF protein expression in erythroid progeny of the edited adult CD34.sup.+ cells.

Example 9: Tracking Edited HSC Contribution to Hematopoiesis Based on Tracking Edited Alleles in their Progeny In Vivo

[0286] DNA lesions created by paired Cas9 WT and nickases (such as the D10A and N863A mutants) can lead to a variety of repair outcomes, including a wide spectrum of insertions and deletions in the region proximal to the nicks (Bothmer 2017). However, in contrast to wild-type Cas9, the repair outcomes induced by paired nickases are more diverse and have a more uniform distribution of frequencies of specific indels (Bothmer 2017).

[0287] The diverse repair outcomes obtained after repair of double strand breaks made by WT CRISPR nucleases or after paired nicking can then be used to estimate the diversity of edited HSCs that are contributing to blood production. This is important because for life-long hematopoiesis from an edited cell pool, multiple edited HSCs must retain their ability to produce blood and self-renew for a human life span. For a specific target site, editing using a CRISPR with a gRNA that has specificity for one target site in the genome suggests that only one site will be modified, presenting the challenge of distinguishing edited alleles among the many HSCs, and thus hard to determine whether multiple HSCs are contributing to hematopoiesis. However, this presents a unique advantage to tracking the edited cells based on subtle differences in DNA repair outcomes that can occur within each allele and in each cell. The unique alleles are distinguished from each other based on indel characteristics including the type and size of the edit (insertion, deletion, insertion/deletion, and number of nucleotides deleted or inserted) and on their relative distance to the cut site and within the amplicon. For example, each deletion or insertion observed when sequencing the cell population can be characterized by its position in the genome, its length, and in the case of insertions, its sequence. The combination of these features can be used as an indel barcode to track the persistence of HSCs and their differentiation into mature blood cells as a measure of diversity after editing (FIG. **19**). Importantly, unlike in other CRISPR indel barcoding approaches or in contrast to gene therapy approaches, the indel barcode is a potentially functional edit at the target locus, requiring no further modification of the genome for purposes of tracking. Although it is possible for different cells to be independently edited in a way that creates the same edit, tracking by indel barcodes can establish a lower bound on the diversity of a population. Because each allele in a diploid cell can receive a different indel, that possibility must also be accounted for in diversity estimates.

[0288] It is hypothesized that if multiple unique indels in hematopoietic stem cells (HSCs) and progeny are detected then edited HSC diversity is maintained after editing. A method of tracking unique edited alleles to determine whether HSC diversity is maintained is disclosed herein. First, an RNA-guided nuclease and guide RNA complexed to form a ribonucleoprotein (RNP) complex for editing is electroporated into cells and several CD34+ cells repair the DNA slightly differently to create unique alleles. Over time after transplantation into an animal, edited HSCs repopulate the blood system and can be collected and sorted based on the different tissues and lineages to evaluate specific unique indels in the long-term engrafted HSCs and in differentiated progeny.

[0289] An HBB locus was used as a model to illustrate this method and determine whether HSC diversity is maintained. CD34+ cells were electroporated with D10A nickase RNPs targeting the HBB locus as described using the methods for electroporation provided in Example 1. Before transplantation, genomic DNA was harvested from an aliquot of the bulk pre-infusion CD34.sup.+ cell product, sequenced, and reads aligned to a reference sequence encompassing the target site at the HBB locus. The remainder (majority bulk) of the CD34.sup.+ cells were transplanted into mice. Four months after transplantation, human cells were purified from the hematopoietic organs of the mice (peripheral blood [PB], spleen, and bone marrow [BM]) and the human cell lineages (myeloid, erythroid, lymphoid, CD34.sup.+ or HSCs) were further purified. The genomic DNA was isolated from all of these human cells derived from the engrafted edited HSCs and sequenced (sequencing reads were aligned to the reference locus). The percentage of each unique edited allele over the total sum of all edited alleles detected was plotted to determine their relative contribution (FIG. **20**). The black bars represent a group of all unique alleles occurring at low frequencies of total edited alleles. White and grey bars correspond to the top five most abundant unique alleles ranked (FIG. **20**). In the bulk CD34.sup.+ cell preinfusion product, the top five most abundant clones together make up less than 10% of total edited alleles, consistent with the diversity and heterogeneity of cell types within the bulk CD34.sup.+ cell population. There are many unique alleles at less than 1% each grouped into the black bar. An analysis of the top five most abundant alleles in mouse 1 (that is, in vivo after transplantation of and long-term engraftment of edited

HSCs derived from the heterogeneous preinfusion product) indicates that there are shared unique HSC alleles across tissues and in different lineages (FIG. **21**). An analysis of the top 5 most abundant clones in mouse 2 indicates that, for the most part, different unique alleles are present in vivo compared to mouse 1, but there are also shared alleles across tissues and lineages (FIG. **21**). These data show that multiple edited HSCs are making blood in vivo and that a repertoire of unique edits are detected across multiple tissues. This demonstrates that there is diversity among the alleles and that no one edited allele is dominant over others. This method provides a means to survey for any unintended effects of alleles at the target site providing a readout on safety of editing. The method also allows tracking of indel diversity over time, which may provide information about the toxicity of a transplanted cell population according to this disclosure, as well as about the efficacy of such transplanted cell population.

Example 10: Lentiviral Screen of Guide RNAs Influencing Fetal Hemoglobin Expression

[0290] A library of approximately 25,000 unique gRNA sequences spanning the beta globin locus (FIG. **22**) was screened to identify cis-regulatory elements involved in the regulation of fetal hemoglobin expression. An immortalized human erythroid progenitor cell line (HUDEP-2, Kurita 2013) was transduced with Cas9 Blasticidin Lentiviral Transduction Particles (Sigma-Aldrich, St. Louis, MO) to generate a cell line that stably expressed *S. pyogenes* Cas9. A guide RNA library comprising the ˜25,000 unique sequences described above, along with 500 non-homologous guide sequences, was designed and packaged in lentiviral vectors as described in Joung 2017. A portion of the lentiviral genome encoding the unique guide RNA sequences is shown below:

TABLE-US-00012 (SEQ ID NO: 1661) AATGATACGG CGACCACCGA GATCTACACT CTTTCCCTAC ACGACGCTCT TCCGATCTTA CGATCGATGG TCCAGAGCTT TATATATCTT GTGGAAAGGA CGAAACACCN NNNNNNNNNN NNNNNNNNNG TTTTAGAGCT AGAAATAGCA AGTTAAAATA AGGCTAGTCC GTTATCAACT TGAAAAAGTG GCACCGAGTC GGAGATCGGA AGAGCACACG TCTGAACTCC AGTCACCAAG GCGAATCTCG TATGCCGTCT TCTGCTTG

(Xs denote a unique 20-mer gRNA targeting sequence; primer binding sequences are underlined.). The lentiviral vectors also encoded puromycin, allowing for the selection of transduced cells carrying the guide RNA expression cassettes.

