



US 20230272115A1

(19) **United States**

(12) **Patent Application Publication**  
**QUIJANO et al.**

(10) **Pub. No.: US 2023/0272115 A1**

(43) **Pub. Date: Aug. 31, 2023**

(54) **COMPOSITIONS AND METHODS FOR  
ENHANCING DONOR  
OLIGONUCLEOTIDE-BASED GENE  
EDITING**

*A61K 31/7088* (2006.01)  
*A61K 35/14* (2006.01)  
*A61K 48/00* (2006.01)  
*C12N 15/10* (2006.01)

(71) Applicant: **Yale University**, New Haven, CT (US)  
(72) Inventors: **Elias QUIJANO**, Durham, CT (US);  
**Audrey TURCHICK**, Chelmsford, MA (US); **Peter GLAZER**, Guilford, CT (US)  
(73) Assignee: **Yale University**, New Haven, CT (US)

(52) **U.S. Cl.**  
CPC ..... *C07K 16/44* (2013.01); *A61K 47/6937* (2017.08); *A61P 43/00* (2018.01); *A61K 31/7088* (2013.01); *A61K 35/14* (2013.01); *A61K 48/0008* (2013.01); *C12N 15/102* (2013.01); *A61K 2039/505* (2013.01)

(21) Appl. No.: **18/054,080**  
(22) Filed: **Nov. 9, 2022**

**Related U.S. Application Data**

(63) Continuation of application No. 17/272,136, filed on Feb. 26, 2021, filed as application No. PCT/US2019/048953 on Aug. 30, 2019.  
(60) Provisional application No. 62/725,920, filed on Aug. 31, 2018.

