SYSTEMS AND METHODS FOR THE TREATMENT OF HEMOGLOBINOPATHIES

REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the benefit of United States Provisional Application Nos. 62/643,168, filed March 14, 2018, 62/767,488, filed November 14, 2018, and 62/773,073, filed November 29, 2018, the contents of which are hereby incorporated by reference in their entirety.

SEQUENCE LISTING

[0002] This application contains a Sequence Listing, which was submitted in ASCII format via EFS-Web, and is hereby incorporated by reference in its entirety. The ASCII copy, created on March 14, 2019, is named 2019-03-14_EM115PCT_Editas_8013WO02_Sequence_Listing.txt and is 811 KB 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 (Aγ)-globin chain (*HBG1*, also known as gamma globin A), and G gamma (Gγ)-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 γ-globin chains are replaced with two delta (Δ)-globin chains. δ-Thal is associated with mutations in the Δ 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^0/β^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, ribonucleoprotein (RNP) complexes, guide RNAs, Cpf1 proteins, including modified Cpf1 proteins (Cpf1 variants), and CRISPR-mediated methods for altering the promoter region of one or more γ–globin genes (e.g., *HBG1*, *HBG2*, or *HBG1* and *HBG2*) and increasing expression of fetal hemoglobin (HbF). In certain embodiments, an RNP complex may include a guide RNA (gRNA) complexed to a wild-type Cpf1 or modified Cpf1 RNA-guided nuclease (modified Cpf1 protein). In certain embodiments, a gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, a gRNA may comprise a gRNA targeting domain. In certain embodiments, a gRNA targeting domain may comprise a sequence selected from the group consisting of SEQ ID NOs:1002, 1254, 1258, 1260, 1262, and 1264. In certain embodiments, a gRNA may comprise a gRNA sequence set forth in **Table 19**. In certain embodiments, a gRNA may comprise a sequence selected from the group consisting of SEQ ID NOs:1022, 1023, 1041-1105. In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, an RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a modified Cpf1 protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**).

[0019] The inventors have discovered herein that delivery of an RNP complex including a gRNA complexed to a modified Cpf1 protein may result in increased editing of a target nucleic acid. In certain embodiments, the modified Cpf1 protein may contain one or more modifications. In certain embodiments, the one or more modifications may include, without limitation, one or more mutations in a wild-type Cpf1 amino acid sequence, one or more mutations in a wild-type Cpf1 nucleic acid sequence, one or more nuclear localization signals (NLS), one or more purification tags (e.g., His tag), or a combination thereof. In certain embodiments, a modified Cpf1 may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). In certain embodiments, the gRNA may be a modified or unmodified gRNA. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a modified Cpf1 protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**). In certain embodiments, an RNP complex comprising a modified Cpf1 protein may increase editing of a target nucleic acid. In certain embodiments, an RNP complex comprising a modified Cpf1 protein may increase editing resulting in an increase of productive indels. In various embodiments, an increase in editing of the target nucleic acid may be assessed by any means known to skilled artisans, such as, but not limited to, PCR amplification of

the target nucleic acid and subsequent sequencing analysis (e.g., Sanger sequencing, next generation sequencing).

[0020] The inventors have also discovered herein that delivery of an RNP complex including a modified gRNA complexed to an unmodified or modified Cpf1 protein may result in increased editing of a target nucleic acid. In certain embodiments, the modified gRNA may comprise one or more modifications including a phosphorothioate linkage modification, a phosphorodithioate (PS2) linkage modification, a 2'-O-methyl modification, one or more or a stretch of deoxyribonucleic acid (DNA) bases (also referred to herein as a "DNA extension"), one or more or a stretch of ribonucleic acid (RNA) bases (also referred to herein as a "RNA extension"), or combinations thereof. In certain embodiments, the DNA extension may comprise a sequence set forth in **Table 24**. For example, in certain embodiments, the DNA extension may comprise a sequence set forth in SEQ ID NOs:1235-1250. In certain embodiments, the RNA extension may comprise a sequence set forth in **Table 24**. For example, in certain embodiments, the RNA extension may comprise a sequence set forth in SEQ ID NOs:1231-1234, 1251-1253. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a modified Cpf1 protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**). In certain embodiments, an RNP complex comprising a modified gRNA may increase editing of a target nucleic acid. In certain embodiments, an RNP complex comprising a modified gRNA may increase editing resulting in an increase of productive indels.

[0021] In certain embodiments, an RNP complex comprising a modified gRNA and a modified Cpf1 protein may increase editing of a target nucleic acid. In certain embodiments, an RNP complex comprising a modified gRNA and a modified Cpf1 protein may increase editing resulting in an increase of productive indels.

[0022] The inventors have also discovered that codelivery of an RNP complex comprising a gRNA complexed to a Cpf1 molecule (e.g., "gRNA-Cpf1-RNP") with a "booster element" may result in increased editing of a target nucleic acid. In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a modified Cpf1 protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**). As used herein, the term "booster element" refers to an element which, when co-delivered with a RNP complex comprising a gRNA complexed to an RNA-guided nuclease ("gRNA-nuclease-RNP"), increases editing of a target nucleic acid compared with editing of the target nucleic acid without the booster element. In certain embodiments, one or more booster elements may be codelivered with a gRNA-nuclease-RNP complex to increase editing of a target nucleic acid. In

certain embodiments, codelivery of a booster element may increase editing resulting in an increase of productive indels. In various embodiments, an increase in editing of the target nucleic acid may be assessed by any means known to skilled artisans, such as, but not limited to, PCR amplification of the target nucleic acid and subsequent sequencing analysis (e.g., Sanger sequencing, next generation sequencing).

[0023] In certain embodiments, a gRNA-nuclease-RNP may comprise a gRNA-Cpf1-RNP. In certain embodiments, a Cpf1 molecule of the gRNA-Cpf1-RNP complex may be a wild-type Cpf1 or modified Cpf1. In certain embodiments, the Cpf1 molecule of the gRNA-Cpf1-RNP may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). In certain embodiments, the gRNA-Cpf1-RNP complex may comprise a gRNA comprising a targeting domain set forth in **Table 13** or **Table 18**. In certain embodiments, the gRNA-Cpf1-RNP complex may comprise a gRNA comprising a sequence set forth in **Table 19**. In certain embodiments, the gRNA may be a modified or unmodified gRNA.

[0024] In certain embodiments, a booster element may comprise a dead RNP comprising a dead gRNA molecule complexed with an RNA-guided nuclease molecule ("dead gRNA-nuclease-RNP"). In certain embodiments, the dead gRNA-nuclease-RNP may comprise a dead gRNA complexed with a wild-type (WT) Cas9 molecule ("dead gRNA-Cas9-RNP"), a dead gRNA complexed with a Cas9 nickase molecule ("dead gRNA-nickase-RNP") or a dead gRNA complexed with an enzymatically inactive (ei) Cas9 molecule ("dead gRNA-eiCas9-RNP"). In certain embodiments, the dead gRNA-nuclease-RNP complex may have decreased activity or lack nuclease activity. In certain embodiments, the dead gRNA of the dead gRNA-nuclease-RNP complex may comprise any of the dead gRNAs set forth herein. For example, the dead gRNA may comprise a targeting domain may be the same as or may differ by no more than 3 nucleotides from a dead gRNA targeting domain set forth in **Table 10** or **Table 15**. In certain embodiments, the dead gRNA may include a targeting domain comprising a truncation of a gRNA targeting domain. In certain embodiments, the gRNA targeting domain to be truncated may be a gRNA targeting domain set forth in **Table 2**, **Table 10**, or **Table 15**. In certain embodiments, the dead gRNA may be a modified or unmodified dead gRNA. As shown herein, codelivery of a gRNA-Cpf1-RNP with a dead gRNA-Cas9-RNP (i.e., an RNP comprising a dead gRNA complexed to a WT Cas9) or codelivery of a gRNA-Cpf1-RNP with a dead gRNA-nickase-RNP (i.e., an RNP comprising a dead gRNA complexed to a Cas9 nickase (i.e., the Cas9 D10A nickase)) resulted in an increase in total editing above levels observed following delivery of gRNA-Cpf1-RNP alone (see, e.g., Examples, 15, 17, 18). Dead gRNA molecules may comprise targeting domains complementary to regions proximal to or within a target region (e.g., the CCAAT box target region, 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 CCAAT box target region, 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in BCL11Ae). In certain embodiments, one or more booster elements may be comprised of one or more dead gRNA-nuclease-RNPs, e.g., dead gRNA-Cas9-RNP, dead gRNA-nickase-RNP, dead gRNA-eiCas9-RNP, to be codelivered with a gRNA-Cpf1-RNP. In certain embodiments, codelivery of a dead gRNA-nuclease-RNP does not alter the indel profile of a gRNA-Cpf1-RNP.

[0025] In certain embodiments, a booster element may comprise an RNP complex comprising a gRNA molecule complexed with an RNA-guided nuclease nickase molecule ("gRNA-nickase-RNP"). In certain embodiments, the RNA-guided nuclease nickase molecule may be a Cas9 nickase molecule, e.g., Cas9 D10A nickase. In certain embodiments, the gRNA of the gRNA-nickase-RNP may comprise any of the gRNAs set forth herein. For example, the gRNA may comprise a gRNA targeting domain set forth in **Table 2**, **Table 10**, or **Table 15**. In certain embodiments, the gRNA may be a modified or unmodified gRNA. As shown herein, codelivery of gRNA-Cpf1-RNP with a gRNA-nickase-RNP complex (RNP comprising a guide RNA complexed to a Cas9D10A nickase molecule) resulted in an increase in total editing above levels observed following delivery of gRNA-Cpf1-RNP alone (see, e.g., Examples 15, 16). Additionally, codelivery of a gRNA-nickase-RNP complex with a gRNA-Cpf1-RNP complex altered the directionality, length, and/or position of the indel profile of gRNA-Cpf1-RNP. In certain embodiments, a booster enhancer may be used to provide a desired editing outcome, for example, to increase the rate of productive indels. In certain embodiments, codelivery of a gRNA-nickase-RNP complex with a gRNA-Cpf1-RNP complex may alter the indel profile of the gRNA-Cpf1-RNP.

[0026] In certain embodiments, a booster element may comprise a single stranded oligodeoxynucleotide (ssODN) or a double stranded oligodeoxynucleotide (dsODN). In certain embodiments, the ssODN may be any ssODN disclosed herein. In certain embodiments, an ssODN may comprise a sequence set forth in **Table 11**. For example, in certain embodiments, an ssODN may comprise the sequence set forth in SEQ ID NO:1040.

[0027] In one aspect, the disclosure relates to an RNP complex comprising a CRISPR from Prevotella and Franciscella 1 (Cpf1) RNA-guided nuclease or a variant thereof and a gRNA, wherein the gRNA is capable of binding to a target site in a promoter of an *HBG* gene in a cell. In certain embodiments, the gRNA may be modified or unmodified. In certain embodiments, the gRNA may comprise one or more modifications including a phosphorothioate linkage modification, a phosphorodithioate (PS2) linkage modification, a 2'-O-methyl modification, a DNA extension, an RNA extension, or combinations thereof. In certain embodiments, the DNA extension may comprise a sequence set forth in **Table 24**. In certain

embodiments, the RNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a Cpf1 variant protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**). In certain embodiments, the Cpf1 variant protein may contain one or more modifications. In certain embodiments, the one or more modifications may include, without limitation, one or more mutations in a wild-type Cpf1 amino acid sequence, one or more mutations in a wild-type Cpf1 nucleic acid sequence, one or more nuclear localization signals (NLS), one or more purification tags (e.g., His tag), or a combination thereof. In certain embodiments, a Cpf1 variant protein may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences).

[0028] In one aspect, the disclosure relates to a method of altering a promoter of an *HBG* gene in a cell comprising contacting the cell with an RNP complex disclosed herein. In certain embodiments, the alteration may comprise an indel within one or more regions set forth in **Table 17**. In certain embodiments, the alteration may comprise an indel within a CCAAT box target region of the promoter of an *HBG* gene. For example, in certain embodiments, the alteration may comprise an indel within Chr 11 (NC_000011.10): 5,249,955 – 5,249,987 (**Table 17**, Region 6), Chr 11 (NC_000011.10): 5,254,879 – 5,254,909 (**Table 17**, Region 16), or a combination thereof. In certain embodiments, the RNP complex may comprise a gRNA and a Cpf1 protein. In certain embodiments, the gRNA may comprise an RNA targeting domain set forth in **Table 19**. In certain embodiments, the gRNA targeting domain may comprise a sequence selected from the group consisting of SEQ ID NOs:1002, 1254, 1258, 1260, 1262, and 1264. In certain embodiments, the gRNA may comprise a gRNA sequence set forth in **Table 19**. In certain embodiments, the gRNA may comprise a sequence selected from the group consisting of SEQ ID NOs:1022, 1023, 1041-1105. In certain embodiments, a gRNA may be configured to provide an editing event at Chr11:5249973, Chr11:5249977 (HBG1); Chr11:5250042, Chr11:5250046 (HBG1); Chr11:5250055, Chr11:5250059 (HBG1); Chr11:5250179, Chr11:5250183 (HBG1); Chr11:5254897, Chr11:5254901 (HBG2); Chr11:5254897, Chr11:5254901 (HBG2); Chr11:5254966, 5254970 (HBG2); Chr11:5254979, 5254983 (HBG2) (**Table 22**, **Table 23**). In certain embodiments, the cell may be further contacted with a booster element. In certain embodiment, a booster element may comprise a single stranded oligodeoxynucleotide (ssODN) or a double stranded oligodeoxynucleotide (dsODN). In certain embodiments, the ssODN may be any ssODN disclosed herein. In certain embodiments, an ssODN may comprise a sequence set forth in **Table 11**. For example, in certain embodiments, an ssODN may comprise the sequence set forth in SEQ ID NO:1040.

[0029] In one aspect, the disclosure relates to an isolated cell comprising an alteration in a promoter of *HBG* gene generated by the delivery of an RNP complex to the cell. In certain embodiments, the RNP complex may comprise a gRNA and a Cpf1 protein. In certain embodiments, the gRNA may be modified or unmodified. In certain embodiments, the gRNA may comprise one or more modifications including a phosphorothioate linkage modification, a phosphorodithioate (PS2) linkage modification, a 2'-O-methyl modification, a DNA extension, an RNA extension, or combinations thereof. In certain embodiments, the DNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the RNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a Cpf1 variant protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**). In certain embodiments, the Cpf1 variant protein may contain one or more modifications. In certain embodiments, the one or more modifications may include, without limitation, one or more mutations in a wild-type Cpf1 amino acid sequence, one or more mutations in a wild-type Cpf1 nucleic acid sequence, one or more nuclear localization signals (NLS), one or more purification tags (e.g., His tag), or a combination thereof. In certain embodiments, a Cpf1 variant protein may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). In certain embodiments, a booster element may be co-delivered with the RNP complex. In certain embodiment, a booster element may comprise a single stranded oligodeoxynucleotide (ssODN) or a double stranded oligodeoxynucleotide (dsODN). In certain embodiments, the ssODN may be any ssODN disclosed herein. In certain embodiments, an ssODN may comprise a sequence set forth in **Table 11**. For example, in certain embodiments, an ssODN may comprise the sequence set forth in SEQ ID NO:1040.

[0030] In one aspect, the disclosure relates to an *ex vivo* method of increasing the level of fetal hemoglobin (HbF) in a human cell by genome editing using an RNP complex comprising a gRNA and a Cpf1 RNA-guided nuclease or a variant thereof to affect an alteration in a promoter of an *HBG* gene, thereby to increase expression of HbF. In certain embodiments, the gRNA may be modified or unmodified. In certain embodiments, the gRNA may comprise one or more modifications including a phosphorothioate linkage modification, a phosphorodithioate (PS2) linkage modification, a 2'-O-methyl modification, a DNA extension, an RNA extension, or combinations thereof. In certain embodiments, the DNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the RNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, the RNP

complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a Cpf1 variant protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**). In certain embodiments, the Cpf1 variant protein may contain one or more modifications. In certain embodiments, the one or more modifications may include, without limitation, one or more mutations in a wild-type Cpf1 amino acid sequence, one or more mutations in a wild-type Cpf1 nucleic acid sequence, one or more nuclear localization signals (NLS), one or more purification tags (e.g., His tag), or a combination thereof. In certain embodiments, a Cpf1 variant protein may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). In certain embodiments, a booster element may be co-delivered with the RNP complex. In certain embodiment, a booster element may comprise a single stranded oligodeoxynucleotide (ssODN) or a double stranded oligodeoxynucleotide (dsODN). In certain embodiments, the ssODN may be any ssODN disclosed herein. In certain embodiments, an ssODN may comprise a sequence set forth in **Table 11**. For example, in certain embodiments, an ssODN may comprise the sequence set forth in SEQ ID NO:1040.

[0031] In one aspect, the disclosure relates to a population of CD34+ or hematopoietic stem cells, wherein one or more cells in the population comprises an alteration in a promoter of an *HBG* gene, which alteration is generated by delivering an RNP complex comprising a gRNA and a Cpf1 RNA-guided nuclease or a variant thereof to the population of CD34+ or hematopoietic stem cells. In certain embodiments, the gRNA may be modified or unmodified. In certain embodiments, the gRNA may comprise one or more modifications including a phosphorothioate linkage modification, a phosphorodithioate (PS2) linkage modification, a 2'-O-methyl modification, a DNA extension, an RNA extension, or combinations thereof. In certain embodiments, the DNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the RNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a Cpf1 variant protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**). In certain embodiments, the Cpf1 variant protein may contain one or more modifications. In certain embodiments, the one or more modifications may include, without limitation, one or more mutations in a wild-type Cpf1 amino acid sequence, one or more mutations in a wild-type Cpf1 nucleic acid sequence, one or more nuclear localization signals (NLS), one or more purification tags (e.g., His tag), or a combination thereof. In certain embodiments, a Cpf1 variant protein may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09

(Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). In certain embodiments, a booster element may be co-delivered with the RNP complex. In certain embodiment, a booster element may comprise a single stranded oligodeoxynucleotide (ssODN) or a double stranded oligodeoxynucleotide (dsODN). In certain embodiments, the ssODN may be any ssODN disclosed herein. In certain embodiments, an ssODN may comprise a sequence set forth in **Table 11**. For example, in certain embodiments, an ssODN may comprise the sequence set forth in SEQ ID NO:1040.

[0032] In one aspect, the disclosure relates to a method of alleviating one or more symptoms of sickle cell disease in a subject in need thereof, the method comprising: a) isolating a population of CD34+ or hematopoietic stem cells from the subject; b) modifying the population of isolated cells ex vivo by delivering an RNP complex comprising a gRNA and a Cpf1 RNA-guided nuclease or a variant thereof to the population of isolated cells, thereby to affect an alteration in a promoter of an *HBG* gene in one or more cells in the population; and c) administering the modified population of cells to the subject, thereby to alleviate one or more symptoms of sickle cell disease in the subject. In certain embodiments, the alteration may comprise an indel within a CCAAT box target region of the promoter of the HBG gene. In certain embodiments, the RNP complex may be delivered using electroporation. In certain embodiments, at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80% or at least about 90% of the cells in the population of cells comprise a productive indel.

[0033] In one aspect, the disclosure relates to a gRNA comprising a 5' end and a 3' end, and comprising a DNA extension at the 5' end and a 2'-O-methyl-3'-phosphorothioate modification at the 3' end, wherein the gRNA includes an RNA segment capable of hybridizing to a target site and an RNA segment capable of associating with a Cpf1 RNA-guided nuclease. In certain embodiments, the DNA extension may comprise a sequence set forth in SEQ ID NOs:1235-1250. In certain embodiments, the gRNA may be modified or unmodified. In certain embodiments, the gRNA may comprise one or more modifications including a phosphorothioate linkage modification, a phosphorodithioate (PS2) linkage modification, a 2'-O-methyl modification, a DNA extension, an RNA extension, or combinations thereof. In certain embodiments, the DNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the RNA extension may comprise a sequence set forth in **Table 24**. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**.

[0034] In one aspect, the disclosure relates to an RNP complex comprising a Cpf1 RNA-guided nuclease as disclosed herein and a gRNA as disclosed herein.

[0035] Also provided herein are genome editing systems, guide RNAs, 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, one or more gRNAs comprising a sequence set forth in **Table 18** or **Table 19** may be used to introduce alterations in the promoter region of the *HBG* gene. In certain embodiments, genome editing systems, guide RNAs, 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"). In certain embodiments, genome editing systems, guide RNAs, and CRISPR-mediated methods may alter a CCAAT box target region that is 5' of the transcription site of the *HBG1*, *HBG2*, or *HBG1* and *HBG2* gene ("CCAAT box target region"). In certain embodiments, the CCAAT box target region may be the region that is at or near the distal CCAAT box and includes the nucleotides of the distal CCAAT box and 25 nucleotides upstream (5') and 25 nucleotides downstream (3') of the distal CCAAT box (i.e., HBG1/2 c.-86 to -140). In certain embodiments, the CCAAT box target region may be the region that is at or near the distal CCAAT box and includes the nucleotides of the distal CCAAT box and 5 nucleotides upstream (5') and 5 nucleotides downstream (3') of the distal CCAAT box (i.e., *HBG1/2* c.-106 to -120). In certain embodiments, the CCAAT box target region may comprise a 18 nt target region, a 13 nt target region, a 11 nt target region, a 4 nt target region, a 1 nt target region, a -117G>A target region, or a combination thereof as disclosed herein. In certain embodiments, the alteration may be a 18 nt deletion, 13 nt deletion, 11 nt deletion, 4 nt deletion, 1 nt deletion, a substitution from G to A at c.-117 of the *HBG1*, *HBG2*, or *HBG1* and *HBG2* gene, or a combination thereof. In certain embodiments, the alteration may be a non-naturally occurring alteration or a naturally occurring alteration. In certain embodiments, one or more gRNAs comprising a targeting domain set forth in SEQ ID NOs:251-901 or 940-942 may be used to introduce alterations in the 13 nt target region. In certain embodiments, one or more gRNAs comprising a sequence set forth in SEQ ID NOs:251-901, 940-942, 996, 997, 970, 971, 1002, or 1003 may be used to introduce alterations in the CCAAT box target region. In certain embodiments, genome editing systems, guide RNAs, 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*.

[0036] Also provided herein in certain embodiments are the use of optional genome editing system components such as template nucleic acids (oligonucleotide donor templates). In certain embodiments, template nucleic acids for use in targeting the CCAAT target region may include, without limitation, template nucleic acids encoding alterations of the CCAAT box target region. In certain embodiments, the CCAAT box target region may comprise a 18 nt target region, a 11 nt target region, a 4 nt target region, a

1 nt target region, or a combination thereof. In certain embodiments, the template nucleic acid may be a single stranded oligodeoxynucleotide (ssODN) or a double stranded oligodeoxynucleotide (dsODN). In certain embodiments, 5' and 3' homology arms, and exemplary full-length donor templates encoding alterations at the CCAAT box target region are also presented below (e.g., SEQ ID NOS: 904-909, 974-995). In certain embodiments, the template nucleic acid may be a positive strand or a negative strand. In certain embodiments, the ssODN may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm. In certain embodiments, the 5' homology arm may be about 25 to about 200 nucleotides or more in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length; the replacement sequence may comprise 0 nucleotides in length; and the 3' homology arm may be about 25 to about 200 nucleotides or more in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the ssODN may comprise one or more phosphorothioates.

[0037] In certain embodiments, the genome editing systems, guide RNAs, and CRISPR-mediated methods for altering one or more γ–globin genes (e.g., *HBG1*, *HBG2*, or *HBG1* and *HBG2*), may include an RNA-guided nuclease. In certain embodiments, the RNA-guided nuclease may a Cas9 or modified Cas 9. In certain embodiments, the RNA-guided nuclease may a Cpf1 or modified Cpf1 as disclosed herein.

[0038] In one aspect, the disclosure relates to compositions including a plurality of cells generated by the methods 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 methods 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 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.

[0039] In one aspect, the disclosure relates to a method of altering a cell, that includes unwinding a chromatin segment within or proximal to a target region of a nucleic acid in a cell and generating a double stranded break (DSB) within the target region of the nucleic acid whereby to alter the target region. In certain embodiments, the step of unwinding the chromatin segment may include contacting the chromatin segment with an RNA-guided helicase. In certain embodiments, the step of unwinding the chromatin

does not include recruiting an exogenous trans-acting factor to the chromatin segment. The RNA-guided helicase may be an RNA-guided nuclease, and the RNA-guided nuclease may be complexed to a dead guide RNA (dgRNA) including a first targeting domain sequence of 15 or fewer nucleotides in length. In certain embodiments, the dgRNA may include modifications at the 5' or 3' end, including, but not limited to, an anti-reverse cap analog (ARCA) at the 5' end of the RNA, a polyA tail at the 3' end of the RNA, or both. In certain embodiments, the RNA-guided helicase may be an enzymatically active RNA-guided nuclease or may be configured to lack nuclease activity. In certain embodiments, the targeting domain sequence of the dgRNA may be complementary to a sequence proximal to the target region. "Proximal to," in some embodiments herein, may mean within 10, 25, 50, 100, or 200 nucleotides of the target region. In certain embodiments, the step of unwinding the chromatin segment may not include forming a single or double-stranded break in the nucleic acid within the chromatin segment. In certain embodiments, the step of generating the DSB within the target region may include contacting the chromatin segment with an RNA-guided nuclease having nuclease activity. In certain embodiments, the RNA-guided nuclease having nuclease activity may be complexed to a gRNA including a targeting domain configured to overlap the target region. In certain embodiments, the RNA-guided nuclease having nuclease activity may be a Cpf1 molecule.

[0040] Another aspect of the disclosure includes a method of inducing accessibility to a target region of a nucleic acid for editing in a cell including contacting the cell with an RNA-guided helicase and a dgRNA and unwinding DNA within or proximal to the target region with the RNA-guided helicase thereby inducing accessibility to the target region for editing. In various cases, the RNA-guided helicase and the dgRNA may be configured to associate within or proximal to the target region. In certain embodiments, the dgRNA may be configured such that it does not provide an RNA-guided nuclease cleavage event. In certain embodiments, the RNA-guided helicase and dgRNA may complex to form a dead ribonucleoprotein (RNP) that lacks cleavage activity. In certain embodiments, the dgRNA may include a targeting domain sequence of 15 or fewer nucleotides in length. In certain embodiments, the RNA-guided helicase may be an RNA-guided nuclease. In certain embodiments, the RNA-guided nuclease and dgRNA are not configured to recruit an exogenous trans-acting factor to the target region. In certain embodiments, the RNA-guided nuclease may be a Cas9 or a Cas9-fusion protein. In certain embodiments, the Cas9 may be an enzymatically active Cas9 or an enzymatically dead Cas9. In certain embodiments, the Cas9 may be a nickase, e.g., Cas9 D10A. In certain embodiments, the step of unwinding the DNA does not comprise forming a single or double-stranded break in the DNA. In certain embodiments, the RNA-guided nuclease having nuclease activity may be complexed to a gRNA including a targeting domain configured to overlap the target region. In certain embodiments, the RNA-guided nuclease having nuclease activity may be a Cpf1 molecule.

[0041] In another aspect, the disclosure relates to a method of increasing a rate of indel formation in a nucleic acid that includes unwinding double stranded DNA within or proximal to a target region of the nucleic acid using an RNA-guided helicase configured to associate within or proximal to the target region and generating a DSB within the target region. In certain embodiments, generating a DSB within the target region results in forming an indel at the target region. In certain embodiments, the DSB may be repaired in a manner forming an indel at the target region. In certain embodiments, the rate of indel formation in the gene achieved using the RNA-guided helicase is increased compared to a rate of indel formation in the gene achieved without using the RNA-guided helicase. In certain embodiments, the RNA-guided helicase may form an RNP complex with a dgRNA configured to associate within or proximal to the target region. In certain embodiments, the dgRNA may include a targeting domain sequence of 15 nucleotides or less in length. In certain embodiments, the RNA-guided helicase may be an RNA-guided nuclease. In certain embodiments, the RNA-guided nuclease may be a Cas9 or a Cas9-fusion protein. In certain embodiments, the Cas9 may be an enzymatically active Cas9 or an enzymatically dead Cas9. In certain embodiments, the Cas9 may be a nickase, e.g., Cas9 D10A. In certain embodiments, the RNA-guided nuclease and the dgRNA are not configured to recruit an exogenous trans-acting factor to the target region. In certain embodiments, the step of unwinding the double stranded DNA does not include forming a single or double-stranded break in the DNA.

[0042] In yet another aspect, this disclosure relates to a method of deleting a segment of a target nucleic acid in a cell that includes contacting the cell with an RNA-guided helicase and generating a DSB within the target region, whereby a segment of the target nucleic acid is deleted. In certain embodiments, the DSB may be repaired in a manner that deletes a segment of the target nucleic acid. In certain embodiments, the RNA-guided helicase may be 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 RNA-guided helicase may form an ribonucleoprotein complex with a dgRNA configured to associate within or proximal to the target region. In certain embodiments, the dgRNA may include a targeting domain sequence of 15 nucleotides or less in length. In certain embodiments, the RNA-guided helicase may be an RNA-guided nuclease. In certain embodiments, the RNA-guided nuclease may be a Cas9 or a Cas9-fusion protein. In certain embodiments, the Cas9 may be an enzymatically active Cas9 or an enzymatically dead Cas9. In certain embodiments, the RNA-guided nuclease and the dgRNA are not configured to recruit an exogenous trans-acting factor to the target region. In certain embodiments, the target nucleic acid may be 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 may be at least about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, or 100 base pairs in length.

[0043] The disclosure also relates to a dead gRNA (dgRNA) molecule including a targeting domain comprising a truncation of a gRNA targeting domain. In certain embodiments, the gRNA targeting domain to be truncated may be a gRNA targeting domain set forth in **Table 2**, **Table 10**, or **Table 15**. In certain embodiments, the gRNA targeting domain may be truncated from a 5' end of the gRNA targeting domain. In certain embodiments, the dgRNA may include a targeting domain sequence of 15 nucleotides or less in length. In certain embodiments, the first targeting domain may be the same as or may differ by no more than 3 nucleotides from a dgRNA targeting domain set forth in Table 10 or Table 15.

[0044] Another aspect of the disclosure relates to compositions including at least one polynucleotide encoding a plurality of gRNAs and an RNA-guided helicase, in which at least one gRNA may be a dgRNA configured such that it does not provide an RNA-guided nuclease cleavage event. In certain embodiments, the dgRNA may include a targeting domain sequence of 15 nucleotides or less in length. In certain embodiments, the RNA-guided helicase may be an RNA-guided nuclease. In certain embodiments, the RNA-guided nuclease may be a Cas9 or a Cas9-fusion protein. In certain embodiments, the Cas9 may be an enzymatically active Cas9 or an enzymatically dead Cas9. In certain embodiments, the Cas9 may be a nickase, e.g., Cas9 D10A. In certain embodiments, the RNA-guided nuclease and the dgRNA are not configured to recruit an exogenous trans-acting factor to the target region. In certain embodiments, the compositions further include a second RNA-guided nuclease configured to provide a cleavage event. In certain embodiments, the compositions further include a second gRNA configured to provide a cleavage event.

[0045] In another aspect, the disclosure relates to genome editing systems that include an RNA-guided nuclease and an RNA-guided helicase configured to associate with a target nucleic acid proximal to a target region of the target nucleic acid and induce a conformational change in the target region thereby promoting accessibility to the target region for the RNA-guided nuclease to form a break in the target region. The disclosure also relates to genome editing systems that include a dgRNA including a targeting domain sequence of 15 nucleotides or less in length, a first RNA-guided nuclease, and an RNA-guided helicase. In certain embodiments, the genome editing system further includes a gRNA. In certain embodiments, the gRNA and the first RNA-guided nuclease may associate with a target region in a target nucleic acid. In certain embodiments, the first RNA-guided nuclease may be a Cpf1 molecule. In certain embodiments, the gRNA and the first RNA-guided nuclease may associate with a first PAM sequence in a target nucleic acid, wherein the first PAM sequence is facing outward. In certain embodiments, the RNA-guided helicase may be a second RNA-guided nuclease. In certain embodiments, the second RNA-guided nuclease and the dgRNA are not configured to recruit an exogenous trans-acting factor to the target region. In certain embodiments, the dgRNA and the second RNA-guided nuclease associate within or proximal to a target region in the target nucleic acid. In certain embodiments, the first RNA-guided

nuclease and second RNA-guided nuclease may be complexed with the gRNA and dgRNAs, respectively, forming first and second ribonucleoprotein complexes.

[0046] In another aspect, the disclosure relates to a genome editing system that includes a dgRNA comprising a targeting domain sequence of 15 nucleotides or less in length, a first RNA-guided nuclease, and an RNA-guided helicase. In certain embodiments, the gRNA and the first RNA-guided nuclease may associate with a target region in a target nucleic acid. In certain embodiments, the first RNA-guided nuclease may be a Cpf1 molecule. In certain embodiments, the gRNA and the first RNA-guided nuclease may associate with a first protospacer adjacent motif (PAM) sequence in a target nucleic acid. In certain embodiments, the first PAM sequence may be facing outward. In certain embodiments, the RNA-guided helicase may be a second RNA-guided nuclease. In certain embodiments, the second RNA-guided nuclease and the dgRNA are not configured to recruit an exogenous trans-acting factor to the target region. In certain embodiments, the dgRNA and the second RNA-guided nuclease may associate within or proximal to a target region in the target nucleic acid. In certain embodiments, the first RNA-guided nuclease and second RNA-guided nuclease may be complexed with the gRNA and dgRNAs, respectively, forming first and second ribonucleoprotein complexes. In certain embodiments, the dgRNA and the second RNA-guided nuclease may associate with a second PAM sequence in a target nucleic acid, wherein the second PAM sequence may be facing outward.

[0047] In another aspect, the disclosure relates to a genome editing system that includes a dgRNA, a first gRNA comprising a second targeting domain sequence greater than 17 nucleotides in length, a first RNA-guided nuclease, and a second RNA-guided nuclease. In certain embodiments, the first RNA-guided nuclease and the dgRNA may be configured to associate within a first target region in a target nucleic acid. In certain embodiments, the second RNA-guided nuclease and the first gRNA may be configured to associate within a second target region and generate a double stranded break (DSB) in the target nucleic acid whereby to create an indel between the first target region and the second target region. In certain embodiments, the second RNA-guided nuclease may be a Cpf1 molecule. In certain embodiments, the dgRNA may comprise a first targeting domain sequence of 15 nucleotides or less in length. In certain embodiments, the dgRNA has reduced or no RNA-guided nuclease cleavage activity. In certain embodiments, the dgRNA may be configured such that it does not provide an RNA-guided nuclease cleavage event. In certain embodiments, the dgRNA and the first RNA-guided nuclease may associate with a first protospacer adjacent motif (PAM) sequence in the target nucleic acid. In certain embodiments, the first PAM sequence may be facing outward. In certain embodiments, the first gRNA and the second RNA-guided nuclease may associate with a second PAM sequence in the target nucleic acid. In certain embodiments, the second PAM sequence may be facing outward.

[0048] In another aspect, the disclosure relates to a method of altering a cell, including contacting the cell with a dgRNA, a first gRNA comprising a second targeting domain sequence greater than 17 nucleotides in length, a first RNA-guided nuclease, and a second RNA-guided nuclease. In certain embodiments, the first RNA-guided nuclease and the dgRNA may be configured to associate within a first target region in a target nucleic acid. In certain embodiments, the second RNA-guided nuclease and the first gRNA may associate within a second target region and generate a double stranded break (DSB) in the target nucleic acid whereby to create an indel between the first target region and the second target region. In certain embodiments, the second RNA-guided nuclease may be a Cpf1 molecule. In certain embodiments, the dgRNA may comprise a first targeting domain sequence of 15 nucleotides or less in length. In certain embodiments, the dgRNA has reduced or no RNA-guided nuclease cleavage activity. In certain embodiments, the dgRNA may be configured such that it does not provide an RNA-guided nuclease cleavage event. In certain embodiments, the dgRNA and the first RNA-guided nuclease may associate with a first protospacer adjacent motif (PAM) sequence in the target nucleic acid. In certain embodiments, the first PAM sequence may be facing outward. In certain embodiments, the first gRNA and the second RNA-guided nuclease may associate with a second PAM sequence in the target nucleic acid. In certain embodiments, the second PAM sequence may be facing outward.

[0049] The disclosure herein also relates to methods of altering a cells, including contacting a cell with any of the genome editing systems disclosed herein. In certain embodiments, the step of contacting the cell may comprise contacting the cell with a solution comprising first and second ribonucleoprotein complexes. In certain embodiments, the step of contacting the cell with the solution further comprises electroporating the cells, thereby introducing the first and second ribonucleoprotein complexes into the cell.

[0050] In another aspect, the disclosure relates to cells that are altered using the methods disclosed herein. Cells that include a productive indel which results in HbF expression are also disclosed herein. In certain embodiments the indel may be produced by contacting the cell with a dgRNA, a first gRNA including a second targeting domain sequence greater than 17 nucleotides in length, a first RNA-guided nuclease, and a second RNA-guided nuclease. In certain embodiments, the first RNA-guided nuclease and the dgRNA may be configured to associate within a first target region in a target nucleic acid. In certain embodiments, the second RNA-guided nuclease and the first gRNA may associate with a second target region and generate a double stranded break (DSB) in the target nucleic acid whereby to create an indel between the first target region and the second target region. In certain embodiments, the cells disclosed herein, may be capable of differentiating into an erythroblast, erythrocyte, or a precursor of an erythrocyte or erythroblast. In certain embodiments, the cell may be a CD34+ cell.