[0291] HUDEP-2 cells were transduced with lentiviral particles encoding the guide RNA library over a range of concentrations and treated with puromycin to determine the viral titer (transducing unit per mL of vector). After the viral titer was determined, lentiviral particles encoding the guide RNA library were applied to *S. pyogenes* Cas9-expressing HUDEP-2 cells at a multiplicity of infection of 0.25 to ensure that most cells would have integrated no more than one copy of the lentiviral genome and thus would express a unique guide RNA. The total number of cells transduced was calculated to ensure that an average of more than 500 cells carried a copy of each guide RNA in the library. Following transduction, the transduced cells were selected using puromycin, expanded and differentiated to become hemoglobinized erythroblasts. The cells were then fixed, permeablized, and stained using a Fluorescein isothiocyanate conjugated antibody against gamma-globin chains (Thorpe 1994) and flow sorted into pools that expressed high or low levels of gamma globin (Canvers 2015) using a SONY cell sorter. Genomic DNA was harvested from both pools and the portion of the lentiviral transgene encoding the guide RNA sequence was PCR amplified and sequenced using next generation sequencing. Transduction, selection, differentiation, sorting and sequencing were repeated across four bioreplicates.

[0292] Guide RNA sequences listed in Table 12, below, (a) were identified as enriched in High-F populations relative to the non-targeting controls, and (b) did not have a perfectly matched cut site located within or proximal to the HBB or HBD genes. The gRNAs were ranked and categorized into five tiers based for prioritization of validation in mPB CD34+ cells (“Tier 1”, “Tier 2”, “Tier 3”, “Tier 4” and “Friend of Tier 1”). Division into four tiers was based on Standard deviation (SD of log 2 enrichment across 4 bioreplicates), Log 2 Enrichment score (average log 2 enrichment

across the 4 bioreplicates), and whether the guide RNA was specific to HBG1 and/or specific to HBG2. Log 2 enrichment values for each replicate are calculated as follows:
[00001]$\log2(\frac{\text{gRNAreadfrequencyinHighHbFpoolfromReplicate}}{\text{gRNAreadfrequencyinLowHbFpoolfromReplicate}})$.
[0293] Guide RNAs of the highest priority were those in Tier 1, which was defined as SD<=0.75, Log 2 Enrichment score >=1.6 or >=1.3 if the guide RNA was specific to either HBG1 or HBG2. Tier 2 was defined as SD<=0.75, Log 2 Enrichment score >=1. Tier 3 was defined as SD<=1, Log 2 Enrichment score >=0.7. Tier 4 was defined as SD<=1.5, Log 2 Enrichment score >=0.5. Friend of Tier 1 was defined as those guide RNAs whose cut site was within 10 nucleotides of a Tier 1 cut site, but were not captured in Tiers 2 to 4 (FIG. **23**).
[0294] Analysis of targeting domains enriched in the screen revealed several regions of interest in which enriched guide RNA cut sites were concentrated. The majority of the HbF inducing gRNAs were mapped to the beta globin locus including HBG, HBD, and HBB (FIG. **24**). Those regions thus enclose regulatory elements that repress HbF expression. gRNA cutting in those regions are likely to induce HbF expression in erythroid cells. gRNAs whose cut-site overlap with HbF regulatory element within those region are likely to yield the highest frequency of HbF inducing indels and thus the highest frequency of high-HbF expressing cells post-editing. These regions of interest are shown in Table 13, below:
TABLE-US-00013 TABLE 13 Regions of interest Genomic Coordinate of HbG* Nucleotides Name of Region Chr 11 TCCTAAAGCT TGGAACACTT Region 1: Downstream of HBG1 (NC_000011.10): TCCCTTCCTT AAGAACCATC 5,247,883- CTTGCTACTC AGCTGCAATC 5,248,186 AATCCAGCCC CCAGGTCTTC ACTGAACCTT TTCCCATCTC TTCCAAAACA TCTGTTTCTG AGAAGTCCTG TCCTATAGAG GTCTTTCTTC CCACCGGATT TCTCCTACAC CATTTACTCC CACTTGCAGA ACTCCCGTGT ACAAGTGTCT TTACTGCTTT TATTTGCTCA TCAAAATGCA CATCTCATAT AAAAATAAAT GAGGAGCATG CACACACCAC AAACACAAAC AGGCATGCAG AAAT (SEQ ID NO: 1640) Chr 11 ATAAAGATGA ACCCATAGTG Region 2: HBG1 Intron 2 - A (NC_000011.10): AGCTGAGAGC TCCAGCCTGG 5,248,509- CCTCCAGATA ACTACACACC 5,249,173 AAGCTTCCAC CCAGAATCAA GCCTATGTTA ACTTCCCTCA AAGCCTGAGA TTTTGCCTTC CCATTAAATG CAGGTAGTTG TTCCCCTTCA AGCACTAGTC ACTGGCCATA ATTTAAATCT TGCTATCTTC TTGCCACCAT GAACCCTGTA TGTTGTAGGC TGAAGACGTT AAAAGAAACA CACGCTGACA CACACACACA CACGCGCGCG CGCACACACA CACACACACA CAGAGCTGAC TTTCAAAATC TACTCCAGCC CAAATGTTTC AATTGTTCCT CACCCCTGGA CATACTTTGC CCCCATCTGG AATTAAAGGA TATAAGTTTG TAATGAAGCA TTAGCAGCAT TTTATATGTG TCCAGCTGAT ATAGGAATAG CCTTAGCAAT GTATGTTTGG CCACCAAAGT TCCCCACTTT GACTGAGCCA ATATATGCCT TCTGCCTGCA TCTTTTTAAC GACCATACTT GTCCTGCCTC CAGATAGATG TTTTAAAACA ACAAAAATGA GGGAAAGATG AAAGTTCTTT CTACTGGAAT CTAATAAAGA AAAGTCATTT TCCTCATTTC CACCTCTCTT TTCTCAAAGT CAAAATTGTC CATCT (SEQ ID NO: 1641) Chr 11 CCCTAAAACA TTACCACTGG Region 3: HBG1 Intron 2 - B (NC_000011.10): GTCTCAGCCC AGTTAGTCCT 5,249,198- CTGCAGTTTC TTCACCCCCA 5,249,362 ACCCCAGTAT CTTCAAACAG CTCACACCCT GCTGTGCTCA GATCAATACT CCGTTGTCTA AGTTGCCTCG AGACTAAAGG CAACAGGGCT GAAACATCTC CTGGA (SEQ ID NO: 1642) Chr 11 CTGTGAGATT GACAAGAACA Region 4: HBG1 Intron 1 (NC_000011.10): GTTTGACAGT CAGAAGGTGC 5,249,591- CACAAATCCT GAGAAGCGAC 5,249,712 CTGGACTTTT GCCAGGCACA GGGTCCTTCC TTCCCTCCCT TGTCCTGGTC ACCAGAGCCT AC (SEQ ID NO: 1643) Chr 11 GCCGCCGGCC CCTGGCCTCA Region 5: HBG1 -60 nt region (NC_000011.10): CTGG (SEQ ID NO: 1644)