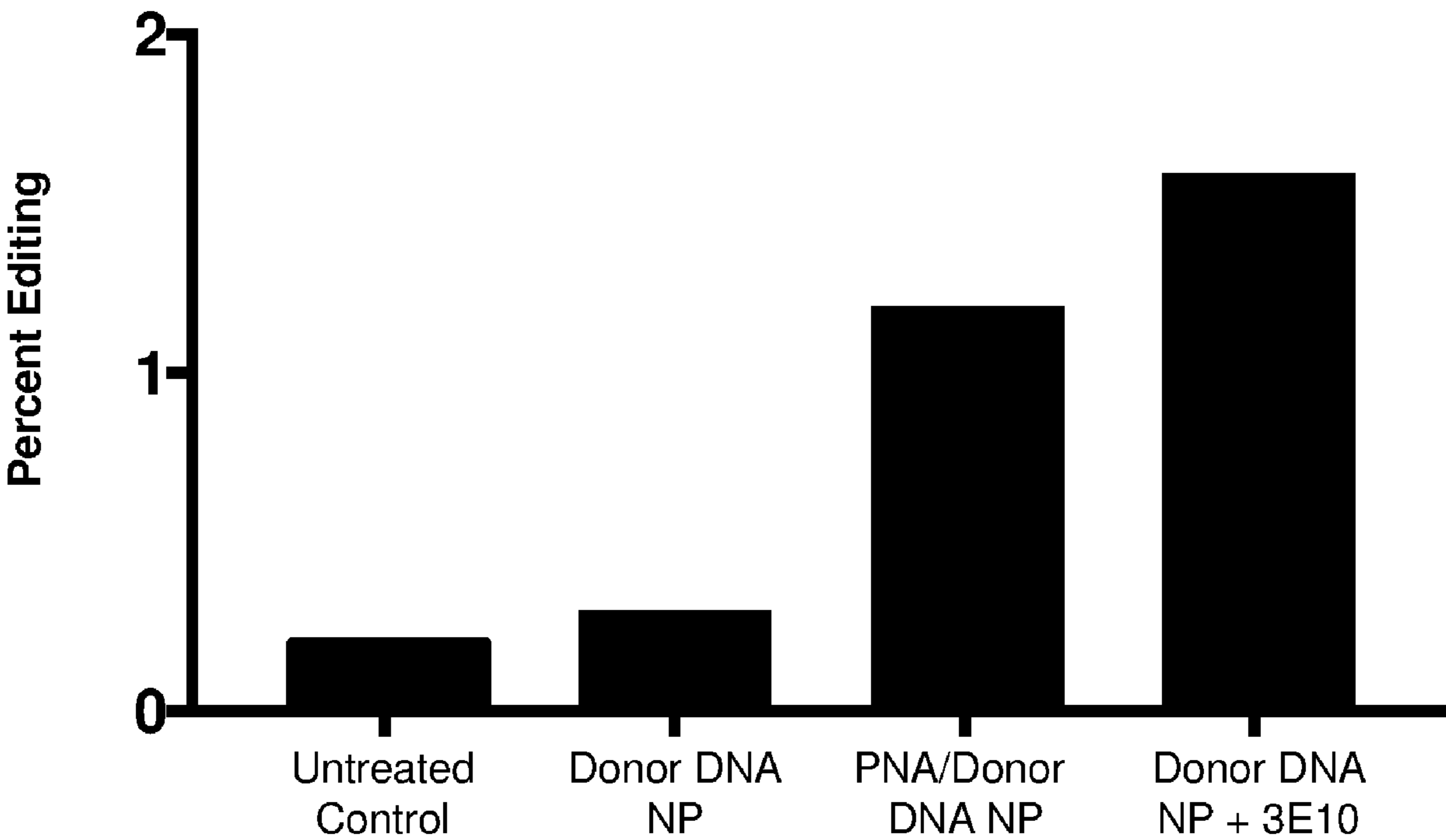
**Publication Classification**

(51) **Int. Cl.**  
*C07K 16/44* (2006.01)  
*A61K 47/69* (2006.01)  
*A61P 43/00* (2006.01)

(57) **ABSTRACT**

Compositions for enhanced gene editing and methods of use thereof are. The composition contains a cell-penetrating antibody and a donor oligonucleotide containing a sequence that can correct a mutation in a cell’s genome. Preferably, the composition does not contain a nuclease, PNA, or nanoparticle. The compositions are used to modify the genome of a cell by contacting the cell with an effective amount of the composition. Genomic modification occurs at a higher frequency both ex vivo and in vivo, when cells are contacted with the cell-penetrating antibody and donor oligonucleotide as compared to the absence of the cell-penetrating antibody.