[0051] A genome editing system or method including any of all of the features described above may include a target nucleic acid comprising a human *HBG1*, *HBG2* gene, or a combination thereof. In certain embodiments, the target region may be a CCAAT box target region of the human *HBG1*, *HBG2* gene, or a combination thereof. In certain embodiments, the first targeting domain sequence may be complementary to a first sequence on a side of a CCAAT box target region of the human *HBG1*, *HBG2* gene, or a combination thereof, in which the first sequence optionally overlaps the CCAAT box target region of the human *HBG1*, *HBG2* gene, or a combination thereof. In certain embodiments, the second targeting domain sequence may be complementary to a second sequence on a side of a CCAAT box target region of the human *HBG1*, *HBG2* gene, or a combination thereof, in which the second sequence optionally overlaps the CCAAT box target region of the human *HBG1*, *HBG2* gene, or a combination thereof. In certain embodiments, the first targeting domain may comprise a truncation of a gRNA targeting domain. In certain embodiments, the gRNA targeting domain may include the gRNAs set forth in **Table 2**, **Table 10**, or **Table 15**, and the gRNA targeting domain has been truncated from a 5' end of the gRNA targeting domain. In certain embodiments, the first targeting domain may be the same as or differs by no more than 3 nucleotides from a dgRNA targeting domain set forth in **Table 10** or **Table 15**. In certain embodiments, the second targeting domain differs by no more than 3 nucleotides from a gRNA targeting domain set forth in **Table 2**, **Table 10**, or **Table 15**. In certain embodiments, the indel may alter the CCAAT box target region indel. In certain embodiments, the indel may be a productive indel resulting in an increased level of fetal hemoglobin expression. In certain embodiments, the gRNA, dgRNA, or both may be *in vitro* synthesized or chemically synthesized.

[0052] In certain embodiments, a cell may include at least one modified allele of the *HBG* locus generated by any of the methods for altering a cell disclosed herein, in which the modified allele of the *HBG* locus comprises an alteration of the human *HBG1* gene, *HBG2*, gene, or a combination thereof.

[0053] In certain embodiments, an isolated population of cells may be modified by any of the methods for altering a cells disclosed herein, wherein the population of cells may include a distribution of indels that may be different from an isolated population of cells or their progenies of the same cell type that have not been modified by the method.

[0054] In certain embodiments, a plurality of cells may be generated by any of the methods for altering a cells disclosed herein, in which at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the cells may include an alteration of a sequence in the CCAAT box target region of the human *HBG1* gene, *HBG2* gene or a combination thereof.

[0055] In certain embodiments, the cells disclosed herein may be used for a medicament. In certain embodiments, the cells may be for use in the treatment of β-hemoglobinopathy. In certain embodiments,

β-hemoglobinopathy may be selected from the group consisting of sickle cell disease and beta-thalassemia.

[0056] In one aspect, the disclosure relates to compositions including a plurality of cells generated by a method including a dgRNA 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 CCAAT box 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 CCAAT box target region of the human *HBG1* or *HBG2*. 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.

[0057] In one aspect, the disclosure relates to a population of cells modified by a genome editing system including a dgRNA described above, wherein 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. The disclosure also relates to a population of cells modified by the genome editing system including a dgRNA described above, wherein 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. In certain embodiments, the indel may be created by a repair mechanism other than microhomology-mediated end joining (MMEJ) repair.

[0058] The disclosure also relates to the use of any of the cells disclosed herein in the manufacture of a medicament for treating β-hemoglobinopathy in a subject.

[0059] In one aspect, the disclosure relates to a method of treating a β-hemoglobinopathy in a subject in need thereof, comprising administering to the subject the cells disclosed herein. In certain embodiments, a method of treating a β-hemoglobinopathy in a subject in need thereof, may include administering a population of modified hematopoietic cells to the subject, wherein one or more cells have been altered according to the methods of altering a cell disclosed herein.

[0060] In one aspect, the disclosure relates to a genome editing system, comprising: an RNA-guided nuclease; and a first guide RNA, in which the first guide RNA may comprise a first targeting domain that

is complementary to a first sequence on a side of a CCAAT box target region of a human *HBG1, HBG2* gene, or a combination thereof, in which the first sequence optionally overlaps the CCAAT box target region of the human *HBG1, HBG2* gene, or a combination thereof. In certain embodiments, the genome editing system may further comprise a template nucleic acid encoding an alteration of the CCAAT box target region of a human *HBG1, HBG2* gene, or a combination thereof. In certain embodiments, the template nucleic acid may be a single stranded oligodeoxynucleotide (ssODN) or a double stranded oligodeoxynucleotide (dsODN). In certain embodiments, the ssODN may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm. In certain embodiments, the homology arms may be symmetrical in length. In certain embodiments, the homology arms may be asymmetrical in length. In certain embodiments, the ssODN may comprise one or more phosphorothioate modifications. In certain embodiments, the one or more phosphorothioate modifications may be at the 5' end, the 3' end or a combination thereof. In certain embodiments, the ssODN may be a positive or negative strand. In certain embodiments, the alteration may be a non-naturally occurring alteration. In certain embodiments, the alteration may comprise a deletion of the CCAAT box target region. In certain embodiments, the deletion may comprise a 18 nt deletion, a 11 nt deletion, a 4 nt deletion, a 1 nt deletion, or a combination thereof. In certain embodiments, the CCAAT box target region may comprise a 18 nt target region, a 11 nt target region, a 4 nt target region, a 1 nt target region, or a combination thereof. In certain embodiments, the 5' homology arm may be about 25 to about 200 or more nucleotides in length, e.g., at about least 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length; the replacement sequence may comprise 0 nucleotides in length; and the 3' homology arm may be about 25 to about 200 or more nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 18 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 18 nt target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of SEQ ID NO:974 or SEQ ID NO:975. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 11 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 11 nt target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of SEQ ID NO:976 or SEQ ID NO:978. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 4 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 4 nt target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of a sequence selected from the group consisting of

SEQ ID NO:984, SEQ ID NO:985, SEQ ID NO:986, SEQ ID NO:987, SEQ ID NO:988, SEQ ID NO:989, SEQ ID NO:990, SEQ ID NO:991, SEQ ID NO:992, SEQ ID NO:993, SEQ ID NO:994, and SEQ ID NO:995. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 1 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 1 nt target region. In certain embodiments, the homology arms may be symmetrical in length. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of SEQ ID NO:982 or SEQ ID NO:983. In certain embodiments, the alteration may be a naturally occurring alteration. In certain embodiments, the alteration may comprise a deletion or mutation of the CCAAT box target region. In certain embodiments, the CCAAT box target region may comprise a 13 nt target region, -117G>A target region, or a combination thereof. In certain embodiments, the alteration may comprise a 13 nt deletion at the 13 nt target region or a substitution from G to A at the -117G>A target region, or a combination thereof. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 13 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 13 nt target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of SEQ ID NO:977 or SEQ ID NO:979. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 13 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 13 nt target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of SEQ ID NO:980 or SEQ ID NO:981. In certain embodiments, the RNA-guided nuclease may be an *S. pyogenes* Cas9. In certain embodiments, the RNA-guided nuclease may be a Cpf1 variant as disclosed herein. In certain embodiments, the first targeting domain may differ by no more than 3 nucleotides from a targeting domain listed in **Table 7**, **Table 18, Table 19** or a gRNA in **Table 12, Table 19**. In certain embodiments, the genome editing system may further comprise a second guide RNA, wherein the second guide RNA may comprise a second targeting domain that may be complementary to a second sequence on a side of a CCAAT box target region of a human *HBG1, HBG2* gene, or a combination thereof, wherein the second sequence optionally overlaps the CCAAT box target region of the human *HBG1, HBG2* gene, or a combination thereof. In certain embodiments, the RNA-guided nuclease may be a nickase, and optionally lacks RuvC activity. In certain embodiments, the genome editing system may comprise first and second RNA-guided nucleases. In certain embodiments, the first and second RNA-guided nucleases may be complexed with the first and second guide RNAs, respectively, forming first and second ribonucleoprotein complexes. In certain embodiments, the genome editing system may further comprise

a third guide RNA; and optionally a fourth guide RNA, wherein the third and fourth guide RNAs may comprise 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, wherein one or both of the third and fourth sequences optionally overlaps the GATA1 binding motif in *BCL11Ae* of the human *BCL11A* gene. In certain embodiments, the genome editing system may further comprise 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 RNA-guided nuclease may be a nickase, and 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 may differ by no more than 3 nucleotides from a targeting domain listed in **Table 9**. In certain embodiments, genome editing system may comprise 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.

[0061] In one aspect, the disclosure relates to a method of altering a cell comprising contacting a cell with a genome editing system. In certain embodiments, the step of contacting the cell with the genome editing system may comprise contacting the cell with a solution comprising first and second ribonucleoprotein complexes. In certain embodiments, the step of contacting the cell with the solution may further comprise electroporating the cells, thereby introducing the first and second ribonucleoprotein complexes into the cell. In certain embodiments, the method of altering a cell may further comprise contacting the cell with a genome editing system, wherein the step of contacting the cell with the genome editing system may comprise contacting the cell with a solution comprising first, second, third, and optionally, fourth ribonucleoprotein complexes. In certain embodiments, the step of contacting the cell with the solution may further comprise 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, erythrocyte, or a precursor of an erythrocyte or erythroblast. In certain embodiments, the cell may be a $CD34^+$ cell.

[0062] In one aspect, the disclosure relates to a CRISPR-mediated method of altering a cell, comprising: introducing a first DNA single strand break (SSB) or double strand break (DSB) within a genome of the

cell between positions c.-106 to -120 of a human *HBG1* or *HBG2* gene; and optionally introducing a second SSB or DSB within the genome of the cell between positions c.-106 to -120 of the human *HBG1* or *HBG2* gene, wherein the first and second SSBs or DSBs may be repaired by the cell in a manner that alters a CCAAT box 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 alteration of a CCAAT box target region of the human *HBG1* or *HBG2* gene. In certain embodiments, the CRISPR-mediated method may further comprise a template nucleic acid encoding the alteration of the CCAAT box target region of a human *HBG1, HBG2* gene, or a combination thereof. In certain embodiments, the template nucleic acid may be a single stranded oligodeoxynucleotide (ssODN). In certain embodiments, the ssODN may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm. In certain embodiments, the ssODNs may be a positive or negative strand. In certain embodiments, the alteration may be a non-naturally occurring alteration. 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 CCAAT box target region of the human *HBG1* or *HBG2* gene. In certain embodiments, the CCAAT box target region may comprise a 18 nt target region, a 11 nt target region, a 4 nt target region, a 1 nt target region, or a combination thereof. In certain embodiments, the 5' homology arm may be about 25 to about 200 nucleotides or more in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length; the replacement sequence may comprise 0 nucleotides in length; and the 3' homology arm may be about 25 to about 200 nucleotides or more in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 18 nt target region, the 11 nt target region, the 4 nt target region, or the 1 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of 18 nt target region, the 11 nt target region, the 4 nt target region, or the 1 nt target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of a sequence selected from the group consisting of SEQ ID NO:974, SEQ ID NO:975, SEQ ID NO:976, SEQ ID NO:978, SEQ ID NO:984, SEQ ID NO:985, SEQ ID NO:986, SEQ ID NO:987, SEQ ID NO:988, SEQ ID NO:989, SEQ ID NO:990, SEQ ID NO:991, SEQ ID NO:992, SEQ ID NO:993, SEQ ID NO:994, SEQ ID NO:995, SEQ ID NO:982 and SEQ ID NO:983. In certain embodiments, the alteration may be a non-naturally occurring alteration. 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 CCAAT box target region of the human *HBG1* or *HBG2* gene. In certain embodiments, the CCAAT box target region may comprise a 13 nt target region, -117G>A target region, or a combination thereof. In certain embodiments, the alteration may comprise a 13 nt deletion at the 13

nt target region or a substitution from G to A at the -117G>A target region, or a combination thereof. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 13 nt target region or the -117G>A target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 13 nt target region or the -117G>A target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of a sequence selected from the group consisting of SEQ ID NO:977 or SEQ ID NO:979. SEQ ID NO:980 or SEQ ID NO:981.

[0063] In one aspect, the disclosure relates to a composition that may comprise a plurality of cells generated by a method of altering a cell disclosed herein, wherein at least 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% of the cells may comprise an alteration of a sequence of a CCAAT box target region of the human *HBG1* gene, *HBG2* gene, or a combination thereof. In certain embodiments, the alteration may comprise a 18 nt deletion, a 11 nt deletion, a 4 nt deletion, a 1 nt deletion, a 13 nt deletion, a substitution from G to A at the -117, of the human *HBG1* gene, *HBG2* gene, or a combination thereof. 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 composition may further comprise a pharmaceutically acceptable carrier.

[0064] In one aspect, the disclosure relates to a cell comprising a synthetic genotype generated by a method of altering a cell disclosed herein, wherein the cell may comprise a 18 nt deletion, a 11 nt deletion, a 4 nt deletion, a 1 nt deletion, a 13 nt deletion, a substitution from G to A at the -117, of the human *HBG1* gene, *HBG2* gene, or a combination thereof.

[0065] In one aspect, the disclosure relates to a cell comprising at least one allele of the *HBG* locus generated by a method of altering a cell disclosed herein, wherein the cell may encode a 18 nt deletion, a 11 nt deletion, a 4 nt deletion, a 1 nt deletion, a 13 nt deletion, a substitution from G to A at the -117, of the human *HBG1* gene, *HBG2* gene, or a combination thereof.

[0066] In one aspect, the disclosure relates to an AAV vector that may comprise a template nucleic acid encoding a non-naturally occurring alteration of a CCAAT box target region of a human *HBG1, HBG2* gene, or a combination thereof. In certain embodiments, the template nucleic acid may be a single stranded oligodeoxynucleotide (ssODN). In certain embodiments, the CCAAT box target region may comprise a 18 nt target region, a 11 nt target region, a 4 nt target region, a 1 nt target region, or a combination thereof. In certain embodiments, the ssODN may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm. In certain embodiments, the 5' homology arm may be about 25 to about 200 or more nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or

200 nucleotides in length; the replacement sequence may comprise 0 nucleotides in length; and the 3' homology arm may be about 25 to about 200 or more nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 18 nt target region, the 11 nt target region, the 4 nt target region, or the 1 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of 18 nt target region, the 11 nt target region, the 4 nt target region, or the 1 nt target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of a sequence selected from the group consisting of SEQ ID NO:974-976, SEQ ID NO:978, SEQ ID NO:982-995.

[0067] In one aspect, the disclosure relates to a nucleotide sequence comprising a template nucleic acid encoding a non-naturally occurring alteration of a CCAAT box target region of a human *HBG1*, *HBG2* gene, or a combination thereof. In certain embodiments, the template nucleic acid may be a single stranded oligodeoxynucleotide (ssODN) or a double stranded oligodeoxynucleotide (dsODN) comprising the alteration. In certain embodiments, the CCAAT box target region may comprise a 18 nt target region, a 11 nt target region, a 4 nt target region, a 1 nt target region, or a combination thereof. In certain embodiments, the ssODN may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm. In certain embodiments, the 5' homology arm may be about 25 to about 200 or more nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length; the replacement sequence may comprise 0 nucleotides in length; and the 3' homology arm may be about 25 to about 200 or more nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 18 nt target region, the 11 nt target region, the 4 nt target region, or the 1 nt target region and the 3' homology arm may comprise about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of 18 nt target region, the 11 nt target region, the 4 nt target region, or the 1 nt target region. In certain embodiments, the ssODN may comprise, may consist essentially of, or may consist of a sequence selected from the group consisting of SEQ ID NO:974-976, SEQ ID NO:978, SEQ ID NO:982-995.

[0068] In one aspect, the disclosure relates to a cell comprising a synthetic genotype, wherein the cell may comprise a 18 nt deletion, a 11 nt deletion, a 4 nt deletion, a 1 nt deletion, a 13 nt deletion, a substitution from G to A at the -117, of the human *HBG1* gene, *HBG2* gene, or a combination thereof.

[0069] In one aspect, the disclosure relates to a composition, comprising a population of cells generated by a method of altering a cell disclosed herein, wherein the cells comprise a higher frequency of an alteration of a sequence of a CCAAT box target region of the human *HBG1* gene, *HBG2* gene, or a combination thereof relative to an unmodified population of cells. In certain embodiments, the higher

frequency is at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% higher. In certain embodiments, the alteration comprises a 18 nt deletion, a 11 nt deletion, a 4 nt deletion, a 1 nt deletion, a 13 nt deletion, a substitution from G to A at the -117, of the human *HBG1* gene, *HBG2* gene, or a combination thereof. In certain embodiments, at least a portion of the population of cells are within an erythroid lineage.

[0070] 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.

BRIEF DESCRIPTION OF THE DRAWINGS

[0071] 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.

[0072] **Fig. 1** depicts, in schematic form, *HBG1* and *HBG2* gene(s) in the context of the β-globin gene cluster on human chromosome 11. **Fig. 1**. Each gene in the β-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_γ and/or G_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). □-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.

[0073] **Figs. 2A-2B** 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.

[0074] **Fig. 3** shows data from gRNA screening for incorporation of the 13 nt deletion in human K562 erythroleukemia cells. **Fig. 3A** 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. 3B** 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. 3C** 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. 3B-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.

[0075] **Fig. 4** depicts results of gene editing in human cord blood (CB) and human adult $CD34^+$ 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. 4A** depicts the percentage of indels detected by T7E1 analysis of *HBG1* and *HBG2* specific PCR products amplified from gDNA extracted from CB $CD34^+$ cells treated with the indicated RNPs or donor matched untreated control cells (n=3 CB $CD34^+$ cells, 3 separate experiments). Data shown represent the mean and error bars correspond to standard deviation across the 3 separate donors/experiments. **Fig. 4B** depicts the percentage of indels detected by T7E1 analysis of *HBG2* specific PCR product amplified from gDNA extracted from CB $CD34^+$ cells or adult $CD34^+$ cells treated with the indicated RNPs or donor matched untreated control cells (n=3 CB $CD34^+$ cells, n=3 mobilized peripheral blood (mPB) $CD34^+$ cells, 3 separate experiments). Data shown represent the mean and error bars correspond to standard deviation across the 3 separate donors/experiments. **Fig. 4C** (Top panel) depicts indels as detected by T7E1 analysis of *HBG2* PCR products amplified from gDNA extracted from human CB $CD34^+$ 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 HBG Sp37 RNP and ssODN. The lower right panel shows the specific types of deletions detected within total deletions.

[0076] **Fig. 5** depicts gene editing of *HBG* in adult human mobilized peripheral blood (mPB) $CD34^+$ cells and induction of fetal hemoglobin in erythroid progeny of RNP treated cells after electroporation of mPB $CD34^+$ cells with HBG Sp37 RNP +/- ssODN encoding the 13 nt deletion. **Fig. 5A** depicts the percentage of indels detected by T7E1 analysis of *HBG2* PCR product amplified from gDNA extracted from mPB $CD34^+$ cells treated with the RNP or donor matched untreated control cells. **Fig. 5B** depicts the fold change in *HBG* mRNA expression in day 7 erythroblasts that were differentiated from RNP treated

and untreated donor matched control mPB $CD34^+$ cells. mRNA levels are normalized to GAPDH and calibrated to the levels detected in untreated controls on the corresponding days of differentiation.

[0077] **Fig. 6** depicts the *ex vivo* differentiation potential of RNP treated and untreated mPB $CD34^+$ cells from the same donor. **Fig. 6A** 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. 6B** 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.

[0078] **Fig. 7A** depicts indels detected byT7E1 analysis of *HBG* PCR product amplified from gDNA extracted from human mPB $CD34^+$ 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. 7B** depicts DNA sequencing analysis for select subset of samples shown in **Fig. 7A**. The indels were subdivided according to the type of indel (insertion, 13 nt deletion, or other deletion).

[0079] **Fig. 8A** depicts the indels at the *HBG* target site after electroporation of mPB $CD34^+$ cells with the indicated pairs of gRNAs complexed in D10A nickase and WT RNP pairs. **Fig. 8B** depicts the large deletion events (e.g. deletion of *HBG2*) after electroporation of mPB $CD34^+$ cells with the indicated pairs of gRNAs complexed in D10A nickase and WT RNPs. **Fig. 8C** depicts DNA sequencing analysis and the subtypes of events (insertions, deletions) detected in gDNA from mPB $CD34^+$ cells treated with paired D10A nickase pairs. **Fig. 8D** depicts DNA sequencing analysis and the subtypes of events (insertions, deletions) detected in gDNA from mPB $CD34^+$ cells treated with paired WT RNP pairs.

[0080] **Fig. 9** depicts the summary of HbF protein and mRNA expression in the progeny of mPB $CD34^+$ 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 $CD34^+$ cells (background levels of HbF detected in donor matched untreated controls were subtracted from the levels detected in progeny of RNP treated $CD34^+$ cells).

[0081] **Fig. 10** depicts the indel frequencies and *ex vivo* and *in vivo* short-term hematopoietic potential of $CD34^+$ cells after treatment with different concentrations (0, 2.5, 3.7 μM) of paired D10A nickase RNPs (SpA+Sp85). Indels were evaluated by T7E1 analysis (**Fig. 10A**) and by Illumina sequencing analysis (insertions and deletions, **Fig. 10B**). **Fig. 10C** depicts the % of HbF protein detected by HPLC analysis (%HbF = 100% x HbF/(HbF+HbA). **Fig. 10D** depicts the hematopoietic activity of the RNP treated and donor matched untreated control $CD34^+$ cells in colony forming cell (CFC) assays. CFCs shown are per

thousand $CD34^+$ cells plated. **Fig. 10E** depicts human blood $CD45^+$ cell reconstitution of the peripheral blood in immunodeficient mice (NSG) 1 month after transplantation with donor matched human mPB $CD34^+$ that were either untreated (0μM), or treated with one of two doses (2.5 and 3.75 μM) of D10A RNP and paired gRNAs. **Fig. 10F** depicts human blood $CD45^+$ cell reconstitution of the peripheral blood in immunodeficient mice (NSG) 2 months after transplantation. **Figs. 10G and 10H** depict the lineage distributions following human $CD45^+$ blood cell reconstitution of NSG mice at 1 month (**Fig. 10G**) and 2 months (**Fig. 10H**).

[0082] **Fig. 11A** 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 $CD34^+$ cells. HbF levels were analyzed in erythroid progeny (day 18) of edited $CD34^+$ cells. HbF protein detected in donor-matched untreated controls were subtracted from edited samples. **Fig. 11B** depicts indel rates overlaid on hematopoietic colony forming cell (CFC) activity associated with $CD34^+$ cells treated with the indicated D10A nickase pairs or untreated controls. **Fig. 11C** depicts human $CD45^+$ blood cell reconstitution of immunodeficient NSG mice one month after transplantation of mPB $CD34^+$ cells treated with indicated D10 RNP nickase pairs at the concentrations given or donor matched untreated controls. **Fig. 11D** depicts the human blood lineage distribution detected in the human $CD45^+$ fraction in mouse peripheral blood one month post-transplant.

[0083] **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).

[0084] **Fig. 13A** depicts the percentage of indels detected by T7E1 endonuclease analysis of *BCL11A* PCR products amplified from gDNA extracted from CB $CD34^+$ 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. 13B** depicts indels detected by T7E1 endonuclease analysis of *BCL11A* PCR products amplified from gDNA extracted from CB $CD34^+$ 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 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.

[0085] **Fig. 14A** depicts the editing frequency of *BCL11Ae* (using single gRNA approach targeting the GATA1 motif) in adult human BM $CD34^+$ cells. **Fig. 14B** depicts the monoallelic and bialleleic editing detected in hematopoietic colonies (GEMMs, clonal progeny of *BCL11Ae* RNP treated $CD34^+$ cells) as

determined by DNA sequencing analysis. **Fig. 14C** depicts the kinetics of erythroblast maturation (enucleation as determined by $DRAQ5^-$ cells detected by flow cytometry analysis). **Fig. 14D** depicts the acquisition of erythroid phenotype (Glycophorin A^+ cells) in differentiated control and RNP-treated BM $CD34^+$ cells, while **Fig. 14E** shows the fold increase in HbF^+ cells as determined by flow cytometry analysis relative to HbF+ cells in untreated donor matched control samples.

[0086] **Fig. 15** depicts gene editing of *BCL11Ae* in adult human mPB $CD34^+$ cells and induction of fetal hemoglobin in erythroid progeny of RNP and ssODN treated cells after electroporation of mPB $CD34^+$ cells with *BCL11Ae* RNP + nonspecific ssODN. **Fig. 15A** depicts the percentage of indels detected by T7E1 analysis of *HBG2* PCR product amplified from gDNA extracted from mPB $CD34^+$ cells treated with the *BCL11Ae* RNP and nonspecific ssODN or donor matched untreated control cells. **Fig. 15B** 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 $CD34^+$ cells (mRNA levels are normalized to GAPDH and calibrated to the levels detected in untreated controls on the corresponding days of differentiation). **Fig. 15C** depicts the percentage of Glycophorin A expressed over the time course of erythroid differentiation of mPB $CD34^+$ cells +/- treatment with *BCL11Ae* RNP and nonspecific ssODN, as determined by flow cytometry analysis at the indicated time points and for the indicated samples.

[0087] **Fig. 16** depicts the percentage of indels detected by next generation sequencing (NGS) of the *HBG* PCR product amplified from gDNA extracted from hematopoietic stem/progenitor cells (HSPCs) treated with Cas9 complexed with the chemically synthesized guide RNA OLI7066 (SEQ ID NO:970, **Table 10**) ("OLI7066-RNP") at a concentration of 16 μM. Various indels were identified including HBG Δ-102:-121, HBG Δ-114:-124, HBG Δ-116, HBG -114+T, HBG -116+G, HBG Δ-112:-115, HBG Δ-113:-115, HBG Δ-114:-115, HBG Δ-115, HBG Δ-102:-114 (the naturally occurring 13 nt deletion).

[0088] **Figs. 17A-D** depict expression levels of G gamma (Gγ)-globin, A gamma (Aγ)-globin chain (or AG gamma (AGγ)-globin resulting from the 4.9kb deletion) as determined by [gamma chain]/[all-gamma chains + beta chain] compared to relative indels carried at *HBG1* or *HBG2* as measured by UPLC analysis in the erythroid progeny of single HSPC that were electroporated with Cas9 complexed with the gRNA OLI7066 (SEQ ID NO:970) ("OLI7066-RNP") (**Table 10**). **Fig. 17A** depicts AgammaT (AγT)-globin chain expression as determined by [AγT-globin chain]/[all-gamma chains + beta chain] for clones carrying the indicated indels on the corresponding *HBG1* allele (HBG1 Δ-115, HBG1 Δ-114:-115, HBG1 Δ-113:-115, HBG1 Δ-112:-115, HBG1 Δ-102:-114, HBG1 Δ-104:-121, HBG1 Δ-116). **Fig. 17B** depicts G gamma (Gγ)-globin chain expression as determined by [Gγ-gamma chain]/[all-gamma chains + beta chain] for clones carrying the indicated indels on an HBG2 allele (HBG2 Δ-115, HBG2 Δ-114:-115, HBG2 Δ-113:-115, HBG2 Δ-112:-115, HBG2 Δ-102:-114, HBG2 Δ-104:-121, HBG2 Δ-116). To insure

that the analysis of G gamma (Gγ)-globin induction is the result of a single edited allele, only clones with monoallelic editing of HBG2, or with a deletion of one of the HBG2 allele were analyzed (resulting from the 4.9kb deletion). **Fig. 17C** depicts AG gammaT (AGγT)-globin chain expression as determined by [AGγT-gamma chain]/[all-gamma chains + beta chain] for clones carrying the indicated indels on the corresponding *HBG1/2* allele (HBG1/2 Δ-115, HBG1/2 Δ-114:-115, HBG1/2 Δ-113:-115, HBG1/2 Δ-112:-115, HBG1/2 Δ-102:-114, HBG1/2 Δ-104:-121, HBG1/2 Δ-116). **Fig. 17D** depicts AGgammaI (AGγI) -globin chain expression as determined by [AGγI-gamma chain]/[all-gamma chains + beta chain] for clones carrying the indicated indels on the corresponding *HBG1/2* allele (HBG1/2 Δ-115, HBG1/2 Δ-114:-115, HBG1/2 Δ-113:-115, HBG1/2 Δ-112:-115, HBG1/2 Δ-102:-114, HBG1/2 Δ-104:-121, HBG1/2 Δ-116).

[0089] **Fig. 18** depicts, in schematic form, *HBG1* and *HBG2* gene(s) in the context of the β-globin gene cluster on human chromosome 11. The schematic shows the CCAAT box target sites at *HBG1* and *HBG2*. Due to the homology within this region, a single guide RNA, such as OLI8394 (SEQ ID NO:971), complexed to an RNA nuclease (e.g., Cas9) will cut at both *HBG1* and *HBG2*. The editing outcomes following delivery of Cas9 complexed to OLI8394 ("OLI8394-RNP") vary and result in different size deletions or insertions. Single stranded oligodeoxynucleotides (ssODN) were designed to provide a template that copies the desired indel signature at the CCAAT box (**Table 11**). The ssODN "encodes" the respective deletion with sequence homology arms flanking the absent sequence to create a perfect deletion.

[0090] **Figs. 19A-G** depict results from gene editing of the CCAAT box target region of *HBG* of adult human CD34+ cells from mPB ("mPB CD34+ cells") electroporated with 2 μM OLI8394-RNP or OLI7066-RNP and 2.5 μM of various ssODNs (**Table 11**). **Fig. 19A** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with OLI8394-RNP and ssODN OLI16413 ("-11 nt + strand") or ssODN OLI16411 ("-11 nt - strand"). **Fig. 19B** depicts the percentage of the precise "-11 nt deletion" to the total indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with OLI8394-RNP and ssODN OLI16413 ("-11 nt + strand") or ssODN OLI16411 ("-11 nt - strand"). **Fig. 19C** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with OLI8394-RNP and ssODN OLI16430 ("-4 nt + strand") or ssODN OLI16424 ("-4 nt - strand"). The percentage of the precise -4 nt deletion (i.e., Δ-112:-115) is distinguished from other indels. **Fig. 19D** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with OLI8394-RNP and ssODN OLI16418 ("-1 nt + strand") or ssODN OLI16417 ("-1 nt - strand"). The percentage of the precise -1 nt deletion (i.e., Δ-116) is distinguished from other indels. **Fig. 19E** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with OLI8394-RNP and ssODN OLI16409 ("-18 nt +

strand") or ssODN OLI16410 ("-18 nt - strand"). The percentage of the precise -18 nt deletion (i.e., Δ-104:-121) is distinguished from other indels. **Fig. 19F** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with OLI7066-RNP and ssODN OLI16414 ("-13 nt + strand") or ssODN OLI16412 ("-13 nt - strand"). The percentage of the precise -13 nt deletion (i.e., Δ-102:-114) is distinguished from other indels **Fig. 19G** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with OLI8394-RNP and ssODN OLI16416 ("-117 G>A + strand") or ssODN OLI16415 ("-117 G>A - strand"). The percentage of reads with the -117 G>A substitution, with or without indels are distinguished from other reads.

[0091] **Figs. 20A-B** depict expression levels of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis on the erythroid progeny of mPB CD34+ cells that were electroporated with OLI8394-RNP or OLI7066-RNP and various ssODNs (**Table 11**). **Fig. 20A** depicts the percentage of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC after electroporation with (i) OLI8394-RNP and OLI7066-RNP alone, (ii) OLI8394-RNP and ssODN OLI16430 ("-4 nt + strand"), ssODN OLI16424 ("-4 nt - strand"), ssODN OLI16413 ("-11 nt + strand"), or ssODN OLI16411 ("-11 nt - strand"), and (iii) OLI7066-RNP and ssODN OLI16414 ("-13 nt + strand") or ssODN OLI16412 ("-13 nt - strand"). **Fig. 20B** depicts the percentage of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC after electroporation with OLI8394-RNP and ssODN OLI16418 ("-1 nt + strand"), ssODN OLI16417 ("-1 nt - strand"), ssODN OLI16416 ("-117 G>A + strand"), ssODN OLI16415 ("-117 G>A - strand"), ssODN OLI16409 ("-18 nt + strand"), or ssODN OLI16410 ("-18 nt - strand").

[0092] **Figs. 21A-D** depict results from gene editing from mPB CD34+ cells electroporated with 2 μM OLI8394-RNP and ssODN OLI16424 ("-4 nt - strand") (**Table 11**) at doses ranging from 0.625 μM to 10 μM. **Fig. 21A** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation. **Fig. 21B** depicts frequency of 4.9kb deletions detected by ddPCR between *HBG1* and *HBG2* after electroporation. **Fig. 21C** depicts the percentage viability of adult human CD34+ cells from mPB 48 hours after electroporation. **Fig. 21D** depicts the percentage of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis of the cell lysates from the erythroid progeny of electroporated cells.

[0093] **Figs. 22A-D** depict results from gene editing from mPB CD34+ cells electroporated with indicated doses of OLI8394-RNP and indicated doses of ssODN OLI16424 ("-4 nt - strand") (**Table 11**). **Fig. 22A** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with indicated doses (2, 4, or 8 μM) of OLI8394-RNP and indicated doses (0, 1.25, 2.5, or 5 μM) of ssODN OLI16424 ("-4 nt - strand"). **Fig. 22B** depicts the percentage of -4 nt deletions ("-

112:-115 deletions") detected by next generation sequencing (NGS) of the *HBG* PCR product after electroporation with indicated doses (2, 4, or 8 μM) of OLI8394-RNP and indicated doses (0, 1.25, 2.5, or 5 μM) of ssODN OLI16424 ("-4 nt - strand"). **Fig. 22C** depicts frequency of 4.9kb deletions between *HBG1* and *HBG2* after electroporation with the indicated doses (2, 4, or 8 μM) of OLI8394-RNP and indicated doses (0, 1.25, 2.5, or 5 μM) of ssODN OLI16424 ("-4 nt - strand"). Deletions were measured via ddPCR. **Fig. 22D** depicts the percentage of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis of the cell lysates from the erythroid progeny of mPB CD34+ cells after electroporation with the indicated doses (2, 4, or 8 μM) of OLI8394-RNP and indicated doses (0, 1.25, 2.5, or 5 μM) of ssODN OLI16424 ("-4 nt - strand").

[0094] **Figs. 23A-B** depict a schematic of and results provided by ssODN templates with symmetrical and asymmetrical homology arms. **Fig. 23A** depicts the CCAAT box target sites at *HBG1* and *HBG2* that is targeted by OLI8394 (SEQ ID NO:971) and OLI7066 (SEQ ID NO:970). ssODNs with symmetrical or asymmetrical arms were designed to provide a template that copies the -4nt deletion (HBG-112:-115) of *HBG1* and *HBG2* (**Table 11**). The ssODN "encodes" the respective deletion with sequence homology arms flanking the absent sequence to create a perfect deletion at HBG-112:-115. **Fig. 23B** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation of mPB CD34+ cells with 2 μM of OLI8394-RNP and 2.5 μM of various ssODNs, OLI16424 ("90/90"), OLI16419 ("40/80") or OLI16421 ("50/50"), that "encode" the 4 nt deletion (HBG-112:-115) (**Table 11**).

[0095] **Fig. 24** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation of mPB CD34+ cells with D10ACas9 complexed with Sp37 and SpA gRNAs (**Table 12**) ("sp37-D10A-RNP+spA-D10A-RNP") alone or with ssODN OLI16424 ("-4 nt - strand") (**Table 11**). The percentage of the precise -4 nt deletion (i.e., Δ-112:-115) is distinguished from other indels.

[0096] **Fig. 25** depicts gene editing of *HBG* of mPB CD34+ cells electroporated with variants of *Acidaminococcus sp.* Cpf1 ("AsCpf1"), namely His-AsCpf1-nNLS (SEQ ID NO:1000) and His-AsCpf1-sNLS-sNLS (SEQ ID NO:1001) complexed with the guide RNA HBG1-1 (OLI13620) (**Table 13**) ("His-AsCpf1-nNLS_HBG1-1 RNP" and "His-AsCpf1-sNLS-sNLS_HBG1-1 RNP"). The RNP were electroporated at 5 μM or 20 μM.

[0097] **Figs. 26A-C** depict gene editing of *HBG* of mPB CD34+ cells electroporated with His-AsCpf1-sNLS-sNLS_HBG1-1 RNP alone or with various ssODNs. **Fig. 26A** depicts the percentage of indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with His-AsCpf1-sNLS-sNLS_HBG1-1 RNP alone or with OLI164324 ("-4nt - strand"), OLI16430 ("-4 nt + strand"), OLI16410 ("-18 nt - strand"), or OLI16409 ("-18 nt + strand"). **Fig. 26B** depicts the percentage of the precise 18 nucleotide deletion indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with His-AsCpf1-sNLS-sNLS_HBG1-1 RNP alone or with OLI16410 ("-18 nt - strand") or OLI16409 ("-

18 nt + strand"). **Fig. 26C** depicts the percentage of the precise 18 nt deletion within all indels detected by sequencing the *HBG* PCR product 72 hours after electroporation with His-AsCpf1-sNLS-sNLS_HBG1-1 RNP alone or with OLI16410 ("-18 nt - strand") or OLI16409 ("-18 nt + strand").

[0098] **Figs. 27A-F** depict schematics of the HBG1-1 target region and *S. Pyogenes* Cas9 gRNA pairs used in combination. **Fig. 27A** shows the target region of HBG1-1 gRNA (comprising the RNA targeting domain set forth in SEQ ID NO:1002, **Table 15**). The distal CCAAT box of *HBG* promoter (i.e., HBG1/2 c.-111 to -115) is indicated by a grey box. **Fig. 27B** shows the target region of HBG1-1, the distal CCAAT box of *HBG*, and the target region SpA gRNA (comprising the targeting domain of SEQ ID NO:941, **Table 15**). **Fig. 27C** shows the target region of HBG1-1, the distal CCAAT box of *HBG*, and the target region SpG gRNA (comprising the targeting domain of SEQ ID NO:359, **Table 15**). **Fig. 27D** shows the target region of HBG1-1, the distal CCAAT box of *HBG*, the target region of tSpA dead gRNA ("dgRNA") (comprising the targeting domain of SEQ ID NO:326, **Table 15**), and the target region of Sp182 dgRNA (comprising the targeting domain of SEQ ID NO:1028, **Table 15**). **Fig. 27E** shows the target region of HBG1-1, the distal CCAAT box of *HBG*, and tSpA dgRNA (comprising the targeting domain of SEQ ID NO:326, **Table 15**). **Fig. 27F** shows the target region of HBG1-1, the distal CCAAT box of *HBG*, and the target region of Sp182 dgRNA (comprising the targeting domain of SEQ ID NO:1028, **Table 15**).