from Transcription Start Site 5,249,904- (TSS) 5,249,927 Chr 11 CCTTGTCAAG GCTATTGGTC Region 6: HBG1 -110 nt region (NC_000011.10): AAGGCAAGGC TGG (SEQ ID from TSS 5,249,955- NO: 1645) 5,249,987 Chr 11 TGAGATAGTG TGGGGAAGGG Region 7: HBG1 -200 nt region (NC_000011.10): GCCCCC AAGAGGATAC (SEQ ID from TSS 5,250,040- NO: 1646) 5,250,075 Chr 11 TATAGCCTTT GCCTTGTTCC Region 8: HBG1 -250 nt region (NC_000011.10): GATTCAGTCA TTCCAGTTTT T from TSS 5,250,089- (SEQ ID NO: 1647) 5,250,129 Chr 11 TCTTCCCTTT AGCTAGTTTC Region 9: HBG1 -333 nt region (NC_000011.10): CTTCTCCCAT CATAGAGGAT from TSS 5,250,141- ACCAGGACTT CTTTTGTCAG 5,250,254 CCGTTTTTTA CCTTCTTGTC TCTAGCTCCA GTGAGGCCTG TAGTTTAAAG CTAA (SEQ ID NO: 1648) Chr 11 CCACAGTTTC AGCGCAGTAA Region 10: HBGI -650 nt region (NC_000011.10): TAGATTAGTG TTACATAATA from TSS 5,250,464- TAAGACCTAA TGCTTACCTC 5,250,549 AATATCTACT TATCCGTACC TATTTG (SEQ ID NO: 1649) Chr 11 TATTCAGGTA TGTATGTATA Region 11: HBG1 -800 nt region (NC_000011.10): CACCAGATGA TGTGTATTTA from TSS 5,250,594- CCACTGGATA AGTGTGTGTG 5,250,735 CTGGCTGATG ACCCAGGGTT TTGGCGTAGC TCTTCTATGC TCAGTAAAGA TGATGGTAGA ATGTTCTTTG GCAGGTACTG TG (SEQ ID NO: 1650) Chr 11 CAATAAAGAT GAACCCATAG Region 12: HBG2 Intron 2 - A (NC_000011.10): TGAGCTGAGA GCTCCAGCCT 5,253,425- GGCCTCCAGA TAACTACACA 5,254,121 CCAAGCTTCC ACCCAGAATC AAGCCTATGT TAACTTCCCT CAAAGCCTGA GATTTTGCTT TCCCATTAAA TGCAGGTAGT TGTTCTTCTT GCAGCACTAG TCACTGGCCA TAATTTAAAT CTTGTTATCT TCTTGCCACC ATGAACCCTG TATGCTGTAG GCTGAAAACG TTAAAAGAAA CACACGCTCT CACACACACA CAAACACACG CGCGCACACA CACACACACA CACACAGAGC TGACTTTCAA AATCTACTCC AGCCCAAATG TTTCAATTGT TCCTCACCCC TGGACATACT TTGCCCCCAT CTGGAATTAA AGGATATAAG TTTGTAATGA AGCATTAGCA GCATTTTATA TGTGTCCAGC TGATATAGGA ATAGCCTTAG CAATGTATGT TTGGCCACCA AAGTTCCCCA CTTTGACTGA GCCAATATAT GCCTTCTGCC TGCATCTTTT TAATGACCAT ACTTGTCCTG CCTCCAGATA GATGTTTTAA AACGAATAAC AAAAATAGGG GAAAGGTGAA AGTTCTTTCT ACCGAAATCT AATAAAGAAA AGTCATTTTC CTCATTTCCA CCTCTCTTTT CTCAAAGTCA AAGTTGTCCA TCTAGATTTT CAGAGGCACT CCTTAGG (SEQ ID NO: 1651) Chr 11 CCCTAAAACA TTGCCACTGG Region 13: HBG2 Intron 2 - B (NC_000011.10): GTCTCAGCCC AGTTAGTCCT 5,254,122 - CTGCAGTTTC TTCACTCCCA 5,254,306 ACCCCAGTAT CTTCAAACAG CTCACACCCT GCTGTGCTCA GATCAATACT CAGTTGTCTA AGTTGCCTCG AGACTAAAGG CAACAGTGCT GAAACATCTC CTGGACTCAC CTTGAAGTTC TCAGG (SEQ ID NO: 1652) Chr 11 AGCCTGTGAG ATTGACAAGA Region 14: HBG2 Intron 1 (NC_000011.10): ACAGTTTGAC AGTCAGAAGG 5,254,511- TGCCACAAAT CCTGAGAAGC 5,254,648 GACCTGGACT TTTGCCAGGC ACAGGGTCCT TCCTTCCCTC CCTTGTCCTG GTCACCAGAG CCTACCTTCC CAGGGTT (SEQ ID NO: 1653) Chr 11 CCGCCGGCCC CTGGCCTCAC Region 15: HBG2 -60 nt region (NC_000011.10): TGGATACTCT AAGACTAT (SEQ from TSS 5,254,829- ID NO: 1654) 5,254,866 Chr 11 CCTTGTCAAG GCTATTGGTC Region 16: HBG2 -110 nt region (NC_000011.10): AAGGCAAGGC T from TSS 5,254,879- (SEQ ID NO: 1655) 5,254,909 Chr 11 CAGGGACCGT TTCAGACAGA Region 17: HBG2 -200 nt region (NC_000011.10): TATTTGCATT GAGATAGTGT from TSS 5,254,935 GGGGAAGGGG CCCCCAAGAG 5,255,009 GATACTGCTG CTTAA (SEQ ID NO: 1656) Chr 11 TTGCCTTGTT CCGATTCAGT Region 18: HBG2 -250 nt region (NC_000011.10):

CATTCCAAT (SEQ ID NO: from TSS 5,255,025- 1657) 5,255,053 Chr 11 TTTAGCTAGT TTTCTTCTCC Region 19: HBG2 -330 nt region (NC_000011.10): CACCATAGAA GATACCAGGA from TSS 5,255,076- CTTCTTTTGT CAGCCGTTTT 5,255,179 TCACCTTCTT GTCTGTAGCT CCAGTGAGGC CTGTAGTTTA AAGT (SEQ ID NO: 1658) Chr 11 GGACACGTCT TAGTCTCATT Region 20: HBG2 -500 nt region (NC_000011.10): TAGTAAGCAT TGGTTTCC (SEQ from TSS 5,255,255- ID NO: 1659) 5,255,292 Chr 11 TTTTTTATAT TCAGGTATGT Region 21: HBG2 -800 nt region (NC_000011.10): ATGTAGGCAC CCGATGATGT from TSS 5,255,518- GTATTTATCA CTGGATAAGT 5,255,641 GTATGTGCTG GCTGATGACC CAGGGTTTTG GTGTAGCTCT TCTATGCTCG GTAAAGATGA TGGT (SEQ ID NO: 1660) *NCBI Reference Sequence NC_000011, "*Homo sapiens* chromosome 11, GRCh38.p12 Primary Assembly," (Version NC_000011.10). All coordinates are Hg38 0-based.