**Specification includes a Sequence Listing.**



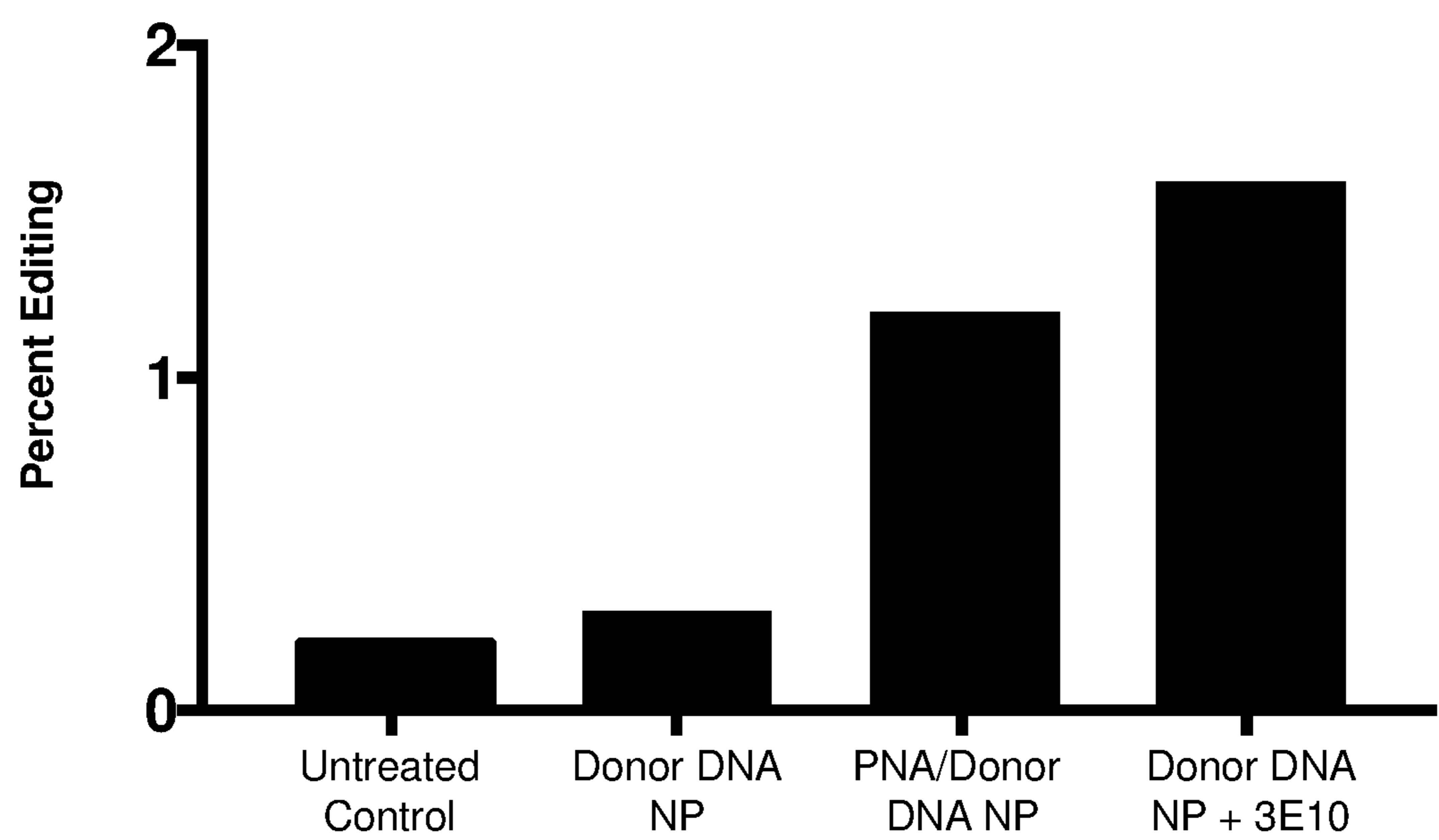