[0099] **Figs. 28A-B** depict HbF expression achieved by *ex vivo* editing of mPB CD34+ cells using HBG1-1-AsCpf1-RNP targeting the *HBG* promotor region. **Fig. 28A** shows results of editing at the *HBG* promoter region following delivery of 5 μM or 20 μM HBG1-1-AsCpf1-RNP ("HBG-1-1") via Amaxa electroporation in mPB CD34+ cells. Delivery of 20 μM HBG1-1-AsCpf1-RNP via Amaxa electroporation results in up to ~43% editing, and 21% HbF induction (above background levels). HbF levels are represented by black circles depicting expression levels of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis on the erythroid progeny of mPB CD34+ cells. Grey bars depict the percentage of indels 72 hours post-electroporation detected by next generation sequencing (NGS) of the *HBG* PCR product. **Fig. 28B** shows results of editing at the *HBG* promoter region following delivery of 5 μM or 20 μM HBG1-1-AsCpf1-RNP ("HBG-1-1") via MaxCyte electroporation in mPB CD34+ cells. Delivery of 20 μM HBG1-1-AsCpf1-RNP via MaxCyte electroporation results in up to ~16% editing, and 7% HbF induction (above background levels). HbF levels are represented by black circles depicting expression levels of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis on the erythroid progeny of mPB CD34+ cells. Grey bars depict the percentage of indels 72 hours post-electroporation detected by NGS of the *HBG* PCR product.

[0100] **Fig. 29** depicts enhanced editing by HBG1-1-AsCpf1H800A-RNP at the *HBG* promotor region on the Maxcyte device by co-delivering various *S. Pyogenes* Cas9 WT or Cas9D10A RNPs. "S.Py D10A" represents the Cas9 D10A nickase protein and "S.Py WT" represents the Cas9 WT protein. RNPs tested include SpA-D10A-RNP, SpG-D10A-RNP, tSpA-Cas9-RNP, Sp182-Cas9-RNP, and tSpA-Cas9-RNP + Sp182-Cas9-RNP (**Table 14**). Total editing at the *HBG* promoter region (grey bars) and the associated HbF protein induction (black circles) following delivery of HBG1-1-AsCpf1H800A-RNP alone ("HBG1-1"), or in combination with *S. Pyogenes* Cas9 RNPs or pairs of RNP, is depicted. HbF levels are represented by black circles depicting expression levels of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis on the erythroid progeny of mPB CD34+ cells. Grey bars depict the percentage of indels detected by NGS of the *HBG* PCR product.

[0101] **Fig. 30** depicts viability of mPB CD34+ cells following MaxCyte delivery of HBG1-1-AsCpf1H800A-RNP alone ("HBG1-1") or in combination with various *S. Pyogenes* Cas9 WT or Cas9D10A RNPs. "S.Py D10A" represents the Cas9 D10A nickase protein and "S.Py WT" represents the Cas9 WT protein. RNPs tested include SpA-D10A-RNP, SpG-D10A-RNP, tSpA-Cas9-RNP, Sp182-Cas9-RNP, and tSpA-Cas9-RNP + Sp182-Cas9-RNP (**Table 14**). Viability was measured by DAPI staining and flow cytometry analysis at 24h post electroporation.

[0102] **Figs. 31A-B** depict the cleavage sites of HBG1-1-AsCpf1H800A-RNP and D10A-Cas9 RNP at the target region and the editing profile resulting from the co-delivery of the HBG1-1-AsCpf1H800A-RNP with a D10A RNP. **Fig. 31A** depicts the position of the HBG1-1-AsCpf1H800A-RNP cut sites on each strand of the target region (light grey arrows), as well as position of the nicking site targeted by the SpG-D10A-RNP and SpA-D10A-RNP (dark arrows) (**Table 14**). **Fig. 31B** depicts the editing profile resulting from the co-delivery of the HBG1-1-AsCpf1H800A-RNP ("HBG1-1 RNP") with either SpG-D10A-RNP ("spG RNP") or SpA-D10A-RNP ("spA RNP") in mPB CD34+ at 72h post-electroporation as detected by NGS analysis of the HBG PCR product (**Table 14**). The X-axis represents genomic position of the center of the indel relative to the HBG1-1-AsCpf1H800A-RNP positive strand cleavage site. The Y axis represents the length of the indel, where deletions are represented as negative values and insertions are represented as positive values. The total frequency of each indel is represented by the area of the symbol. Indels occurring at frequency equal or above 0.1% are depicted. The SpG and SpA target sites are indicated by dotted lines.

[0103] **Figs. 32A-B** depict that the co-delivery of Sp182-Cas9-RNP with HBG1-1-AsCpf1H800A-RNP results in a boost in total indels and in distal CCAAT box disrupting indels with no substantial alteration of the indel profile, as detected by NGS analysis of the HBG PCR product at 72h post-electroporation. **Fig. 32A** shows the indel profiles following editing with HBG1-1-AsCpf1H800A-RNP ("HBG1-1 RNP")

alone, or in combination with Sp182-Cas9-RNP ("sp182 RNP"). The X-axis represents genomic position of the center of the indel relative to the HBG1-1-AsCpf1H800A-RNP positive strand cleavage site. The Y axis represents the length of the indel, where deletions are represented as negative values and insertions are represented as positive values. The total frequency of each indel is represented by the area of the symbol. Indels occurring at frequency equal or above 0.1% are depicted. The Sp182 target site is indicated by a dotted line. **Fig. 32B** depicts the frequency of indels disrupting either none, 1nt, 2nt, 3nt, 4nt, or the entire 5nt of the distal CCAAT box sequence.

[0104] **Fig. 33** depicts that optimal doses of HBG1-1-AsCpf1H800A-RNP co-delivered with Sp182-Cas9-RNP result in an increase in total editing, and HbF production (**Table 14**). Delivery of HBG1-1-AsCpf1H800A-RNP ("HBG1-1") as an RNP pair alongside Sp182-Cas9-RNP ("Sp182"), achieved >92% editing (grey bars) at the HBG promoter region with up to 34% HbF induction (above background) (black circles). No editing was observed when Sp182-Cas9-RNP was delivered alone at 12μM. HbF levels are represented by black circles depicting expression levels of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis on the erythroid progeny of mPB CD34+ cells. Grey bars depict the percentage of indels detected by NGS of the *HBG* PCR product.

[0105] **Fig. 34** depicts the distribution of levels of gamma chain expression over total beta-like chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC in the clonal erythroid progeny of single human mPB CD34+ cells edited at the *HBG* promotor region with HBG1-1-AsCpf1H800A-RNP in combination with Sp182-Cas9-RNP (**Table 14**). Each black circle represents the gamma-globin protein level detected in a clonal erythroid population derived from a single cell, isolated by FACS sorting at 48h post electroporation.

[0106] **Figs. 35A-C** depicts total editing, HbF production, viability, and colony forming potential after co-delivery of RNP containing modified HBG1-1 gRNA (SEQ ID NO:1041, **Table 14**) complexed to His-AsCpf1-sNLS-sNLS H800A (SEQ ID NO:1032, **Table 14**) ("His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP," represented as "HBG1-1" in **Figs. 35A-C**) with increasing concentrations of ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) (represented as "OLI16431" in **Figs. 35A-C**) (**Table 11**). **Fig. 35A** depicts editing at the distal CAATT box (grey bars) and HbF induction (black circles) after co-delivery of 6 μM His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP and increasing concentrations of ssODN OLI16431. HbF levels are represented by black circles depicting expression levels of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis on the erythroid progeny of mPB CD34+ cells. Grey bars depict the percentage of indels detected by NGS of the *HBG* PCR product. **Fig. 35B** depicts viability of mPB CD34+ cells following delivery of His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP alone or in combination with

increasing doses of ssODN OLI16431. Viability was measured by DAPI exclusion at 72 hours post electroporation. **Fig. 35C** depicts the hematopoietic activity of the "HBG1-1" RNP and ssODN OLI16431 treated and donor matched untreated control $CD34^+$ cells in colony forming cell (CFC) assays. CFCs shown are per 800 $CD34^+$ cells plated. The number and subtype of colonies are indicated (GEMM: granulocyte-erythroid-monocyte-macrophage colony (black), GM: granulocyte-macrophage colony (dark grey), E: erythroid colony (light grey)).

[0107] **Figs. 36A-B** depicts total editing, HbF production, and viability results using different concentrations of RNP containing unmodified HBG1-1 gRNA (SEQ ID NO:1022, **Table 14**) complexed to His-AsCpf1-sNLS-sNLS H800A (SEQ ID NO:1032, **Table 14**) ("His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP," represented as "HBG1-1" in **Figs. 36A-B**) co-delivered with different concentrations of ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) (represented as "OLI16431" in **Figs. 36A-B**) (**Table 11**). **Fig. 36A** depicts editing at the distal CAATT box (black bars (48 hours) and light grey bars (14 days of erythroid culture)) and HbF induction (black circles) after co-delivery of His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP ("HBG1-1") and ssODN OLI16431 at varying concentrations. HbF levels are represented by black circles depicting expression levels of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) as measured by UPLC analysis on the erythroid progeny of mPB CD34+ cells. Bars depict the percentage of indels detected by NGS of the *HBG* PCR product at 48 hours (black) and 14 day erythroid culture (light grey) timepoints. **Fig. 36B** depicts viability of mPB CD34+ cells following delivery of His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP alone ("HBG1-1"), ssODN OLI16431 alone, or in combination with varying doses of HBG1-1 and OLI16431. Viability was measured by DAPI exclusion at 48 hours post electroporation and after 14 days in erythroid culture. Following the editing of mPB CD34+ cells, *ex vivo* differentiation into the erythroid linage was performed for 18 days (Giarratana 2011). At day 14 of culture, a subset of cells were isolated and viability (DAPI exclusion) and editing (NGS of the HBG PCR product) measurements were taken.

[0108] **Fig. 37** depicts editing by RNP in mPB CD34+ cells. RNPs included gRNA complexed with Cpf1 protein as set forth in **Table 21**. Illumina sequencing was performed on isolated genomic DNA at 72 hours post electroporation.

[0109] **Figs. 38A-B** depict editing in bulk CD34+ cell population (black bars), progenitor cells (light grey bars), and HSCs (dark grey bars), as determined by Illumina sequencing 48 hours post electroporation. **Fig. 38A** depicts RNP33 (**Table 21**) delivered alone or co-delivered with ssODN OLI16431 (SEQ ID NO:1040, **Table 11**). **Fig. 38B** depicts RNP33 (**Table 21**) delivered alone or co-delivered with Sp182 RNP (dead gRNA comprising SEQ ID NO:1027 (**Table 14**) complexed with *S. pyogenes* Cas9 (SEQ ID NO:1033)).

[0110] **Fig. 39** depicts editing in bulk CD34+ cell population (black bars), progenitor cells (light grey bars), and HSCs (dark grey bars), as determined by Illumina sequencing 48 hours post electroporation. RNP34, RNP33, and RNP43 (**Table 21**) were delivered alone or in combination with Sp182 RNP (dead gRNA comprising SEQ ID NO:1027 (**Table 14**) complexed with *S. pyogenes* Cas9 (SEQ ID NO:1033)) or ssODN OLI16431 (SEQ ID NO:1040, **Table 11**).

[0111] **Fig. 40** depicts editing in CD34+ cells as determined by Illumina sequencing 72 hours post electroporation. RNP64, RNP63, and RNP45 (**Table 21**) were delivered at a stoichiometry (gRNA:Cpf1 complexation ratio) of either 2, or 4, where the gRNA is in a molar excess.

[0112] **Fig. 41** depicts editing in CD34+ cells as determined by Illumina sequencing 72 hours post electroporation. RNP33, RNP64, RNP63, and RNP45 (**Table 21**) were delivered alone or in combination with Sp182 RNP (dead gRNA comprising SEQ ID NO:1027 (**Table 14**) complexed with *S. pyogenes* Cas9 (SEQ ID NO:1033)) or ssODN OLI16431 (SEQ ID NO:1040, **Table 11**).

[0113] **Fig. 42** depicts editing in CD34+ cells as determined by Illumina sequencing. RNPs comprising Cpf1 (SEQ ID:1094) complexed to gRNAs with various 5' DNA extensions (**Table 21**), were delivered alone or in combination with 8 μM OLI16431 (SEQ ID NO:1040, **Table 11**).

[0114] **Fig. 43** depicts editing in CD34+ cells, as determined by Illumina sequencing. RNPs comprising gRNAs with matched 5' ends (RNP49 vs RNP58 and RNP59 vs RNP60, **Table 21**) were delivered to CD34+ cells to assess the impact of 3' modifications. In both comparisons, gRNAs with 3' PS-OMe outperformed the unmodified 3' version at 24 hours post electroporation.

[0115] **Figs. 44A-B** depict editing in CD34+ cells as determined by Illumina sequencing 24 and 48 hours post electroporation. RNP58 (**Table 21**) was delivered to CD34+ cells at a stoichiometry (gRNA:Cpf1 complexation ratio) of either 2:1, 1:1 or 0.5:1 molar ratios. At all doses tested, editing was best when RNP was complexed at 2:1 ratio.

[0116] **Figs. 45A-B** depict editing in CD34+ cells as determined by Illumina sequencing. Figs. 45A and 45B depict RNPs comprising gRNAs with matched 5' ends, but different 3' modifications (**Table 21**) delivered to CD34+ cells to assess the impact of 3' modifications or extensions.

[0117] **Figs. 46A-C** depict editing in CD34+ cells and their erythroid progeny, and HbF levels in the erythroid progeny following delivery of RNPs targeting various cut sites within the HBG locus. **Fig 46A** depicts RNPs comprising guide RNAs containing an unmodified 5' and 1 x PS-Ome at the 3' end (**Table 21**). **Fig 46B** depicts RNPs comprising guide RNAs containing 2PS +20 DNA extension at the 5' and 1 x PS-Ome at the 3' end (**Table 21**). **Fig 46C** depicts RNPs comprising guide RNAs containing a 25 DNA extension at the 5' and 1 x PS-Ome at the 3' end (**Table 21**).

[0118] **Fig. 47** depicts editing and HbF levels in erythroid progeny of CD34+ cells following delivery of RNP58 at 1 μM, 2 μM, and 4 μM.

[0119] **Fig. 48** depicts editing in CD34+ cells following Maxcyte electroporation of RNPs. RNP58, RNP26, RNP27, and RNP28 comprising gRNA SEQ ID:1051 complexed to different Cpf1 proteins (SEQ IDs: 1094, 1096, 1107, 1108) (**Table 21**) were delivered into CD34+ cells. Editing was determined by Illumina-sequencing 24 and 48 hours post electroporation.

[0120] **Fig. 49** depicts editing in CD34+ cells following Maxcyte electroporation of RNPs. RNP58, RNP29, RNP30, and RNP31 comprising Cpf1 protein SEQ ID: 1094 complexed to guide RNAs with various 5' extensions (**Table 21**) were delivered into CD34+ cells. Editing was determined by Illumina-seq 24 and 48 hours post electroporation. RNP30 was not tested (nt) at 1 μM due to limiting cell numbers.

[0121] **Fig. 50** depicts editing in bulk CD34+ cell population (black bars), progenitor cells (dark grey bars), and HSCs (light grey bars), as determined by Illumina sequencing 48 hours post electroporation. RNP58, RNP27, and RNP26 (**Table 21**) were delivered to CD34+ cells at 2 μM or 4 μM.

[0122] **Fig. 51** depicts editing in bulk CD34+ cell population (black bars), progenitor cells (dark grey bars), and HSCs (light grey bars), as determined by Illumina sequencing 48 hours post electroporation. RNP61, RNP62, and RNP34 (**Table 21**) (8 μM) were co-delivered to CD34+ cells with ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) (8 μM).

[0123] **Fig. 52** depicts editing in bulk CD34+ cell population (black bars), progenitor cells (dark grey bars), and HSCs (light grey bars), as determined by Illumina sequencing 48 hours post electroporation. RNP58 and RNP32 (**Table 21**) were delivered to CD34+ cells at 2 μM.

[0124] **Fig. 53** depicts editing in bulk CD34+ cell population (black bars), progenitor cells (dark grey bars), and HSCs (light grey bars), as determined by Illumina sequencing 48 hours post electroporation. RNP58 and RNP1 (**Table 21**) were delivered to CD34+ cells at 2 μM, 4 μM, or 8 μM. The cells edited with RNP1 at 2 μM were not sorted (N.S), and thus editing data is not available.

[0125] **Fig. 54** depicts the indels of engrafted mPB CD34+ cells from BM of "NBSGW" mice 8 weeks post infusion of electroporated cells. RNP34 and RNP33 (**Table 21**) (8 μM) were co-delivered to CD34+ cells with ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) (6 μM).

[0126] **Figs. 55A-B** depict the indels of engrafted mPB CD34+ cells and HbF expression by erythroid cells derived from chimeric BM of "NBSGW" mice 8 weeks post infusion of electroporated cells. RNP33 or RNP34 (**Table 21**) was co-delivered with Sp182 RNP (dead gRNA comprising SEQ ID NO:1027 (**Table 14**) complexed with *S. pyogenes* Cas9 (SEQ ID NO:1033)) (16 μM total RNP) to CD34+ cells. **Fig. 55A** depicts the indel frequency in unfractionated bone marrow or flow-sorted individual populations of CD15+, CD19+, GlyA+, and Lin-CD34+ cells in mock-transfected (no RNP added) or RNP transfected cells. Lin-CD34+ cells are defined as CD34+ cells that are negative for CD3, CD14, CD15, CD16, CD19, CD20, and CD56) from bone marrow (BM) of nonirradiated NOD,B6.SCID

Il2rγ-/- Kit(W41/W41) ("NBSGW") mice infused with mock (no RNP) or RNP transfected mPB CD34+ cells. Indels were determined for each cell population by Illumina sequencing. **Fig. 55B** depicts the HbF expression, calculated by UPLC as gamma/beta-like (%) from erythroid cell lysates following an 18-day erythroid differentiation culture from total chimeric BM.

[0127] **Figs. 56A-B** depict the indels of engrafted mPB CD34+ cells and HbF expression by erythroid cells derived from chimeric BM of "NBSGW" mice 8 weeks post infusion of electroporated cells. RNP61 or RNP62 (**Table 21**) (8 μM) was co-delivered with ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) (8 μM) to CD34+ cells. **Fig. 56A** depicts the indels of unfractionated bone marrow or flow-sorted individual populations of CD15+, CD19+, GlyA+, and Lin-CD34+ cells in mock-transfected (no RNP added) or RNP transfected cells. Lin-CD34+ cells are defined as CD34+ cells that are negative for CD3, CD14, CD15, CD16, CD19, CD20, and CD56) from bone marrow (BM) of nonirradiated NBSGW mice infused with mock (no RNP) or RNP transfected mPB CD34+ cells. Indels were determined for each cell population by Illumina sequencing. **Fig. 56B** depicts the HbF expression, calculated by UPLC as gamma/beta-like (%) by erythroid cells following an 18-day erythroid differentiation culture from total chimeric BM.

[0128] **Fig. 57** depicts human chimerism within bone marrow 8 weeks post infusion with mock (no RNP added) mPB CD34+ cells, or mPB CD34+ cells edited with RNP1 (4 or 8 μM) or RNP58 (2, 4 or 8 μM) (table 21). Human chimerism and lineage reconstitution (CD45+, CD14+, CD19+, glycophorin A (GlyA, CD235a+), lineage, and CD34+, and mouse CD45+ marker expression) in BM was determined by flow cytometry.

[0129] **Fig. 58** depicts indels within unsorted bulk bone marrow 8 weeks post infusion with mock (no RNP added) mPB CD34+ cells, or mPB CD34+ cells edited with RNP1 (4 or 8 μM) or RNP58 (2, 4 or 8 μM) (table 21). Indels were determined by Illumina sequencing.

[0130] **Fig. 59** depicts the indel frequency in unfractionated bone marrow or flow-sorted individual populations of CD15+, CD19+, GlyA+, and Lin-CD34+ cells in mock-transfected (no RNP added) or RNP transfected cells. Lin-CD34+ cells are defined as CD34+ cells that are negative for CD3, CD14, CD15, CD16, CD19, CD20, and CD56 from bone marrow (BM) of nonirradiated NOD,B6.SCID Il2rγ-/- Kit(W41/W41) ("NBSGW") mice infused with mock (no RNP) or RNP transfected mPB CD34+ cells. Indels were determined for each cell population by Illumina sequencing.

[0131] **Fig. 60** depicts HbF from GlyA+ fraction isolated from bone marrow at 8 weeks post infusion with mock (no RNP added) mPB CD34+ cells, or mPB CD34+ cells edited with RNP58 (2, 4 or 8 μM) (table 21).

[0132] **Fig. 61** depicts colony forming potential of cells from bone marrow flushes taken 8 weeks post infusion of mock or edited human mobilized CD34+ cells. The number and subtype of colonies are

indicated (GEMM: granulocyte-erythroid-monocyte-macrophage colony (black), GM: granulocyte-macrophage colony (dark grey), E: erythroid colony (light grey)).

[0133] **Fig. 62** depicts the sequences of Cpf1 protein variants set forth in **Table 20**. Nuclear localization sequences are shown as bolded letters, six-histidine sequences are shown as underlined letters. Additional permutations of the identity and N-terminal/C-terminal positions of NLS sequences, e.g., appending two or more nNLS sequences or combinations of nNLS and sNLS sequences (or other NLS sequences) to either the N-terminal/C-terminal positions, as well as sequences with and without purification sequences, e.g., six-histidine sequences, are within the scope of the instantly disclosed subject matter.

DETAILED DESCRIPTION

Definitions and Abbreviations

[0134] Unless otherwise specified, each of the following terms has the meaning associated with it in this section.

[0135] 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.

[0136] The conjunctions "or" and "and/or" are used interchangeably as non-exclusive disjunctions.

[0137] "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.

[0138] 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.

[0139] The term "booster element" refers to an element which, when co-delivered with a ribonucleoprotein (RNP) complex comprising a gRNA complexed to an RNA-guided nuclease ("gRNA-

nuclease-RNP"), increases editing of a target nucleic acid compared with editing of the target nucleic acid without the booster element. In certain embodiments, co-delivery may be sequential or simultaneous. In certain embodiments, a booster element may be an RNP complex comprised of a dead guide RNA complexed with a WT Cas9 protein, a Cas9 nickase protein (*e.g.,* Cas9 D10A protein), or an enzymatically inactive Cas9 (eiCas9) protein. In certain embodiments, a booster element may be an RNP complex comprised of a guide RNA complexed with a Cas9 nickase protein (*e.g.,* Cas9 D10A protein) or an enzymatically inactive Cas9 (eiCas9) protein. In certain embodiments, a booster element may be a single-or double stranded donor template DNA. In certain embodiments, one or more booster elements may be codelivered with a gRNA-nuclease-RNP to increase editing of a target nucleic acid. In certain embodiments, a booster element may be co-delivered with an RNP comprising a gRNA complexed to a Cpf1 molecule ("gRNA-Cpf1-RNP") to increase editing of a target nucleic acid.

[0140] "Productive indel" refers to an indel (deletion and/or insertion) that results in HbF expression. In certain embodiments, a productive indel may induce HbF expression. In certain embodiments, a productive indel may result in an increased level of HbF expression.

[0141] 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.

[0142] "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.

[0143] 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.

[0144] "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.

[0145] "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.

[0146] Unless indicated otherwise, the term "HDR" as used herein encompasses both canonical HDR and alt-HDR.

[0147] "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).

[0148] "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.

[0149] "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.

[0150] "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.

[0151] "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.

[0152] 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 guide RNA complexed or able to complex with an RNA-guided nuclease, and accompanied by (e.g. suspended in, or suspendable in) a pharmaceutically acceptable carrier. In certain embodiments, the kit may include a booster element. 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.

[0153] 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.

[0154] Conventional IUPAC notation is used in nucleotide sequences presented herein, as shown in Table 1, below (*see also* Cornish-Bowden A, Nucleic Acids Res. 1985 May 10; 13(9):3021-30, 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 1: IUPAC nucleic acid notation

Character	Base
A	Adenine
T	Thymine or Uracil
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

[0155] 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.

[0156] The notation "CCAAT box 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. CCAAT boxes are highly conserved motifs within the promoter region of α-like and β-like globin genes. The regions within or near the CCAAT box play important roles in globin gene regulation. For example, the γ-globin distal CCAAT box is associated with hereditary persistence of fetal hemoglobin. A number of transcription factors have been reported to bind to the duplicated CCAAT box region of the γ-globin promoter, e.g., NF-Y, COUP-TFII (NF-E3), CDP, GATA1/NF-E1 and DRED (Martyn 2017). While not wishing to be bound by theory, it is believed that the binding sites of the transcriptional activator NF-Y overlaps with transcriptional repressors at the γ-globin promoter. HPFH mutations present within the distal γ-globin promoter region, e.g., within or near the CCAAT box, may alter the competitive binding of those factors and thus contribute to the increased γ-globin expression and elevated levels of HbF. Genomic locations provided herein for HBG1

and HBG2 are based on the coordinates provided in NCBI Reference Sequence NC_000011, "Homo sapiens chromosome 11, GRCh38.p12 Primary Assembly," (Version NC_000011.10). The distal CCAAT box of *HBG1* and *HBG2* is positioned at *HBG1* and *HBG2* c.-111 to -115 (Genomic location is Hg38 Chr11:5,249,968 to Chr11:5,249,972 and Hg38 Chr11:5,254,892 to Chr11:5,254,896, respectively). The *HBG1* c.-111 to -115 region is exemplified in SEQ ID NO:902 (*HBG1*) at positions 2823-2827, and the *HBG2* c.-111 to -115 region is exemplified in SEQ ID NO:903 (*HBG2*) at positions 2747-2751. In certain embodiments, the "CCAAT box target region" denotes the region that is at or near the distal CCAAT box and includes the nucleotides of the distal CCAAT box and 25 nucleotides upstream (5') and 25 nucleotides downstream (3') of the distal CCAAT box (i.e., *HBG1/2* c.-86 to -140) (Genomic location is Hg38 Chr11:5249943 to Hg38 Chr11:5249997 and Hg38 Chr11:5254867 to Hg38 Chr11:5254921, respectively). The *HBG1* c.-86 to -140 region is exemplified in SEQ ID NO:902 (*HBG1*) at positions 2798-2852, and the *HBG2* c.-86 to -140 region is exemplified in SEQ ID NO:903 (*HBG2*) at positions 2723-2776. In other embodiments, the "CCAAT box target region" denotes the region that is at or near the distal CCAAT box and includes the nucleotides of the distal CCAAT box and 5 nucleotides upstream (5') and 5 nucleotides downstream (3') of the distal CCAAT box (i.e., *HBG1/2* c.-106 to -120 (Genomic location is Hg38 Chr11:5249963 to Hg38 Chr11:5249977 (HGB1 and Hg38 Chr11:5254887 to Hg38 Chr11:5254901, respectively)). The *HBG1* c.-106 to -120 region is exemplified in SEQ ID NO:902 (*HBG1*) at positions 2818-2832, and the *HBG2* c.-106 to -120 region is exemplified in SEQ ID NO:903 (*HBG2*) at positions 2742-2756. The term "CCAAT box target site alteration" and the like refer to alterations (e.g., deletions, insertions, mutations) of one or more nucleotides of the CCAAT box target region. Examples of exemplary CCAAT box target region alterations include, without limitation, the 1 nt deletion, 4 nt deletion, 11nt deletion, 13 nt deletion, and 18 nt deletion, and -117 G>A alteration. As used herein, the terms "CCAAT box" and "CAAT box" can be used interchangeably.

[0157] The notations "c.-114 to -102 region," "c.-102 to -114 region," "-102:-114," "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 at the genomic location Hg38 Chr11:5,249,959 to Hg38 Chr11:5,249,971 and Hg38 Chr11:5,254,883 to Hg38 Chr11:5,254,895, respectively. The *HBG1* c.-102 to -114 region is exemplified in SEQ ID NO:902 (*HBG1*) at positions 2824-2836 and the *HBG2* c.-102 to -114 region is exemplified in SEQ ID NO:903 (*HBG2*) at positions 2748-2760. The term "13 nt deletion" and the like refer to deletions of the 13 nt target region.

[0158] The notations "c.-121 to -104 region," "c.-104 to -121 region," "-104:-121," "18 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 at the genomic location Hg38 Chr11:5,249,961 to Hg38 Chr11:5,249,978 and Hg38

Chr11:5,254,885 to Hg38 Chr11: 5,254,902, respectively. The *HBG1* c.-104 to -121 region is exemplified in SEQ ID NO:902 (*HBG1*) at positions 2817-2834, and the *HBG2* c.-104 to -121 region is exemplified in SEQ ID NO:903 (*HBG2*) at positions 2741-2758. The term "18 nt deletion" and the like refer to deletions of the 18 nt target region.

[0159] The notations "c.-105 to -115 region," "c.-115 to -105 region," "-105:-115," "11 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 at the genomic location Hg38 Chr11:5,249,962 to Hg38 Chr11:5,249,972 and Hg38 Chr11:5,254,886 to Hg38 Chr11:5,254,896, respectively. The *HBG1* c.-105 to -115 region is exemplified in SEQ ID NO:902 (*HBG1*) at positions 2823-2833, and the *HBG2* c.-105 to -115 region is exemplified in SEQ ID NO:903 (*HBG2*) at positions 2747-2757. The term "11 nt deletion" and the like refer to deletions of the 11 nt target region.

[0160] The notations "c.-115 to -112 region," "c.-112 to -115 region," "-112:-115," "4 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 at the genomic location Hg38 Chr11:5,249,969 to Hg38 Chr11:5,249,972 and Hg38 Chr11:5,254,893 to Hg38 Chr11:5,254,896, respectively. The *HBG1* c.-112 to -115 region is exemplified in SEQ ID NO:902 at positions 2823-2826, and the *HBG2* c.-112 to -115 region is exemplified in SEQ ID NO:903 (*HBG2*) at positions 2747-2750. The term "4 nt deletion" and the like refer to deletions of the 4 nt target region.

[0161] The notations "c.-116 region," "HBG-116," "1 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 at the genomic location Hg38 Chr11:5,249,973 and Hg38 Chr11:5,254,897, respectively. The *HBG1* c.-116 region is exemplified in SEQ ID NO:902 at position 2822, and the *HBG2* c.-116 region is exemplified in SEQ ID NO:903 (*HBG2*) at position 2746. The term "1 nt deletion" and the like refer to deletions of the 1 nt target region.

[0162] The notations "c.-117 G>A region," "HBG-117 G>A," "-117 G>A 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 at the genomic location Hg38 Chr11:5,249,974 to Hg38 Chr11:5,249,974 and Hg38 Chr11:5,254,898 to Hg38 Chr11:5,254,898, respectively. The *HBG1* c.-117 G>A region is exemplified by a substitution from guanine (G) to adenine (A) in SEQ ID NO:902 at position 2821, and the *HBG2* c.-117 G>A region is exemplified by a substitution from G to A in SEQ ID NO:903 (*HBG2*) at position 2745. The term "-117 G>A alteration" and the like refer to a substitution from G to A at the -117G>A target region.

[0163] 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.

[0164] 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.

[0165] 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

[0166] 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 Aγ and Gγ subunits of hemoglobin, respectively. In certain embodiments, increased expression of one or more γ–globin genes (e.g., HBG1, HBG2) using the methods provided herein results in preferential formation of HbF over HbA and/or increased HbF levels as a percentage of total hemoglobin. In certain embodiments, the disclosure generally relates to the use of RNP complexes comprising a gRNA complexed to a Cpf1 molecule. In certain embodiments, the gRNA may be unmodified or modified, the Cpf1 molecule may be a wild-type Cpf1 protein or a modified Cpf1 protein. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, a modified Cpf1 may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID

NO:1051 and a modified Cpf1 protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**).

[0167] It has previously been shown that patients with the condition Hereditary Persistence of Fetal Hemoglobin (HPFH) contain mutations in an γ–globin regulatory element that results in fetal γ–globin expression throughout life, rather than being repressed around the time of birth (Martyn 2017). This results in elevated fetal hemoglobin (HbF) expression. HPFH mutations may be deletional or non-deletional (e.g., point mutations). Subjects with HPFH exhibit lifelong expression of HbF, i.e., they do not undergo or undergo only partial globin switching, with no symptoms of anemia.

[0168] HbF expression can be induced through point mutations in an γ–globin regulatory element that is associated with a naturally occurring HPFH variant, including, for example, HBG1 c.-114 C>T; c.-117 G>A; c.-158 C>T; c.-167 C>T; c.-170 G>A; c.-175 T>G; c.-175 T>C; c.-195 C>G; c.-196 C>T; c.-197 C>T; c.-198 T>C; c.-201 C>T; c.-202 C>T; c.-211 C>T, c.-251 T>C; or c.-499 T>A; or HBG2 c.-109 G>T; c.-110 A>C; c.-114 C>A; c.-114 C>T; c.-114 C>G; c.-157 C>T; c.-158 C>T; c.-167 C>T; c.-167 C>A; c.-175 T>C; c.-197 C>T; c.-200+C; c.-202 C>G; c.-211 C>T; c.-228 T>C; c.-255 C>G; c.-309 A>G; c.-369 C>G; or c.-567 T>G.

[0169] Naturally occurring mutations at the distal CCAAT box motif found within the promoter of the *HBG1* and/or *HBG2* genes (i.e., HBG1/2 c.-111 to -115) have also been shown to result in continued γ–globin expression and the HPFH condition. It is thought that alteration (mutation or deletion) of the CCAAT box may disrupt the binding of one or more transcriptional repressors, resulting in continued expression of the γ–globin gene and elevated HbF expression (Martyn 2017). For example, a naturally occurring 13 base pair del c.-114 to -102 ("13 nt deletion") has been shown to be associated with elevated levels of HbF (Martyn 2017). The distal CCAAT box likely overlaps with the binding motifs within and surrounding the CCAAT box of negative regulatory transcription factors that are expressed in adulthood and repress *HBG* (Martyn 2017).

[0170] A gene editing strategy disclosed herein is to increase HbF expression by disrupting one or more nucleotides in the distal CCAAT box and/or surrounding the distal CCAAT box. In certain embodiments, the "CCAAT box target region" may be the region that is at or near the distal CCAAT box and includes the nucleotides of the distal CCAAT box and 25 nucleotides upstream (5') and 25 nucleotides downstream (3') of the distal CCAAT box (i.e., *HBG1/2* c.-86 to -140). In other embodiments, the "CCAAT box target region" may be the region that is at or near the distal CCAAT box and includes the nucleotides of the distal CCAAT box and 5 nucleotides upstream (5') and 5 nucleotides downstream (3') of the distal CCAAT box (i.e., HBG1/2 c.-106 to -120). Unique, non-naturally occurring alterations of the CCAAT box target region are disclosed herein that induce *HBG* expression including, without limitation, HBG del c. -104 to -121 ("18 nt deletion"), HBG del c.-105 to -115 ("11 nt deletion"), HBG

del c.-112 to -115 ("4 nt deletion"), and HBG del c.-116 ("1 nt deletion"). In certain embodiments, genome editing systems disclosed herein may be used to introduce alterations into the CCAAT box target region of *HBG1* and/or *HBG2*. In certain embodiments, the genome editing systems may include one or more of a DNA donor template that encodes an alteration (such as a deletion, insertion, or mutation) in the CCAAT box target region. In certain embodiments, the alterations may be non-naturally occurring alterations or naturally occurring alterations. In certain embodiments, the donor templates may encode the 1 nt deletion, 4 nt deletion, 11 nt deletion, 13 nt deletion, 18 nt deletion, or c.-117 G>A alteration. In certain embodiments, the genome editing systems may include an RNA guided nuclease including a Cas9, modified Cas 9, a Cpf1, or modified Cpf1. In certain embodiments, the genome editing systems may include an RNP comprising a gRNA and a Cpf1 molecule. In certain embodiments, a gRNA may be unmodified or modified, the Cpf1 molecule may be a wild-type Cpf1 protein or a modified Cpf1 protein, or a combination thereof. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. In certain embodiments, a modified Cpf1 may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a modified Cpf1 protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**).

[0171] HbF expression can also be induced through targeted disruption of the erythroid cell specific expression of a transcriptional repressor, *BCL11A*, which encodes a repressor that silences *HBG1* and *HBG2* (Canvers 2015). Another gene editing strategy disclosed herein is to increase HbF expression by targeting disruption the of the erythroid specific enhancer of *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). 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*, the CCAAT box target region, the 13 nt target region of *HBG1* and/or *HBG2*, or a combination thereof.

[0172] 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.

[0173] A variety of approaches to the introduction of mutations into the CCAAT box target region, 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 the CCAAT box target region, 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 CCAAT box target region, 13 nt target region and/or the GATA1 binding motif in *BCL11Ae*.

[0174] 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 Aγ or Gγ 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.

[0175] In certain embodiments, alterations that result in induction of Aγ and/or Gγ 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 CCAAT box 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 CCAAT box 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 CCAAT box target region or a region proximate thereto. Examples of suitable gRNAs and gRNA targeting domains directed to the CCAAT box 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, 970, 971, 996, 997, 1002, and 1004.

[0176] In certain embodiments, alterations that result in induction of Aγ and/or Gγ 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 gRNAs and gRNA 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, 970, 971, 996, 997, 1002, and 1004.

[0177] 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.

[0178] 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.

[0179] 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.

[0180] 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).

[0181] 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 the CCAAT box target region, 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.

[0182] 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.

[0183] 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.*, $\geq 2 \times 10^6$ cells/kg, $\geq 3 \times 10^6$ cells/kg, $\geq 4 \times 10^6$ cells/kg, $\geq 5 \times 10^6$ cells/kg, 2×10^6 cells/kg to 3×10^6 cells/kg, 3×10^6 cells/kg to 4×10^6 cells/kg, 4×10^6 cells/kg to 5×10^6 cells/kg) to be reinfused back into the subject, along with a dose of unedited autologous CD34+ HSPCs/kg for backup storage (*e.g.*, $\geq 1.5 \times 10^6$ 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.

[0184] 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.*, $\geq 2 \times 10^6$ cells/kg, $\geq 3 \times 10^6$ cells/kg, $\geq 4 \times 10^6$ cells/kg, $\geq 5 \times 10^6$ cells/kg, 2×10^6 cells/kg to 3×10^6 cells/kg, 3×10^6 cells/kg to 4×10^6 cells/kg, 4×10^6 cells/kg to 5×10^6 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+ cells. Neutrophil engraftment may be defined as three consecutive measurements of ANC $\geq 0.5 \times 10^9$/L.

[0185] 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.

[0186] 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.

[0187] Functionally, alteration of the CCAAT box target region, 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 Aγ and/or Gγ 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 Aγ and/or Gγ subunit expression relative to unmodified controls. This induction of protein expression is generally the result of alteration of the CCAAT box target region, 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 the CCAAT box target region, 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in *BCL11Ae*.