[0295] HUDEP-2 cells were individually electroporated with RNPs complexed with the gRNAs listed in Table 12 and *S. pyogenes* Cas9 protein at a concentration of 5 μM. After electroporation, HUDEP-2 cells were pooled (2 replicate per pool) as detailed in Table 14. The pooling of electroporated samples was perform based on the cut-site position of the included RNPs to allow for PCR amplification and NGS analysis of each pool with a single primer pair per pool. Each pool of cell was differentiated in erythroid cells, and sorted based on gamma globin expression in a "high HbF" fraction and a "low HbF" fraction. Genomic DNA from sorted populations was prepared, PCR amplified, and sequenced. The amount of gDNA to be amplified and amount of PCR product to be sequenced was adjusted for each pool based on the number of individual electroporated samples (corresponding to the number of gRNAs tested) initially pooled. Sequence reads were mapped to the reference amplicon sequence of the human genome (Hg38) to identify insertions or deletions (indels) (>35 million total aligned reads). Frequencies of individual indels were calculated and indels with average frequencies across samples that were below a cut-off adjusted for each pool were eliminated from further analysis (cutoff: 0.1/[number of electroporation sample included in the pool]). Average HbF enrichment scores of individual indels were calculated (as average of

[00002]$\log2\left(\frac{\text{indelread - frequencyinHighHbFpoolfromReplicate}}{\text{indelread - frequencyinLowHbFpoolfromReplicate}}\right)$

across each bioreplicate) and their position was determined relative to either HBG1 (FIG. **25**A) or HBG2 (FIG. **25**B). When the sequence homology between the HBG1 and HBG2 locus did not allow differentiation of reads originating from one or the other, the indels are displayed with both potential coordinates on FIG. **25**A and FIG. **25**B. Multiple clusters of fetal globin inducing indels were identified in HUDEP-2 cells, likely overlapping with HBG repressing elements.

TABLE-US-00014 TABLE 14 Pool of electroporated samples, list of gRNA included in each pool. Pool gRNA targeting domain (DNA) Pool 1 CTTCCTTCCCTCCCTTGTCC (SEQ ID NO: 1084) CAGGACAAGGGAGGGAAGGA (SEQ ID NO: 1311) TAGTCTTAGAGTATCCAGTG (SEQ ID NO: 982) CTGGTGACCAGGACAAGGGA (SEQ ID NO: 983) TCTGGTGACCAGGACAAGGG (SEQ ID NO: 984) GGCTCTGGTGACCAGGACAA (SEQ ID NO: 1035) AGGCTCTGGTGACCAGGACA (SEQ ID NO: 1312) GAAGGTAGGCTCTGGTGACC (SEQ ID NO: 1233) AACCCTGGGAAGGTAGGCTC (SEQ ID NO: 1085) CGAGTGTGTGGAACTGCTGA (SEQ ID NO: 1036) GAGTGTGTGGAACTGCTGAA (SEQ ID NO: 1234) CCTAGCCAGCCGCCGGCCCC (SEQ ID NO: 1314) TGAGGCCAGGGGCCGGCGGC (SEQ ID NO: 1313) CCGCCGGCCCCTGGCCTCAC (SEQ ID NO: 1316) CCAGTGAGGCCAGGGGCCGG (SEQ ID NO: 1037) TATCCAGTGAGGCCAGGGGC (SEQ ID NO: 1086) AGAGTATCCAGTGAGGCCAG

(SEQ ID NO: 985) TAGAGTATCCAGTGAGGCCA (SEQ ID NO: 1315) TTAGAGTATCCAGTGAGGCC (SEQ ID NO: 1235) TAGTCTTAGAGTATCCAGTG (SEQ ID NO: 986) TGGTCAAGTTTGCCTTGTCA (SEQ ID NO: 919) GTTTGCCTTGTCAAGGCTAT (SEQ ID NO: 913) CTTGTCAAGGCTATTGGTCA (SEQ ID NO: 917) CTTGACCAATAGCCTTGACA (SEQ ID NO: 915) CAAGGCTATTGGTCAAGGCA (SEQ ID NO: 911) GCTATTGGTCAAGGCAAGGC (SEQ ID NO: 1038) Pool 2 TATCTGTCTGAAACGGTCCC (SEQ ID NO: 1039) ATATTTGCATTGAGATAGTG (SEQ ID NO: 1317) TATTTGCATTGAGATAGTGT (SEQ ID NO: 945) ATTTGCATTGAGATAGTGTG (SEQ ID NO: 1318) GCATTGAGATAGTGTGGGGA (SEQ ID NO: 1040) CATTGAGATAGTGTGGGGAA (SEQ ID NO: 987) ATTGAGATAGTGTGGGGAAG (SEQ ID NO: 988) GTGGGGAAGGGGCCCCCAAG (SEQ ID NO: 1319) AAGCAGCAGTATCCTCTTGG (SEQ ID NO: 1001) TAAGCAGCAGTATCCTCTTG (SEQ ID NO: 1049) TTAAGCAGCAGTATCCTCTT (SEQ ID NO: 1326) ATTAAGCAGCAGTATCCTCT (SEQ ID NO: 1327) ACTGAATCGGAACAAGGCAA (SEQ ID NO: 989) GGAATGACTGAATCGGAACA (SEQ ID NO: 990) AAAAACTGGAATGACTGAAT (SEQ ID NO: 1320) AAAAATTGGAATGACTGAAT (SEQ ID NO: 1328) GGAGAAGGAAACTAGCTAAA (SEQ ID NO: 1041) GGGAGAAGGAAACTAGCTAA (SEQ ID NO: 1087) GGGAGAAGAAAACTAGCTAA (SEQ ID NO: 1050) GTTTCCTTCTCCCATCATAG (SEQ ID NO: 1042) GTATCCTCTATGATGGGAGA (SEQ ID NO: 991) CTCCCATCATAGAGGATACC (SEQ ID NO: 1321) CTCCCACCATAGAAGATACC (SEQ ID NO: 1264) GTCCTGGTATCCTCTATGAT (SEQ ID NO: 1088) GTCCTGGTATCTTCTATGGT (SEQ ID NO: 1002) AGTCCTGGTATCCTCTATGA (SEQ ID NO: 1043) AGTCCTGGTATCTTCTATGG (SEQ ID NO: 1003) AGAAGTCCTGGTATCTTCTA (SEQ ID NO: 1263) ACGGCTGACAAAAGAAGTCC (SEQ ID NO: 992) AGAGACAAGAAGGTAAAAAA (SEQ ID NO: 1089) ACAGACAAGAAGGTGAAAAA (SEQ ID NO: 1051) CACTGGAGCTAGAGACAAGA (SEQ ID NO: 1090) TCTTGTCTCTAGCTCCAGTG (SEQ ID NO: 1091) TCTTGTCTGTAGCTCCAGTG (SEQ ID NO: 1118) CTTTAAACTACAGGCCTCAC (SEQ ID NO: 1092) Pool 3 CTATTACTGCGCTGAAACTG (SEQ ID NO: 1044) TAGATATTGAGGTAAGCATT (SEQ ID NO: 1236) GTACGGATAAGTAGATATTG (SEQ ID NO: 1045) TATACATACATACCTGAATA (SEQ ID NO: 1237) AGATGATGTGTATTTACCAC (SEQ ID NO: 1322) AGTGGTAAATACACATCATC (SEQ ID NO: 1094) CAGCACACACACTTATCCAG (SEQ ID NO: 993) TGTGTGCTGGCTGATGACCC (SEQ ID NO: 994) GTGTGCTGGCTGATGACCCA (SEQ ID NO: 995) TGGCTGATGACCCAGGGTTT (SEQ ID NO: 1323) AAGAGCTACGCCAAAACCCT (SEQ ID NO: 996) GAAGAGCTACGCCAAAACCC (SEQ ID NO: 997) TCTATGCTCAGTAAAGATGA (SEQ ID NO: 998) ATGTTCTTTGGCAGGTACTG (SEQ ID NO: 1238) AATGCTAGGTTCACTTCTCA (SEQ ID NO: 1239) CATGGAAAACAACTCTAAAG (SEQ ID NO: 1095) AAACAACTCTAAAGAGGCAA (SEQ ID NO: 1096) Pool 4 AATGAGAACTTAAGAGATAA (SEQ ID NO: 1266) TAAAGCAACAGTTTCAGTGC (SEQ ID NO: 1267) GATAAGTAGATATTGAAGTA (SEQ ID NO: 1268) TTATATTCAGGTATGTATGT (SEQ ID NO: 1119) AGTGATAAATACACATCATC (SEQ ID NO: 1269) CAGTGATAAATACACATCAT (SEQ ID NO: 1270) TATGTGCTGGCTGATGACCC (SEQ ID NO: 1004) ATGTGCTGGCTGATGACCCA (SEQ ID NO: 1005) TGGCTGATGACCCAGGGTTT (SEQ ID NO: 1323) AAGAGCTACACCAAAACCCT (SEQ ID NO: 1053)