FIG. 1

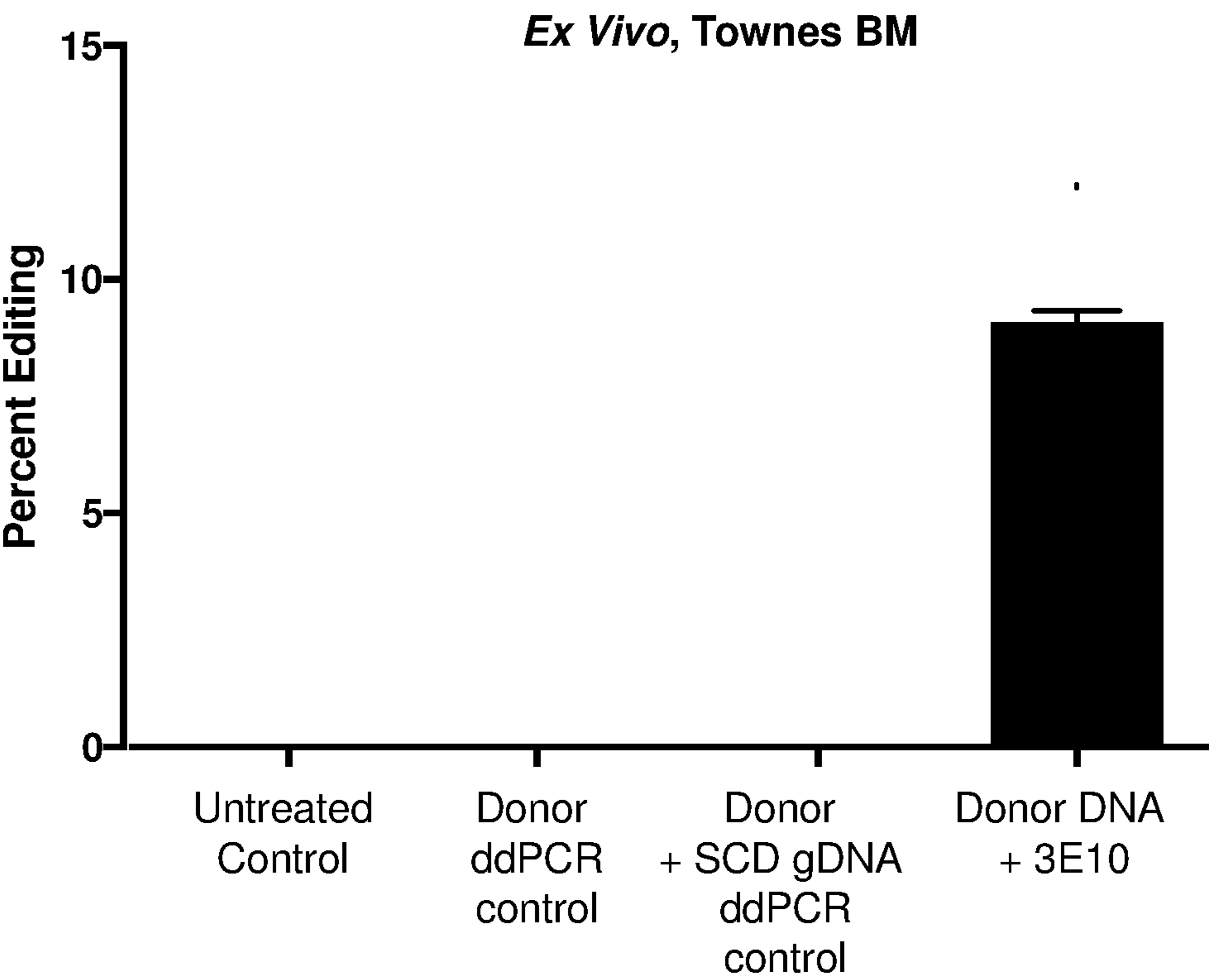
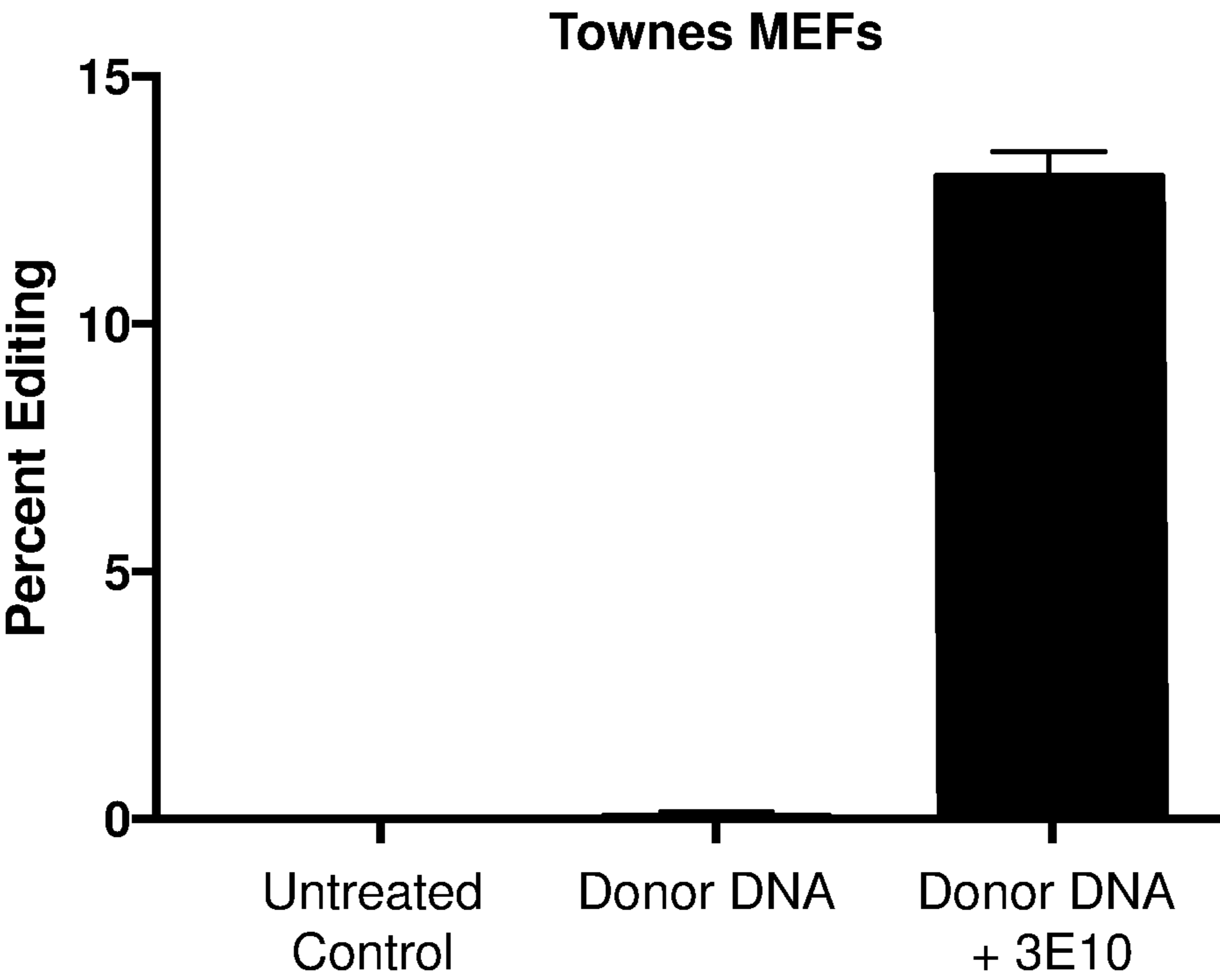