[0188] 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 in 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 Aγ and Gγ globin chains as is known in the art.

[0189] It should be noted that the rate at which the CCAAT box target region (e.g., 18 nt, 11 nt, 4 nt, 1 nt, c.-117 G>A target regions), 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). In certain embodiments, donor templates for use in targeting the 18 nt target region may include, without limitation, donor templates encoding alterations (e.g., deletions) of HBG1 c.-104 to -121, HBG2 c.-104 to -121, or a combination thereof. Exemplary full-length donor templates encoding deletions such as c.-104 to -121 include SEQ ID NOs:974 and 975. In certain embodiments, donor templates for use in targeting the 11 nt target region may include, without limitation, donor templates encoding alterations (e.g., deletions) of HBG1 c.-105 to -115, HBG2 c.-105 to -115, or a combination thereof. Exemplary full-length donor templates encoding deletions such as c.-105 to -115 include SEQ ID NOs:976 and 978. In certain embodiments, donor templates for use in targeting the 4 nt target region may include, without limitation, donor templates encoding alterations (e.g., deletions) of HBG1 c.-112 to -115, HBG2 c.-112 to -115, or a combination thereof. Exemplary full-length donor templates encoding deletions such as c.-112 to -115 include SEQ ID NOs:984-995. In certain embodiments, donor templates for use in targeting the 1 nt target region may include, without limitation, donor templates encoding alterations (e.g., deletions) of HBG1 c.-116, HBG2 c.-116, or a combination thereof. Exemplary full-length donor templates encoding deletions such as c.-116 include SEQ ID NOs:982 and 983. In certain embodiments, donor templates for use in targeting the c.-117 G>A target region may include, without limitation, donor templates encoding alterations (e.g., deletions) of HBG1 c.-117 G>A, HBG2 c.-117 G>A, or a combination thereof. Exemplary full-length donor templates encoding deletions such as c.-117 G>A include SEQ ID NOs:980 and 981. In certain embodiments, the donor template may be a positive strand or a negative strand.

[0190] 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 alterations (e.g., deletions) of *BCL11Ae*, including, without limitation, the GATA1 motif in *BCL11Ae*.

[0191] The embodiments described herein may be used in all classes of vertebrate including, but not limited to, primates, mice, rats, rabbits, pigs, dogs, and cats.

[0192] 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.

RNA-guided helicases, guide RNAs and dead guide RNAs

[0193] 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 CCAAT box target region, 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.

[0194] 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).

[0195] 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.

[0196] A 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.

[0197] 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 CCAAT box target region, 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 whereby the target region is altered. In certain embodiments, the DSB may be 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.

[0198] 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 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 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.

[0199] 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 CCAAT box target region, 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. In certain embodiments, the RNA-guided nuclease is not configured to recruit an exogenous trans-acting factor to the desired target region to be edited (*e.g.*, the CCAAT box target region, 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in *BCL11Ae*).

[0200] Turning to dead gRNAs, these include any of the dead gRNAs discussed herein and *infra* under the heading entitled "Dead gRNA molecules." Dead gRNAs (also referred to herein as "dgRNAs") 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. In certain embodiments, a dgRNA may be generated by truncating the 5' end of any one of a gRNA targeting domain sequence disclosed herein in **Table 2** or **Table 10**. Dead guide RNA molecules according to the present disclosure include dead guide RNA molecules that have reduced, low, or undetectable cleavage activity. The targeting domain sequences of dead guide RNAs may be shorter in length by 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides compared to the targeting domain sequence of active guide RNAs. Dead gRNA molecules may comprise targeting domains complementary to regions proximal to or within a target region (*e.g.*, the CCAAT box target region, 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 CCAAT box target region, 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 CCAAT box target region, 13 nt target region, proximal HBG1/2 promoter target sequence, and/or the GATA1 binding motif in *BCL11Ae*).

[0201] 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 CCAAT box target region, 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.

[0202] 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 CCAAT box target region, 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.

[0203] 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.

[0204] 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.

[0205] Also provided herein are genome editing systems and methods of altering one or more indels (e.g., indel signature) generated by an active guide. As the inventors have discovered herein, pairing of a dead RNP (dRNP) (i.e., a dead guide RNA complexed with an RNA-guided nuclease) and an active RNP (i.e., an active guide RNA complexed with an RNA-guided nuclease) can result in a change of the directionality of the indels (e.g., indel signature) generated by the active RNP alone (without a dRNP). As shown in the examples below, the use of the dead guide RNA may result in an increased frequency of larger deletions extending from the active guide RNA cut site toward the dead guide RNA binding site. Thus, the dead guide RNA may be used to effectively "orient" deletion editing toward a desired target site. In certain embodiments, the use of the dead guide RNA with an active guide RNA may increase the frequency of deletions that are not associated with micro-homologies.

[0206] Although the examples disclosed in the Examples section below are directed to alterations of the CCAAT box target region, skilled artisans would contemplate that the genome editing systems, methods, cells and compositions described herein may be used to alter any other target region, for example, without limitation, to increase the frequency of deletions at the target region, increase the frequency of deletions at the target region that are not associated with micro-homologies (e.g., not repaired via MMEJ).

[0207] 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

[0208] 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.

[0209] 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 (dgRNA) 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.

[0210] 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.

[0211] 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.

[0212] 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.

[0213] 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, 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.

[0214] 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*.

[0215] 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 & Maizels 2014 (describing Alt-HDR); Frit 2014 (describing Alt-NHEJ); Iyama & Wilson 2013 (describing canonical HDR and NHEJ pathways generally)).

[0216] 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.

[0217] 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. In certain embodiments, a genome editing system may include an RNA-guided helicase

that unwinds DNA within or proximal to the target sequence, without causing single- or double-stranded breaks. For example a genome editing system may include an RNA-guided helicase configured to associate within or near the target sequence to unwind DNA and induce accessibility to the target sequence. In certain embodiments, the RNA-guided helicase may be complexed to a dead guide RNA that is configured to lack cleavage activity allowing for unwinding of the DNA without causing breaks in the DNA.

Guide RNA (gRNA) molecules

[0218] 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 (see, e.g., Briner 2014, which is incorporated by reference; 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.

[0219] 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).

[0220] 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 2013, 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.

[0221] 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 2014; 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.

[0222] 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.

[0223] 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 2015, 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). Exemplary targeting domains of Cpf1 gRNAs are set forth in **Table 13** and **Table 18**.

[0224] 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.

[0225] 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

[0226] 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). Each of these references is 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.

[0227] 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.

[0228] 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).

[0229] 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 2: gRNA targeting domain sequences for HBG1/2 target sites

	HBG1	**HBG2**
S. pyogenes Tier 1	251-256	760-764
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

	HBG1	**HBG2**
S. aureus Tier 1	367-376	784-791
Tier 2	343, 377-393	778, 792-803
Tier 3	357, 365, 394-461	357, 365, 394-461

Tier 4	252-254, 256, 268, 272-274, 292, 295, 347, 348, 353, 360-362, 366, 598-759	292, 295, 347, 348, 353, 360-362, 366, 462-468 476-481, 489-587, 601-607, 614-620, 640-666, 674-679, 687-693, 708-714, 733-753, 762-764, 775, 779-781, 804-901

[0230] Additional gRNA sequences that were designed to target alteration of the CCAAT box target region include, but are not limited to, the sequences set forth in SEQ ID NOs:970 and 971. Targeting domain sequences of gRNAs that were designed to target disruption of the CCAAT box target region include, but are not limited to, SEQ ID NO:1002. Targeting domain sequences plus PAM (UUUG) of gRNAs that were designed to target disruption of the CCAAT box target region include, but are not limited to, SEQ ID NO:1004. In certain embodiments, gRNAs comprising the sequence set forth in SEQ ID NOs:1002 and/or 1004 may be complexed with a Cpf1 protein or modified Cpf1 protein to generate alterations at the CCAAT box target region. In certain embodiments, gRNAs comprising any of the Cpf1 gRNAs set forth in **Table 15**, **Table 18,** or **Table 19** may be complexed with a Cpf1 protein or modified Cpf1 protein forming an RNP ("gRNA-Cpf1-RNP") to generate alterations at the CCAAT box target region. In certain embodiments, the modified Cpf1 protein may be His-AsCpf1-nNLS (SEQ ID NO: 1000) or His-AsCpf1-sNLS-sNLS (SEQ ID NO:1001). In certain embodiments, the Cpf1 molecule of the gRNA-Cpf1-RNP may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021 (Cpf1 polynucleotide sequences).

[0231] gRNAs may be designed to target the erythroid specific enhancer of *BCL11A* (*BCL11Ae*) to disrupt expression of a transcriptional repressor, BCL11A (described in Friedland, which is incorporated by reference herein). 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. Targeting domain sequences plus PAM (NGG) 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:960-963.

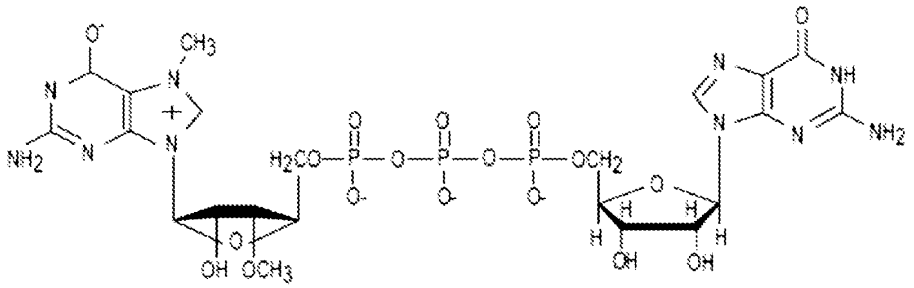
gRNA modifications

[0232] 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.

[0233] 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.

[0234] 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:



The cap or cap analog can be included during either chemical synthesis or *in vitro* transcription of the gRNA.

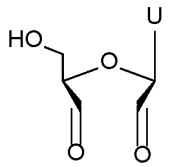
[0235] Along similar lines, the 5' end of the gRNA can lack a 5' triphosphate group. For instance, *in vitro* transcribed gRNAs can be phosphatase-treated (e.g., using calf intestinal alkaline phosphatase) to remove a 5' triphosphate group.

[0236] 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.

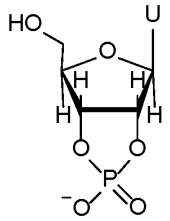
[0237] It should be noted that the modifications described herein can be combined in any suitable manner, e.g. a gRNA, whether transcribed *in vivo* from a DNA vector, or *in vitro* transcribed gRNA, can include either or both of a 5' cap structure or cap analog and a 3' polyA tract.

[0238] 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:



wherein "U" can be an unmodified or modified uridine.

[0239] The 3' terminal U ribose can be modified with a 2'3' cyclic phosphate as shown below:



wherein "U" can be an unmodified or modified uridine.

[0240] 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.

[0241] 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_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.

[0242] 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_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroarylamino, ethylenediamine, or polyamino) and aminoalkoxy or $O(CH_2)_n$-amino (wherein amino can be, e.g., NH_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroarylamino, ethylenediamine, or polyamino).

[0243] 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'→2')).

[0244] 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.

[0245] 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.

[0246] In certain embodiments, gRNAs as used herein may be modified or unmodified gRNAs. In certain embodiments, a gRNA may include one or more modifications. In certain embodiments, the one or more modifications may include a phosphorothioate linkage modification, a phosphorodithioate (PS2) linkage modification, a 2'-O-methyl modification, or combinations thereof. In certain embodiments, the one or more modifications may be at the 5' end of the gRNA, at the 3' end of the gRNA, or combinations thereof.

[0247] In certain embodiments, a gRNA modification may comprise one or more phosphorodithioate (PS2) linkage modifications.

[0248] In some embodiments, a gRNA used herein includes one or more or a stretch of deoxyribonucleic acid (DNA) bases, also referred to herein as a "DNA extension." In some embodiments, a gRNA used herein includes a DNA extension at the 5' end of the gRNA, the 3' end of the gRNA, or a combination thereof. In certain embodiments, the DNA extension may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 DNA bases long. For example, in certain embodiments, the DNA extension may be 1, 2, 3, 4, 5, 10, 15, 20, or 25 DNA bases long. In certain embodiments, the DNA extension may include one or more DNA bases selected from adenine (A), guanine (G), cytosine (C), or thymine (T). In certain embodiments, the DNA extension includes the same DNA bases. For example, the DNA extension may include a stretch of adenine (A) bases. In certain embodiments, the DNA extension may include a stretch of thymine (T) bases. In certain embodiments, the DNA extension includes a combination of different DNA bases. In certain embodiments, a DNA extension may comprise a sequence set forth in **Table 24**. For example, a DNA extension may comprise a sequence set forth in SEQ ID NOs:1235-1250. In certain embodiments, a gRNA used herein includes a DNA extension as well as one or more phosphorothioate linkage modifications, one or more phosphorodithioate (PS2) linkage modifications, one or more 2'-O-methyl modifications, or combinations thereof. In certain embodiments, the one or more modifications may be at the 5' end of the gRNA, at the 3' end of the gRNA, or combinations thereof. In certain embodiments, a gRNA including a DNA extension may comprise a sequence set forth in **Table 19** that includes a DNA extension. In a particular embodiment, a gRNA including a DNA extension may

comprise the sequence set forth in SEQ ID NO:1051. In certain embodiments, a gRNA including a DNA extension may comprise a sequence selected from the group consisting of SEQ ID NOs:1046-1060, 1067, 1068, 1074, 1075, 1078, 1081-1084, 1086-1087, 1089-1090, 1092-1093, 1098-1102, and 1106. Without wishing to be bound by theory, it is contemplated that any DNA extension may be used herein, so long as it does not hybridize to the target nucleic acid being targeted by the gRNA and it also exhibits an increase in editing at the target nucleic acid site relative to a gRNA which does not include such a DNA extension.

[0249] In some embodiments, a gRNA used herein includes one or more or a stretch of ribonucleic acid (RNA) bases, also referred to herein as an "RNA extension." In some embodiments, a gRNA used herein includes an RNA extension at the 5' end of the gRNA, the 3' end of the gRNA, or a combination thereof. In certain embodiments, the RNA extension may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 RNA bases long. For example, in certain embodiments, the RNA extension may be 1, 2, 3, 4, 5, 10, 15, 20, or 25 RNA bases long. In certain embodiments, the RNA extension may include one or more RNA bases selected from adenine (rA), guanine (rG), cytosine (rC), or uracil (rU), in which the "r" represents RNA, 2'-hydroxy. In certain embodiments, the RNA extension includes the same RNA bases. For example, the RNA extension may include a stretch of adenine (rA) bases. In certain embodiments, the RNA extension includes a combination of different RNA bases. In certain embodiments, an RNA extension may comprise a sequence set forth in **Table 24**. For example, an RNA extension may comprise a sequence set forth in 1231-1234, 1251-1253. In certain embodiments, a gRNA used herein includes an RNA extension as well as one or more phosphorothioate linkage modifications, one or more phosphorodithioate (PS2) linkage modifications, one or more 2'-O-methyl modifications, or combinations thereof. In certain embodiments, the one or more modifications may be at the 5' end of the gRNA, at the 3' end of the gRNA, or combinations thereof. In certain embodiments, a gRNA including a RNA extension may comprise a sequence set forth in **Table 19** that includes an RNA extension. gRNAs including an RNA extension at the 5' end of the gRNA may comprise a sequence selected from the group consisting of SEQ ID NOs:1042-1045, 1103-1105. gRNAs including an RNA extension at the 3' end of the gRNA may comprise a sequence selected from the group consisting of SEQ ID NOs:1070-1075, 1079, 1081, 1098-1100.

[0250] It is contemplated that gRNAs used herein may also include an RNA extension and a DNA extension. In certain embodiments, the RNA extension and DNA extension may both be at the 5' end of the gRNA, the 3' end of the gRNA, or a combination thereof. In certain embodiments, the RNA

extension is at the 5' end of the gRNA and the DNA extension is at the 3' end of the gRNA. In certain embodiments, the RNA extension is at the 3' end of the gRNA and the DNA extension is at the 5' end of the gRNA.

[0251] In some embodiments, a gRNA which includes both a phosphorothioate modification at the 3' end as well as a DNA extension at the 5' end is complexed with a RNA-guided nuclease, e.g., Cpf1, to form an RNP, which is then employed to edit a hematopoietic stem cell (HSC) or a CD34+ cell *ex vivo* (i.e., outside the body of a subject from whom such a cell is derived), at the HBG locus.

[0252] An example of a gRNA as used herein comprises the sequence set forth in SEQ ID NO:1051.

Dead gRNA molecules

[0253] Dead guide RNA (dgRNA) molecules according to the present disclosure include dead guide RNA molecules that comprise reduced, low, or undetectable cleavage activity. The targeting domain sequences of dead guide RNAs are shorter in length by 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides compared to the targeting domain sequence of active guide RNAs. In certain embodiments, dead guide RNA molecules may comprise a targeting domain comprising 15 nucleotides or fewer in length, 14 nucleotides or fewer in length, 13 nucleotides or fewer in length, 12 nucleotides or fewer in length, or 11 nucleotides or fewer in length. In some embodiments, dead guide RNAs are configured such that they do not provide an RNA guided-nuclease cleavage event. Dead guide RNAs may be generated by removing the 5' end of a gRNA targeting domain sequence, which results in a truncated targeting domain sequence. For example, if a gRNA sequence, configured to provide a cleavage event (i.e., 17 nucleotides or more in length), 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. For example, dgRNAs used herein may comprise a targeting domain set forth in Table 2 or Table 10 that has been truncated from the 5' end of the gRNA sequence and comprises 15 nucleotides or fewer in length. 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 CCAAT box target region, 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 or CCAAT box 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 CCAAT box target region, 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.

[0254] In certain embodiments, the use of a dead guide RNA with the genome editing systems and methods disclosed herein may increase the total editing level of an active guide RNA. In certain embodiments, the use of a dead guide RNA with the genome editing systems disclosed herein and methods thereof may increase the frequency of deletions. In certain embodiments, the deletions may extend from the cut site of the active guide RNA toward the dead guide RNA binding site. In this way the dead guide RNA can change the directionality of an active guide RNA and orient editing toward a desired target region.

[0255] As used herein, the terms "dead gRNA" and "truncated gRNA" are used interchangeably.

RNA-guided nucleases

[0256] 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. 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.

[0257] 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.

[0258] 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. Cpf1, on the other hand, generally recognizes PAM sequences that are 5' of the protospacer.

[0259] 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, wherein the N residues are immediately 3' of the region recognized by the gRNA targeting domain. *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.

[0260] 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 and in Ran & Hsu 2013, incorporated by reference herein), or that do not cut at all.

Cas9

[0261] 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).

[0262] 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.

[0263] 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).

[0264] 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, and 14.

Cpf1

[0265] 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 by 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.

[0266] 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.

[0267] In certain embodiments, a Cpf1 protein may be a modified Cpf1 protein. In certain embodiments, a modified Cpf1 protein may include one or more modifications. In certain embodiments the modifications may be, without limitation, one or more mutations in a Cpf1 nucleotide sequence or Cpf1 amino acid sequence, one or more additional sequences such as a His tag or a nuclear localization signal (NLS), or a combination thereof. In certain embodiments, a modified Cpf1 may also be referred to herein as a Cpf1 variant.

[0268] In certain embodiments, the Cpf1 protein may be derived from a Cpf1 protein selected from the group consisting of *Acidaminococcus sp.* strain BV3L6 Cpf1 protein (AsCpf1), *Lachnospiraceae bacterium* ND2006 Cpf1 protein (LbCpf1), and *Lachnospiraceae bacterium* MA2020 (Lb2Cpf1). In certain embodiments, the Cpf1 protein may comprise a sequence selected from the group consisting of SEQ ID NOs. 1016-1018, having the codon-optimized nucleic acid sequences of SEQ ID NOs. 1019-1021, respectively.

[0269] In certain embodiments, the modified Cpf1 protein may comprise a nuclear localization signal (NLS). For example, but not by way of limitation, NLS sequences useful in connection with the methods

and compositions disclosed herein will comprise an amino acid sequence capable of facilitating protein import into the cell nucleus. NLS sequences useful in connection with the methods and compositions disclosed herein are known in the art. Examples of such NLS sequences include the nucleoplasmin NLS having the amino acid sequence: KRPAATKKAGQAKKKK (SEQ ID NO. 1006) and the simian virus 40 "SV40" NLS having the amino acid sequence PKKKRKV (SEQ ID NO. 1007).

[0270] In certain embodiments, the NLS sequence of the modified Cpf1 protein is positioned at or near the C-terminus of the Cpf1 protein sequence. For example, but not by way of limitation, the modified Cpf1 protein can be selected from the following: His-AsCpf1-nNLS (SEQ ID NO. 1000); His-AsCpf1-sNLS (SEQ ID NO. 1008) and His-AsCpf1-sNLS-sNLS (SEQ ID NO. 1001), where "His" refers to a six-histidine purification sequence, "AsCpf1" refers to the *Acidaminococcus sp.* Cpf1 protein sequence, "nNLS" refers to the nucleoplasmin NLS, and "sNLS" refers to the SV40 NLS. Additional permutations of the identity and C-terminal positions of NLS sequences, e.g., appending two or more nNLS sequences or combinations of nNLS and sNLS sequences (or other NLS sequences), as well as sequences with and without purification sequences, e.g., six-histidine sequences, are within the scope of the instantly disclosed subject matter.

[0271] In certain embodiments, the NLS sequence of the modified Cpf1 protein may be positioned at or near the N-terminus of the Cpf1 protein sequence. For example, but not by way of limitation, the modified Cpf1 protein may be selected from the following: His-sNLS-AsCpf1 (SEQ ID NO. 1009), His-sNLS-sNLS-AsCpf1 (SEQ ID NO. 1010), and sNLS-sNLS-AsCpf1 (SEQ ID NO. 1011). Additional permutations of the identity and N-terminal positions of NLS sequences, e.g., appending two or more nNLS sequences or combinations of nNLS and sNLS sequences (or other NLS sequences), as well as sequences with and without purification sequences, e.g., six-histidine sequences, are within the scope of the instantly disclosed subject matter.

[0272] In certain embodiments, the modified Cpf1 protein may comprise NLS sequences positioned at or near both the N-terminus and C-terminus of the Cpf1 protein sequence. For example, but not by way of limitation, the modified Cpf1 protein may be selected from the following: His-sNLS-AsCpf1-sNLS (SEQ ID NO. 1012) and His-sNLS-sNLS-AsCpf1-sNLS-sNLS (SEQ ID NO. 1013). Additional permutations of the identity and N-terminal/C-terminal positions of NLS sequences, e.g., appending two or more nNLS sequences or combinations of nNLS and sNLS sequences (or other NLS sequences) to either the N-terminal/C-terminal positions, as well as sequences with and without purification sequences, e.g., six-histidine sequences, are within the scope of the instantly disclosed subject matter.

[0273] In certain embodiments, the modified Cpf1 protein may comprise an alteration (e.g., a deletion or substitution) at one or more cysteine residues of the Cpf1 protein sequence. For example, but not by way of limitation, modified Cpf1 protein may comprise an alteration at a position selected from the group consisting of: C65, C205, C334, C379, C608, C674, C1025, and C1248. In certain embodiments, the modified Cpf1 protein may comprise a substitution of one or more cysteine residues for a serine or alanine. In certain embodiments, the modified Cpf1 protein may comprise an alteration selected from the group consisting of: C65S, C205S, C334S, C379S, C608S, C674S, C1025S, and C1248S. In certain embodiments, the modified Cpf1 protein may comprise an alteration selected from the group consisting of: C65A, C205A, C334A, C379A, C608A, C674A, C1025A, and C1248A. In certain embodiments, the modified Cpf1 protein may comprise alterations at positions C334 and C674 or C334, C379, and C674. In certain embodiments, the modified Cpf1 protein may comprise the following alterations: C334S and C674S, or C334S, C379S, and C674S. In certain embodiments, the modified Cpf1 protein may comprise the following alterations: C334A and C674A, or C334A, C379A, and C674A. In certain embodiments, the modified Cpf1 protein may comprise both one or more cysteine residue alteration as well as the introduction of one or more NLS sequences, e.g., His-AsCpf1-nNLS Cys-less (SEQ ID NO. 1014) or His-AsCpf1-nNLS Cys-low (SEQ ID NO. 1015). In various embodiments, the Cpf1 protein comprising a deletion or substitution in one or more cysteine residues exhibits reduced aggregation.

[0274] In certain embodiments, other modified Cpf1 proteins known in the art may be used with the methods and systems described herein. For example, in certain embodiments, the modified Cpf1 may be Cpf1 containing the mutation S542R/K548V/N552R ("Cpf1 RVR"). Cpf1 RVR has been shown to cleave target sites with a TATV PAM. In certain embodiments, the modified Cpf1 may be Cpf1 containing the mutation S542R/K607R ("Cpf1 RR"). Cpf1 RR has been shown to cleave target sites with a TYCV/CCCC PAM.

[0275] In some embodiments, a Cpf1 variant is used herein, wherein the Cpf1 variant comprises mutations at one or more residues of AsCpf1 (Acidaminococcus sp. BV3L6) selected from the group consisting of 11, 12, 13, 14, 15, 16, 17, 34, 36, 39, 40, 43, 46, 47, 50, 54, 57, 58, 111, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 642, 643, 644, 645, 646, 647, 648, 649, 651, 652, 653, 654, 655, 656, 676, 679, 680, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 707, 711, 714, 715, 716, 717, 718, 719, 720, 721, 722,

739, 765, 768, 769, 773, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, or 1048 or the corresponding position of an AsCpf1 orthologue, homologue, or variant.

[0276] In certain embodiments, a Cpf1 variant as used herein may include any of the Cpf1 proteins described in International Publication Number WO 2017/184768 A1 by Zhang et al. ("'768 Publication"), which is incorporated by reference herein.

[0277] In certain embodiments, a modified Cpf1 protein (also referred to as a Cpf1 variant) used herein may be encoded by any of the sequences set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). **Table 20** sets forth exemplary Cpf1 variant amino acid and nucleotide sequences. These sequences are set forth in **Fig. 62**, which details the positions of six-histidine sequences (underlined letters) and NLS sequences (bolded letters). Additional permutations of the identity and N-terminal/C-terminal positions of NLS sequences, e.g., appending two or more nNLS sequences or combinations of nNLS and sNLS sequences (or other NLS sequences) to either the N-terminal/C-terminal positions, as well as sequences with and without purification sequences, e.g., six-histidine sequences, are within the scope of the instantly disclosed subject matter.

[0278] In certain embodiments, any of the Cpf1 proteins or modified Cpf1 proteins disclosed herein may be complexed with one or more gRNA comprising the targeting domain set forth in SEQ ID NOs 1002 and/or 1004 to alter a CCAAT box target region. In certain embodiments, any of the Cpf1 proteins or modified Cpf1 proteins disclosed herein may be complexed with one or more gRNA comprising a sequence set forth in **Table 18** or **Table19**. In certain embodiments, the modified Cpf1 protein may be His-AsCpf1-nNLS (SEQ ID NO: 1000) or His-AsCpf1-sNLS-sNLS (SEQ ID NO:1001). In certain embodiments, a modified Cpf1 protein used herein may be encoded by any of the sequences set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). In certain embodiments, the modified Cpf1 protein may comprise the sequence set forth in SEQ ID NO:1097.

Modifications of RNA-guided nucleases

[0279] 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.

[0280] 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 & Hsu 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).

[0281] On the other hand, inactivation of a Cas9 HNH domain results in a nickase that cleaves the bottom or non-complementary strand.

[0282] Modifications of PAM specificity relative to naturally occurring Cas9 reference molecules has been described by Kleinstiver et al. for both *S. pyogenes* (Kleinstiver 2015a) and *S. aureus* (Kleinstiver 2015b). Kleinstiver et al. have also described modifications that improve the targeting fidelity of Cas9 (Kleinstiver 2016). Each of these references is incorporated by reference herein.

[0283] RNA-guided nucleases have been split into two or more parts, as described by Zetsche 2015 and Fine 2015 (both incorporated by reference herein).

[0284] 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, incorporated by reference herein for all purposes.

[0285] 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.

[0286] 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.

Nucleic acids encoding RNA-guided nucleases

[0287] 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).

[0288] 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.

[0289] 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.

[0290] 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

[0291] 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)

[0292] 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.

[0293] 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.

[0294] Two non-limiting examples of DSF assay conditions are set forth below:

[0295] To determine the best solution to form RNP complexes, a fixed concentration (e.g. 2 μM) of Cas9 in water+10x 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.

[0296] The second assay consists of mixing various concentrations of gRNA with fixed concentration (e.g. 2 μM) 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 + 10x 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

[0297] 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.

[0298] 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.

[0299] 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.

[0300] 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 & Hsu 2013 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).

[0301] 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.

[0302] 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.

[0303] 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.

[0304] 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

[0305] 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.

[0306] 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 CCAAT box target region. In certain embodiments, one or more gRNAs comprising a sequence set forth in SEQ ID NOs:970, 971, 996, 997 may be used to introduce alterations in the CCAAT box target region. In certain embodiments, one or more gRNAs comprising a targeting domain set forth in SEQ ID NOs:1002, 1004 may be used to introduce alterations in the CCAAT box target region. 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*, the CCAAT box target region, the 13 nt target region of *HBG1* and/or *HBG2*, or a combination thereof. 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*. 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 or gRNAs comprising a targeting domain set forth in SEQ ID NOs:970, 971, 996, 997 may be used to introduce alterations in the CCAAT box target region. 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:1002, 1004 may be used to introduce alterations in the CCAAT box target region.

[0307] In certain embodiments, multiple gRNAs and an RNA-guided nuclease may be used in genome editing systems to introduce alterations (e.g., deletions, insertions) into the CCAAT box target region of *HBG1* and/or *HBG2*. In certain embodiments, the RNA-guided nuclease may be a Cas9, modified Cas9 (e.g., D10A), Cpf1, or modified Cpf1.

Donor template design

[0308] 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.

[0309] 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:

[5' homology arm] — [replacement sequence] —- [3' homology arm].

[0310] 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.

[0311] Replacement sequences in donor templates have been described elsewhere, including in Cotta-Ramusino. 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.

[0312] 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 80-200 nucleotides (e.g., 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, or 200 nucleotides).

[0313] 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.

[0314] 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.

[0315] In certain embodiments, silent, non-pathogenic SNPs may be included in the ssODN donor template to allow for identification of a gene editing event.

[0316] 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.

[0317] A donor template or template nucleic acid, as that term is used herein, refers to a nucleic acid sequence which can be used in conjunction with an RNA nuclease molecule and one or more gRNA molecules to alter (e.g., delete, disrupt, or modify) a target DNA sequence. In certain embodiments, the template nucleic acid results in an alteration (e.g., deletion) at the CCAAT box target region of *HBG1* and/or *HBG2*. In certain embodiments, the alteration is a non-naturally occurring alteration. In certain embodiments, the non-naturally occurring alteration at the CCAAT box target region of *HBG1* and/or *HBG2* may comprise the 18 nt target region, the 11 nt target region, the 4 nt target region, or the 1 nt target region, or a combination thereof. In certain embodiments, the alteration is a naturally occurring alteration. In certain embodiments, the naturally occurring alteration at the CCAAT box target region of *HBG1* and/or *HBG2* may comprise the 13 nt target region, the c.-117 G>A target region, or a combination thereof. In certain embodiments, the template nucleic acid is an ssODN. In certain embodiments, the ssODN is a positive strand or a negative strand.

[0318] For example, a template nucleic acid for introducing the 18 nt deletion at the 18 nt target region (HBG1 c.-104 to -121, HBG2 c.-104 to -121, or a combination thereof) may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm, where the replacement sequence is 0 nucleotides or 0 bp. In certain embodiments, the 5' homology arm may be about 25 to about 200 nucleotides or more in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 18 nt target region. In certain embodiments, the 3' homology arm may be about 25 to about 200 nucleotides or more in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 3' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 18 nt target region. In certain embodiments, the 5' and 3' homology arms are symmetrical in length. In certain embodiments, the 5' and 3' homology arms are asymmetrical in length. In certain embodiments, the template nucleic acid is an ssODN. In certain embodiments, the ssODN is a positive strand. In certain embodiments, the ssODN is a negative strand. In certain embodiments, the ssODN comprises, consists essentially of, or consists of SEQ ID NO:974 (OLI16409) or SEQ ID NO:975 (OLI16410).

[0319] In certain embodiments, a template nucleic acid for introducing the 11 nt deletion at the 11 nt target region (HBG1 c.-105 to -115, HBG2 c.-105 to -115, or a combination thereof) may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm, where the replacement sequence is 0 nucleotides or 0 bp. In certain embodiments, the 5' homology arm may be about 25 to about 200

nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 11 nt target region. In certain embodiments, the 3' homology arm may be about 25 to about 200 nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 3' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 11 nt target region. In certain embodiments, the 5' and 3' homology arms are symmetrical in length. In certain embodiments, the 5' and 3' homology arms are asymmetrical in length. In certain embodiments, the template nucleic acid is an ssODN. In certain embodiments, the ssODN is a positive strand. In certain embodiments, the ssODN is a negative strand. In certain embodiments, the ssODN comprises, consists essentially of, or consists of SEQ ID NO:976 (OLI16411) or SEQ ID NO:978 (OLI16413).

[0320] In certain embodiments, a template nucleic acid for introducing the 4 nt deletion at the 4 nt target region (HBG1 c.-112 to -115, HBG2 c.-112 to -115, or a combination thereof) may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm, where the replacement sequence is 0 nucleotides or 0 bp. In certain embodiments, the 5' homology arm may be about 25 to about 200 nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 4 nt target region. In certain embodiments, the 3' homology arm may be about 25 to about 200 nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 3' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 4 nt target region. In certain embodiments, the 5' and 3' homology arms are symmetrical in length. In certain embodiments, the 5' and 3' homology arms are asymmetrical in length. In certain embodiments, the template nucleic acid is an ssODN. In certain embodiments, the ssODN is a positive strand. In certain embodiments, the ssODN is a negative strand. In certain embodiments, the ssODN comprises, consists essentially of, or consists of SEQ ID NO:984 (OLI16419), SEQ ID NO:985 (OLI16420), SEQ ID NO:986 (OLI16421), SEQ ID NO:987 (OLI16422), SEQ ID NO:988 (OLI16423), SEQ ID NO:989 (OLI16424), SEQ ID NO:990 (OLI16425), SEQ ID NO:991 (OLI16426), SEQ ID NO:992 (OLI16427), SEQ ID NO:993 (OLI16428), SEQ ID NO:994 (OLI16429), or SEQ ID NO:995 (OLI16430).

[0321] In certain embodiments, a template nucleic acid for introducing the 1 nt deletion at the 1 nt target region (HBG1 c.-116, HBG2 c.-116, or a combination thereof) may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm, where the replacement sequence is 0 nucleotides or 0 bp. In certain embodiments, the 5' homology arm may be about 25 to about 200 nucleotides in length, e.g., at

least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 1 nt target region. In certain embodiments, the 3' homology arm may be about 25 to about 200 nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 3' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 1 nt target region. In certain embodiments, the 5' and 3' homology arms are symmetrical in length. In certain embodiments, the 5' and 3' homology arms are asymmetrical in length. In certain embodiments, the template nucleic acid is an ssODN. In certain embodiments, the ssODN is a positive strand. In certain embodiments, the ssODN is a negative strand. In certain embodiments, the ssODN comprises, consists essentially of, or consists of SEQ ID NO:982 (OLI16417) or SEQ ID NO:983 (OLI16418).

[0322] In certain embodiments, the alteration at the CCAAT box target region recapitulates or is similar to a naturally occurring alteration, such as a 13 nt deletion. In certain embodiments, a template nucleic acid for introducing the 13 nt deletion at the 13 nt target region (HBG1 c.-116, HBG2 c.-116, or a combination thereof) may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm, where the replacement sequence is 0 nucleotides or 0 bp. In certain embodiments, the 5' homology arm may be about 25 to about 200 nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the 13 nt target region. In certain embodiments, the 3' homology arm may be about 25 to about 200 nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 3' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the 13 nt target region. In certain embodiments, the 5' and 3' homology arms are symmetrical in length. In certain embodiments, the 5' and 3' homology arms are asymmetrical in length. In certain embodiments, the template nucleic acid is an ssODN. In certain embodiments, the ssODN is a positive strand. In certain embodiments, the ssODN is a negative strand. In certain embodiments, the ssODN comprises, consists essentially of, or consists of SEQ ID NO:979 (OLI16414) or SEQ ID NO:977 (OLI16412).

[0323] In certain embodiments, the alteration at the CCAAT box target region recapitulates or is similar to a naturally occurring alteration, such as a substitution from G to A at the -117G>A target region. In certain embodiments, a template nucleic acid for introducing the -117G>A substitution at the -117G>A target region (HBG1 c.-117 G>A, HBG2 c.-117 G>A, or a combination thereof) may comprise a 5' homology arm, a replacement sequence, and a 3' homology arm, where the replacement sequence is 0

nucleotides or 0 bp. In certain embodiments, the 5' homology arm may be about 100 to about 200 nucleotides in length, e.g., at least about 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 5' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 5' of the -117G>A target region. In certain embodiments, the 3' homology arm may be about 25 to about 200 nucleotides in length, e.g., at least about 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides in length. In certain embodiments, the 3' homology arm comprises about 50 to 100 bp, e.g., 55 to 95, 60 to 90, 70 to 90, or 80 to 90 bp, homology 3' of the -117G>A target region. In certain embodiments, the 5' and 3' homology arms are symmetrical in length. In certain embodiments, the 5' and 3' homology arms are asymmetrical in length. In certain embodiments, the template nucleic acid is an ssODN. In certain embodiments, the ssODN is a positive strand. In certain embodiments, the ssODN is a negative strand. In certain embodiments, the ssODN comprises, consists essentially of, or consists of SEQ ID NO:980 (OLI16415) or SEQ ID NO:981 (OLI16416).

[0324] In certain embodiments, the 5' homology arm comprises a 5' phosphorothioate (PhTx) modification. In certain embodiments, the 3' homology arm comprises a 3' PhTx modification. In certain embodiments, the template nucleic acid comprises a 5' and 3' PhTx modification.

[0325] In certain embodiments, the ssODNs for introducing alterations (e.g., deletions) at the CCAAT box target region may be used in conjunction with an RNA nuclease and one or more gRNAs that target the CCAAT target region, for example, the gRNAs disclosed in **Table 7**, **Table 10**, **Table 12**, **Table 13, Table 18**, **Table 19**.