GAAGAGCTACACCAAAACCC (SEQ ID NO: 1006) TCTATGCTCGGTAAAGATGA (SEQ ID NO: 1007)

[0296] To further identify which nucleotides of the genomic regions targeted by the gRNAs led to HbF induction, we quantified the number of "high HbF-enriched indels" (as defined by average enrichment score >1.75). Using the same cut-off when comparing replicate fraction <0.4% of indels were identified as enriched in one replicate over the other covering each position of the target regions within the human genome (FIG. **26**). To achieve a fine analysis, all deletions spanning more than 10 nt were excluded from the analysis. The fraction of "enriched indels" covering each position over count of all covering indels (positions with less than 10 total covering indels were excluded from the analysis) was quantified. Each position for which 70% of the covering indels were characterized as "high HbF-enriched indels" are listed in Table 15.

TABLE-US-00015 TABLE 15 Genomic position whose disruption by indels lead to HbF expression. Postion Start* Position End* 5249683 5249684 5249686 5249687 5249961 5249962 5249963 5249964 5249964 5249965 5249965 5249966 5249966 5249967 5249967 5249968 5249968 5249969 5249969 5249970 5249970 5249971 5249971 5249972 5249972 5249973 5249973 5249974 5249974 5249975 5249975 5249976 5250049 5250050 5250050 5250051 5250051 5250052 5250052 5250053 5250053 5250054 5250054 5250055 5250107 5250108 5250108 5250109 5250109 5250110 5250110 5250111 5250111 5250112 5250112 5250113 5250113 5250114 5250114 5250115 5250115 5250116 5250183 5250184 5250184 5250185 5250185 5250186 5250186 5250187 5250187 5250188 5250188 5250189 5250189 5250190 5254607 5254608 5254610 5254611 5254885 5254886 5254887 5254888 5254888 5254889 5254889 5254890 5254890 5254891 5254891 5254892 5254892 5254893 5254893 5254894 5254894 5254895 5254895 5254896 5254896 5254897 5254897 5254898 5254898 5254899 5254899 5254900 5254973 5254974 5254974 5254975 5254975 5254976 5254976 5254977 5254977 5254978 5254978 5254979 5255031 5255032 5255032 5255033 5255033 5255034 5255034 5255035 5255035 5255036 5255036 5255037 5255037 5255038 5255038 5255039 5255039 5255040 5255107 5255108 5255108 5255109 5255109 5255110 5255110 5255111 5255111 5255112 5255112 5255113 5255113 5255114 5255629 5255630 5255630 5255631 5255631 5255632 5255632 5255633 5255633 5255634 5255635 5255636 *NCBI Reference Sequence NC_000011, "Homo sapiens chromosome 11, GRCh38.p12 Primary Assembly," (Version NC_000011.10). All coordinates are Hg38 0-based.

[0297] Since the analysis performed here is not exhaustive (the genomic position evaluated here is limited by the gRNAs used and the region covered by indels they generated), it is likely that other genomic positions also induce HbF expression, in particular genomic positions very close to the position listed in Table 15. The contiguous genomic positions identified in Table 15 (allowing for a gap of 2 nt) were joined to defined genomic regions within which disruption of one or multiple indels would lead to HbF expression in erythroid cells, listed in Table 16.

TABLE-US-00016 TABLE 16 Regions within which disruption of one or multiple indels would lead to HbF expression in erythoid cells. Genomic Coordinate of HbG Regions* Chr 11 (NC_000011.10): 5249683-5249687 Chr 11 (NC_000011.10): 5249961-5249976 Chr 11 (NC_000011.10): 5250049-5250055 Chr 11 (NC_000011.10): 5250107-5250116 Chr 11 (NC_000011.10): 5250183-5250190 Chr 11 (NC_000011.10): 5254607-5254611 Chr 11 (NC_000011.10): 5254885-5254900 Chr 11 (NC_000011.10): 5254973-5254979 Chr 11 (NC_000011.10): 5255031-5255040 Chr 11 (NC_000011.10): 5255107-5255124 Chr 11 (NC_000011.10): 5255629-5255636 *NCBI Reference Sequence NC_000011, "Homo sapiens chromosome 11, GRCh38.p12 Primary Assembly," (Version NC_000011.10). All coordinates are Hg38 0-based.

[0298] The coordinates of those domains very likely overlap closely with coordinates of HbF repressing motifs. It is expected than a gRNA targeting close or within those regions would lead to HbF expression. gRNAs whose cut site is within those coordinates would yield high frequency of

high HbF expressing cells since most cells would carry HbF inducing indels.

[0299] HUDEP-2 cells were transfected with RNPs made with individual gRNAs listed in Table 12 complexed with *S. pyogenes* Cas9 protein at a concentration of 5 μM. After transfection, HUDEP-2 cells were differentiated and RNA was extracted from cell pellets using TaqMan® Gene Expression Cells-to-CT™ Kit from Life Technologies kit. HBG1 and RSP18 mRNA levels were then measured by qRT-PCR according to BioRad PrimePCR™ Probe Assay instructions and fold changes in HBG1 expression were calculated. When HBG1 expression fold changes were plotted against the HbF enrichment score from lentiviral-mediated screen, a positive correlation was seen (FIG. **27**), demonstrating that HbF enrichment score based on lentiviral transduction is predictive of the HbF induction by RNP transfection.

[0300] Mobilized peripheral blood CD34+ cells were transfected with 4-8 M RNPs made by complexing *S. pyogenes* Cas9 protein with a subset of Tier 1 and Tier 2 single gRNAs from Table 12. CD34+ cells were differentiated into erythroid cells and lysed by repeated freeze-thaw in water. Cell lysates were cleared by centrifugation followed by filtration. Relative ratios of individual globin chains in the cell lysates were determined by reverse phase ultra performance liquid chromatography (Chang 2017). HbF level was calculated as ((Aγ-globin+ Gγ-globin)/(Aγ-globin+ Gγ-globin+β-globin) %). Increased HbF levels were observed as compared to mock transfected samples as provided in Table 17 demonstrating the disruption of repressive elements identified through lenti-mediated screen could lead to fetal globin induction in erythroid cells derived from primary human hematopoietic stem and progenitor cells.