FIG. 2



**FIG. 3**

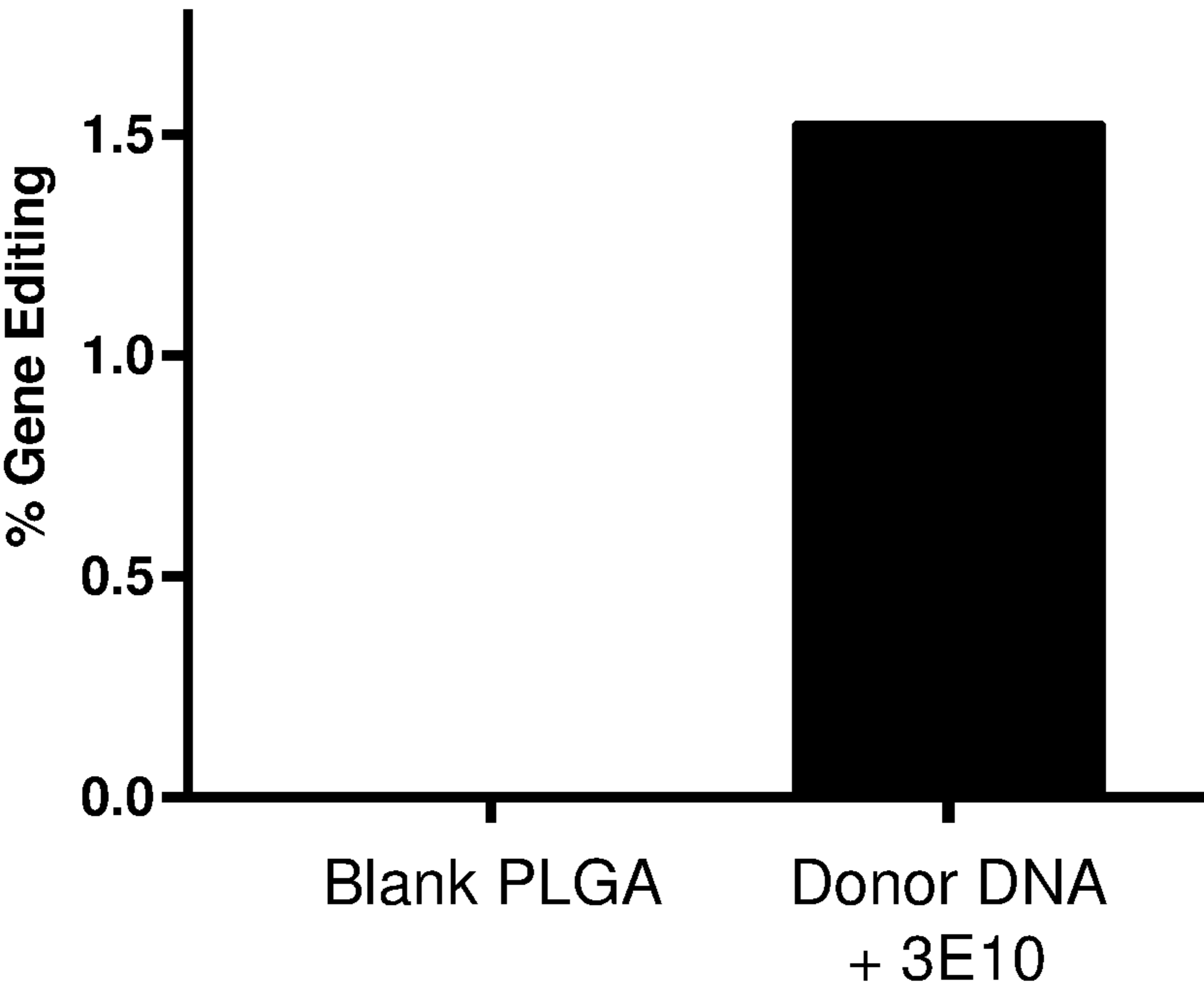


FIG. 4

# COMPOSITIONS AND METHODS FOR ENHANCING DONOR OLIGONUCLEOTIDE-BASED GENE EDITING

## CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of and priority to U.S. Ser. No. 62/725,920, filed Aug. 31, 2018, which is specifically incorporated by reference herein in its entirety.

## STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

[0002] This invention was made with government support under CA168733 and CA197574 awarded by National Institutes of Health. The Government has certain rights in the invention.

## REFERENCE TO THE SEQUENCE LISTING

[0003] The Sequence Listing submitted as a text file named "YU\_7503\_PCT" created on Aug. 28, 2019, and having a size of 61,341 bytes is hereby incorporated by reference pursuant to 37 C.F.R. § 1.52(e)(5).

## FIELD OF THE INVENTION

[0004] The invention is generally related to the field of gene editing technology, and more particularly to compositions and uses of potentiating agents such as cell-penetrating anti-DNA antibodies, and donor oligonucleotides in methods of ex vivo and in vivo gene editing.

## BACKGROUND OF THE INVENTION

[0005] Gene editing provides an attractive strategy for treatment of inherited genetic disorders such as sickle cell anemia and  $\beta$ -thalassemia. Genes can be selectively edited by several methods, including targeted nucleases such as zinc finger nucleases (ZFNs) (Haendel, et al., *Gene Ther.*, 11:28-37 (2011)) and CRISPRs (Yin, et al., *Nat. Biotechnol.*, 32:551-553 (2014)), short fragment homologous recombination (SFHR) (Goncz, et al., *Oligonucleotides*, 16:213-224 (2006)), or triplex-forming oligonucleotides (TFOs) (Vasquez, et al., *Science*, 290:530-533 (2000)). It is generally thought that a DNA break in a target gene is needed for high efficiency gene editing with a donor DNA. Hence, there has been widespread focus on targeted nucleases