Target cells

[0326] 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*.

[0327] 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.

[0328] 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.

[0329] 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

[0330] 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 3

Genome Editing System Components			
RNA-guided Nuclease	**gRNA**	**Donor 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 DNA		A first DNA or DNA vector encoding an RNA-guided nuclease and a donor template, and a second DNA or DNA vector encoding a gRNA
	DNA RNA		A DNA or DNA vector encoding 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.

[0331] **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 4

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

	Biological liposomes: Erythrocyte Ghosts and Exosomes	YES	Transient	NO	Nucleic Acids

Nucleic acid-based delivery of genome editing systems

[0332] 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.

[0333] 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.

[0334] 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).

[0335] 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.

[0336] 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.

[0337] 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) nonparticles 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 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-propanaminium bromide	GAP-DLRIE	Cationic
Cetyltrimethylammonium bromide	CTAB	Cationic
6-Lauroxyhexyl ornithinate	LHON	Cationic
1-(2,3-Dioleoyloxypropyl)-2,4,6-trimethylpyridinium	2Oc	Cationic
2,3-Dioleyloxy-*N*-[2(sperminecarboxamido-ethyl]-*N,N*-dimethyl-1-propanaminium trifluoroacetate	DOSPA	Cationic
1,2-Dioleyl-3-trimethylammonium-propane	DOPA	Cationic
N-(2-Hydroxyethyl)-*N,N*-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide	MDRIE	Cationic
Dimyristooxypropyl dimethyl hydroxyethyl ammonium bromide	DMRI	Cationic
3β-[*N*-(*N',N'*-Dimethylaminoethane)-carbamoyl]cholesterol	DC-Chol	Cationic
Bis-guanidium-tren-cholesterol	BGTC	Cationic
1,3-Diodeoxy-2-(6-carboxy-spermyl)-propylamide	DOSPER	Cationic
Dimethyloctadecylammonium bromide	DDAB	Cationic

Dioctadecylamidoglicylspermidin	DSL	Cationic
rac-[(2,3-Dioctadecyloxypropyl)(2-hydroxyethyl)]-dimethylammonium chloride	CLIP-1	Cationic
rac-[2(2,3-Dihexadecyloxypropyl-oxymethyloxy)ethyl]trimethylammonium bromide	CLIP-6	Cationic
Ethyldimyristoylphosphatidylcholine	EDMPC	Cationic
1,2-Distearyloxy-*N,N*-dimethyl-3-aminopropane	DSDMA	Cationic
1,2-Dimyristoyl-trimethylammonium propane	DMTAP	Cationic
O,O'-Dimyristyl-N-lysyl aspartate	DMKE	Cationic
1,2-Distearoyl-sn-glycero-3-ethylphosphocholine	DSEPC	Cationic
N-Palmitoyl D-erythro-sphingosyl carbamoyl-spermine	CCS	Cationic
N-t-Butyl-*N*0-tetradecyl-3-tetradecylaminopropionamidine	diC14-amidine	Cationic
Octadecenolyoxy[ethyl-2-heptadecenyl-3 hydroxyethyl] imidazolinium chloride	DOTIM	Cationic
*N*1-Cholesteryloxycarbonyl-3,7-diazanonane-1,9-diamine	CDAN	Cationic
2-(3-[Bis(3-amino-propyl)-amino]propylamino)-*N*-ditetradecylcarbamoylme-ethyl-acetamide	RPR209120	Cationic
1,2-dilinoleyloxy-3- dimethylaminopropane	DLinDMA	Cationic
2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]- dioxolane	DLin-KC2-DMA	Cationic
dilinoleyl- methyl-4-dimethylaminobutyrate	DLin-MC3-DMA	Cationic

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

[0338] 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.

[0339] 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

[0340] 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.

[0341] *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

[0342] 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.

[0343] 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.

[0344] 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.

[0345] 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.

[0346] 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.

[0347] 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

[0348] 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 non-simultaneously. 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.

[0349] 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.

[0350] 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.

[0351] 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.

[0352] 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.

[0353] 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.

[0354] 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.

[0355] 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.

[0356] 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.

[0357] In certain embodiments, the second mode of delivery comprises a relatively transient element, e.g., an RNA or protein.

[0358] 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.

[0359] 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.

[0360] 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.

[0361] 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.

[0362] 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

[0363] 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

[0364] 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 (20A) 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 $CD34^+$ cells, RNPs were first delivered to K562 cells by electroporation (Amaxa Nucleofector).

[0365] 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 $CD34^+$ cells (<10% editing in $CD34^+$ 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. 3A**).

Table 7: Selected gRNAs for screening in K562 cells or $CD34^+$ cells

gRNA ID	Targeting domain sequence (RNA)	Targeting domain sequence (DNA)	Targeting domain sequence plus PAM (NGG) (RNA)	Targeting domain sequence plus PAM (NGG) (DNA)	Sense
Sp9	GGCUAUUGGU CAAGGCA (SEQ ID NO:277)	GGCTATTGGTCA AGGCA (SEQ ID NO:910)	GGCUAUUGGUCA AGGCAAGG (SEQ ID NO:920)	GGCTATTGGTCA AGGCAAGG (SEQ ID NO:930)	Antisense
Sp36	CAAGGCUAUU GGUCAAGGCA (SEQ ID NO:338)	CAAGGCTATTGG TCAAGGCA (SEQ ID NO:911)	CAAGGCUAUUGG UCAAGGCAAGG (SEQ ID NO:921)	CAAGGCTATTGG TCAAGGCAAGG (SEQ ID NO:931)	Antisense
Sp40	UGCCUUGUCA AGGCUAU (SEQ ID NO:327)	TGCCTTGTCAAG GCTAT (SEQ ID NO:912)	UGCCUUGUCAAGG CUAUUGG (SEQ ID NO:922)	TGCCTTGTCAAG GCTATTGG (SEQ ID NO:932)	Antisense
Sp42	GUUUGCCUUG UCAAGGCUAU (SEQ ID NO:299)	GTTTGCCTTGTC AAGGCTAT (SEQ ID NO:913)	GUUUGCCUUGUCA AGGCUAUUGG (SEQ ID NO:923)	GTTTGCCTTGTC AAGGCTATTGG (SEQ ID NO:933)	Antisense
Sp38	GACCAAUAGC CUUGACA (SEQ ID NO:276)	GACCAATAGCCT TGACA (SEQ ID NO:914)	GACCAAUAGCCUU GACAAGG (SEQ ID NO:924)	GACCAATAGCCT TGACAAGG (SEQ ID NO:934)	Sense
Sp37	CUUGACCAAU AGCCUUGACA (SEQ ID NO:333)	CTTGACCAATAG CCTTGACA (SEQ ID NO:915)	CUUGACCAAUAGC CUUGACAAGG (SEQ ID NO:925)	CTTGACCAATAG CCTTGACAAGG (SEQ ID NO:935)	Sense
Sp43	GUCAAGGCUA	GTCAAGGCTATT	GUCAAGGCUAUU	GTCAAGGCTATT	Antisense

	UUGGUCA (SEQ ID NO:278)	GGTCA (SEQ ID NO:916)	GGUCAAGG (SEQ ID NO:926)	GGTCAAGG (SEQ ID NO:936)	
Sp35	CUUGUCAAGG CUAUUGGUCA (SEQ ID NO:339)	CTTGTCAAGGCT ATTGGTCA (SEQ ID NO:917)	CUUGUCAAGGCUA UUGGUCAAGG (SEQ ID NO:927)	CTTGTCAAGGCT ATTGGTCAAGG (SEQ ID NO:937)	Antisense
Sp41	UCAAGUUUGC CUUGUCA (SEQ ID NO:310)	TCAAGTTTGCCT TGTCA (SEQ ID NO:918)	UCAAGUUUGCCUU GUCAAGG (SEQ ID NO:928)	TCAAGTTTGCCT TGTCAAGG (SEQ ID NO:938)	Antisense
Sp34	UGGUCAAGUU UGCCUUGUCA (SEQ ID NO:340)	TGGTCAAGTTTG CCTTGTCA (SEQ ID NO:919)	UGGUCAAGUUUG CCUUGUCAAGG (SEQ ID NO:929)	TGGTCAAGTTTG CCTTGTCAAGG (SEQ ID NO:939)	Antisense
Sp85	AGUAUCCAGU GAGGCCA (SEQ ID NO:940)	AGTATCCAGTGA GGCCA (SEQ ID NO:943)	AGUAUCCAGUGA GGCCAGGG (SEQ ID NO:946)	AGTATCCAGTGA GGCCAGGG (SEQ ID NO:949)	Antisense
SpA	GGCAAGGCUG GCCAACCCAU (SEQ ID NO:941)	GGCAAGGCTGG CCAACCCAT (SEQ ID NO:944)	GGCAAGGCUGGCC AACCCAUGGG (SEQ ID NO:947)	GGCAAGGCTGG CCAACCCATGGG (SEQ ID NO:950)	Sense
SpB	UAUUUGCAUU GAGAUAGUGU (SEQ ID NO:942)	TATTTGCATTGA GATAGTGT (SEQ ID NO:945)	UAUUUGCAUUGA GAUAGUGUGGG (SEQ ID NO:948)	TATTTGCATTGA GATAGTGTGGG (SEQ ID NO:951)	Sense

[0366] 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. 3B)** for *HBG1*. At least five of the eight sgRNAs also supported targeted deletion of the 13 nt in *HBG2* promoter region **(Fig. 3C)**. 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. 3B-3C** 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

[0367] Of the RNPs containing different gRNAs tested in human cord blood (CB) $CD34^+$ 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 $CD34^+$ cells from a three cord blood donors **(Fig. 4A)**. 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).

[0368] 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 $CD34^+$ cells (n=3 donors) and adult mobilized peripheral blood (mPB) $CD34^+$ 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 $CD34^+$ cells across 3 donors and 3 separate experiments **(Fig. 4B)**.

[0369] 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 8: Single strand deoxynucleotide donor repair templates (ssODN)

ssODN ID	SEQ ID NO	Sequence
ssODN1 5' homology arm	904	**GGGTGCTTCCTTTTATTCTTCATCCCTAGCCAGCCGCCGGCCCCTG GCCTCACTGGATACTCTAAGACTATTGGTCAAGTTTGCCTT**
ssODN1 3' homology arm	905	GTCAAGGCAAGGCTGGCCAACCCATGGGTGGAGTTTAGCCAGGGACCG TTTCAGACAGATATTTGCATTGAGATAGTGTGGGGAAGGGG
ssODN1	906	**GGGTGCTTCCTTTTATTCTTCATCCCTAGCCAGCCGCCGGCCCCTG GCCTCACTGGATACTCTAAGACTATTGGTCAAGTTTGCCTT**GTCAAG GCAAGGCTGGCCAACCCATGGGTGGAGTTTAGCCAGGGACCGTTTCAG ACAGATATTTGCATTGAGATAGTGTGGGGAAGGGG
PhTx ssODN1 5' homology arm	907	**G*GGTGCTTCCTTTTATTCTTCATCCCTAGCCAGCCGCCGGCCCCTG GCCTCACTGGATACTCTAAGACTATTGGTCAAGTTTGCCTT**
PhTx ssODN1 3' homology arm	908	GTCAAGGCAAGGCTGGCCAACCCATGGGTGGAGTTTAGCCAGGGACCG TTTCAGACAGATATTTGCATTGAGATAGTGTGGGGAAGGG*G
PhTx ssODN1	909	**G*GGTGCTTCCTTTTATTCTTCATCCCTAGCCAGCCGCCGGCCCCTG GCCTCACTGGATACTCTAAGACTATTGGTCAAGTTTGCCTT**GTCAAG GCAAGGCTGGCCAACCCATGGGTGGAGTTTAGCCAGGGACCGTTTCAG ACAGATATTTGCATTGAGATAGTGTGGGGAAGGG*G
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.		

[0370] 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 $CD34^+$ 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. 4C**). 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. 4C**, 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 of the CAAT box in the proximal promoter), the absence of which is associated with elevation of HbF expression (**Fig. 4C**, 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 $CD34^+$ 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.

[0371] 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 $CD34^+$ cells, human adult $CD34^+$ cells from mobilized peripheral blood (mPB) were electroporated with the RNPs. Briefly, mPB $CD34^+$ 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 $CD34^+$ cells treated with RNP complexed to Sp37 while no editing was detected for cells that were treated with RNP complexed to Sp35 (**Fig. 5A**).

[0372] 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. 5A**). To determine whether editing *HBG* increases production of fetal hemoglobin in erythroid progeny of edited adult $CD34^+$ 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 $CD34^+$ cells and donor matched negative (untreated) controls. By day 7 of differentiation, erythroblast progeny of human $CD34^+$ 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. 5B**). Importantly, $CD34^+$ 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 $CD34^+$ cell negative control), as determined in hematopoietic colony forming cell (CFC) assays (**Fig. 6A**). Furthermore, the erythroblasts differentiated from RNP treated $CD34^+$ 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^+ cells) (**Fig. 6B**). 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

[0373] To determine whether co-delivery of paired nickase RNPs targeting *HBG* would increase targeted disruption of the proximal *HBG* promoter, mPB $CD34^+$ 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 $CD34^+$ 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. 7A)**. DNA sequencing analysis was performed on limited samples shown in **Fig. 7A**. 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. 7B)**. 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+ cells supports induction of HbF protein in erythroid progeny.

[0374] To further optimize editing conditions in mPB $CD34^+$ cells at the target site and to evaluate editing in additional human cell donors, human mPB $CD34^+$ 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. 8A).** 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. 8B)**. Illumina sequencing analysis of indels correlated with indels determined by T7E1 analysis **(Fig. 8C-8D)**.

[0375] To determine whether $CD34^+$ 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 $CD34^+$ cells was maintained during erythroid differentiation and that edited $CD34^+$ cells maintain erythroid differentiation potential.

[0376] 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.

[0377] 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 $CD34^+$ cells (data not shown).

Example 6: Increasing the dose of RNP increases total editing efficiency in human adult $CD34^+$ cells at the *HBG* locus.

[0378] 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 $CD34^+$ cells by electroporation. The increased RNP concentration supported an increase in indels at the *HBG* target site to >30% (**Fig. 10A**) as determined by T7E1 endonuclease analysis of the *HBG* PCR product amplified for gDNA extracted 3 days after electroporation of $CD34^+$ cells. Sequencing analysis indicated that increasing the RNP concentration increased insertions (**Fig. 10B**). Erythroid progeny of RNP treated $CD34^+$ cells also had an increase in HbF protein production (**Fig. 10C**). Importantly, the hematopoietic colony forming potential was maintained after editing (**Fig. 10D**). These cells were then transplanted into immunodeficient mice and their engraftment 1 month (**Fig. 10E**) and 2 months (**Fig. 10F**) after transplantation was evaluated by sampling the peripheral blood and measuring the percentage of human $CD45^+$ 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 $CD45^+$ fraction among cohorts at indicated time points (**Fig. 10G-H**).

[0379] Two additional D10A nickase pairs were also tested in RNP dose response studies in adult mPB $CD34^+$ cells (Sp37+SpA, Sp37+SpB). Here, mPB $CD34^+$ cells were electroporated with D10A paired nickases delivered at 0, 2.5, and 3.75 μM 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 $CD34^+$ cells was plotted with the HbF levels detected in erythroid progeny in order to correlate editing and HbF induction **(Fig. 11A)**. RNP treated and untreated control mPB $CD34^+$ 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 $CD34^+$ cells **(Fig. 11B)**. There was no difference in the percentage of human $CD45^+$ cells in the mouse peripheral blood 1 month after transplantation and no difference in blood lineage distribution **(Fig. 11C-D)** for cells exposed to different D10A RNP pairs at different doses compared to untreated donor matched control $CD34^+$ 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 $CD34^+$ cells and supports induction of fetal hemoglobin expression in erythroid progeny

[0380] 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.

[0381] In this experiment, CB $CD34^+$ 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. 13A)**. 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) in CB $CD34^+$ cells. T7E1 analysis of *BCL11A* erythroid enhancer PCR product from gDNA extracted from CB $CD34^+$ cells treated with *BCL11Ae* RNP indicated that ~5% indels was achieved **(Fig. 13A)**. 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 $CD34^+$

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. 13B**).

Table 9: Select gRNA sequences targeting *BCL11A* erythroid enhancer for screening in CD34$^+$ cells

gRNA ID	Targeting domain sequence (RNA)	Targeting domain sequence (DNA)	Targeting domain sequence plus PAM (NGG) (RNA)	Targeting domain sequence plus PAM (NGG) (DNA)	Sense
SpK	CUAACAGUUG CUUUUAUCAC (SEQ ID NO:952)	CTAACAGTTG CTTTTATCAC (SEQ ID NO:956)	CUAACAGUUGC UUUUAUCACAG G (SEQ ID NO:960)	CTAACAGTTGC TTTTATCACAGG (SEQ ID NO:964)	Antisense
SpM	GGGCGUGGGU GGGGUAGAA G (SEQ ID NO:953)	GGGCGTGGG TGGGGTAGA AG (SEQ ID NO:957)	GGGCGUGGGUG GGGUAGAAGAG G (SEQ ID NO:961)	GGGCGTGGGTG GGGTAGAAGAG G (SEQ ID NO:965)	Antisense
SpN	CUCUUAGACA UAACACACCA (SEQ ID NO:954)	CTCTTAGACA TAACACACC A (SEQ ID NO:958)	CUCUUAGACAU AACACACCAGG G (SEQ ID NO:962)	CTCTTAGACAT AACACACCAGG G (SEQ ID NO:966)	Antisense
SpC	AUCAGAGGCC AAACCCUUCC (SEQ ID NO:955)	ATCAGAGGC CAAACCCTTC C (SEQ ID NO:959)	AUCAGAGGCCA AACCCUUCCUG G (SEQ ID NO:963)	ATCAGAGGCCA AACCCTTCCTG G (SEQ ID NO:967)	Sense

[0382] Next, human adult bone marrow CD34$^+$ cells were electroporated with the *BCL11Ae* RNP. DNA sequencing analysis of the *BCL11A* PCR product amplified from gDNA extracted from marrow CD34$^+$ cells indicated 15% gene editing comprised of insertions and deletions (**Fig. 14A**). 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$^+$ cells were plated into colony forming assays and the mixed hematopoietic colonies (GEMMs), which correspond to CD34$^+$ 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$^+$ 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+ cell population (**Fig. 14B**). 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 heterogenous CD34$^+$ cells versus the other lineages present, but not captured/differentiated in the short-

term CFC assays. The RNP treated marrow $CD34^+$ cells also maintained similar kinetics of erythroid maturation (enucleation, **Fig. 14C**) and differentiation (phenotype acquisition, **Fig. 14D**) compared to donor matched untreated control cells. Erythroid progeny of edited marrow $CD34^+$ cells exhibited ~5-fold increase in HbF induction as determined by flow cytometry analysis (**Fig. 14E**).

[0383] Gene editing and induction of fetal hemoglobin was also evaluated in human adult mPB $CD34^+$ cells. Co-delivery of *BCL11Ae* RNP and nonspecific ssODN supported ~20% indels at the target site (**Fig. 15A**). To evaluate early induction of fetal hemoglobin in erythroid progeny of edited cells, mPB $CD34^+$ cells were differentiated into erythroblasts and induction of fetal hemoglobin transcription (*HBG* mRNA) was evaluated by qRT-PCR analysis. The erythroid progeny of *BCL11A*e RNP treated $CD34^+$ cells exhibited a 2-fold induction of *HBG* mRNA compared to untreated controls, suggesting induction of fetal hemoglobin expression (**Fig. 15B**). The RNP treated marrow $CD34^+$ cells also maintained similar kinetics of differentiation (phenotype acquisition, **Fig. 15C**) compared to donor matched untreated control cells.

Example 8: Electroporation of Cas9 RNP targeting the distal CCAAT box at the *HBG* promoter in human hematopoietic stem/progenitor cells generates several deletions that promote *HBG* expression after erythroid differentiation

[0384] Hereditary persistence of fetal hemoglobin (HPFH phenotype) is observed in patients carrying a 13 nt deletion overlapping with the *HBG* distal CCAAT box. Cas9-RNP targeting the *HBG* distal CCAAT box can be used in hematopoietic stem/progenitor cells (HSPCs) to reproduce the HPFH phenotype, likely by disrupting the binding sites of transcription factors repressing *HBG* expression. DNA double strand breaks (DSBs) created by Cas9 RNP can lead to a variety of repair outcomes, including insertions and deletions proximal to the RNP cut site. Some of the introduced indels may disrupt the binding of repressing factors less efficiently. It was envisioned that ssODN donor templates could be used to improve the frequency of indels reproducing the HPFH phenotype by directing the repair towards specific deletions that leads to *HBG* gene expression.

[0385] To evaluate the repair outcome of Cas9-RNP targeting the distal CCAAT box at the *HBG* promoter, Cas9 was complexed with the chemically synthesized guide RNA OLI7066 (SEQ ID NO:970, **Table 10**) ("OLI7066-RNP") and RNP were electroporated into HSPCs at 16 μM. Sequencing analysis (next generation sequencing) performed at day 2 post-electroporation indicated that 23.7% of the alleles carried the 13 nt deletion identical to the naturally occurring HPFH mutation (**Fig. 16,** 13 nt deletion indicated by "Δ-102:-114"). Several other frequent deletions were also observed around the OLI7066-RNP cut site (**Fig. 16**).

Table 10: Sequences of chemically synthesized gRNA targeting the CCAAT box, with or without end modifications.

Name	Oli-ID	gRNA sequence (RNA)	gRNA sequence (DNA)
Sp35 unmodified	OLI7066	CUUGUCAAGGCUAUUGGUCAG UUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCA CCGAGUCGGUGCUUUU (SEQ ID NO:970)	CTTGTCAAGGCTATTGGTCA GTTTTAGAGCTAGAAATAGC AAGTTAAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCTTTT (SEQ ID NO:972)
Sp35 5'-3' PSOMe	OLI8394	mC*mU*mU*GUCAAGGCUAUUG GUCAGUUUUAGAGCUAGAAAU AGCAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAAG UGGCACCGAGUCGGUGCmU*m U*mU*U (SEQ ID NO:971)	mC*mT*mT*GTCAAGGCTATT GGTCAGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGC TAGTCCGTTATCAACTTGAA AAAGTGGCACCGAGTCGGTG CmT*mT*mT*T (SEQ ID NO:973)
*: Represents phosphorothioate modification m: Represents 2-o-methyl modification			

[0386] A single cell experiment was performed to evaluate the level of HbF expression induced by the most frequent deletions generated by delivery of RNP (Cas9 complexed with the gRNA OLI7066 (SEQ ID NO:970) ("OLI7066-RNP")) targeting the distal CCAAT box. Briefly, single HSPC electroporated with OLI7066-RNP complexes targeting the distal CCAAT box were plated in single wells and differentiated into erythroid cells by culturing for 18 days in the presence of human cytokines (erythropoietin, SCF, IL3), human plasma (Octoplas), and other supplements (hydrocortisone, heparin, transferrin, insulin). A fraction of the erythroid progeny from each single cell was split after 14 days for gDNA extraction. Sequencing analysis of the PCR product from *HBG1* and *HBG2* was performed to identify the genotype of each clonal population. ddPCR analysis was also performed on the genomic DNA of each clonal populations to detect deletions of the 4.9kb fragment between the guide RNA target sites in *HBG1* and *HBG2*. At day 18, the erythroid cells deriving from each clone were lysed and the relative expression of the globin chains was determined by ultra-performance liquid chromatography (UPLC). Finally, the level of G-gamma (Gγ)-globin, A-gamma (Aγ)-globin chain expression (or AG-gamma (AGγ)-globin resulting from the 4.9kb deletion) as determined by [gamma chain]/[all-gamma chains + beta chain] was compared relative to the indels carried at *HBG1* and *HBG2* respectively. The 13 nt deletion that reproduces the HPFH genotype was shown to induce HBG expression. In addition several unique non-naturally occurring deletions that induced *HBG* expression at comparable levels were identified. These include HBG Δ-112:-115 ("4 nt deletion"), HBG Δ-104:-121 ("18 nt deletion"), HBG Δ-116 ("1 nt deletion") **(Figs. 17A-D)**.

Example 9: Co-delivery of Cas9 RNP targeting the region at or near the distal CCAAT box with ssODN donors supports precise gene editing in human hematopoietic stem/progenitor cells and increased gamma-globin expression in the erythroid progeny

[0387] To improve the frequency of HbF inducing indels in HSPC, single strand deoxynucleotide donor repair templates (ssODNs) "encoding" identified HbF inducing deletions (e.g., the "1nt" (HBG Δ-116) and "4nt" (HBG Δ-112:-115) deletions) were designed (**Fig. 18**). ssODNs used to induce these deletions were as follows: "1nt" ssODNs: OLI16417-18; "4nt" ssODNs: OLI16424, OLI16430, **Table 11**). ssODNs were designed with 90 nt homology arms flanking this absent sequence to create a perfect deletion. The ssODNs were modified to contain phosphorothioates (PhTx) at the 5' and 3' ends. Additional ssODNs were designed to encode naturally occurring mutations observed in patients with HPFH (e.g., HBGΔ-102:-114 ("13 nt" deletion, **Fig. 18**) and HBG-117 G>A mutation ("HBG-117 G>A"). ssODNs used to induce these deletions were as follows: "13nt" ssODN: OLI16412, OLI16414 and "-117 G>A" ssODN: OLI16415-16, **Table 11**). Finally, to facilitate the evaluation of gene correction, ssODNs (OLI16411, OLI16413, **Table 11**) were designed to encode an 11 nt deletion overlapping the distal CCAAT box that occurs at low frequency after electroporation of OLI8394-RNP alone (0.1%, see **Fig. 19B**, **Fig. 18**).

Table 11: Single strand deoxynucleotide donor repair templates "encoding" deletions or mutations at or near the CCAAT box

ssODN ID	OLI-ID	Sequence
Ptx ssODN - Positive Strand -18nt (180)	OLI16409	G*GTGCTTCCTTTTATTCTTCATCCCTAGCCAGCCGCCGGCCCC TGGCCTCACTGGATACTCTAAGACTATTGGTCAAGTTTGCCTT GTCAAGGCTGGCCAACCCATGGGTGGAGTTTAGCCAGGGACC GTTTCAGACAGATATTTGCATTGAGATAGTGTGGGGAAGGGG CCCCCAAGA*G (SEQ ID NO:974)
Ptx ssODN - Negative Strand -18nt (180)	OLI16410	C*TCTTGGGGGCCCCTTCCCCACACTATCTCAATGCAAATATC TGTCTGAAACGGTCCCTGGCTAAACTCCACCCATGGGTTGGCC AGCCTTGACAAGGCAAACTTGACCAATAGTCTTAGAGTATCC AGTGAGGCCAGGGGCCGGCGGCTGGCTAGGGATGAAGAATA AAAGGAAGCAC*C (SEQ ID NO:975)
Ptx ssODN - Negative Strand -11nt (180)	OLI16411	G*GCCCCTTCCCCACACTATCTCAATGCAAATATCTGTCTGAA ACGGTCCCTGGCTAAACTCCACCCATGGGTTGGCCAGCCTTGC CTTGAGACAAGGCAAACTTGACCAATAGTCTTAGAGTATCCA GTGAGGCCAGGGGCCGGCGGCTGGCTAGGGATGAAGAATAA AAGGAAGCACC*C (SEQ ID NO:976)
Ptx ssODN - Negative Strand -13nt (180)	OLI16412	G*GCCCCTTCCCCACACTATCTCAATGCAAATATCTGTCTGAA ACGGTCCCTGGCTAAACTCCACCCATGGGTTGGCCAGCCTTGC CTTGACAAGGCAAACTTGACCAATAGTCTTAGAGTATCCAGT GAGGCCAGGGGCCGGCGGCTGGCTAGGGATGAAGAATAAAA GGAAGCACCCT*T (SEQ ID NO:977)

Ptx ssODN - Positive Strand -11nt (180)	OLI16413	G*GGTGCTTCCTTTTATTCTTCATCCCTAGCCAGCCGCCGGCCC CTGGCCTCACTGGATACTCTAAGACTATTGGTCAAGTTTGCCT TGTCTCAAGGCAAGGCTGGCCAACCCATGGGTGGAGTTTAGC CAGGGACCGTTTCAGACAGATATTTGCATTGAGATAGTGTGG GGAAGGGGC*C (SEQ ID NO:978)
Ptx ssODN - Positive Strand -13nt (180)	OLI16414	A*AGGGTGCTTCCTTTTATTCTTCATCCCTAGCCAGCCGCCGGC CCCTGGCCTCACTGGATACTCTAAGACTATTGGTCAAGTTTGC CTTGTCAAGGCAAGGCTGGCCAACCCATGGGTGGAGTTTAGC CAGGGACCGTTTCAGACAGATATTTGCATTGAGATAGTGTGG GGAAGGGGC*C (SEQ ID NO:979)
Ptx ssODN - Negative Strand 117: G>A (180)	OLI16415	G*GCCCCTTCCCCACACTATCTCAATGCAAATATCTGTCTGAA ACGGTCCCTGGCTAAACTCCACCCATGGGTTGGCCAGCCTTGC CTTAACCAATAGCCTTGACAAGGCAAACTTGACCAATAGTCTT AGAGTATCCAGTGAGGCCAGGGGCCGGCGGCTGGCTAGGGAT GAAGAATAA*A (SEQ ID NO:980)
Ptx ssODN - Positive Strand 117: G>A (180)	OLI16416	T*TTATTCTTCATCCCTAGCCAGCCGCCGGCCCCTGGCCTCACT GGATACTCTAAGACTATTGGTCAAGTTTGCCTTGTCAAGGCTA TTGGTTAAGGCAAGGCTGGCCAACCCATGGGTGGAGTTTAGC CAGGGACCGTTTCAGACAGATATTTGCATTGAGATAGTGTGG GGAAGGGGC*C (SEQ ID NO:981)
Ptx ssODN - Negative Strand -1nt (180)	OLI16417	G*GCCCCTTCCCCACACTATCTCAATGCAAATATCTGTCTGAA ACGGTCCCTGGCTAAACTCCACCCATGGGTTGGCCAGCCTTGC CTTGCCAATAGCCTTGACAAGGCAAACTTGACCAATAGTCTTA GAGTATCCAGTGAGGCCAGGGGCCGGCGGCTGGCTAGGGATG AAGAATAAA*A (SEQ ID NO:982)
Ptx ssODN - Positive Strand -1nt (180)	OLI16418	T*TTATTCTTCATCCCTAGCCAGCCGCCGGCCCCTGGCCTCAC TGGATACTCTAAGACTATTGGTCAAGTTTGCCTTGTCAAGGCT ATTGGCAAGGCAAGGCTGGCCAACCCATGGGTGGAGTTTAGC CAGGGACCGTTTCAGACAGATATTTGCATTGAGATAGTGTGG GGAAGGGGC*C (SEQ ID NO:983)
Ptx ssODN - Negative Strand -4nt 40/80 (120)	OLI16419	T*GGCTAAACTCCACCCATGGGTTGGCCAGCCTTGCCTTGATA GCCTTGACAAGGCAAACTTGACCAATAGTCTTAGAGTATCCA GTGAGGCCAGGGGCCGGCGGCTGGCTAGGGATGAA*G (SEQ ID NO:984)
Ptx ssODN - Negative Strand -4nt 30/70 (100)	OLI16420	C*CACCCATGGGTTGGCCAGCCTTGCCTTGATAGCCTTGACAA GGCAAACTTGACCAATAGTCTTAGAGTATCCAGTGAGGCCAG GGGCCGGCGGCTGGC*T (SEQ ID NO:985)
Ptx ssODN - Negative Strand -4nt (100)	OLI16421	A*AACGGTCCCTGGCTAAACTCCACCCATGGGTTGGCCAGCCT TGCCTTGATAGCCTTGACAAGGCAAACTTGACCAATAGTCTTA GAGTATCCAGTGAG*G (SEQ ID NO:986)
Ptx ssODN - Negative Strand -4nt (120)	OLI16422	T*ATCTGTCTGAAACGGTCCCTGGCTAAACTCCACCCATGGGT TGGCCAGCCTTGCCTTGATAGCCTTGACAAGGCAAACTTGACC AATAGTCTTAGAGTATCCAGTGAGGCCAGGGGCC*G (SEQ ID NO:987)
Ptx ssODN - Negative Strand -4nt (140)	OLI16423	T*CAATGCAAATATCTGTCTGAAACGGTCCCTGGCTAAACTCC ACCCATGGGTTGGCCAGCCTTGCCTTGATAGCCTTGACAAGGC AAACTTGACCAATAGTCTTAGAGTATCCAGTGAGGCCAGGGG CCGGCGGCTGGC*T (SEQ ID NO:988)

Ptx ssODN - Negative Strand -4nt (180)	OLI16424	G*GCCCCTTCCCCACACTATCTCAATGCAAATATCTGTCTGAA ACGGTCCCTGGCTAAACTCCACCCATGGGTTGGCCAGCCTTGC CTTGATAGCCTTGACAAGGCAAACTTGACCAATAGTCTTAGA GTATCCAGTGAGGCCAGGGGCCGGCGGCTGGCTAGGGATGAA GAATAAAAGG*A (SEQ ID NO:989)
Ptx ssODN - Positive Strand -4nt 40/80 (120)	OLI16425	T*ACTCTAAGACTATTGGTCAAGTTTGCCTTGTCAAGGCTATC AAGGCAAGGCTGGCCAACCCATGGGTGGAGTTTAGCCAGGGA CCGTTTCAGACAGATATTTGCATTGAGATAGTGTG*G (SEQ ID NO:990)
Ptx ssODN - Positive Strand -4nt 30/70 (100)	OLI16426	C*TATTGGTCAAGTTTGCCTTGTCAAGGCTATCAAGGCAAGGC TGGCCAACCCATGGGTGGAGTTTAGCCAGGGACCGTTTCAGA CAGATATTTGCATTG*A (SEQ ID NO:991)
Ptx ssODN - Positive Strand -4nt (100)	OLI16427	C*CTCACTGGATACTCTAAGACTATTGGTCAAGTTTGCCTTGTC AAGGCTATCAAGGCAAGGCTGGCCAACCCATGGGTGGAGTTT AGCCAGGGACCGTT*T (SEQ ID NO:992)
Ptx ssODN - Positive Strand -4nt (120)	OLI16428	C*GGCCCCTGGCCTCACTGGATACTCTAAGACTATTGGTCAAG TTTGCCTTGTCAAGGCTATCAAGGCAAGGCTGGCCAACCCATG GGTGGAGTTTAGCCAGGGACCGTTTCAGACAGAT*A (SEQ ID NO:993)
Ptx ssODN - Positive Strand -4nt (140)	OLI16429	A*GCCAGCCGCCGGCCCCTGGCCTCACTGGATACTCTAAGACT ATTGGTCAAGTTTGCCTTGTCAAGGCTATCAAGGCAAGGCTGG CCAACCCATGGGTGGAGTTTAGCCAGGGACCGTTTCAGACAG ATATTTGCATTG*A (SEQ ID NO:994)
Ptx ssODN - Positive Strand -4nt (180)	OLI16430	T*CCTTTTATTCTTCATCCCTAGCCAGCCGCCGGCCCCTGGCCT CACTGGATACTCTAAGACTATTGGTCAAGTTTGCCTTGTCAAG GCTATCAAGGCAAGGCTGGCCAACCCATGGGTGGAGTTTAGC CAGGGACCGTTTCAGACAGATATTTGCATTGAGATAGTGTGG GGAAGGGGC*C (SEQ ID NO:995)
Ptx ssODN - Positive Strand -18nt (100)	OLI16431	C*CCTGGCCTCACTGGATACTCTAAGACTATTGGTCAAGTTTG CCTTGTCAAGGCTGGCCAACCCATGGGTGGAGTTTAGCCAGG GACCGTTTCAGACAG*A (SEQ ID NO:1040)
* : Represents phosphorothioate modification		

[0388] The ssODN donors, OLI16409 – OLI16418, OLI16424 and OLI16430 (**Table 11**), were co-delivered to adult mPB CD34+ cells with Cas9 protein precomplexed to OLI8394 ("OLI8394-RNP") or Cas9 protein precomplexed to OLI7066 ("OLI7066-RNP") targeting the HBG distal CCAAT box (**Table 10**). Briefly, mPB CD34+ cells were pre-stimulated for 2 days with human cytokines in X-Vivo-10 and then electroporated with a mixture composed of an ssODN donor at 2.5 μM and OLI8394-RNP or OLI7066-RNP at 2μM. After three days post electroporation the genomic DNA was extracted, and next-generation sequencing was performed on the *HBG* PCR products. The level of editing was increased from 62.1% when using OLI8394-RNP alone to up to 78.73% when co-delivering "-11" ssODN (i.e., OLI16411 or OLI16413) (**Fig. 19A**). The level of gene correction mediated by the co-delivery of the

ssODN templates to human HSPC was shown to contribute to up to 42.4% of total indels as measured by the frequency of the "-11" deletion within total indels detected by sequencing of the HBG PCR product from HSPC electroporated with OLI8394-RNP + "11" ssODN (**Fig. 19B**). Efficient gene correction was observed across other ssODN templates and OLI8394-RNP (**Figs. 19C-E, G**) or OLI7066-RNP (**Fig. 19F**).

[0389] The mPB CD34+ cells electroporated with OLI8394-RNP or OLI7066-RNP co-delivered with ssODN templates were differentiated into erythroid cells to evaluate the level of gamma-globin expression resulting from the gene editing. To determine whether co-delivering RNP with ssODNs increased the production of fetal hemoglobin in the erythroid progeny of edited adult CD34+ cells, the cells were differentiated into erythroid cells by culture for 18 days in the presence of human cytokines (erythropoietin, SCF, IL3), human plasma (Octoplas), and other supplements (hydrocortisone, heparin, transferrin, insulin). At day 18, the relative expression levels of gamma-globin chains over total beta-like globin chains (gamma chains/[gamma chains + beta chain]) was measured by UPLC. Up to 39.1% of gamma-globin was detected after co-delivery of ssODN instead of 29% when using OLI8394-RNP alone (**Figs. 20A-B**).