TABLE-US-00017 TABLE 17 HbF Expression in Erythroid Progeny Derived from Transfected mPB CD34+ Cells Name of gRNA sequence (RNA) Tier Targeting domain (%) HbF level No gRNA — — 7.57 ± 2.36 (Control) Sp35 CUUGUCAAGGCUAUUGGUCA 1 30.31 ± 1.33 (SEQ ID NO: 339) #2 GGAAUGACUGAAUCGGAACA 1 19 (SEQ ID NO: 294) #3 GCAUUGAGAUAGUGUGGGGA 2 22.50 ± 3.54 (SEQ ID NO: 295) #4 CAUUGAGAUAGUGUGGGGAA 1 25.50 ± 2.12 (SEQ ID NO: 332) #5 AUUGAGAUAGUGUGGGGAAG 1 33.70 ± 3.28 (SEQ ID NO: 354) Sp36 CAAGGCUAUUGGUCAAGGCA 1 22.08 ± 6.82 (SEQ ID NO: 338)

[0301] The experiment described above was repeated for gRNAs #3, #4 and #5, using guide RNAs synthesized from different vendors, RNP doses comprised within 4-16 M and gRNA to protein molar stoichiometry comprised within 2:1-4:1. Globin chain analysis in erythroid cells derived from electroporated mPB CD34+ cells demonstrated HbF levels reaching up to 42.80% in RNP treated samples versus 11.99% in mock treated samples (Table 18). Those results, showing clinically relevant levels of fetal hemoglobin expression, support the use of such gRNAs, complexed with *S. pyogenes* Cas9 protein, and electroporated into human HSPCs to provide a therapeutic cell population for the treatment of beta-hemoglobinopathies.

TABLE-US-00018 TABLE 18 Therapeutic levels of HbF Expression in Erythroid Progeny Derived from Transfected mPB CD34+ Cells Name of gRNA sequence (RNA) Tier Targeting domain (%) HbF level No gRNA — — 11.99 ± 3.60 (Control) #3 GCAUUGAGAUAGUGUGGGGA 2 41.23 ± 5.54 (SEQ ID NO: 295) #4 CAUUGAGAUAGUGUGGGGAA 1 36.83 ± 1.44 (SEQ ID NO: 332) #5 AUUGAGAUAGUGUGGGGAAG 1 42.80 ± 2.19 (SEQ ID NO: 354)

Example 11: Infusion of Edited mPB CD34+ Cells into NOD,B6.SCID Il2rγ−/− Kit(W41/W41) Mice Results in Long Term Engraftment and HbF Induction

[0302] To determine whether delivery of RNP #3 (comprising gRNA #3 targeting domain (SEQ ID NO:295, Table 18), complexed with *S. pyogenes* wildtype Cas9) achieves edits in long term repopulating hematopoietic stem cells, human adult CD34.sup.+ cells from mobilized peripheral blood (mPB) were infused into nonirradiated NOD,B6.SCID Il2rγ−/− Kit(W41/W41) (Jackson lab stock name: NOD.Cg-Kit<W-41J> Tyr<+> Prkdc<scid> Il2rg<tm1Wjl>/ThomJ) (“NBSGW”)

mice. Briefly, 62.5×10.sup.6/mL mPB CD34+ cells were electroporated via MaxCyte electroporation with RNP #3 at a dose of 16 μM with a complexation ratio of 4:1 (gRNA:Cas9 protein) following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. After 24 hours, mCD34+ cells were cryopreserved. One day later, mock-transfected (no gRNA added) or RNP #3-transfected mPB CD34+ cells were thawed and infused into NBSGW mice at 1 million cells per mouse via intravenous tail vein injection. Eight weeks later, mice were euthanized and bone marrow (BM) was collected from femurs, tibias, and pelvic bones. Human chimerism and lineage reconstitution (CD45+, CD15+, CD19+, glycophorin A (GlyA, CD235a+), lineage, and CD34+, and mouse CD45+ marker expression) in BM was determined by flow cytometry and analyzed. FIG. **28**A depicts the frequency of individual populations in the BM. Human chimerism was defined as human CD45/(human CD45+mCD45). The frequency of GlyA+ cells was calculated as GlyA+ cells/total cells in BM. All other markers were calculated as marker+ cells/human CD45+ cells.

[0303] Similar chimerism and lineage distributions were achieved 8-weeks post-transplant by RNP #3-transfected mPB CD34+ cells compared to mock-transfected mPB CD34+ cells demonstrating that editing is comparable with retaining the engraftment potential of hematopoietic stem cells. FIG. **28**B depicts the indels, as determined by next generation sequencing, of unfractionated BM, or flow-sorted individual populations. Approximately 80% of human alleles from the RNP-treated group were found to carry indels, suggesting hematopoietic stem cells were successfully edited. A similar indel frequency was observed across total unfractionated BM and individual lineages, suggesting that the editing at this site does not cause lineage skewing.

[0304] Lastly, long term HbF induction by RNP #3 edited CD235a+ (GlyA+) erythroid cells was analyzed. Briefly, unfractionated BM cells extracted from mice 8 weeks after infusion were placed in erythroid culture conditions for 18 days. FIG. **28**C depicts the HbF expression, calculated as gamma/beta-like chains (%) by erythroid cells. Erythroid cells from mock-transfected group expressed approximately 21% HbF whereas those from RNP-treated group expressed approximately 42% HbF, significantly higher than mock-transfected group and potentially clinically relevant. These data demonstrate that robust long-term HbF induction is achieved by RNP #3 editing of human CD34+ cells.

Example 12: Treatment of R-Hemoglobinopathy Using Edited Hematopoietic Stem Cells

[0305] The methods and genome editing systems disclosed herein may be used for the treatment of a β-hemoglobinopathy, such as sickle cell disease or beta-thalassemia, in a patient in need thereof. For example, genome editing may be performed on cells derived from the patient in an autologous procedure. Correction of the patient's cells ex-vivo and reintroduction of the cells into the patient may result in increased HbF expression and treatment of the β-hemoglobinopathy.

[0306] For example, HSCs may be extracted from the bone marrow of a patient with a β-hemoglobinopathy using techniques that are well-known to skilled artisans. The HSCs may be modified using methods disclosed herein for genome editing. For example, RNPs comprised of guide RNAs that target one or more regions in Table 13 complexed with an RNA-guided nuclease may be used to edit the HSCs. In certain embodiments, the gRNAs may be one or more gRNAs set forth in Table 12. In certain embodiments, modified HSCs have an increase in the frequency or level of an indel in the human HBG1 gene, HBG2 gene, or both, relative to unmodified HSCs. In certain embodiments, the modified HSCs can differentiate into erythroid cells that express an increased level of HbF. A population of the modified HSCs may be selected for reintroduction into the patient via transfusion or other methods known to skilled artisans. The population of modified HSCs for reintroduction may be selected based on, for example, increased HbF expression of the erythroid progeny of the modified HSCs or increased indel frequency of the modified HSCs. In some embodiments, any form of ablation prior to reintroduction of the cells may be used to enhance engraftment of the modified HSCs. In other embodiments, peripheral blood stem cells (PBSCs) can be extracted from a patient with a β-hemoglobinopathy using techniques that are well-

known to skilled artisans (e.g., apheresis or leukapheresis) and stem cells can be removed from the PBSCs. The genome editing methods described above can be performed on the stem cells and the modified stem cells can be reintroduced into the patient as described above.