such as CRISPR/Cas9 technology because of its ease of use and facile reagent design (Doudna, et al., *Science*, 346:1258096 (2014)). However, like ZFNs, the CRISPR approach introduces an active nuclease into cells, which can lead to off-target cleavage in the genome (Cradick, et al., *Nucleic Acids Res.*, 41:9584-9592 (2013)), a problem that so far has not been eliminated.

[0006] In addition, the efficiency of gene modification could be low, especially in the context of CRISPR/Cas-mediated editing in primary stem cells. For example, in an attempt to correct the CFTR locus in cystic fibrosis patient derived stem cells, approximately 0.3% of treated organoids (3 to 6/1400) had the desired modification (Schwank, et al., *Cell Stem Cell.*, 13:653-658 (2013)).

[0007] Accordingly, there remains a need for compositions and methods for improved gene editing.

[0008] It is therefore an object of the invention to provide compositions and methods for achieving on-target modification with reduced or low off-target modification.

[0009] It is another object of the invention to provide compositions and methods for achieving an increased frequency of gene modification.

[0010] It is a further object of the invention to provide compositions and methods for gene modification that improve one or more symptoms of a disease or disorder in a subject.

## SUMMARY OF THE INVENTION

[0011] It has been discovered that treatment of cells with a cell-penetrating, anti-DNA, lupus enhances targeted gene editing by donor DNAs alone, in the absence of a nuclease or a PNA. As described in the Examples, it has been discovered that eMab 3E10 boosts gene editing with nanoparticles containing donor DNA alone. It was also discovered that 3E10 promotes gene editing by naked donor DNAs alone in cells in culture and in mice in vivo, without any associated nuclease or PNA and without encapsulation in a nanoparticle for delivery.