[0390] The effect of the dose of ssODN was evaluated using ssODN OLI16424. mPB CD34+ cells were pre-stimulated for 2 days with human cytokines in X-Vivo-10 and then electroporated with a mixture composed of OLI16424-ssODN at doses ranging from 0.625 μM to 10 μM and OLI8394-RNP at 2μM. After three days post electroporation the genomic DNA was extracted, and next-generation sequencing was performed on the *HBG* PCR products. Increasing the dose of ssODN up to 5 μM resulted in increased gene correction and reduced frequency of 4.9kb deletions between *HBG1* and *HBG2* (as measured by ddPCR), while maintaining overall editing level and without affecting cell viability (**Figs. 21A-C**). The level of gamma-globin increased up to 39.5% of total beta-like chains at the 5 μM dose of ssODN (**Fig. 21D**), as measured by UPLC analysis of the cell lysates after 18 days of erythroid culture.

[0391] The co-delivery of the OLI8394-RNP and the OLI16424 ssODN was further optimized by varying the relative dose of ssODN and RNP. mPB CD34+ cells were pre-stimulated for 2 days with human cytokines in X-Vivo-10 and then electroporated with a mixture composed of OLI16424-ssODN at doses ranging from 1.25 μM to 5 μM and OLI8394-RNP at 2 μM to 8 μM. After three days post electroporation the genomic DNA was extracted, and next-generation sequencing was performed on the *HBG* PCR products. An improved gene editing outcome, i.e., a lower frequency of 4.9kb deletions between the HBG1 and HBG2 genes and a higher level of gene correction, was achieved with a lower RNP dose combined with a higher ssODN dose (**Fig. 22A-C**). This resulted in the highest level of gamma-globin expression after erythroid differentiation (**Fig. 22D**).

Example 10: Gene correction in human hematopoietic stem/progenitor cells can be achieved with short ssODN templates with symmetrical and asymmetrical homology arms

[0392] To determine whether gene correction could be achieved using templates shorter than the previously tested 180 nt lengths used in Example 9, ssODNs encoding the "4 nt" deletion (HBGΔ-112:-115) were designed and synthesized with shorter homology arms either symmetrical or asymmetrical relative to the deleted sequence (OLI16419- OLI16423, and OLI16425 – OLI16429, **Table 11**, **Fig. 23A**). Briefly, mPB CD34+ cells were pre-stimulated for 2 days with human cytokines in X-Vivo-10 and then electroporated with a mixture composed of an ssODN donor (OLI16424, OLI16419 or OLI16421) at 2.5 μM and Cas9 protein precomplexed to OLI8394 (OLI8394-RNP) at 2 μM. After three days post electroporation the genomic DNA was extracted, and next-generation sequencing was performed on the *HBG* PCR products. Total editing and gene correction level were similar across ssODNs having symmetrical 90 nt homology arms, symmetrical 50 nt homology arms or 40/80 asymmetrical arms **(Fig. 23B)**.

Example 11: Co-delivery of ssODN donors with paired D10A-Cas9 RNPs to adult CD34+ cells supports gene correction at the *HBG* distal CCAAT box

[0393] To determine if ssODN-mediated gene correction to introduce CCAAT box disrupting deletions (such as, the "4nt" deletion) in mPB $CD34^+$ cells could be supported by D10A nickase pairs, CD34+ cells were electroporated with D10A Cas9 RNPs targeting *HBG,* with or without ssODN (OLI16424) (**Table 11**). Briefly, Sp37 and SpA gRNA were chemically synthesized (OLI7075 and OLI7074, respectively, **Table 12**) and complexed with D10A-Cas9 nickase mutant protein. Human adult mPB CD34+ cells were pre-stimulated for 48h in medium supplemented with human cytokines. Next, the D10A-Cas9 RNP pair comprising Sp37+SpA (2 μM + 2 μM) was delivered to the CD34+ cells, alone or in combination with OLI16424, and genomic DNA was extracted 3 days post-electroporation for Illumina sequencing analysis.

Table 12: gRNA sequences for targeting the CCAAT box in CD34$^+$ cells with D10A Cas9

gRNA ID	OLI-ID	gRNA sequence (RNA)	gRNA sequence (DNA)
Sp37	OLI17075	CUUGACCAAUAGCCUUGACA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU (SEQ ID NO:996)	CTTGACCAATAGCCTTGACAGTTT TAGAGCTAGAAATAGCAAGTTAA AATAAGGCTAGTCCGTTATCAAC TTGAAAAAGTGGCACCGAGTCGG TGCTTTT (SEQ ID NO:998)
SpA	OLI17074	GGCAAGGCUGGCCAACCCAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU (SEQ ID NO:997)	GGCAAGGCTGGCCAACCCATGTT TTAGAGCTAGAAATAGCAAGTTA AAATAAGGCTAGTCCGTTATCAA CTTGAAAAAGTGGCACCGAGTCG GTGCTTTT (SEQ ID NO:999)

[0394] Co-delivery of the ssODN supported ~65% indels, instead of ~51% when the D10A-Cas9 RNP pair was delivered alone, as determined by sequencing of the *HBG* PCR product from genomic DNA (**Fig. 24**). Detailed sequencing analysis also demonstrated that 16% of the alleles carried the precise 4 nt deletion (HBG:Δ-112:-115) when the OLI16424 ssODN donor was co-delivered, whereas this deletion was undetected when the RNP pair was delivered alone. This indicated that ~25% of indels occurred by precise gene correction in the presence of the ssODN.

Example 12: Cpf1 RNP containing gRNA targeting the CCAAT box supports gene editing in human hematopoietic stem/progenitor cells

[0395] A Cpf1 guide RNA, HBG1-1 (i.e., OLI13620 (**Table 13**)), was designed to target the CCAAT box. To determine optimal nuclear localization signal configuration for *Acidaminococcus sp.* Cpf1 ("AsCpf1") delivery in CD34+ cells, HBG1-1 gRNA was complexed to two nuclear localization signal (NLS) variants of AsCpf1, namely His-AsCpf1-nNLS (SEQ ID NO: 1000) and His-AsCpf1-sNLS-sNLS (SEQ ID NO:1001). "His" refers to a six-histidine purification sequence, "AsCpf1" refers to the *Acidaminococcus sp.* Cpf1 protein sequence, "nNLS" refers to the nucleoplasmin NLS, and "sNLS" refers to the SV40 NLS.

[0396] Briefly, mPB CD34+ cells were pre-stimulated for 2 days with human cytokines in X-Vivo-10 and then electroporated with the RNPs at 5 μM or 20 μM. The genomic DNA was extracted three days post electroporation and next-generation sequencing was performed on the *HBG* PCR products. HBG1-1 gRNA complexed to either of the Cpf1 NLS variants tested ("His-AsCpf1-sNLS-sNLS_HBG1-1 RNP" or "His-AsCpf1-nNLS_HBG1-1 RNP"), supported editing of CD34+ cells at the 13 nt target site. His-

AsCpf1-sNLS-sNLS_HBG1-1 RNP generated 60.6% edited alleles and His-AsCpf1-nNLS_HBG1-1 RNP generated 51.1% edited alleles at the highest dose tested (**Fig. 25A**).

Table 13: HBG1-1 gRNA sequence for targeting the CCAAT box in CD34$^+$ cells

gRNA ID	OLI-ID	Targeting domain sequence (RNA)	Targeting domain sequence (DNA)	Targeting domain sequence plus PAM (UUUG) (RNA)	Targeting domain sequence plus PAM (TTTG) (DNA)	Sense
HBG1-1	OLI13620	CCUUGUC AAGGCUA UUGGUC (SEQ ID NO:1002)	CCTTGTC AAGGCTA TTGGTC (SEQ ID NO:1003)	UUUGCCUUGU CAAGGCUAUU GGUC (SEQ ID NO:1004)	TTTGCCTTG TCAAGGCT ATTGGTC (SEQ ID NO:1005)	Antisense

Example 13: Co-delivery of Cpf1 RNP targeting the CCAAT box with ssODN donors supports gene editing in human hematopoietic stem/progenitor cells

[0397] RNP comprising HBG1-1 gRNA complexed to the His-AsCpf1-sNLS-sNLS variant ("His-AsCpf1-sNLS-sNLS_HBG1-1 RNP") were co-delivered by electroporation with single stranded oligodeoxynucleotide donor repair templates (ssODNs) to mPB CD34+ cells. As discussed in the examples above, OLI16430 and OLI16424 ssODNs were designed to "encode" a 4 nucleotide deletion and OLI16409 and OLI16410 ssODNs were designed to "encode" a 18 nucleotide deletion (**Table 11**). Both the 4 nt and 18 nt deletions disrupt the HBG distal CCAAT box and are associated with induction of *HBG* expression. The ssODNs include 90 nucleotide-long homology arms flanking the encoded absent sequence to create perfect deletion. The ssODNs were modified to contain phosphorothioates (PhTx) at the 5' and 3' ends (OLI16430, OLI16424, OLI16409, and OLI16410, **Table 11**). Briefly, human adult mPB CD34+ cells pre-stimulated for two days in medium supplemented with human cytokines were electroporated with 5 μM RNP comprising the His-AsCpf1-sNLS-sNLS protein complexed to HBG1-1 gRNA ("His-AsCpf1-sNLS-sNLS_HBG1-1 RNP") either alone, or in combination with 2.5 μM of one of the ssODN donors (OLI16430, OLI164324, OLI16409, or OLI16410). Co-delivery of the RNP and ssODN donor encoding the 18 nt deletion with positive strand homology arms (OLI16409) enhanced the editing frequency from 39.5% without donor to 73.3%, as determined by sequencing analysis of the *HBG* PCR product from genomic DNA extracted at 72 hours post-electroporation (**Fig. 26A**).

[0398] Further analysis of the specific type and size of deletions at the target site revealed that in the presence of the 18 nt positive strand donor (OLI16409), 57.3% of alleles carried the 18 nt deletion compared to 7.5% of alleles when the His-AsCpf1-sNLS-sNLS_HBG1-1 RNP was delivered alone (**Fig.**

26B). The ssODN co-delivery mediated precise repair of the DNA DSB because the 18 nt deletion represented 71.2% of all the indels generated with co-delivery of ssODN OLI16409 and His-AsCpf1-sNLS-sNLS_HBG1-1 RNP, whereas the 18 nt deletion represented only 19.0% of all the indels were generated when delivering His-AsCpf1-sNLS-sNLS_HBG1-1 RNP alone (**Fig. 26C**).

[0399] Although the data in **Figures 26A-26C** originally suggested that codelivery of Cpf1 RNP with ssODN donors increased precise gene editing of genomic DNA resulting in the 18 nt deletion, subsequent testing repeating the experiment in Example 13 confirmed that these data were an artifact of the experiment. Nonetheless, data acquired from the subsequent repeated experiments testing the AsCpf1 RNP using the guide RNA HBG1-1 and ssODN OLI16409 indicated that codelivery of an ssODN donor supports increased total editing in human mPB-CD34+ cells, although not associated with precise repair of the DNA DSB toward the "-18nt deletion" (see **Example 21**).

Example 14: Cpf1 RNP containing gRNA targeting the distal CCAAT box region of the *HBG* promoter supports gene editing in human hematopoietic stem/progenitor cells which promotes induction of HbF protein expression in the erythroid progeny

[0400] Guide RNA HBG1-1 (**Table 14**) having a targeting domain comprising SEQ ID NO:1002 (**Table 15**), targets a site within the *HBG* promoter (**Fig. 27A**). The HBG1-1 gRNA (SEQ ID NO:1022) was complexed to wild-type AsCpf1 (AsCpf1-sNLS-sNLS, SEQ ID NO:1001) to form an RNP ("AsCpf1-HBG1-1-RNP", **Table 14**). This complex (5μM or 20μM) was then electroporated into mobilized peripheral blood (mPB) derived CD34+ cells using either the Amaxa nucleofector (Lonza), or the MaxCyte GT (MaxCyte, Inc.) electroporation device. The level of insertions / deletions at the target site was analyzed by Illumina sequencing (NGS) of the PCR amplified target site three days after electroporation. AsCpf1-HBG1-1-RNP supported on-target editing of 43% and 17%, respectively, on the Amaxa and MaxCyte electroporation systems (**Figs. 28A** (Amaxa) **and 28B** (MaxCyte)). Following the editing of mPB CD34+ cells, *ex vivo* differentiation into the erythroid linage was performed for 18 days (Giarratana 2011). Then, relative expression levels of gamma-globin chains (over total beta-like globin chains) was measured by UPLC (**Figs. 28A** (Amaxa) and **28B** (MaxCyte)). AsCpf1-HBG1-1-RNP led to an increase of gamma-globin expression by up to 21% above levels detected in cells derived from mock electroporated mPB CD34+ cells (**Fig. 28A**).

Table 14: RNP components

gRNA ID	Protein variant	RNP components	RNP ID	gRNA modification	gRNA sequence (RNA)
HBG1-1 gRNA	AsCpf1-sNLS-sNLS (SEQ ID NO:1001)	gRNA + AsCpf1 WT	HBG1-1-AsCpf1-RNP	Synthetic unmodified	UAAUUUCUACUC UUGUAGAUCCUU GUCAAGGCUAUU GGUC (SEQ ID NO:1022)
HBG1-1 gRNA	His-AsCpf1-sNLS-sNLS H800A (SEQ ID NO:1032)	gRNA + AsCpf1 H800A	HBG1-1-AsCpf1H800A-RNP	Synthetic unmodified	UAAUUUCUACUC UUGUAGAUCCUU GUCAAGGCUAUU GGUC (SEQ ID NO:1022)
HBG1-1 gRNA	His-AsCpf1-sNLS-sNLS H800A (SEQ ID NO:1032)	gRNA + AsCpf1 H800A	HBG1-1-AsCpf1H800A-RNP	5'OMe-PS2OMe	UAAUUUCUACUC UUGUAGAUCCUU GUCAAGGCUAUU GGmU/PS2/mC (SEQ ID NO:1041)
tSpA dead gRNA	SpCas9WT (SEQ ID NO:1033)	dead gRNA + SpCas9 WT	tSpA-Cas9-RNP	Synthetic 5'-3' 3xPSOMe	mG*mG*mC*UGG CCAACCCAUGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCmU* mU*mU*U (SEQ ID NO:1024)
SpA gRNA	His-NLS-SpCas9D10A (SEQ ID NO:1034)	gRNA + SpCas9 nickase	SpA-D10A-RNP	Synthetic 5'-3' 3xPSOMe	mG*mG*mC*AAG GCUGGCCAACC CAUGUUUUAGA GCUAGAAAUAG CAAGUUAAAAU AAGGCUAGUCC GUUAUCAACUU GAAAAAGUGGC ACCGAGUCGGU GCmU*mU*mU*U (SEQ ID NO:1025)

SpG gRNA	His-NLS-SpCas9D10A (SEQ ID NO:1034)	gRNA + SpCas9 nickase	SpG-D10A-RNP	Synthetic 5'-3' 3xPSOMe	mU*mA*mG*UCU UAGAGUAUCCA GUGGUUUUAGA GCUAGAAAUAG CAAGUUAAAAU AAGGCUAGUCC GUUAUCAACUU GAAAAAGUGGC ACCGAGUCGGU GCmU*mU*mU*U (SEQ ID NO:1026)
Sp182 dead gRNA	SpCas9WT (SEQ ID NO:1033)	dead gRNA + SpCas9 WT	Sp182-Cas9-RNP	Synthetic 5'-3' 3xPSOMe	mU*mU*mA*GAG UAUCCAGUGGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCmU* mU*mU*U (SEQ ID NO:1027)
"His" refers to a six-histidine purification sequence "AsCpf1" refers to the *Acidaminococcus sp.* Cpf1 protein sequence "nNLS" refers to the nucleoplasmin NLS, and "sNLS" refers to the SV40 NLS "SpCas9" refers to the *S. Pyogenes* Cas9 protein sequence "*" represents phosphorothioate modification "PS" represents phosphorothioate modification "PS2" represents phosphorodithioate modification "OMe" represents a 2-o-methyl modification "m" represents a 2-o-methyl modification					

Table 15: Cpf1 guide RNA targeting sequences

gRNA ID	gRNA Targeting domain sequence (RNA)	gRNA Targeting domain sequence (DNA)
HBG1-1 gRNA	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)	CCTTGTCAAGGCTATTGGTC (SEQ ID NO:1003)
Sp182 dgRNA	UUAGAGUAUCCAGUG (SEQ ID NO:1028)	TTAGAGTATCCAGTG (SEQ ID NO:1029)
SpG gRNA	UAGUCUUAGAGUAUCCAGUG (SEQ ID NO:359)	TAGTCTTAGAGTATCCAGTG (SEQ ID NO:1030)
SpA gRNA	GGCAAGGCUGGCCAACCCAU (SEQ ID NO:941)	GGCAAGGCTGGCCAACCCAT (SEQ ID NO:944)
tSpA dgRNA	GGCUGGCCAACCCAU (SEQ ID NO:326)	GGCTGGCCAACCCAT (SEQ ID NO:1031)

Example 15: Cpf1 RNP editing efficiency at the CCAAT box region of the HBG promoter is enhanced when co-delivered with *S. Pyogenes* Cas9 RNP with no impact on viability

[0401] It was hypothesized that the co-delivery of Cpf1-RNP with proximally targeting *S. Pyogenes* Cas9 RNP, either catalytically inactive or introducing single nicks, could enhance editing levels at the target site. In an attempt to enhance Cpf1-RNP mediated editing at the distal CCAAT box region of the HBG promoter, HBG1-1-AsCpf1H800A-RNP (composed of His-AsCpf1-sNLS-sNLS-H800A (SEQ ID NO:1032) complexed to HBG1-1 gRNA with a 3' modification as shown in **Table 14** (SEQ ID NO:1041) was co-delivered into mPB CD34+ cells with: (1) *S. Pyogenes* Cas9 D10A RNP containing a Cas9 D10A protein (His-NLS-SpCas9D10A, SEQ ID NO:1034) complexed to a full length guide RNA (100mer) selected from: (a) SpA gRNA (**Tables 14**, **15**; **Fig. 27B**) ("SpA-D10A-RNP", **Table 14**) or (b) SpG gRNA (**Tables 14**, **15**; **Fig. 27C**) ("SpG-D10A-RNP", **Table 14**)); (2) *S. Pyogenes* WT Cas9 RNP containing a WT Cas9 protein (SpCas9WT, SEQ ID NO:1033) complexed to a truncated dead guide RNA (95mer - with a shortened protospacer region) selected from (a) tSpA dgRNA (**Tables 14, 15**; **Fig. 27E**) ("tSpA-Cas9-RNP", **Table 14**); (b) Sp182 dgRNA (**Tables 14**, **15**; **Fig. 27F**) ("Sp182-Cas9-RNP", **Table 14**); or (c) tSpA-Cas9-RNP and Sp182-Cas9-RNP (**Tables 14**, **15**, **Fig. 27D**). The RNP complexes were electroporated using the MaxCyte GT device (MaxCyte, Inc). The level of insertions / deletions at the target site was then analyzed by Illumina sequencing (NGS) of the PCR amplified target site three days after electroporation. In all combinations tested, irrespective of the *S. Pyogenes* Cas9 enzyme used (D10A or WT) and of PAM orientation, total editing was increased above levels observed following Maxcyte delivery of HBG1-1-AsCpf1H800A-RNP alone (**Fig. 29** and **Table 16**). In addition there was no detrimental effect on viability of mPB CD34+ cells when *S. Pyogenes* Cas9 RNP were co-delivered (**Fig. 30**).

Table 16: Summary of indels and gamma globin expression with each RNP combination at the maximum RNP dose tested

Cpf1 RNP	Cas9 RNP1	Cas9 RNP2	Indels (NGS)	Gamma Globin (over total beta-like chains)
-	-	-	0%	13.5%
HBG1-1-AsCpf1H800A-RNP	-	-	8.0%	20.0%
HBG1-1-AsCpf1H800A-RNP	SpG-D10A-RNP	-	88.6%	30.2%

HBG1-1-AsCpf1H800A-RNP	SpA-D10A-RNP	-	42.6%	32.5%
HBG1-1-AsCpf1H800A-RNP	tSpA-Cas9-RNP	-	32.2%	19.8%
HBG1-1-AsCpf1H800A-RNP	Sp182-Cas9-RNP	-	85.4%	42.8%
HBG1-1-AsCpf1H800A-RNP	tSpA-Cas9-RNP	Sp182-Cas9-RNP	80.1%	33.7%

Example 16: Cpf1 editing profile can be manipulated with co-delivery of *S. Pyogenes* Cas9-D10A-RNP

[0402] In addition to the increase in total editing observed when HBG1-1-AsCpf1H800A-RNP was co-delivered to mPB CD34+ cells with *S. Pyogenes* Cas9-D10A-RNP (**Fig. 29**), changes to the indel profile were also observed. The introduction of a single strand break proximal to the Cpf1-RNP target site by the D10A enzyme (**Fig. 31A**) altered the directionality, length and/or position of the indels (**Fig. 31B**). For example, Sp182-D10A-RNP strongly shifted the profile toward the nicking site introduced by the D10A RNP, with deletions of various length extending from the Cpf1 cut site toward the upstream nicking site (Sp182-D10A-RNP, **Fig. 31B**). While the SpA-D10A-RNP also promoted the generation of deletions extending from the HBG1-1-AsCpf1H800A-RNP cut site to the nicking site introduced by the D10A enzyme (here located downstream), in this case, additional deletions apparently originating from the nicking site, but not extending fully to the HBG1-1-AsCpf1H800A-RNP cut site, were also produced (SpA-D10A-RNP, **Fig. 31B**).

[0403] It is probable that the orientation, target strand, and distance of the *S. Pyogenes* Cas9-D10A-RNP target site to the Cpf1-RNP target site leads to differences in the position and length of the mutations promoted by the additional DNA nick (**Fig. 31A**). It should be noted that in certain applications this directional manipulation of the indel profile, *i.e.*, an increase in the frequency of indels occurring between the Cpf1-RNP and *S. Pyogenes* Cas9 D10A-RNP binding sites, could be used to favor a desired editing outcome, to increase the rate of productive indels (*e.g.*, indels disrupting a targeted site). In the case set-out within, to disrupt the dCCAAT target region and induce HbF protein expression, the co-delivery of HBG1-1-AsCpf1H800A-RNP with *S. Pyogenes* D10A-RNP led to an increase of gamma globin levels by 16.7% (SpG-D10A-RNP) or 19.0% (SpA-D10A-RNP) above levels detected in mock electroporated cells, as detected by UPLC analysis after 18 days of erythroid differentiation post electroporation (**Fig. 29** and **Table. 16**). The frequency of productive indels was higher when pairing HBG1-1-AsCpf1H800A-

RNP with SpA-D10A-RNP (versus SpG-D10A-RNP) as a higher increase in HbF levels (19% with Cas9D10A-SpA-RNP vs 16.7% with SpG-D10A-RNP) was achieved from a lower editing level (42.6% with SpA-D10A-RNP vs 88.6% with SpG-D10A-RNP) (**Fig. 29** and **Table 16**).

Example 17: Cpf1 editing is increased without altering the editing profile in the presence of *S. Pyogenes* Cas9 WT RNP containing a dead guide RNA

[0404] When co-delivering HBG1-1-AsCpf1H800A-RNP (at a fixed dose of 6 μM) to mPB CD34+ cells with increasing doses of Sp182-Cas9-RNP (complexed with a truncated, 95mer dead gRNA, Sp182), using the MaxCyte electroporator device, an editing boost has been observed up to greater than 10-fold (85.4% editing with the highest dose of Sp182-Cas9-RNP versus 8.0% when HBG1-1-AsCpf1H800A-RNP was delivered alone) (**Fig. 29**). The levels of editing achieved here by co-delivering Sp182-Cas9-RNP also greatly surpassed the levels of editing obtained using the MaxCyte electroporator device when HBG1-1-AsCpf1H800A-RNP was delivered alone at doses as high as 20uM (**Fig. 28B**). In contrast to the results observed by co-delivering a D10A-RNP (in Example 16), the increase in editing achieved by co-delivering HBG1-1-AsCpf1H800A-RNP with Sp182-Cas9-RNP (Cas9-WT complexed with a truncated gRNA) was not associated with an apparent impact on the indel profile. The indels detected in electroporated cells were centered around the Cpf1 cut site in an approximal symmetrical fashion and no indels were detected at the Sp182-Cas9-RNP target site (**Fig. 32A**). Of note, 96% of the total indels were disrupting the distal-CCAAT box, and 79.5% of total indels disrupted 3 or more nucleotides of the CCAAT box motif (**Fig. 32B**). Further dosing optimization using HBG1-1-AsCpf1H800A-RNP (composed of His-AsCpf1-sNLS-sNLS-H800A (SEQ ID NO:1032) complexed to HBG1-1 gRNA (SEQ ID NO:1022)) and Sp182-Cas9-RNP enabled editing levels of up to 92% in mPB-CD34+ cells at 120 hours post-electroporation using the MaxCyte device, and lead to gamma chain expression levels 32% above background in erythroid derived cells (41% of gamma chains in treated cells, 8% in mock treated cells) (**Fig. 33**). Those results demonstrate that an RNP composed of WT *S. Pyogenes* Cas9 protein complexed with a truncated gRNA can increase the editing from a proximally binding Cpf1-RNP without introducing detectable levels of editing at its own binding site nor noticeably affecting the length nor directionality of indels generated by the Cpf1-RNP. Here, the editing enhancement provided by the Sp182-Cas9-RNP enables high editing of the HBG1-1-AsCpf1H800A-RNP at the HBG promoter, with a high frequency of indels disrupting the distal CCAAT box target motif and leading to therapeutically relevant levels of gamma chain expression in the bulk erythroid progeny of the electroporated cells.

Example 18: Clonal HbF distribution within cell population edited with Cpf1 gRNA targeting the distal CCAAT box

[0405] A single cell experiment was next performed to evaluate the distribution of gamma chain expression levels in erythroid cells derived from mPB CD34+ cells electroporated with the HBG1-1-AsCpf1H800A-RNP (composed of His-AsCpf1-sNLS-sNLS-H800A (SEQ ID NO:1032) complexed to HBG1-1 gRNA with a 3' modification as shown in **Table 14** (SEQ ID NO:1041)) + Sp182-Cas9-RNP combination (**Fig. 27F**, **Table 14**). After being left to recover during 48 hours post-electroporation, the cells were sorted by Fluorescence Activated Cell Sorting (FACS) at 1-cell / well in non-tissue culture treated 384-well plates. The cells were then differentiated and expanded clonally for 18 days into the erythroid lineage (adapted from Giarratana 2011). UPLC analysis was performed to determine the distribution of gamma chain expression levels (percentage of gamma chains/[total beta-like chains]) in the clonal erythroid progeny of cells derived from the total population of mPB-CD34+ cells initially edited. It is considered that in order to achieve a functional benefit and alleviate symptoms associated with sickle cell disease, ~30% of the erythroid cells should have fetal hemoglobin levels greater than 30%. Of the clones analyzed, 83.1% had gamma chains levels above the median level detected in erythroid clones derived from mock electroporated mPB CD34+ cells (median value = 18.8%), 64.2% had gamma chains levels exceeding 30% of total beta-like and 35.8% had gamma chains levels exceeding 48.8% of total beta-like (30% +18.8% median level in control cells) (**Fig. 34**).

Example 19: Screen of Cpf1 gRNAs targeting the HBG promoter region

[0406] To identify other AsCpf1 gRNA that could be used as a component of a single RNP or in combination with a "booster element" to increase editing of the HBG promoter region in CD34+ cells and induce fetal globin expression in the erythroid progeny of modified cells, His-AsCpf1-NLS-NLS ("AsCpf1," SEQ ID NO:1000); AsCpf1 S542R/K607R ("AsCpf1 RR," SEQ ID NO:1036); or AsCpf1 S542R/K548V/N552R ("AsCpf1 RVR," SEQ ID NO:1037) gRNA sequences targeting several domains of the HBG promoter (**Table 17**) were designed (listed in **Table 18**). AsCpf1 RR and AsCpf1 RVR are engineered AsCpf1 variants which recognize TYCV/ACCC/CCCC and TATV/RATR PAMs, respectively (Gao 2017).

Table 17: Subdomains of the HBG genomic region

Genomic Coordinate of HbG*	Nucleotides	Name of Region
Chr 11 (NC_000011.10): 5,247,883-5,248,186	TCCTAAAGCT TGGAACACTT TCCCTTCCTT AAGAACCATC CTTGCTACTC AGCTGCAATC 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:1118)	Region 1: Downstream of *HBG1*
Chr 11 (NC_000011.10): 5,248,509 – 5,249,173	ATAAAGATGA ACCCATAGTG AGCTGAGAGC TCCAGCCTGG CCTCCAGATA ACTACACACC 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	Region 2: *HBG1* Intron 2 - A

	AAAGTTCTTT CTACTGGAAT CTAATAAAGA AAAGTCATTT TCCTCATTTC CACCTCTCTT TTCTCAAAGT CAAAATTGTC CATCT (SEQ ID NO:1119)	
Chr 11 (NC_000011.10): 5,249,198 – 5,249,362	CCCTAAAACA TTACCACTGG GTCTCAGCCC AGTTAGTCCT CTGCAGTTTC TTCACCCCCA ACCCCAGTAT CTTCAAACAG CTCACACCCT GCTGTGCTCA GATCAATACT CCGTTGTCTA AGTTGCCTCG AGACTAAAGG CAACAGGGCT GAAACATCTC CTGGA (SEQ ID NO:1120)	Region 3: *HBG1* Intron 2 - B
Chr 11 (NC_000011.10): 5,249,591– 5,249,712	CTGTGAGATT GACAAGAACA GTTTGACAGT CAGAAGGTGC CACAAATCCT GAGAAGCGAC CTGGACTTTT GCCAGGCACA GGGTCCTTCC TTCCCTCCCT TGTCCTGGTC ACCAGAGCCT AC (SEQ ID NO:1121)	Region 4: *HBG1* Intron 1
Chr 11 (NC_000011.10): 5,249,904 – 5,249,927	GCCGCCGGCC CCTGGCCTCA CTGG (SEQ ID NO:1122)	Region 5: *HBG1* -60 nt region from Transcription Start Site (TSS)
Chr 11 (NC_000011.10): 5,249,955 – 5,249,987	CCTTGTCAAG GCTATTGGTC AAGGCAAGGC TGG (SEQ ID NO:1123)	Region 6: *HBG1* -110 nt region from TSS
Chr 11 (NC_000011.10): 5,250,040 – 5,250,075	TGAGATAGTG TGGGGAAGGG GCCCCC AAGAGGATAC (SEQ ID NO:1124)	Region 7: *HBG1* -200 nt region from TSS
Chr 11 (NC_000011.10): 5,250,089 – 5,250,129	TATAGCCTTT GCCTTGTTCC GATTCAGTCA TTCCAGTTTT T (SEQ ID NO:1125)	Region 8: *HBG1* -250 nt region from TSS
Chr 11 (NC_000011.10): 5,250,141 – 5,250,254	TCTTCCCTTT AGCTAGTTTC CTTCTCCCAT CATAGAGGAT ACCAGGACTT CTTTTGTCAG CCGTTTTTTA CCTTCTTGTC TCTAGCTCCA GTGAGGCCTG TAGTTTAAAG CTAA (SEQ ID NO:1126)	Region 9: *HBG1* -333 nt region from TSS
Chr 11 (NC_000011.10): 5,250,464 – 5,250,549	CCACAGTTTC AGCGCAGTAA TAGATTAGTG TTACATAATA TAAGACCTAA TGCTTACCTC AATATCTACT TATCCGTACC TATTTG (SEQ ID NO:1127)	Region 10: *HBG1* -650 nt region from TSS
Chr 11	TATTCAGGTA TGTATGTATA	Region 11: *HBG1* -800 nt region

(NC_000011.10): 5,250,594 – 5,250,735	CACCAGATGA TGTGTATTTA CCACTGGATA AGTGTGTGTG CTGGCTGATG ACCCAGGGTT TTGGCGTAGC TCTTCTATGC TCAGTAAAGA TGATGGTAGA ATGTTCTTTG GCAGGTACTG TG (SEQ ID NO:1128)	from TSS
Chr 11 (NC_000011.10): 5,253,425 – 5,254,121	CAATAAAGAT GAACCCATAG TGAGCTGAGA GCTCCAGCCT GGCCTCCAGA TAACTACACA 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:1129)	Region 12: *HBG2* Intron 2 - A
Chr 11 (NC_000011.10): 5,254,122 – 5,254,306	CCCTAAAACA TTGCCACTGG GTCTCAGCCC AGTTAGTCCT CTGCAGTTTC TTCACTCCCA ACCCCAGTAT CTTCAAACAG CTCACACCCT GCTGTGCTCA GATCAATACT CAGTTGTCTA AGTTGCCTCG AGACTAAAGG CAACAGTGCT GAAACATCTC	Region 13: *HBG2* Intron 2 - B

	CTGGACTCAC CTTGAAGTTC TCAGG (SEQ ID NO:1130)	
Chr 11 (NC_000011.10): 5,254,511 – 5,254,648	AGCCTGTGAG ATTGACAAGA ACAGTTTGAC AGTCAGAAGG TGCCACAAAT CCTGAGAAGC GACCTGGACT TTTGCCAGGC ACAGGGTCCT TCCTTCCCTC CCTTGTCCTG GTCACCAGAG CCTACCTTCC CAGGGTT (SEQ ID NO:1131)	Region 14: *HBG2* Intron 1
Chr 11 (NC_000011.10): 5,254,829 – 5,254,866	CCGCCGGCCC CTGGCCTCAC TGGATACTCT AAGACTAT (SEQ ID NO:1132)	Region 15: *HBG2* -60 nt region from TSS
Chr 11 (NC_000011.10): 5,254,879 – 5,254,909	CCTTGTCAAG GCTATTGGTC AAGGCAAGGC T (SEQ ID NO:1133)	Region 16: *HBG2* -110 nt region from TSS
Chr 11 (NC_000011.10): 5,254,935 5,255,009	CAGGGACCGT TTCAGACAGA TATTTGCATT GAGATAGTGT GGGGAAGGGG CCCCCAAGAG GATACTGCTG CTTAA (SEQ ID NO:1134)	Region 17: *HBG2* -200 nt region from TSS
Chr 11 (NC_000011.10): 5,255,025 – 5,255,053	TTGCCTTGTT CCGATTCAGT CATTCCAAT (SEQ ID NO:1135)	Region 18: *HBG2* -250 nt region from TSS
Chr 11 (NC_000011.10): 5,255,076 – 5,255,179	TTTAGCTAGT TTTCTTCTCC CACCATAGAA GATACCAGGA CTTCTTTTGT CAGCCGTTTT TCACCTTCTT GTCTGTAGCT CCAGTGAGGC CTGTAGTTTA AAGT (SEQ ID NO:1136)	Region 19: *HBG2* -330 nt region from TSS
Chr 11 (NC_000011.10): 5,255,255 – 5,255,292	GGACACGTCT TAGTCTCATT TAGTAAGCAT TGGTTTCC (SEQ ID NO:1137)	Region 20: *HBG2* -500 nt region from TSS
Chr 11 (NC_000011.10): 5,255,518 – 5,255,641	TTTTTTATAT TCAGGTATGT ATGTAGGCAC CCGATGATGT GTATTTATCA CTGGATAAGT GTATGTGCTG GCTGATGACC CAGGGTTTTG GTGTAGCTCT TCTATGCTCG GTAAAGATGA TGGT (SEQ ID NO:1138)	Region 21: *HBG2* -800 nt region from TSS
*NCBI Reference Sequence NC_000011, the coordinates are reported using the One-based coordinate system, "Homo sapiens chromosome 11, GRCh38.p12 Primary Assembly," (Version NC_000011.10).		