SEQUENCES

[0307] Genome editing system components according to the present disclosure (including without limitation, RNA-guided nucleases, guide RNAs, donor template nucleic acids, nucleic acids encoding nucleases or guide RNAs, and portions or fragments of any of the foregoing), are exemplified by the nucleotide and amino acid sequences presented in the Sequence Listing. The sequences presented in the Sequence Listing are not intended to be limiting, but rather illustrative of certain principles of genome editing systems and their component parts, which, in combination with the instant disclosure, will inform those of skill in the art about additional implementations and modifications that are within the scope of this disclosure. A list of the sequences presented is provided in the following Table 19.

TABLE-US-00019 TABLE 19 Sequences presented in the Sequence Listing: SEQ ID NOS: Description 1-2, 4-6, 12, 14 Cas9 polypeptides 3, 7-11, 13 Cas9 coding sequences 15-23, 52-123 Cas9 RuvC-like domains 24-28, 124-198 Cas9 HNH-like domains 29-31, 38-51 Full-length modular and unimolecular gRNAs 32-37 gRNA proximal and tail domains 199-205 PAM sequences 251-901, 940-942, gRNA targeting domains (RNA)- see 952-955, 1329-1639 Tables 2, 7, 9, 10, 12, 17, 18 910-919, 943-945, gRNA targeting domains (DNA)- see 956-959, 978-1328 Tables 7, 9, 12 920-929, 946-948, gRNA targeting domains plus PAM 960-963 (NGG) (RNA) - see Tables 7, 9 930-939, 949-951, gRNA targeting domains plus PAM 964-967 (NGG) (DNA) - see Tables 7, 9 970, 972, 974, 976 dgRNA targeting domains (RNA) - see Table 10 971, 973, 975, 977 dgRNA targeting domains (DNA) - see Table 10 902, 903 Human HBG1, 2 promoter sequences including HPFH deletion site 904-909 Oligonucleotide donor sequences and homology arms - see Table 8 968-969 BCL11Ae sequences 1640-1660 Genomic Coordinates of HbG Regions of interest, Table 13

INCORPORATION BY REFERENCE

[0308] All publications, patents, and patent applications mentioned herein are hereby incorporated by reference in their entirety as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

EQUIVALENTS

[0309] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments described herein. Such equivalents are intended to be encompassed by the following claims.

REFERENCES

[0310] Aliyu et al. Am J Hematol 83:63-70 (2008) [0311] Angastiniotis & Modell Ann N Y Acad Sci 850:251-269 (1998) [0312] Anders et al. Nature 513(7519):569-573 (2014) [0313] Bae et al. Bioinformatics 30(10):1473-1475 (2014) [0314] Bothmer et al. Nat Commun 8:13905 (2017) [0315] Bouva Hematologica 91(1):129-132 (2006) [0316] Briner et al. Mol Cell 56(2):333-339 (2014) [0317] Brousseau Am J Hematol 85(1):77-78 (2010) [0318] Canvers et al. Nature 527(12):192-197 (2015) [0319] Chang et al. Mol Ther Methods Clin Dev 4:137-148 (2017) [0320] Chen et al. Nat Commun 8:14958 (2017) [0321] Cong et al. Science 399(6121):819-823 (2013) [0322] Cornish-Bowden Nucleic Acids Res 13(9):3021-3030 (1985) [0323] Davis & Maizels Proc Natl Acad Sci USA 111(10):E924-E932 (2014) [0324] Fine et al. Sci Rep 5:10777 (2015) [0325] Frit et al. DNA Repair (Amst.) 17:81-97 (2014) [0326] Fu et al. Nat Biotechnol 32(3):279-284 (2014) [0327] Guilinger et al. Nat Biotechnol 32(6):577-582 (2014) [0328] Heigwer et al. Nat Methods 11(2):122-123 (2014) [0329] Hinz et al. J Biol Chem 291(48):24851-24856 (2016) [0330] Hsu et al. Nat Biotechnol 31(9):827-832 (2013) [0331] Iyama & Wilson DNA Repair (Amst.) 12(8):620-636 (2013) [0332] Jiang et al. Nat Biotechnol 31(3):233-239 (2013) [0333] Jinek et al.

Science 337(6096):816-821 (2012) [0334] Jinek et al. Science 343(6176):1247997 (2014) [0335] Joung et al., Nat Protocols 12: 828-863 (2017) [0336] Kleinstiver et al. Nature 523(7561):481-485 (2015a) [0337] Kleinstiver et al. Nat Biotechnol 33(12):1293-1298 (2015b) [0338] Kleinstiver et al. Nature 529(7587):490-495 (2016) [0339] Komor et al. Nature 533(7603):420-424 (2016) [0340] Kurita et al. PLoS ONE 8(3):e59890 (2013) [0341] Lee et al. Nano Lett 12(12):6322-6327 (2012) [0342] Lewis “Medical-Surgical Nursing: Assessment and Management of Clinical Problems” (2014) [0343] Makarova et al. Nat Rev Microbiol 9(6):467-477 (2011) [0344] Mali et al. Science 339(6121):823-826 (2013) [0345] Nishimasu et al. Cell 156(5):935-949 (2014) [0346] Nishimasu et al. Cell 162(5):1113-1126 (2015) [0347] Ran et al. Cell 154(6):1380-1389 (2013) [0348] Ran et al. Nature 520(7546):186-191 (2015) [0349] Richardson et al. Nat Biotechnol 34(3):339-344 (2016) [0350] Shmakov et al. Mol Cell 60:385-397 (2015) [0351] Thein Hum Mol Genet 18(R2):R216-223 (2009) [0352] Thorpe Br J Haematol. 87(1):125-132 (1994) [0353] Tsai et al. Nat Biotechnol 34(5):483 (2016) [0354] Wang et al. Cell 153(4):910-918 (2013) [0355] Xiao et al. Bioinformatics 30(8):1180-1182 (2014) [0356] Yamano et al. Cell 165(4):949-962 (2016) [0357] Zetsche et al. Nat Biotechnol 33(2):139-142 (2015a) [0358] Zetsche et al. Cell 163(3):759-771 (2015b)

Claims

1-**43**. (canceled)

44. A synthetic guide RNA (gRNA) comprising a targeting domain complementary to a sequence located within one or more regions selected from the group consisting of Chr 11 (NC_000011.10): 5,249,198-5,249,362 (Region 3); 5,249,591-5,249,712 (Region 4); 5,250,464-5,250,549 (Region 10); 5,250,594-5,250,735 (Region 11); 5,253,425-5,254,121 (Region 12); 5,254,122-5,254,306 (Region 13); 5,254,511-5,254,648 (Region 14); 5,255,255-5,255,292 (Region 20); and 5,255,518-5,255,641 (Region 21).