[0012] Thus, compositions for enhancing targeted gene editing and methods of use thereof are disclosed. In some embodiments, the composition contains a potentiating agent such as a cell-penetrating antibody, and a donor oligonucleotide containing a sequence that can correct a mutation in a cell's genome. Preferably, the composition does not contain a nuclease (e.g., transcription activator-like effector nuclease (TALEN), ZFN, Cas9), peptide nucleic acid (PNA), or nanoparticle. Typically, the donor oligonucleotide is not covalently linked to the cell-penetrating antibody. The oligonucleotide (e.g., DNA) may be single stranded or double stranded. Preferably, the oligonucleotide is single stranded DNA.

[0013] In some embodiments, the oligonucleotide sequence corresponds to the wild type sequence of a mutated gene underlying a disease or disorder (e.g., hemophilia, muscular dystrophy, globinopathies, cystic fibrosis, xeroderma pigmentosum, lysosomal storage diseases, immune deficiency syndromes such as X-linked severe combined immunodeficiency and ADA deficiency, tyrosinemia, Fanconi anemia, the red cell disorder spherocytosis, alpha-1-anti-trypsin deficiency, Wilson's disease, Leber's hereditary optic neuropathy, or chronic granulomatous disorder). Exemplary genes include genes encoding coagulation factor VIII, coagulation factor IX, dystrophin, beta-globin, CFTR, XPC, XPD, DNA polymerase eta, Fanconi anemia genes A through L, SPTA1 and other spectrin genes, ANK1 gene, SERPINA1 gene, ATP7B gene, interleukin 2 receptor gamma (IL2RG) gene, ADA gene, FAH gene, and genes linked to chronic granulomatous disease including the CYBA, CYBB, NCF1, NCF2, or NCF4 genes, etc.

[0014] The potentiating agent typically increases gene editing by the donor oligonucleotide. In preferred embodiments, the potentiating agent is a cell-penetrating antibody. In some embodiments, the potentiating agent is an anti-RAD51 factor. The cell-penetrating antibody may be an anti-DNA antibody which is transported into the cytoplasm and/or nucleus of the cell without the aid of a carrier or conjugate.

**[0015]** In some embodiments, the cell-penetrating anti-DNA antibody is isolated or derived from a subject with systemic lupus erythematosus or an animal model thereof (e.g., such as a mouse or rabbit). In a preferred embodiment, the cell-penetrating anti-DNA antibody is the monoclonal anti-DNA antibody 3E10, or a variant, fragment, or humanized form thereof that binds the same epitope(s) as 3E10. A particularly preferred variant is a 3E10 variant incorporating a D31N substitution in the heavy chain. The cell-penetrating anti-DNA antibody may have the same or different epitope specificity as monoclonal antibody 3E10 produced by ATCC No. PTA 2439 hybridoma.

**[0016]** In some embodiments, the antibody has

**[0017]** (i) the CDRs of any one of SEQ ID NO:1-6, 12, or 13 in combination with the CDRs of any one of SEQ ID NO:7-11, or 15;

**[0018]** (ii) first, second, and third heavy chain CDRs selected from SEQ ID NOS:15-23 in combination with first, second and third light chain CDRs selected from SEQ ID NOS:24-30;

**[0019]** (iii) humanized forms of (i) or (ii);

**[0020]** (iv) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:1 or 2 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:7 or 8;

**[0021]** (v) a humanized form or (iv); or

**[0022]** (vi) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:3-6 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:9-11.

**[0023]** Preferably, the antibody can bind directly to RAD51. In some embodiments, the anti-DNA antibody has the paratope of monoclonal antibody 3E10. The anti-DNA antibody may be a single chain variable fragment of an anti-DNA antibody, or conservative variant thereof. For example, the anti-DNA antibody can be a monovalent, divalent, or multivalent single chain variable fragment of 3E10 (3E10 Fv), or a variant, for example a conservative variant, thereof. In some embodiments, the anti-DNA antibody is a monovalent, divalent, or multivalent single chain variable fragment of 3E10 (3E10 Fv) incorporating a D31N substitution in the heavy chain.

**[0024]** Also provided is a pharmaceutical composition containing a potentiating agent, such as a cell-penetrating antibody, and donor oligonucleotides in a pharmaceutically acceptable excipient. The compositions may be used to modify the genome of a cell by contacting the cell with an effective amount of the composition.