Table 18: Cpf1 guide RNAs

gRNA targeting domain sequence ID*	gRNA Targeting domain sequence (RNA)	gRNA Targeting domain sequence (DNA)	Genomic coordinates at HbG1**	Genomic coordinates at HbG2**	% Editing	Strand
AsCpf1 HBG1 Promoter-1	AGACAGAUAU UUGCAUUGAG (SEQ ID NO:1139)	AGACAGATAT TTGCATTGAG (SEQ ID NO:1140)	Chr11:525002 4:5250043	Chr11:52549 48:5254967	5.43	+
AsCpf1 HBG1 Promoter-2	CAUUGAGAUA GUGUGGGGA A (SEQ ID NO:1141)	CATTGAGATA GTGTGGGGAA (SEQ ID NO:1142)	Chr11:525003 7:5250056	Chr11:52549 61:5254980	8.30	+
AsCpf1 HBG1 Promoter-3	UAGCCUUUGC CUUGUUCCGA (SEQ ID NO:1143)	TAGCCTTTGC CTTGTTCCGA (SEQ ID NO:1144)	Chr11:525009 1:5250110	Chr11:52550 19:5255038	0.23	+
AsCpf1 HBG1 Promoter-4	CCUUGUUCCG AUUCAGUCAU (SEQ ID NO:1145)	CCTTGTTCCG ATTCAGTCAT (SEQ ID NO:1146)	Chr11:525010 0:5250119	Chr11:52550 28:5255047	1.15	+
AsCpf1 HBG1 Promoter-5	UCUAAUUUAU UCUUCCCUUU (SEQ ID NO:1147)	TCTAATTTATT CTTCCCTTT (SEQ ID NO:1148)	Chr11:525013 1:5250150	Chr11:52550 59:5255078	0.16	+
AsCpf1 HBG1 Promoter-6	CUUCUCCCAU CAUAGAGGAU (SEQ ID NO:1149)	CTTCTCCCATC ATAGAGGAT (SEQ ID NO:1150)	Chr11:525016 1:5250180		12.73	+
AsCpf1 HBG2 Promoter-7	UUCUCCCACC AUAGAAGAU A (SEQ ID NO:1151)	TTCTCCCACC ATAGAAGATA (SEQ ID NO:1152)		Chr11:52550 90:5255109	8.11	+
AsCpf1 HBG1 Promoter-8	CCACUGGAUA AGUGUGUGU G (SEQ ID NO:1153)	CCACTGGATA AGTGTGTGTG (SEQ ID NO:1154)	Chr11:525063 4:5250653		13.33	+
AsCpf1 HBG1 Promoter-9	GCGUAGCUCU UCUAUGCUCA (SEQ ID NO:1155)	GCGTAGCTCT TCTATGCTCA (SEQ ID NO:1156)	Chr11:525067 7:5250696		13.48	+
AsCpf1 HBG1 Promoter-10	CUGAGCAUAG AAGAGCUACG (SEQ ID NO:1157)	CTGAGCATAG AAGAGCTACG (SEQ ID NO:1158)	Chr11:525067 8:5250697		10.73	-
AsCpf1 HBG2 Promoter-11	UCACUGGAUA AGUGUAUGU G (SEQ ID NO:1159)	TCACTGGATA AGTGTATGTG (SEQ ID NO:1160)		Chr11:52555 65:5255584	0.43	+
AsCpf1 HBG2 Promoter-12	GUGUAGCUCU UCUAUGCUCG	GTGTAGCTCT TCTATGCTCG		Chr11:52556 08:5255627	5.78	+

	(SEQ ID NO:1161)	(SEQ ID NO:1162)				
AsCpf1 HBG2 Promoter-13	CCGAGCAUAG AAGAGCUACA (SEQ ID NO:1163)	CCGAGCATAG AAGAGCTACA (SEQ ID NO:1164)		Chr11:52556 09:5255628	3.24	-
HBG1-1 AsCpf1	CCUUGUCAAG GCUAUUGGUC (SEQ ID NO:1002)	CCTTGTCAAG GCTATTGGTC (SEQ ID NO:1003)	Chr11:524995 5:5249974	Chr11:52548 79:5254898	17.96	+
AsCpf1 RR HBG1 Promoter-1	GACAGAUAUU UGCAUUGAGA (SEQ ID NO:1167)	GACAGATATT TGCATTGAGA (SEQ ID NO:1168)	Chr11:525002 5:5250044	Chr11:52549 49:5254968	8.48	+
AsCpf1 RR HBG1 Promoter-2	ACACUAUCUC AAUGCAAAUA (SEQ ID NO:1169)	ACACTATCTC AATGCAAATA (SEQ ID NO:1170)	Chr11:525003 1:5250050	Chr11:52549 55:5254974	0.09	-
AsCpf1 RR HBG1 Promoter-3	CACACUAUCU CAAUGCAAAU (SEQ ID NO:1171)	CACACTATCT CAATGCAAAT (SEQ ID NO:1172)	Chr11:525003 2:5250051	Chr11:52549 56:5254975	2.10	-
AsCpf1 RR HBG1 Promoter-4	CCACACUAUC UCAAUGCAAA (SEQ ID NO:1173)	CCACACTATC TCAATGCAAA (SEQ ID NO:1174)	Chr11:525003 3:5250052	Chr11:52549 57:5254976	2.52	-
AsCpf1 RR HBG1 Promoter-5	UUCCCCACAC UAUCUCAAUG (SEQ ID NO:1175)	TTCCCCACAC TATCTCAATG (SEQ ID NO:1176)	Chr11:525003 7:5250056	Chr11:52549 61:5254980	0.05	-
AsCpf1 RR HBG1 Promoter-6	GAUUCAGUCA UUCCAGUUUU (SEQ ID NO:1177)	GATTCAGTCA TTCCAGTTTT (SEQ ID NO:1178)	Chr11:525010 9:5250128		0.77	+
AsCpf1 RR HBG1 Promoter-7	AUUCAGUCAU UCCAGUUUUU (SEQ ID NO:1179)	ATTCAGTCAT TCCAGTTTTT (SEQ ID NO:1180)	Chr11:525011 0:5250129		0.24	+
AsCpf1 RR HBG1 Promoter-8	GUCAUUCCAG UUUUUCUCUA (SEQ ID NO:1181)	GTCATTCCAG TTTTTCTCTA (SEQ ID NO:1182)	Chr11:525011 5:5250134		1.00	+
AsCpf1 RR HBG1 Promoter-9	AGUUUUUCUC UAAUUUAUUC (SEQ ID NO:1183)	AGTTTTTCTCT AATTTATTC (SEQ ID NO:1184)	Chr11:525012 3:5250142		0.15	+
AsCpf1 RR HBG1 Promoter-10	GUUUUUCUCU AAUUUAUUCU (SEQ ID NO:1185)	GTTTTTCTCTA ATTTATTCT (SEQ ID NO:1186)	Chr11:525012 4:5250143		0.15	+

AsCpf1 RR HBG2 Promoter-11	GAUUCAGUCA UUCCAAUUUU (SEQ ID NO:1187)	CAAGAGGATA CTGCTGCTTA (SEQ ID NO:1188)		Chr11:52549 89:5255008	***	+
AsCpf1 RR HBG2 Promoter-12	AUUCAGUCAU UCCAAUUUUU (SEQ ID NO:1189)	AAGAGGATAC TGCTGCTTAA (SEQ ID NO:1190)		Chr11:52549 90:5255009	***	+
AsCpf1 RR HBG2 Promoter-13	GUCAUUCCAA UUUUUCUCUA (SEQ ID NO:1191)	GATTCAGTCA TTCCAATTTT (SEQ ID NO:1192)		Chr11:52550 37:5255056	0.25	+
AsCpf1 RR HBG2 Promoter-14	AAUUUUUCUC UAAUUUAUUC (SEQ ID NO:1193)	ATTCAGTCAT TCCAATTTTT (SEQ ID NO:1194)		Chr11:52550 38:5255057	0.13	+
AsCpf1 RR HBG2 Promoter-15	AUUUUUCUCU AAUUUAUUCU (SEQ ID NO:1195)	GTCATTCCAA TTTTTCTCTA (SEQ ID NO:1196)		Chr11:52550 43:5255062	0.32	+
AsCpf1 RR HBG2 Promoter-16	UUCUCCCAUC AUAGAGGAU A (SEQ ID NO:1197)	AATTTTTCTCT AATTTATTC (SEQ ID NO:1198)		Chr11:52550 51:5255070	0.14	+
AsCpf1 RR HBG2 Promoter-17	AUCAUAGAGG AUACCAGGAC (SEQ ID NO:1199)	ATTTTTCTCTA ATTTATTCT (SEQ ID NO:1200)		Chr11:52550 52:5255071	0.10	+
AsCpf1 RR HBG1 Promoter-18	ACCAUAGAAG AUACCAGGAC (SEQ ID NO:1201)	TTCTCCCATC ATAGAGGATA (SEQ ID NO:1202)	Chr11:525016 2:5250181		1.40	+
AsCpf1 RR HBG1 Promoter-19	CAGUACCUGC CAAAGAACAU (SEQ ID NO:1203)	ATCATAGAGG ATACCAGGAC (SEQ ID NO:1204)	Chr11:525016 9:5250188		7.88	+
AsCpf1 RR HBG2 Promoter-20	UAGUAUCUGG UAAAGAGCAU (SEQ ID NO:1205)	ACCATAGAAG ATACCAGGAC (SEQ ID NO:1206)		Chr11:52550 97:5255116	13.03	+
AsCpf1 RR HBG1 Promoter-21	UCAAUGCAAA UAUCUGUCUG (SEQ ID NO:1207)	CAGTACCTGC CAAAGAACAT (SEQ ID NO:1208)	Chr11:525071 4:5250733		13.31	-
AsCpf1 RR HBG2 Promoter-22	CUCUUGGGGG CCCCUUCCCC (SEQ ID NO:1209)	TAGTATCTGG TAAAGAGCAT (SEQ ID NO:1210)		Chr11:52556 45:5255664	4.07	-
AsCpf1 RVR HBG1	GAUUCAGUCA UUCCAAUUUU	TCAATGCAAA TATCTGTCTG	Chr11:525002 3:5250042	Chr11:52549 47:5254966	0.15	-

Promoter-1	(SEQ ID NO:1211)	(SEQ ID NO:1212)				
AsCpf1 RVR HBG1 Promoter-2	AUUCAGUCAU UCCAAUUUUU (SEQ ID NO:1213)	CTCTTGGGGG CCCCTTCCCC (SEQ ID NO:1214)	Chr11:525005 1:5250070	Chr11:52549 75:5254994	1.09	-
AsCpf1 RVR HBG1 Promoter-3	AAAAAAAUU AGCAGUAUCC U (SEQ ID NO:1215)	AAAAAAATTA GCAGTATCCT (SEQ ID NO:1216)	Chr11:525006 9:5250088		***	-
AsCpf1 RVR HBG1 Promoter-4	GCCUUUGCCU UGUUCCGAUU (SEQ ID NO:1217)	GCCTTTGCCTT GTTCCGATT (SEQ ID NO:1218)	Chr11:525009 3:5250112	Chr11:52550 21:5255040	3.96	+
AsCpf1 RVR HBG2 Promoter-5	AAAAAAAUU AAGCAGCAGU A (SEQ ID NO:1219)	AAAAAAATTA AGCAGCAGTA (SEQ ID NO:1220)		Chr11:52549 97:5255016	***	-
AsCpf1 RVR HBG1 Promoter-6	CUCAGUAAAG AUGAUGGUA G (SEQ ID NO:1221)	CTCAGTAAAG ATGATGGTAG (SEQ ID NO:1222)	Chr11:525069 3:5250712		5.32	+
AsCpf1 RVR HBG2 Promoter-7	ACUGGAUAAG UGUAUGUGCU (SEQ ID NO:1223)	ACTGGATAAG TGTATGTGCT (SEQ ID NO:1224)		Chr11:52555 67:5255586	9.78	+
AsCpf1 RVR HBG2 Promoter-8	UGCUGGCUGA UGACCCAGGG (SEQ ID NO:1225)	TGCTGGCTGA TGACCCAGGG (SEQ ID NO:1226)	Chr11:525065 2:5250671	Chr11:52555 83:5255602	0.24	+
AsCpf1 RVR HBG2 Promoter-9	CUCGGUAAAG AUGAUGGUA G (SEQ ID NO:1227)	CTCGGTAAAG ATGATGGTAG (SEQ ID NO:1228)		Chr11:52556 24:5255643	5.75	+
AsCpf1 RVR HBG2 Promoter-10	UGGUAAAGA GCAUUCUACC A (SEQ ID NO:1229)	TGGTAAAGAG CATTCTACCA (SEQ ID NO:1230)		Chr11:52556 38:5255657	8.55	-

* the gRNA ID name provides the particular Cpf1 molecule used in the RNP complex
**NCBI Reference Sequence NC_000011, the coordinates are reported using the One-based coordinate system, "Homo sapiens chromosome 11, GRCh38.p12 Primary Assembly," (Version NC_000011.10).
*** represents gRNAs that were not tested.

[0407] RNPs (5 μM) containing AsCpf1 protein (SEQ ID NO:1000), AsCpf1 RR protein (SEQ ID NO:1036), or AsCpf1 RVR (SEQ ID NO:1037) complexed with single gRNAs comprising gRNA targeting domains from **Table 18** (see gRNA ID name for the particular Cpf1 molecule used) were delivered to mobilized peripheral blood (mPB) CD34+ cells using the Amaxa electroporator device

(Lonza). After 72 hours, genomic DNA was extracted from cells and the level of insertions / deletions at the target site was then analyzed by Illumina sequencing (NGS) of the PCR amplified target site. The percentage of editing (indels= deletions and insertions) for each gRNA is shown in **Table 18** above. In certain embodiments, Cpf1 RNPs comprising one or more of the gRNAs set forth in **Table 18** may be used to target the regions listed in **Table 17** to induce HbF expression and may be co-delivered with a "booster element" to achieve higher editing levels compared to the editing level of the Cpf1 RNP alone.

Example 20: Co-delivery of HBG1-1-Cpf1 RNP targeting the CCAAT box with ssODN supports an increase in gene editing in human hematopoietic stem/progenitor cells

[0408] Having demonstrated that a 100 nt ssODN (50/50 homology) donor template encoding the "4 nt deletion" (HBGΔ-112:-115) co-delivered with Cas9 RNP generated a comparable outcome to the longer 180 nt templates (**Example 10** and **Fig. 23**), a 100 nt ssODN generating the "18 nt deletion" (HBG Δ-104:-121) (*i.e.,* ssODN OLI16431 (SEQ ID NO:1040), **Table 11**) was co-delivered with Cpf1 RNP to further investigate the editing outcome.

[0409] Briefly, human adult mPB CD34+ cells pre-stimulated for two days in medium supplemented with human cytokines were electroporated with RNP comprising the His-AsCpf1-sNLS-sNLS H800A protein (SEQ ID NO:1032, **Table 14**) complexed to modified HBG1-1 gRNA (SEQ ID NO:1041, **Table 14**) ("His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP") either alone at 6 μM or in combination with OLI16431 (0.5 – 6 μM). Co-delivery of the RNP and ssODN donor encoding the 18 nt deletion with positive strand homology arms (OLI16431) enhanced the editing frequency from 8.0% without donor to 75.6%, as determined by sequencing analysis of the *HBG* PCR product from genomic DNA extracted at 72 hours post-electroporation (**Fig. 35A**). This level of editing enabled an increase in HbF levels of ~24% above background (**Fig. 35A**). Additionally, it should be noted that there was no decrease observed in cell viability (**Fig. 35B**) as measured by DAPI exclusion or of colony forming potential (**Fig. 35C**) at any of the doses tested.

Example 21: Optimization of HBG1-1-Cpf1 RNP and OLI16431 dosing to maximize editing at the RNP targeting distal CCAAT box site

[0410] Having demonstrated increased editing when co-delivering ssODN with RNP (Examples 9-11, 13, 20), the same methodology was used to optimize the dosing of each component in order to maximize total editing. Briefly, a dosing matrix was set up with RNP comprising the His-AsCpf1-sNLS-sNLS H800A protein (SEQ ID NO:1032, **Table 14**) complexed to unmodified HBG1-1 gRNA (SEQ ID NO:1022, **Table 14**) ("His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP") being co-delivered at 0 – 12 μM with OLI16431 (SEQ ID NO:1040, **Table 11**) at 0 – 12 μM. Using a dose of 8 μM His-AsCpf1-

sNLS-sNLS H800A_HBG1-1 RNP co-delivered with 8 μM OLI16431, a maximum of 89.4% indels was achieved when measured by sequencing analysis of the *HBG* PCR product from genomic DNA extracted from cells following 14 days erythroid culture (**Fig. 36A**). This coincided with an increase in HbF levels to >32% above background under these dosing conditions (**Fig. 36A**). The viability of this sample (His-AsCpf1-sNLS-sNLS H800A_HBG1-1 RNP [8 μM] + OLI16431 [8 μM]) was comparable to that of the untreated control sample at both 48 hours post electroporation and after 14 days in erythroid culture (**Fig. 36B**).

[0411] However, as eluded to in **Example 13**, an artifact was observed when > 8 μM OLI16431 was electroporated into the cells alone, without RNP. Under these conditions, there was a false positive result when PCR amplification was performed 48 hour post electroporation, likely due to excessive ssODN within the system (see **Fig. 36A**, 8 μM and 12 μM for OLI16431 alone at 48 hours (black bars)). At lower doses of ssODN, and at later timepoints, this false positive was no longer apparent (**Fig. 36A**, 4 μM and 6 μM for OLI16431 alone at 48 hours (black bars) and 4 μM, 6 μM, 8 μM, and 12 μM for OLI16431 alone at 14 days (light grey bars)). In addition, the decrease in cell viability observed with the ssODN alone groups (**Fig. 36B**, dropping to ~57% with 12 μM OLI16431 alone at 48 hours (black bars)) may be a compounding factor. That there was no false positive result at ssODN doses of 4 μM and 6 μM suggests that this artifact may be due to excessive ssODN. These factors, along with the baseline HbF levels observed with the 8 μM and 12 μM OLI16431 alone groups (**Fig. 36A**, grey dots as compared to negative control), provide confidence that the substantial and sustained boost in RNP editing with the addition of ssODN is real and not an artifact.

Example 22: RNPs containing various Cpf1 and gRNA targeting the *HBG* promoter region support gene editing in human hematopoietic stem/progenitor cells which promotes induction of HbF protein expression in the erythroid progeny

[0412] Guide RNAs comprising SEQ ID NOs:1022, 1023, 1041-1093, 1098-1106 (**Table 19**) were complexed to various Cpf1 variant enzymes (SEQ ID NOs:1032, 1094-1097, 1107-1109, **Table 20**) to form various RNP complexes (**Table 21**). The RNPs contained gRNAs with modifications to the 5' end and/or modifications to the 3' end of the gRNA (**Table 19**). Guide RNAs comprising SEQ ID NOs:1022, 1023, 1041-1084, 1098-1106 (**Table 19**) have the same expected cleavage site at the distal CCAAT box target region, the related targeting domains contain the sequences set forth in SEQ ID NO:1002 (HBG1-1), SEQ ID NO:1254 (HBG1-1 – 21mer), SEQ ID NO:1256 (HBG1-1 – 22mer), SEQ ID NO:1258 (HBG1-1 – 23mer), for gRNA comprising a 20 mer, 21 mer, 22 mer, or 23 mer protospacer sequence respectively (**Table 22, Table 23**). In some cases gRNA targeting other positions within the HBG promoter were also tested, including guide RNAs SEQ ID NOs:1085-1096, comprising targeting domains containing the sequences set forth in SEQ ID NOs:1260 (AsCpf1 HBG1 Promoter-1 (21mer)), SEQ ID

NO:1262 (AsCpf1 HBG1 Promoter-2 (21mer)) or SEQ ID NO:1264 (AsCpf1 HBG1 Promoter-6 (21mer)) (**Table 22, Table 23**). **Table 21** provides a listing of each RNP tested in Examples 22-24 and the SEQ ID NO of the gRNA and the SEQ ID NO of the Cpf1 variant that form each RNP complex. Additional information about each gRNA and Cpf1 variant may be found in **Table 19** and **Table 20**, respectively.

[0413] The gRNAs used in Examples 22 and 23 were chemically synthesized. Chemicals for oligonucleotide synthesis were purchased from BioAutomation, Glen Research, Millipore Sigma, Sigma-Aldrich, ChemGenes, and Thermo Fisher Scientific. The solid support used for synthesis was either a Unylinker 2000 Å CPG resin, a 2'-TBDMS rU 2000 Å CPG resin or a 2'-O-methyl adenosine (N-Bz) Icaa 2000 Å CPG resin from ChemGenes. RNA (TBDMS-protected) and DNA phosphoramidites were obtained from Thermo Fisher Scientific. In certain embodiments, phosphorothioates were introduced during a sulfurization step with a solution of DDTT (3-((dimethylamino-methylidene)amino)-3H-1,2,4-dithiazole-3-thione) from Glen Research. Oligonucleotides were synthesized using standard RNA and DNA phosphoramidite chemistry on either a BioAutomation MerMade 12 synthesizer or on a GE Äkta OligoPilot 100 synthesizer. Following synthesis, the oligonucleotides were cleaved from the solid support and deprotected in a two-step process using ammonium hydroxide/methylamine and TEA-3HF. After desalting, the oligonucleotides were purified using reversed-phase chromatography on a preparative HPLC.

[0414] First, the effect of editing using RNPs comprising gRNAs with modifications to the 5' end of the gRNAs was tested. Briefly, RNP complexes (6.0 μM and 12 μM) were delivered to 1 x 10^6 mPB CD34+ cells via MaxCyte electroporation, following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. The level of insertions / deletions at the target site was analyzed by Illumina sequencing (NGS) of the PCR amplified target site at 72 hours after electroporation. Results demonstrate that RNPs comprising gRNAs with no modification (RNP33, also referred to as "HBG1-1-AsCpf1 RNP," **Table 14**) or modifications to the 5' end of the gRNA including the addition of 5 nt RNA (RNP37), 10 nt RNA (RNP38), 25 nt RNA (RNP39), 60 nt RNA (RNP40), 5 nt DNA (RNP41), 10 nt DNA (RNP42), 25 nt DNA (RNP43), and 60 nt DNA (RNP44) (**Table 21**) supports on-target editing (**Fig. 37**).

[0415] Next, the effect of co-delivery of Cpf1 RNP with Sp182 dead RNP (dead gRNA comprising SEQ ID NO:1027 (**Table 14**) complexed with *S. pyogenes* Cas9 (SEQ ID NO:1033)) or ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) was tested. Briefly, a dosing matrix with RNP33 (no 5' or 3' gRNA modification, **Table 21**) complexes at varying concentrations (6 μM, 8 μM, 8 μM, and 12 μM) were co-delivered with Sp182 RNP (8 μM, 8 μM, 6 μM, and 4 μM) or ssODN OLI16431 (8 μM, 8 μM, 6 μM, and 4 μM) to 5.25 x 10^6 mPB CD34+ cells via MaxCyte electroporation, following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. Following electroporation, cells were placed back to culture for a further 48 hours. Then, a fraction of the CD34 cells were split for

gDNA extraction and indel quantification in the bulk cell population. In addition, to investigate the editing in phenotypic progenitors and phenotypic hematopoietic stem cells (HSCs), HSPC subpopulations were characterized (amongst the remainder CD34 cells) by immune-phenotyping (Notta 2011) and separated by fluorescence Activated Cell Sorting (FACS). Immune phenotyping at 48 hours post electroporation was performed by staining cells with antibodies against hCD34, hCD38, hCD45RA, hCD90, and hCD123. Phenotypic HSCs were defined as hCD34bright hCD38 hCD90+ hCD45RA-, and progenitors were defined as hCD34bright CD38+. These two populations were sorted by FACS and DNA was extracted to determine the editing levels in these sub-populations. The level of insertions / deletions at the target site was analyzed by Illumina sequencing (NGS) of the PCR amplified target. Results demonstrate that RNP33 co-delivered with the "booster elements" Sp182 RNP or ssODN OLI16431 support on-target editing **(Figs. 38A-B)** at levels higher than those observed when delivering RNP33 alone **(Fig. 37)**.

[0416] Next, the effect of co-delivery of RNP33 (no 5' or 3' gRNA modification**, Table 21**), RNP43 (+25 DNA 5' gRNA modification, **Table 21**) or RNP34 (1xPS2-OMe + 1xOMe 3' gRNA modifications, **Table 21**) with Sp182 dgRNA (**Table 15**) or ssODN OLI16431 (**Table 11**) was tested. Briefly, RNP33, RNP34, or RNP43 complexes (8 μM) were co-delivered with Sp182 RNP (8 μM) or ssODN OLI16431 (8 μM) to 5.25 x 10^6 mPB CD34+ cells via MaxCyte electroporation, following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. Following electroporation, cells were placed back to culture for a further 48 hours prior to sorting into phenotypic progenitor and phenotypic HSC fractions for indel quantification. The level of insertions / deletions at the target site was analyzed by Illumina sequencing (NGS) of the PCR amplified target site at 48 hours after electroporation. Results demonstrate that RNPs containing gRNAs with 5' modification or 3' modifications co-delivered with the "booster elements" Sp182 RNP or ssODN OLI16431 support on-target editing **(Figs. 39)**. Both booster elements tested provided comparable editing levels to RNP34 and RNP33. When compared to editing levels obtained without the booster element, while RNP33 editing was enhanced by both booster elements, RNP43 editing was however only increased with the addition of Sp182 RNP **(Figs. 37, 39)**.

[0417] Next the effect of different Cpf1 proteins on RNP editing was tested. A stoichiometric comparison (gRNA:Cpf1) with RNPs comprising various Cpf1 proteins was performed. Briefly, RNPs (RNP64, RNP63, RNP45, **Table 21**) were delivered at a stoichiometry (gRNA:Cpf1 complexation ratio) of either 2 or 4, where the gRNA is in a molar excess. All RNPs were delivered via MaxCyte electroporation at 8 μM to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. Following electroporation, cells were placed back to culture for a further 72 hours prior to indel quantification. Results demonstrated that there was no difference in editing rates with altered stoichiometry at the concentration tested **(Fig. 40)**. RNP45 (comprising Cpf1

protein SEQ ID NO:1094) outperforms RNP63 (comprising Cpf1 protein SEQ ID NO:1095) and RNP64 (comprising Cpf1 SEQ ID NO:1109).

[0418] Next, the effect of Sp182 RNP or ssODN OLI16431 co-delivery with various RNPs containing different Cpf1 proteins was tested. Briefly, RNPs (RNP33, RNP64, RNP63, RNP45, **Table 21**) were delivered alone or in combination with Sp182 RNP or ssODN OLI16431. All reagents were delivered via MaxCyte electroporation at 8 μM to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. Following electroporation, cells were placed back to culture for a further 72 hours prior to indel quantification. Results indicate the "booster elements" Sp182 RNP or ssODN OLI16431 enhanced editing for all RNPs tested (**Fig. 41**).

[0419] The effect of editing using RNPs containing gRNAs with various 5' DNA extensions, or RNP without a 5' DNA extension (RNP45) co-delivered with or without ssODN OLI16431 was tested. Briefly, following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3, RNPs (RNP45 (no 5' modification), RNP46, RNP47, RNP48, RNP49, RNP50, RNP51, RNP52, RNP53, RNP54, RNP55, RNP56, and RNP57, **Table 21**) at a concentration of 8 μM were delivered alone or co-delivered with 8 μM ssODN OLI16431 via MaxCyte electroporation to 1 x 10^6 CD34+ cells. Following electroporation, cells were placed back to culture prior to indel quantification. Results demonstrate that all RNPs support on-target editing (**Fig. 42**).

[0420] The effect of editing using RNPs including gRNAs having the same 5' DNA extension but different 3' gRNA modifications was tested to assess the impact of 3' gRNA modifications. Briefly, RNPs comprising gRNAs with matched 5' ends, but different 3' gRNA modifications (RNP49 vs. RNP58 and RNP59 v. RNP60) were delivered at a concentration of 8 μM via MaxCyte electroporation to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. Following electroporation, cells were placed back to culture prior to indel quantification. In both comparisons, RNPs containing gRNA with 3' PS-OMe outperformed the unmodified 3' version at 24h post electroporation (**Fig. 43**).

[0421] Next, different concentrations of Cpf1 and gRNA for RNP58 were tested. Briefly, RNP58 (+25 DNA 5' gRNA modification and 1xPS-OMe 3' gRNA modification, **Table 21**) was delivered via MaxCyte electroporation to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation at a stoichiometry (gRNA:Cpf1 complexation ratio) of either 2:1, 1:1 or 0.5:1 molar ratios. Following electroporation, cells were placed back to culture prior to indel quantification. At all doses tested, editing was best when RNP was complexed at a 2:1 ratio (**Figs. 44A-B**).

[0422] The effect of editing using RNPs including gRNAs having the same 5' DNA extension but different 3' gRNA modifications was further tested to assess the impact of 3' gRNA modifications. Briefly, RNPs comprising gRNAs with matched 5' ends, but different 3' gRNA modifications (RNP58,

RNP2, RNP3, RNP4, RNP5, RNP6, RNP7, RNP8, RNP9, RNP10) were delivered at varying concentrations (1 μM, 2 μM, 4 μM) via MaxCyte electroporation to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation. Following electroporation, cells were placed back to culture prior to indel quantification. Results indicate that all RNPs support on-target editing (**Figs. 45A-B**).

[0423] Next, editing by RNPs that include gRNAs with 3' gRNA modifications or 5' and 3' modifications that target various regions of the HBG promoter were tested. Those include guide RNAs SEQ ID NOs :1085-1096, comprising targeting domains SEQ ID NOs:1260 (AsCpf1 HBG1 Promoter-1 (21mer)), 1262 (AsCpf1 HBG1 Promoter-2 (21mer)), 1264 (AsCpf1 HBG1 Promoter-6 (21mer)) (**Table 22**, **Table 23**). Instead of the distal CCAAT box target region, those gRNAs are configured to provide an editing event within regions selected from **Table 17**. Briefly, RNPs including gRNAs containing an unmodified 5' gRNA and a 1xPS-OMe 3' gRNA modification (RNP11, RNP16, RNP19, and RNP22, **Table 21**), RNPs including gRNAs containing a +20 DNA +2xPS 5' gRNA modification and a 1xPS-OMe 3' gRNA modification (RNP12, RNP21, and RNP24, **Table 21**), RNPs including gRNAs containing a +25 DNA 5' gRNA modification and a 1xPS-OMe 3' gRNA modification (RNP58 and RNP20, **Table 21**) were delivered at a concentration of 8 μM via MaxCyte electroporation to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. Following the editing of mPB CD34+ cells, *ex vivo* differentiation into the erythroid lineage was performed for 18 days (Giarratana 2011), with gDNA being isolated on day 14 of culture for indel analysis. Briefly, RNP was delivered to CD34+ cells, as described above. Following 48 hours recovery in X-Vivo 10 media supplemented with SCF, TPO and FLT3, the treated cells were counted and transferred to erythroid differentiation media, with cell counts and feeds occurring on days 4, 7, 10 and 14, with erythroid collection at day 18. These CD34+ derived erythroid cells were then counted and lysed in HPLC grade water, before filtering to removed cell debris. Then, relative expression levels of gamma-globin chains (over total beta-like globin chains) was measured for each sample by UPLC, with protein separation being achieved by gradually increasing the ratio of acetonitrile with 0.1% trifluoroacetic acid, to water with 0.1% trifluoroacetic acid (**Figs. 46A-C**). **Figs 46A-C** showed all RNP supported on-target editing. However only editing at certain target sites give rise to increased HBF expression. The same experiment was performed with varying concentrations (0, 1 μM, 2 μM, 4 μM) of RNP58 (**Fig. 47**).

[0424] Next the effect of different Cpf1 proteins on RNP editing was tested. Briefly, RNP58, RNP26, RNP27, and RNP28 (**Table 21**) including gRNAs comprising SEQ ID NO:1051 (+25 DNA 5' gRNA modification and a 1xPS-OMe 3' gRNA modification, **Table 19**) complexed with varying Cpf1 proteins were delivered via MaxCyte electroporation at 8 μM to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. Results demonstrated that all RNPs support on-target editing (**Fig. 48**).

[0425] Next the effect of different RNPs containing gRNAs with various 5' gRNA modifications and the same 3' modification was tested. Briefly, RNPs (RNP58 (+25 DNA 5' gRNA modification and a 1xPS-OMe 3' gRNA modification), RNP29 (+25 DNA + 2xPS 5' gRNA modification and a 1xPS-OMe 3' gRNA modification), RNP30 (PolyA RNA + 2xPS 5' gRNA modification and a 1xPS-OMe 3' gRNA modification), and RNP31 (PolyU RNA + 2xPS 5' gRNA modification and a 1xPS-OMe 3' gRNA modification) (**Table 21**)) were delivered via MaxCyte electroporation at 1 μM, 2 μM, or 4 μM to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3 (RNP30 was not tested at 1 μM due to availability of cells). Results demonstrated that all RNPs support on-target editing (**Fig. 49**).

[0426] Next the effect of different Cpf1 proteins on RNP editing was tested. Briefly, RNP58, RNP27, and RNP26 (**Table 21**) including gRNAs comprising SEQ ID NO:1051 (+25 DNA 5' gRNA modification and a 1xPS-OMe 3' gRNA modification, **Table 19**) complexed with varying Cpf1 proteins were delivered via MaxCyte electroporation at 2 μM or 4 μM to 1 x 10^6 CD34+ cells following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. To investigate the editing in bulk CD34, phenotypic progenitors, and phenotypic hematopoietic stem cells (HSCs), HSPC subpopulations were characterized. Thus, following electroporation, cells were placed back to culture for a further 48 hours prior to sorting into progenitor and HSC fractions for indel quantification. The level of insertions / deletions at the target site was analyzed by Illumina sequencing (NGS) of the PCR amplified target site at 48 hours after electroporation. Results demonstrated that all RNPs support on-target editing in bulk CD34, phenotypic progenitors, and phenotypic hematopoietic stem cells (**Fig. 50**).

[0427] Next, the effect of co-delivery of RNP with Sp182 RNP (dead gRNA comprising SEQ ID NO:1027 (**Table 14**) complexed with *S. pyogenes* Cas9 (SEQ ID NO:1033)) or ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) on RNP containing different Cpf1 proteins was tested. Briefly, RNP61, RNP62, RNP34 (**Table 21**) were co-delivered at 8 μM with Sp182 RNP (8 μM) or ssODN OLI16431 (8 μM) to 25 x 10^6 mPB CD34+ cells via MaxCyte electroporation, following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. To investigate the editing in bulk CD34, phenotypic progenitors, and phenotypic hematopoietic stem cells (HSCs), HSPC subpopulations were characterized. Thus, following electroporation, cells were placed back to culture for a further 48 hours prior to sorting into progenitor and HSC fractions for indel quantification. The level of insertions / deletions at the target site was analyzed by Illumina sequencing (NGS) of the PCR amplified target site at 48 hours after electroporation. Results demonstrate that RNP co-delivered with the "booster elements" Sp182 RNP or ssODN OLI16431 support on-target editing (**Fig. 51**). A fraction of these cells were also cryopreserved 24 hours post electroporation to be further characterized in an *in vivo* engraftment model (see **Example 23**).

[0428] Next, the effect of RNP58 and RNP32 editing was tested. Briefly, RNP58 and RNP32 (**Table 21**) were delivered at 2 μM to 6 x 10^6 mPB CD34+ cells via MaxCyte electroporation, following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. To investigate the editing in bulk CD34, phenotypic progenitors, and phenotypic HSCs, HSPC subpopulations were characterized. Thus, following electroporation, cells were placed back to culture for a further 48 hours prior to sorting into progenitor and HSC fractions for indel quantification. The level of insertions / deletions at the target site was analyzed by Illumina sequencing (NGS) of the PCR amplified target site at 48 hours after electroporation. Results demonstrate that RNP58 and RNP32 support on-target editing (**Fig. 52**).

[0429] RNP58 and RNP1 (**Table 21**) were delivered at concentrations of 2 μM to 8 μM to 25 x 10^6 mPB CD34+ cells via MaxCyte electroporation, following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. To investigate the editing in bulk CD34, phenotypic progenitors, and phenotypic hematopoietic stem cells (HSCs), HSPC subpopulations were characterized. Thus, following electroporation, cells were placed back to culture for a further 48 hours prior to sorting into progenitor and HSC fractions for indel quantification. The level of insertions / deletions at the target site was analyzed by Illumina sequencing (NGS) of the PCR amplified target site at 48 hours after electroporation. Results demonstrate that the RNP tested support on-target editing (**Fig. 53**). A fraction of these cells were also cryopreserved 24 hours post electroporation for further characterized in an *in vivo* engraftment model (see **Example 23**).

Example 23: Infusion of edited mPB CD34+ cells into NOD,B6.SCID Il2rγ-/- Kit(W41/W41) mice results in long term engraftment and HbF induction.

[0430] To determine whether delivery of RNP34 and RNP33 (**Table 21**) co-delivered with ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) achieves edits in long term repopulating hematopoietic stem cells, human adult CD34+ 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, mPB CD34+ cells at 62.5 x 10^6/mL were electroporated via MaxCyte electroporation with RNP at a dose of 8 μM and 6 μM ssODN OLI16431 following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. After 24 hours, mCD34+ cells were cryopreserved. Five days later, 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. Bone marrow sub-populations of cells respectively identified as CD15+, CD19+, glycophorin A (GlyA, CD235a+), and lineage-negative CD34+ were isolated by FACS, and DNA was extracted to determine the editing levels in each of these fractions. **Fig. 54** depicts the frequency of

indels, as determined by next generation sequencing, of unfractionated BM, or flow-sorted individual BM sub-populations.

[0431] Next, to determine whether delivery of RNP34 or RNP33 **(Table 21)** co-delivered with Sp182 RNP (dead gRNA comprising SEQ ID NO:1027 **(Table 14)** complexed with *S. pyogenes* Cas9 (SEQ ID NO:1033)) achieves edits in long term repopulating hematopoietic stem cells, human adult CD34+ cells from mobilized peripheral blood (mPB) were infused into nonirradiated NBSGW mice. Briefly, mPB CD34+ cells at 62.5 x 10^6/mL were electroporated via MaxCyte electroporation with RNP at varying doses and varying doses of Sp182 RNP following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. After 24 hours, mCD34+ cells were cryopreserved. Five days later, 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. Bone marrow sub-populations of cells respectively identified as CD15+, CD19+, glycophorin A (GlyA, CD235a+), and lineage-negative CD34+ were isolated by FACS, and DNA was extracted to determine the editing levels in each of these fractions. **Fig. 55A** depicts the frequency of indels, as determined by next generation sequencing, of unfractionated BM, or flow-sorted individual BM sub-populations.

[0432] Lastly, long term HbF induction by BM-derived CD34+ cells was analyzed. An aliquot of BM cells were cultured under erythroid differentiation conditions for 18 days (Giarratana 2011), and evaluated for HbF expression by UPLC. Briefly, unfractionated BM cells extracted from mice 8 weeks after infusion were placed in erythroid culture conditions for 18 days. Cell counts and feeds occurred on days 7, 10 and 14, with erythroid collection at day 18. These bone marrow derived erythroid cells were then counted and lysed in HPLC grade water before filtering to removed cell debris. Then, relative expression levels of gamma-globin chains (over total beta-like globin chains) was measured for each sample by UPLC, with protein separation being achieved by gradually increasing the ratio of acetonitrile with 0.1% trifluoroacetic acid, to water with 0.1% trifluoroacetic acid **(Fig. 54B)**. These data demonstrate that robust long-term HbF induction is achieved by RNP34 and RNP33 co-delivered with Sp182 RNP editing of human CD34+ cells.

[0433] Next, to determine whether delivery of RNP61 or RNP62 **(Table 21)** co-delivered with ssODN OLI16431 (SEQ ID NO:1040, **Table 11**) achieves edits in long term repopulating hematopoietic stem cells, human adult CD34+ cells from mPB were infused into nonirradiated NBSGW mice. Briefly, mPB CD34+ cells at 62.5 x 10^6/mL were electroporated via MaxCyte electroporation with 8 μM RNP and 8 μM ssODN OLI16431 following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. After 24 hours, mCD34+ cells were cryopreserved. Four days later, 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. Bone marrow sub-populations of cells respectively identified as CD15+, CD19+, glycophorin A (GlyA, CD235a+), and lineage-negative CD34+ were isolated by FACS, and DNA was extracted to determine the editing levels in each of these fractions. **Fig. 56A** depicts the frequency of indels, as determined by next generation sequencing, of unfractionated BM, or flow-sorted individual BM sub-populations.

[0434] Lastly, long term HbF induction by BM-derived CD235a+ (GlyA+) erythroid cells was analyzed. An aliquot of BM cells were cultured under erythroid differentiation conditions for 18 days and evaluated for HbF expression by UPLC. Briefly, unfractionated BM cells extracted from mice 8 weeks after infusion were placed in erythroid culture conditions for 18 days. **Fig. 56B** depicts the HbF expression, calculated as gamma/beta-like chains (%) by erythroid cells. These data demonstrate that robust long-term HbF induction is achieved by RNP61 and RNP62 co-delivered with ssODN OLI16431 editing of human CD34+ cells.