45. The synthetic gRNA of claim 44 comprising: (a) the targeting domain differing by no more than 3 nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1331-44, 1346-47, 1350-77, 1380, 1382-83, 1401-02, 1405-15, 1421, 1442-44, 1446, 1547, 1552-54, 1556-63, 1586, 1592-93, 1633-39; or (b) the targeting domain consisting of positions 5-20 of a sequence selected from the group consisting of SEQ ID NOs: 1331-44, 1346-47, 1350-77, 1380, 1382-83, 1401-02, 1405-15, 1421, 1442-44, 1446, 1547, 1552-54, 1556-63, 1586, 1592-93, 1633-39.

46. The synthetic gRNA of claim 45, further comprising (a) the targeting domain differing by no more than 3 nucleotides from a Tier 1 or Tier 2 targeting domain sequence selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402, 1414, and 1446; or (b) the targeting domain consisting of positions 5-20 of a Tier 1 or Tier 2 targeting domain sequence selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402, 1414, and 1446.

47. The synthetic gRNA of claim 46, wherein the targeting domain comprises a Tier 1 or Tier 2 gRNA selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402, 1414, and 1446.

48. The synthetic gRNA of claim 47, wherein the gRNA comprises one or more 2-o-methyl modifications, one or more phosphorothioate modifications, or a combination thereof.

49. A genome editing system, comprising: an RNA-guided nuclease; and at least one guide RNA (gRNA) comprising a targeting domain complementary to a sequence within one or more regions selected from the group consisting of Chr 11 (NC_000011.10): 5,249,198-5,249,362 (Region 3); 5,249,591-5,249,712 (Region 4); 5,250,464-5,250,549 (Region 10); 5,250,594-5,250,735 (Region 11); 5,253,425-5,254,121 (Region 12); 5,254,122-5,254,306 (Region 13); 5,254,511-5,254,648 (Region 14); 5,255,255-5,255,292 (Region 20); and 5,255,518-5,255,641 (Region 21).

50. The genome editing system of claim 49, wherein the at least one gRNA comprises: (a) the

targeting domain differing by no more than 3 nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1331-44, 1346-47, 1350-77, 1380, 1382-83, 1401-02, 1405-15, 1421, 1442-44, 1446, 1547, 1552-54, 1556-63, 1586, 1592-93, 1633-39; or (b) the targeting domain consisting of positions 5-20 of a sequence selected from the group consisting of SEQ ID NOs: 1331-44, 1346-47, 1350-77, 1380, 1382-83, 1401-02, 1405-15, 1421, 1442-44, 1446, 1547, 1552-54, 1556-63, 1586, 1592-93, 1633-39.

51. The genome editing system of claim 50, further comprising (a) the targeting domain differing by no more than 3 nucleotides from a Tier 1 or Tier 2 targeting domain sequence selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402, 1414, and 1446; or (b) the targeting domain consisting of positions 5-20 of a Tier 1 or Tier 2 targeting domain sequence selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402, 1414, and 1446.

52. The genome editing system of claim 49, wherein the targeting domain comprises a Tier 1 or Tier 2 targeting domain sequence selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402, 1414, and 1446.

53. The genome editing system of claim 49, wherein the RNA-guided nuclease is a nickase, and optionally lacks RuvC activity.

54. The genome editing system of claim 49, further comprising a second guide RNA.

55. A CRISPR-mediated method of altering a cell, comprising: introducing a first DNA single strand break (SSB) or double strand break (DSB) within a region selected from the group consisting of Chr 11 (NC_000011.10): 5,249,198-5,249,362 (Region 3); 5,249,591-5,249,712 (Region 4); 5,250,464-5,250,549 (Region 10); 5,250,594-5,250,735 (Region 11); 5,253,425-5,254,121 (Region 12); 5,254,122-5,254,306 (Region 13); 5,254,511-5,254,648 (Region 14); 5,255,255-5,255,292 (Region 20); and 5,255,518-5,255,641 (Region 21).

56. The CRISPR-mediated method of claim 55, wherein the first SSB or DSB is repaired by the cell in a manner that alters the regulation of an HBG1 gene or an HBG2 gene.

57. The CRISPR-mediated method of claim 55, wherein the first SSB or DSB results in the formation of at least one indel, insertion or deletion in the region selected from the group consisting of Chr 11 (NC_000011.10): 5,249,198-5,249,362 (Region 3); 5,249,591-5,249,712 (Region 4); 5,250,464-5,250,549 (Region 10); 5,250,594-5,250,735 (Region 11); 5,253,425-5,254,121 (Region 12); 5,254,122-5,254,306 (Region 13); 5,254,511-5,254,648 (Region 14); 5,255,255-5,255,292 (Region 20); and 5,255,518-5,255,641 (Region 21).

58. The CRISPR-mediated method of claim 55, wherein the SSB or DSB is formed via an editing event by a genome editing system, comprising: an RNA-guided nuclease; and at least one guide RNA (gRNA) comprising a targeting domain complementary to a sequence within one or more regions selected from the group consisting of Chr 11 (NC_000011.10): 5,249,198-5,249,362 (Region 3); 5,249,591-5,249,712 (Region 4); 5,250,464-5,250,549 (Region 10); 5,250,594-5,250,735 (Region 11); 5,253,425-5,254,121 (Region 12); 5,254,122-5,254,306 (Region 13); 5,254,511-5,254,648 (Region 14); 5,255,255-5,255,292 (Region 20); and 5,255,518-5,255,641 (Region 21).

59. The CRISPR-mediated method of claim 58, wherein the at least one gRNA comprises: (a) the targeting domain differing by no more than 3 nucleotides from a sequence selected from the group consisting of SEQ ID NOs: 1331-44, 1346-47, 1350-77, 1380, 1382-83, 1401-02, 1405-15, 1421, 1442-44, 1446, 1547, 1552-54, 1556-63, 1586, 1592-93, 1633-39; or (b) the targeting domain consisting of positions 5-20 of a sequence selected from the group consisting of SEQ ID NOs: 1331-44, 1346-47, 1350-77, 1380, 1382-83, 1401-02, 1405-15, 1421, 1442-44, 1446, 1547, 1552-54, 1556-63, 1586, 1592-93, 1633-39.

60. The CRISPR-mediated method of claim 59, further comprising (a) the targeting domain differing by no more than 3 nucleotides from a Tier 1 or Tier 2 targeting domain sequence selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402,

1414, and 1446; or (b) the targeting domain consisting of positions 5-20 of a Tier 1 or Tier 2 targeting domain sequence selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402, 1414, and 1446.

61. The CRISPR-mediated method of claim 60, wherein the targeting domain comprises a Tier 1 or Tier 2 targeting domain sequence selected from the group consisting of SEQ ID NOs: 1331-47, 1350-62, 1364-77, 1380, 1382-83, 1402, 1414, and 1446.

62. The CRISPR-mediated method of claim 58, wherein the RNA-guided nuclease is a nickase, and optionally lacks RuvC activity.

63. The CRISPR-mediated method of claim 58, further comprising a second guide RNA.