[0435] Next, to determine whether delivery of RNP1 or RNP58 **(Table 21)** achieves edits in long term repopulating hematopoietic stem cells, human adult CD34+ cells from mPB were infused into nonirradiated NBSGW mice. Briefly, mPB CD34+ cells at 62.5 x 10^6/mL were electroporated via MaxCyte electroporation with 4 μM or 8 μM RNP1 or 2 μM, 4 μM, or 8 μM RNP58 following 48 hours pre-stimulation in X-Vivo 10 media supplemented with SCF, TPO and FLT3. After 24 hours, mCD34+ cells were cryopreserved. Four days later, 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+, CD14+, CD19+, glycophorin A (GlyA, CD235a+), lineage, and CD34+, and mouse CD45+ marker expression) in BM was determined by flow cytometry and analyzed (**Fig. 57**). The frequency of GlyA+ cells was calculated as GlyA+ cells/total cells in BM. Human chimerism was defined as human CD45/(human CD45+mCD45).

[0436] Similar chimerism and lineage distributions were achieved 8-weeks post-transplant by RNP-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. 58** depicts the indels, as determined by next generation sequencing, of unfractionated BM, at 8 weeks post-infusion. **Fig. 59** shows the frequency of indel rate in BM, CD15+, CD19+, glycophorin A (GlyA, CD235a+), and lin-CD34+ cells with 2 μM, 4 μM, or 8 μM of RNP58.

[0437] Long term HbF induction by CD235a+ (GlyA+) erythroid cells, derived from edited CD34+ cells was also analyzed. GlyA+ cells obtained from bone marrow, were isolated by fluorescence Activated Cell Sorting (FACS), collected and lysed in HPLC grade water. Lysates were then evaluated for HbF

expression by UPLC. **Fig.60** depicts the HbF expression, calculated as gamma/beta-like chains (%) by GlyA+ cells. These data demonstrate that robust long-term HbF induction is achieved RNP58 editing of human CD34+ cells at various RNP58 concentrations.

[0438] Importantly, $CD34^+$ cells that were electroporated with varying concentrations of RNP1 or RNP58 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 $CD34^+$ cell negative control), as determined in hematopoietic colony forming cell (CFC) assays (**Fig. 61**).

Example 24: Treatment of β-hemoglobinopathy using edited hematopoietic stem cells.

[0439] 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.

[0440] 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 (gRNA) that target one or more regions in the *HBG* gene complexed with an RNA-guided nuclease may be used to edit the HSCs. In certain embodiments, the RNA-guided nuclease may be a Cpf1 protein. In certain embodiments, the Cpf1 protein may be a modified Cpf1 protein. In certain embodiments, the modified Cpf1 protein may be encoded by a sequence set forth in SEQ ID NOs:1000, 1001, 1008-1018, 1032, 1035-39, 1094-1097, 1107-09 (Cpf1 polypeptide sequences) or SEQ ID NOs:1019-1021, 1110-17 (Cpf1 polynucleotide sequences). For example, the modified Cpf1 protein may be encoded by the sequence set forth in SEQ ID NO:1097. In certain embodiments, the gRNA may be a modified or unmodified gRNA. In certain embodiments, the gRNA may comprise a sequence set forth in **Table 13**, **Table 18**, or **Table 19**. For example, in certain embodiments, the gRNA may comprise the sequence set forth in SEQ ID NO:1051. In certain embodiments, the RNP complex may comprise an RNP complex set forth in **Table 21**. For example, the RNP complex may include a gRNA comprising the sequence set forth in SEQ ID NO:1051 and a modified Cpf1 protein encoded by the sequence set forth in SEQ ID NO:1097 (RNP32, **Table 21**). 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.

Table 20: Cpf1 Protein Variants

Cpf1 Variant ID	Cpf1 Variant Amino Acid SEQ ID NO	Cpf1 Variant Nucleotide SEQ ID NO
His-AsCpf1-sNLS-sNLS H800A	SEQ ID NO:1032	SEQ ID NO:1110
Cpf1-1	SEQ ID NO:1094	SEQ ID NO:1111
Cpf1-2	SEQ ID NO:1095	SEQ ID NO:1112
Cpf1-3	SEQ ID NO:1096	SEQ ID NO:1113
Cpf1-4	SEQ ID NO:1097	SEQ ID NO:1114
Cpf1-5	SEQ ID NO:1107	SEQ ID NO:1115
Cpf1-6	SEQ ID NO:1108	SEQ ID NO:1116
Cpf1-7	SEQ ID NO:1109	SEQ ID NO:1117

Table 21: RNP complexes

RNP Name	gRNA SEQ ID NO*	Cpf1 Amino Acid SEQ ID NO**
RNP64	SEQ ID NO: 1022	SEQ ID NO: 1109
RNP1	SEQ ID NO: 1051	SEQ ID NO: 1095
RNP2	SEQ ID NO: 1098	SEQ ID NO: 1094
RNP3	SEQ ID NO: 1099	SEQ ID NO: 1094
RNP4	SEQ ID NO: 1100	SEQ ID NO: 1094
RNP5	SEQ ID NO: 1101	SEQ ID NO: 1094
RNP6	SEQ ID NO: 1068	SEQ ID NO: 1094
RNP7	SEQ ID NO: 1102	SEQ ID NO: 1094
RNP8	SEQ ID NO: 1081	SEQ ID NO: 1094
RNP9	SEQ ID NO: 1082	SEQ ID NO: 1094
RNP10	SEQ ID NO: 1083	SEQ ID NO: 1094
RNP11	SEQ ID NO: 1069	SEQ ID NO: 1094
RNP12	SEQ ID NO: 1084	SEQ ID NO: 1094
RNP16	SEQ ID NO: 1085	SEQ ID NO: 1094
RNP19	SEQ ID NO: 1088	SEQ ID NO: 1094
RNP20	SEQ ID NO: 1089	SEQ ID NO: 1094
RNP21	SEQ ID NO: 1090	SEQ ID NO: 1094
RNP22	SEQ ID NO: 1091	SEQ ID NO: 1094
RNP23	SEQ ID NO: 1048	SEQ ID NO: 1097

RNP24	SEQ ID NO: 1093	SEQ ID NO: 1094
RNP26	SEQ ID NO: 1051	SEQ ID NO: 1096
RNP27	SEQ ID NO: 1051	SEQ ID NO: 1107
RNP28	SEQ ID NO: 1051	SEQ ID NO: 1108
RNP29	SEQ ID NO: 1103	SEQ ID NO: 1094
RNP30	SEQ ID NO: 1104	SEQ ID NO: 1094
RNP31	SEQ ID NO: 1105	SEQ ID NO: 1094
RNP32	SEQ ID NO: 1051	SEQ ID NO: 1097
RNP33	SEQ ID NO: 1022	SEQ ID NO: 1032
RNP34	SEQ ID NO: 1041	SEQ ID NO: 1032
RNP37	SEQ ID NO: 1042	SEQ ID NO: 1032
RNP38	SEQ ID NO: 1043	SEQ ID NO: 1032
RNP39	SEQ ID NO: 1044	SEQ ID NO: 1032
RNP40	SEQ ID NO: 1045	SEQ ID NO: 1032
RNP41	SEQ ID NO: 1046	SEQ ID NO: 1032
RNP42	SEQ ID NO: 1047	SEQ ID NO: 1032
RNP43	SEQ ID NO: 1048	SEQ ID NO: 1032
RNP44	SEQ ID NO: 1049	SEQ ID NO: 1032
RNP45	SEQ ID NO: 1022	SEQ ID NO: 1094
RNP46	SEQ ID NO: 1046	SEQ ID NO: 1094
RNP47	SEQ ID NO: 1047	SEQ ID NO: 1094
RNP48	SEQ ID NO: 1052	SEQ ID NO: 1094
RNP49	SEQ ID NO: 1048	SEQ ID NO: 1094
RNP50	SEQ ID NO: 1053	SEQ ID NO: 1094
RNP51	SEQ ID NO: 1054	SEQ ID NO: 1094
RNP52	SEQ ID NO: 1055	SEQ ID NO: 1094
RNP53	SEQ ID NO: 1056	SEQ ID NO: 1094
RNP54	SEQ ID NO: 1057	SEQ ID NO: 1094
RNP55	SEQ ID NO: 1058	SEQ ID NO: 1094
RNP56	SEQ ID NO: 1059	SEQ ID NO: 1094
RNP57	SEQ ID NO: 1060	SEQ ID NO: 1094
RNP58	SEQ ID NO: 1051	SEQ ID NO: 1094
RNP59	SEQ ID NO: 1067	SEQ ID NO: 1094
RNP60	SEQ ID NO: 1068	SEQ ID NO: 1094
RNP61	SEQ ID NO: 1041	SEQ ID NO: 1095
RNP62	SEQ ID NO: 1041	SEQ ID NO: 1094
RNP63	SEQ ID NO: 1022	SEQ ID NO: 1095
*See Table 19 **See Table 20		

Table 22: Cpf1 HBG1 targeting domains and expected cleavage sites

gRNA ID	gRNA Targeting	gRNA Targeting	Targeting Domain	Expected cleavage site	Strand	Targeting

	Domain (RNA)	Domain (DNA)	coordinates at HBG1 *	coordinates at HBG1#		Domain Length
HBG1-1	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)	CCTTGTCAAGGCTATTGGTC (SEQ ID NO:1003)	Chr11:5249954–5249974	Chr11:5249973, Chr11:5249977	+	20
HBG1-1 (21mer)	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)	CCTTGTCAAGGCTATTGGTCA (SEQ ID NO:1255)	Chr11:5249954–Chr11:5249975	Chr11:5249973, Chr11:5249977	+	21
HBG1-1 (22mer)	CCUUGUCAAGGCUAUUGGUCAA (SEQ ID NO:1256)	CCTTGTCAAGGCTATTGGTCAA (SEQ ID NO:1257)	Chr11:5249954–Chr11:5249976	Chr11:5249973, Chr11:5249977	+	22
HBG1-1 (23mer)	CCUUGUCAAGGCUAUUGGUCAAG (SEQ ID NO:1258)	CCTTGTCAAGGCTATTGGTCAAG (SEQ ID NO:1259)	Chr11:5249954–Chr11:5249976	Chr11:5249973, Chr11:5249977	+	23
AsCpf1 HBG1 Promoter-1	AGACAGAUAUUUGCAUUGAG (SEQ ID NO:1139)	AGACAGATATTTGCATTGAG (SEQ ID NO:1140)	Chr11:5250023-Chr11:5250043	Chr11:5250042, Chr11:5250046	+	20
AsCpf1 HBG1 Promoter-1 (21mer)	AGACAGAUAUUUGCAUUGAGA (SEQ ID NO:1260)	AGACAGATATTTGCATTGAGA (SEQ ID NO:1261)	Chr11:5250023-Chr11:5250044	Chr11:5250042, Chr11:5250046	+	21
AsCpf1 HBG1 Promoter-2	CAUUGAGAUAGUGUGGGGAA (SEQ ID NO:1141)	CATTGAGATAGTGTGGGGAA (SEQ ID NO:1142)	Chr11:5250036-Chr11:5250056	Chr11:5250055, Chr11:5250059	+	20
AsCpf1 HBG1 Promoter-2 (21mer)	CAUUGAGAUAGUGUGGGGAAG (SEQ ID NO:1262)	CATTGAGATAGTGTGGGGAAG (SEQ ID NO:1263)	Chr11:5250036-Chr11:5250057	Chr11:5250055, Chr11:5250059	+	21
AsCpf1 HBG1 Promoter-6	CUUCUCCCAUCAUAGAGGAU (SEQ ID NO:1149)	CTTCTCCCATCATAGAGGAT (SEQ ID NO:1150)	Chr11:5250160-Chr11:5250180	Chr11:5250179, Chr11:5250183	+	20
AsCpf1 HBG1 Promoter-6 (21mer)	CUUCUCCCAUCAUAGAGGAUA (SEQ ID NO:1264)	CTTCTCCCATCATAGAGGATA (SEQ ID NO:1265)	Chr11:5250160-Chr11:5250181	Chr11:5250179, Chr11:5250183	+	21
*NCBI Reference Sequence NC_000011, the coordinates are reported using the One-based coordinate system, "Homo sapiens chromosome 11, GRCh38.p12 Primary Assembly," (Version NC_000011.10).						

[#] Expected cleavage sites based on Zetsche et al, 2015, coordinates are reported using zero-based coordinates.

Table 23: Cpf1 HBG2 targeting domains and expected cleavage sites

gRNA ID	gRNA Targeting Domain (RNA)	gRNA Targeting Domain (DNA)	Targeting Domain coordinates at HBG2 *	Expected cleavage site coordinates at HBG2[#]	Strand	Targeting Domain Length
HBG1-1	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)	CCTTGTCAAGGCTATTGGTC (SEQ ID NO:1003)	Chr11:5254878 Chr11:5254898	Chr11:5254897, Chr11:5254901	+	20
HBG1-1 21mer	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)	CCTTGTCAAGGCTATTGGTCA (SEQ ID NO:1255)	Chr11:5254878 Chr11:5254899	Chr11:5254897, Chr11:5254901	+	21
HBG1-1 22mer	CCUUGUCAAGGCUAUUGGUCAA (SEQ ID NO:1256)	CCTTGTCAAGGCTATTGGTCAA (SEQ ID NO:1257)	Chr11:5254878 Chr11:5254900	Chr11:5254897, Chr11:5254901	+	22
HBG1-1 23mer	CCUUGUCAAGGCUAUUGGUCAAG (SEQ ID NO:1258)	CCTTGTCAAGGCTATTGGTCAAG (SEQ ID NO:1259)	Chr11:5254878 Chr11:5254901	Chr11:5254897, Chr11:5254901	+	23
AsCpf1 HBG1 Promoter-1	AGACAGAUAUUUGCAUUGAG (SEQ ID NO:1139)	AGACAGATATTTGCATTGAG (SEQ ID NO:1140)	Chr11:5254947 Chr11:5254967	Chr11:5254966 Chr11:5254970	+	20
AsCpf1 HBG1 Promoter-1 (21mer)	AGACAGAUAUUUGCAUUGAGA (SEQ ID NO:1260)	AGACAGATATTTGCATTGAGA (SEQ ID NO:1261)	Chr11:5254947 Chr11:5254968	Chr11:5254966 Chr11:5254970	+	21
AsCpf1 HBG1 Promoter-2	CAUUGAGAUAGUGUGGGGAA (SEQ ID NO:1141)	CATTGAGATAGTGTGGGGAA (SEQ ID NO:1142)	Chr11:5254960 Chr11:5254980	Chr11:5254979 Chr11:5254983	+	20
AsCpf1 HBG1 Promoter-2 (21mer)	CAUUGAGAUAGUGUGGGGAAG (SEQ ID NO:1262)	CATTGAGATAGTGTGGGGAAG (SEQ ID NO:1263)	Chr11:5254960 Chr11:5254981	Chr11:5254979 Chr11:5254983	+	21
*NCBI Reference Sequence NC_000011, the coordinates are reported using the One-based coordinate system, "Homo sapiens chromosome 11, GRCh38.p12 Primary Assembly," (Version NC_000011.10). [#] Expected cleavage sites based on (Zetsche et al, 2015), coordinates are reported using zero-based						

coordinates.

Table 24: gRNA 5' Extensions

5' extension Sequence ID No:	5' extension sequence	5' modification
1231	rCrUrUrUrU	+5 RNA
1232	rArArGrArCrCrUrUrUrU	+10 RNA
1233	rArUrGrUrGrUrUrUrUrUrGrUrCrArArArArGrArCrCrUrUrUrU	+25 RNA
1234	rArGrGrCrCrArGrCrUrUrGrCrCrGrGrUrUrUrUrUrUrArGrUrCrGrUrGrCrUrGrCrUrUrCrArUrGrUrGrUrUrUrUrUrGrUrCrArArArArGrArCrCrUrUrUrU	+60 RNA
1235	CTTTT	+5 DNA
1236	AAGACCTTTT	+10 DNA
1237	ATGTGTTTTTGTCAAAAGACCTTTT	+25 DNA
1238	AGGCCAGCTTGCCGGTTTTTTAGTCGTGCTGCTTCATGTGTTTTTGTCAAAAGACCTTTT	+60 DNA
1239	TTTTTGTCAAAAGACCTTTT	+20 DNA
1240	GCTTCATGTGTTTTTGTCAAAAGACCTTTT	+30 DNA
1241	GCCGGTTTTTTAGTCGTGCTGCTTCATGTGTTTTTGTCAAAAGACCTTTT	+50 DNA
1242	TAGTCGTGCTGCTTCATGTGTTTTTGTCAAAAGACCTTTT	+40 DNA
1243	C*C*GAAGTTTTCTTCGGTTTT	+20 DNA + 2xPS
1244	T*T*TTTCCGAAGTTTTCTTCGGTTTT	+25 DNA + 2xPS
1245	A*A*CGCTTTTTCCGAAGTTTTCTTCGGTTTT	+30 DNA + 2xPS
1246	G*C*GTTGTTTTCAACGCTTTTTCCGAAGTTTTCTTCGGTTTT	+41 DNA + 2xPS
1247	G*G*CTTCTTTTGAAGCCTTTTTGCGTTGTTTTCAACGCTTTTTCCGAAGTTTTCTTCGGTTTT	+62 DNA + 2xPS
1248	A*T*GTGTTTTTGTCAAAAGACCTTTT	+25 DNA + 2xPS
1249	AAAAAAAAAAAAAAAAAAAAAAAAA	+25 A
1250	TTTTTTTTTTTTTTTTTTTTTTTTT	+25 T
1251	mA*mU*rGrUrGrUrUrUrUrGrUrCrArArArArGrArCrCrUrUrUrU	+25 RNA + 2xPS
1252	mA*mA*rArArArArArArArArArArArArArArArArArArArArArArArA	PolyA RNA + 2xPS
1253	mU*mU*rUrUrUrUrUrUrUrUrUrUrUrUrUrUrUrUrUrUrUrUrUrU	PolyU RNA + 2xPS

All bases are in upper case
Lowercase "r" represents RNA, 2'-hydroxy; bases not modified by an "r" are DNA
All bases are linked via standard phosphodiester bonds except as noted:
"*" represents phosphorothioate modification
"PS" represents phosphorothioate modification

Table 19: Cpf1 guide RNAs

gRNA Sequence SEQ ID NO.	gRNA Sequence	5' mod.**	3' mod.	Length of crRNA + gRNA targeting domain	Length of gRNA targeting domain	gRNA Targeting Domain (RNA)
1022	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	-	-	40	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1023	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUmC/PS2/mA	-	1xPS2-OMe + 1xOMe	41	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1041	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGmU/PS2/mC	-	1xPS2-OMe + 1xOMe	40	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1042	rCrUrUrUrUrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+5 RNA	-	45	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1043	rArArGrArCrCrUrUrUrUrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+10 RNA	-	50	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1044	rArUrGrUrGrUrUrUrUrUrGrUrCrArArArArGrArCrCrUrUrUrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrU	+25 RNA	-	65	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)

	rCrArArGrGrCrUrAr UrUrGrGrUrC					
1045	rArGrGrCrCrArGrCr UrUrGrCrCrGrGrUrU rUrUrUrUrArGrUrCr GrUrGrCrUrGrCrUrU rCrArUrGrUrGrUrUr UrUrUrGrUrCrArArA rArGrArCrCrUrUrUr UrUrArArUrUrUrCrU rArCrUrCrUrUrGrUr ArGrArUrCrCrUrUrG rUrCrArArGrGrCrUr ArUrUrGrGrUrC	+60 RNA	-	100	20	CCUUGUCAA GGCUAUUGG UC (SEQ ID NO:1002)
1046	CTTTTrUrArArUrUr UrCrUrArCrUrCrUrU rGrUrArGrArUrCrCr UrUrGrUrCrArArGrG rCrUrArUrUrGrGrUr C	+5 DNA	-	45	20	CCUUGUCAA GGCUAUUGG UC (SEQ ID NO:1002)
1047	AAGACCTTTTrUrAr ArUrUrUrCrUrArCrU rCrUrUrGrUrArGrAr UrCrCrUrUrGrUrCrA rArGrGrCrUrArUrUr GrGrUrC	+10 DNA	-	50	20	CCUUGUCAA GGCUAUUGG UC (SEQ ID NO:1002)
1048	ATGTGTTTTTGTCA AAAGACCTTTTrUr ArArUrUrUrCrUrArC rUrCrUrUrGrUrArGr ArUrCrCrUrUrGrUrC rArArGrGrCrUrArUr UrGrGrUrC	+25 DNA	-	65	20	CCUUGUCAA GGCUAUUGG UC (SEQ ID NO:1002)
1049	AGGCCAGCTTGCC GGTTTTTTAGTCGT GCTGCTTCATGTG TTTTTGTCAAAAG ACCTTTTrUrArArUr UrUrCrUrArCrUrCrU rUrGrUrArGrArUrCr CrUrUrGrUrCrArArG rGrCrUrArUrUrGrGr UrC	+60 DNA	-	100	20	CCUUGUCAA GGCUAUUGG UC (SEQ ID NO:1002)

1050	ATGTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUmC/PS2/mA	+25 DNA	1xPS2-OMe + 1xOMe	66	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1051	ATGTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC*mA	+25 DNA	1xPS-OMe	66	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1052	TTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+20 DNA	-	60	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1053	GCTTCATGTGTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+30 DNA	-	70	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1054	GCCGGTTTTTTAGTCGTGCTGCTTCATGTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+50 DNA	-	90	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1055	TAGTCGTGCTGCTTCATGTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrAr	+40 DNA	-	80	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)

	UrUrGrGrUrC					
1056	C*C*GAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+20 DNA + 2xPS	-	60	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1057	T*T*TTTCCGAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+25 DNA + 2xPS	-	65	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1058	A*A*CGCTTTTTCCGAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+30 DNA + 2xPS	-	70	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1059	G*C*GTTGTTTTCAACGCTTTTTCCGAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+41 DNA + 2xPS	-	81	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1060	G*G*CTTCTTTTGAAGCCTTTTTGCGTTGTTTTCAACGCTTTTTCCGAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+62 DNA + 2xPS	-	102	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1061	mUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGr	1xOMe	-	40	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID

	UrCrArArGrGrCrUrArUrUrGrGrUrC					NO:1002)
1062	mU*rArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	1xPS-OMe	-	40	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1063	mUmArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	2xOMe	-	40	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1064	mU*mA*rArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	2xPS-OMe	-	40	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1065	mUmAmArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	3xOMe	-	40	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1066	mU*mA*mA*rUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	3xPS-OMe	-	40	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1067	A*T*GTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC	+25 DNA + 2xPS	-	65	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1068	A*T*GTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC*mA	+25 DNA + 2xPS	1xPS-OMe	66	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)

1069	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC*mA	-	1xPS-OMe	41	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1070	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCmA*rU	-	1xPS-OMe + rU	42	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1071	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrU	-	rU	41	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1072	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrArU	-	rU	42	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1073	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrArArGrU	-	rU	44	23	CCUUGUCAAGGCUAUUGGUCAAG (SEQ ID NO:1258)
1074	A*T*GTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrU	+25 DNA + 2xPS	rU	66	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1075	A*T*GTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrArArGrU	+25 DNA + 2xPS	rU	69	23	CCUUGUCAAGGCUAUUGGUCAAG (SEQ ID NO:1258)

1076	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrA	-	-	41	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1077	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrArArG	-	-	43	23	CCUUGUCAAGGCUAUUGGUCAAG (SEQ ID NO:1258)
1078	A*T*GTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrArArG	+25 DNA + 2xPS	-	68	23	CCUUGUCAAGGCUAUUGGUCAAG (SEQ ID NO:1258)
1079	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrUrUrUrU	-	4xrU	44	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1080	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCmU*mU*mU*rU	-	3xPS-OMe-U + rU	44	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1081	A*T*GTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrCrUrUrUrU	+25 DNA + 2xPS	4xrU	69	20	CCUUGUCAAGGCUAUUGGUC (SEQ ID NO:1002)
1082	AAAAAAAAAAAAAAAAAAAAArUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUr	+25 A	1xPS-OMe	66	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)

	ArUrUrGrGrUrC*mA					
1083	TTTTTTTTTTTTTTTTTTTTTTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC*mA	+25 T	1xPS-OMe	66	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1084	C*C*GAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUrC*mA	+20 DNA + 2xPS	1xPS-OMe	61	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1085	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrArGrArCrArGrArUrArUrUrUrGrCrArUrUrGrArG*mA	-	1xPS-OMe	41	21	AGACAGAUAUUUGCAUUGAGA (SEQ ID NO:1260)
1086	ATGTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrArGrArCrArGrArUrArUrUrUrGrCrArUrUrGrArG*mA	+25 DNA	1xPS-OMe	66	21	AGACAGAUAUUUGCAUUGAGA (SEQ ID NO:1260)
1087	C*C*GAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrArGrArCrArGrArUrArUrUrUrGrCrArUrUGrArG*mA	+20 DNA + 2xPS	1xPS-OMe	61	21	AGACAGAUAUUUGCAUUGAGA (SEQ ID NO:1260)
1088	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrArUrUrGrArGrArUrArGrUrGrUrGrGrArA*mG	-	1xPS-OMe	41	21	CAUUGAGAUAGUGUGGGGAAG (SEQ ID NO:1262)
1089	ATGTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGr	+25 DNA	1xPS-OMe	66	21	CAUUGAGAUAGUGUGGGGAAG (SEQ ID

	ArUrCrArUrUrGrArGrArUrArGrUrGrUrGrGrGrArA*mG					NO:1262)
1090	C*C*GAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrArUrUrGrArGrArUrArGrUrGrUrGrUrGrGrGrArA*mG	+20 DNA + 2xPS	1xPS-OMe	61	21	CAUUGAGAUAGUGUGGGGAAG (SEQ ID NO:1262)
1091	rUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrUrUrCrUrCrCrCrArUrCrArUrArGrArGrGrArU*mA	-	1xPS-OMe	41	21	CUUCUCCCAUCAUAGAGGAUA (SEQ ID NO:1264)
1092	ATGTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrUrUrCrUrCrCrCrArUrCrArUrArGrArGrGrArU*mA	+25 DNA	1xPS-OMe	66	21	CUUCUCCCAUCAUAGAGGAUA (SEQ ID NO:1264)
1093	C*C*GAAGTTTTCTTCGGTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrUrUrCrUrCrCrCrArUrCrArUrArGrArGrGrArU*mA	+20 DNA + 2xPS	1xPS-OMe	61	21	CUUCUCCCAUCAUAGAGGAUA (SEQ ID NO:1264)
1098	A*T*GTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUmA*rU	+25 DNA + 2xPS	1xPS-OMe + rU	67	21	CCUUGUCAAGGCUAUUGGUCA (SEQ ID NO:1254)
1099	A*T*GTGTTTTTGTCAAAAGACCTTTTrUrArArUrUrUrCrUrArCrUrCrUrUrGrUrArGrArUrCrCrUrUrGrUrCrArArGrGrCrUrArUrUrGrGrUmA*m	+25 DNA + 2xPS	2xPS-OMe + rU	68	22	CCUUGUCAAGGCUAUUGGUCAA (SEQ ID NO:1256)

	A*rU					
1100	A*T*GTGTTTTTGT CAAAAGACCTTTTr UrArArUrUrUrCrUrA rCrUrCrUrUrGrUrAr GrArUrCrCrUrUrGrU rCrArArGrGrCrUrAr UrUrGrGrUmC*rU	+25 DNA + 2xPS	1xPS-OMe + rU	66	20	CCUUGUCAA GGCUAUUGG UC (SEQ ID NO:1002)
1101	A*T*GTGTTTTTGT CAAAAGACCTTTTr UrArArUrUrUrCrUrA rCrUrCrUrUrGrUrAr GrArUrCrCrUrUrGrU rCrArArGrGrCrUrAr UrUrGrGrU*mC	+25 DNA + 2xPS	1xPS-OMe	65	20	CCUUGUCAA GGCUAUUGG UC (SEQ ID NO:1002)
1102	A*T*GTGTTTTTGT CAAAAGACCTTTTr UrArArUrUrUrCrUrA rCrUrCrUrUrGrUrAr GrArUrCrCrUrUrGrU rCrArArGrGrCrUrAr UrUrGrGrUrC*mA* mA	+25 DNA + 2xPS	2xPS-OMe	67	22	CCUUGUCAA GGCUAUUGG UCAA (SEQ ID NO:1256)
1103	mA*mU*rGrUrGrUr UrUrUrUrGrUrCrArA rArArGrArCrCrUrUr UrUrUrArArUrUrUrC rUrArCrUrCrUrUrGr UrArGrArUrCrCrUrU rGrUrCrArArGrGrCr UrArUrUrGrGrUrC* mA	+25 RNA + 2xPS	1xPS-OMe	66	21	CCUUGUCAA GGCUAUUGG UCA (SEQ ID NO:1254)
1104	mA*mA*rArArArAr ArArArArArArArA rArArArArArArArAr ArArUrArArUrUrUrC rUrArCrUrCrUrUrGr UrArGrArUrCrCrUrU rGrUrCrArArGrGrCr UrArUrUrGrGrUrC* mA	PolyA RNA + 2xPS	1xPS-OMe	66	21	CCUUGUCAA GGCUAUUGG UCA (SEQ ID NO:1254)
1105	mU*mU*rUrUrUrUr UrUrUrUrUrUrUrU rUrUrUrUrUrUrUr	PolyU RNA + 2xPS	1xPS-OMe	66	21	CCUUGUCAA GGCUAUUGG UCA (SEQ ID

	UrUrUrArArUrUrUrC rUrArCrUrCrUrUrGr UrArGrArUrCrCrUrU rGrUrCrArArGrGrCr UrArUrUrGrGrUrC* mA					NO:1254)
All bases are in upper case Lowercase "r" represents RNA, 2'-hydroxy; bases not modified by an "r" are DNA All bases are linked via standard phosphodiester bonds except as noted: "*" represents phosphorothioate modification "PS" represents phosphorothioate modification "PS2" represents phosphorodithioate modification "OMe" represents a 2'-o-methyl modification "m" represents a 2'-o-methyl modification **__Table 24__ provides a listing of the sequences of the gRNA 5' extensions						

SEQUENCES

[0441] 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 25**.

Table 25: 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, 1002	gRNA targeting domains (RNA)– see Table 2, 19, 22, 23
910-919, 943-945, 956-959,	gRNA targeting domains (DNA)– see Tables 7, 9

SEQ ID NOS:	DESCRIPTION
1003	
920-929, 946-948, 960-963, 1004	gRNA targeting domains plus PAM (NGG) (RNA) – see Tables 7, 9
930-939, 949-951, 964-967, 1005	gRNA targeting domains plus PAM (NGG) (DNA) – see Tables 7, 9
970, 971, 996, 997	gRNA sequences (DNA)– see Tables 10, 12
972, 973, 998, 999	gRNA sequences (DNA)– see Tables 10, 12
902, 903	Human *HBG1* and *HBG2* promoter sequences including HPFH deletion site
904-909, 974-995	Oligonucleotide donor sequences and homology arms – see Tables 8, 11
968-969	*BCL11Ae* sequences
1006, 1007	Nuclear Localization Signal sequences
1000, 1001, 1008-1018 1032-1039 1094-1097, 1107-1109	Cpf1 polypeptide sequences
1019-1021 1110-1117	Cpf1 nucleotide sequences

INCORPORATION BY REFERENCE

[0442] 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

[0443] 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

Ahern et al., Br J Haematol 25(4):437-444 (1973)

Akinbami Hemoglobin 40:64-65 (2016)
Aliyu et al. Am J Hematol 83:63-70 (2008)
Anders et al. Nature 513(7519):569-573 (2014)
Angastiniotis & Modell Ann N Y Acad Sci 850:251-269 (1998)
Bae et al. Bioinformatics 30(10):1473-1475 (2014)
Barbosa et al. Braz J Med Bio Res 43(8):705-711 (2010)
Bouva Hematologica 91(1):129-132 (2006)
Briner et al. Mol Cell 56(2):333-339 (2014)
Brousseau Am J Hematol 85(1):77-78 (2010)
Caldecott Nat Rev Genet 9(8):619-631 (2008)
Canvers et al. Nature 527(12):192-197 (2015)
Chang et al. Mol Ther Methods Clin Dev 4:137-148 (2017)
Chassanidis Ann Hematol 88(6):549-555 (2009)
Chylinski et al. RNA Biol 10(5):726–737 (2013)
Cong et al. Science 399(6121):819-823 (2013)
Costa et al., Cad Saude Publica 18(5):1469-1471 (2002)
Davis & Maizels 2 Proc Natl Acad Sci USA 111(10):E924-932 (2014)
Fine et al. Sci Rep. 5:10777 (2015)
Frit et al. DNA Repair (Amst) 17:81-97 (2014)
Fu et al. Nat Biotechnol 32:279-284 (2014)
Gao et al. Nat Biotechnol.: 35(8):789-792 (2017)
Giarratana et al. Blood:118, 5071-5079 (2011)
Guilinger et al. Nat Biotechnol 32:577–582 (2014)
Heigwer et al. Nat Methods 11(2):122-123 (2014)
Hsu et al. Nat Biotechnol 31(9):827–832 (2013)
Iyama & Wilson DNA Repair (Amst) 12(8):620-636 (2013)
Jiang et al. Nat Biotechnol 31(3):233–239 (2013)
Jinek et al. Science 337(6096):816-821 (2012)
Jinek et al. Science 343(6176):1247997 (2014)
Kleinstiver et al. Nature 523(7561):481-485 (2015a)
Kleinstiver et al. Nat Biotechnol 33(12):1293–1298 (2015b)
Kleinstiver et al. Nature 529(7587):490-495 (2016)
Komor et al. Nature 533(7603):420–424 (2016)
Lee et al. Nano Lett 12(12):6322-6327 (2012)

Lewis "Medical-Surgical Nursing: Assessment and Management of Clinical Problems" (2014)
Li Cell Res 18(1):85-98 (2008)
Makarova et al. Nat Rev Microbiol 9(6):467–477 (2011)
Mali et al. Science 339(6121):823-826 (2013)
Mantovani et al. Nucleic Acids Res 16(16):7783-7797 (1988)
Marteijn et al. Nat Rev Mol Cell Biol 15(7):465-481 (2014)
Martyn et al., Biochim Biophys Acta 1860(5):525-536 (2017)
Nishimasu et al. Cell 156(5):935-949 (2014)
Nishimasu et al. Cell 162:1113-1126 (2015)
Notta et al. Science 333(6039):218-21 (2011)
Ran & Hsu Cell 154(6):1380-1389 (2013)
Richardson et al. Nat Biotechnol 34:339–344 (2016)
Shmakov et al. Molecular Cell 60(3):385–397 (2015)
Sternberg et al. Nature 507(7490):62-67 (2014)
Superti-Furga et al. EMBO J 7(10):3099-3107 (1988)
Thein Hum Mol Genet 18(R2):R216-223 (2009)
Tsai et al. Nat Biotechnol 34(5): 483 (2016)
Waber et al. Blood 67(2):551-554 (1986)
Wang et al. Cell 153(4):910-918 (2013)
Xiao et al. Bioinformatics 30(8):1180-1182 (2014)
Xu et al. Genes Dev 24(8):783-798 (2010)
Yamano et al. Cell 165(4): 949–962 (2016)
Zetsche et al. Nat Biotechnol 33(2):139-42 (2015)

CLAIMS

1. An RNP complex comprising a CRISPR from Prevotella and Franciscella 1 (Cpf1) RNA-guided nuclease or a variant thereof and a gRNA, wherein the gRNA is capable of binding to a target site in a promoter of an *HBG* gene in a cell.

2. The RNP complex of claim 1, wherein the gRNA is modified or unmodified.

3. A method of altering a promoter of an *HBG* gene in a cell comprising contacting the cell with the RNP complex of claim 1.

4. The method of claim 3, wherein the alteration comprises an indel within a CCAAT box target region of the promoter of an *HBG* gene.

5. The RNP complex of claim 1, wherein the target site comprises nucleotides located between Chr 11 (NC_000011.10) 5,249,904 –5,249,927 (**Table 17**, Region 6); Chr 11 (NC_000011.10) 5,254,879 – 5,254,909 (**Table 17**, Region 16); or a combination thereof.

6. An isolated cell comprising an alteration in a promoter of *HBG* gene generated by the delivery of an RNP complex to the cell according to claim 1.

7. An *ex vivo* method of increasing the level of fetal hemoglobin (HbF) in a human cell by genome editing using an RNP complex comprising a gRNA and a Cpf1 RNA-guided nuclease or a variant thereof to affect an alteration in a promoter of an *HBG* gene, thereby to increase expression of HbF.

8. A population of CD34+ or hematopoietic stem cells, wherein one or more cells in the population comprises an alteration in a promoter of an *HBG* gene, which alteration is generated by delivering an RNP complex comprising a gRNA and a Cpf1 RNA-guided nuclease or a variant thereof to the population of CD34+ or hematopoietic stem cells.

9. A method of alleviating one or more symptoms of sickle cell disease in a subject in need thereof, the method comprising:

a) isolating a population of CD34+ or hematopoietic stem cells from the subject;

b) modifying the population of isolated cells ex vivo by delivering an RNP complex comprising a gRNA and a Cpf1 RNA-guided nuclease or a variant thereof to the population of isolated cells, thereby to affect an alteration in a promoter of an *HBG* gene in one or more cells in the population; and

c) administering the modified population of cells to the subject,

thereby to alleviate one or more symptoms of sickle cell disease in the subject.

10. The method of claim 9, wherein the alteration comprises an indel within a CCAAT box target region of the promoter of the HBG gene.

11. The method of claim 9, wherein the RNP complex is delivered using electroporation.

12. The population of cells modified according to claim 9, wherein at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80% or at least about 90% of the cells in the population of cells comprise a productive indel.

13. A gRNA comprising a 5' end and a 3' end, and comprising a DNA extension at the 5' end and a 2'-O-methyl-3'-phosphorothioate modification at the 3' end, wherein the gRNA includes an RNA segment capable of hybridizing to a target site and an RNA segment capable of associating with a Cpf1 RNA-guided nuclease.

14. The gRNA of claim 13, wherein the DNA extension comprises a sequence set forth in SEQ ID NOs:1235-1250.

15. An RNP complex comprising a Cpf1 RNA-guided nuclease and a gRNA according to claim 14.