**[0025]** Also provided is a method of modifying the genome of a cell by contacting the cell with an effective amount of (i) a potentiating agent such as a cell-penetrating antibody, and (ii) a donor oligonucleotide containing a sequence that can correct a mutation in a cell's genome. Genomic modification may occur at a higher frequency when cells are contacted with both (i) and (ii), than when contacted with (ii) in the absence of (i). Preferably, the method does not involve contacting the cell with a nuclease (e.g., ZFN, Cas9) or peptide nucleic acid (PNA).

**[0026]** The donor oligonucleotide(s) may be separately encapsulated in nanoparticles. The nanoparticles may be formed from polyhydroxy acids (e.g., poly(lactic-co-gly-

colic acid) (PLGA)). The nanoparticles may be prepared by double emulsion. In some embodiments, a targeting moiety, a cell penetrating peptide, or a combination thereof is associated with, linked, conjugated, or otherwise attached directly or indirectly to the nanoparticle.

**[0027]** In some embodiments, the cells (e.g., hematopoietic stem cells) are contacted ex vivo and the cells may further be administered to a subject in need thereof. The cells may be administered to the subject in an effective amount to treat one or more symptoms of a disease or disorder. In other embodiments, the cells are contacted in vivo following administration of the cell-penetrating antibody and donor oligonucleotide to a subject.

**[0028]** The subject may have a disease or disorder such as hemophilia, muscular dystrophy, globinopathies, cystic fibrosis, xeroderma pigmentosum, lysosomal storage diseases, immune deficiency syndromes such as X-linked severe combined immunodeficiency and ADA deficiency, tyrosinemia, Fanconi anemia, the red cell disorder spherocytosis, alpha-1-anti-trypsin deficiency, Wilson's disease, Leber's hereditary optic neuropathy, or chronic granulomatous disorder. In such embodiments, gene modification may occur in an effective amount to reduce one or more symptoms of the disease or disorder in the subject.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0029]** FIG. 1 is a bar graph showing gene correction of the IVS2-654 (C->T) mutation within the  $\beta$ -globin/GFP fusion gene in MEFs treated with 1) NPs containing donor DNA alone, 2) NPs containing donor DNA and PNA, and 3) NPs containing donor DNA alone and 3E10. Percent gene editing was determined by the fraction of GFP<sup>+</sup> MEFs assessed by flow cytometry.

**[0030]** FIG. 2 is a bar graph showing the percentage of gene editing in the beta globin gene following ex vivo treatment of bone marrow cells from Townes mice with eMab (3E10) and donor DNA. Gene editing frequencies were determined by droplet digital PCR (ddPCR) analysis of genomic DNA.

**[0031]** FIG. 3 is a bar graph showing the percentage of gene editing in the beta globin gene following in vitro treatment of MEFs from Townes mice with eMab (3E10) and donor DNA. Gene editing frequencies were determined by droplet digital PCR (ddPCR) analysis of genomic DNA. Editing in MEFs treated with 3E10/donor DNA was significantly higher than MEFs treated with donor DNA alone.

**[0032]** FIG. 4 is a bar graph showing the percentage of gene editing in the beta globin gene following in vivo treatment of Townes mice with 3E10 and donor DNA delivered by intraperitoneal injection. Gene editing was assessed in bone marrow cells 2 months post injection. As compared to blank PLGA nanoparticles, mice treated with 3E10/donor DNA demonstrated significantly higher levels of gene editing.

#### DETAILED DESCRIPTION OF THE INVENTION

##### I. Definitions

**[0033]** As used herein, the term "single chain Fv" or "scFv" as used herein means a single chain variable fragment that includes a light chain variable region (VL) and a heavy chain variable region (VH) in a single polypeptide

chain joined by a linker which enables the scFv to form the desired structure for antigen binding (i.e., for the VH and VL of the single polypeptide chain to associate with one another to form a Fv). The VL and VH regions may be derived from the parent antibody or may be chemically or recombinantly synthesized.

**[0034]** As used herein, the term “variable region” is intended to distinguish such domain of the immunoglobulin from domains that are broadly shared by antibodies (such as an antibody Fc domain). The variable region includes a “hypervariable region” whose residues are responsible for antigen binding. The hypervariable region includes amino acid residues from a “Complementarity Determining Region” or “CDR” (i.e., typically at approximately residues 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the light chain variable domain and at approximately residues 27-35 (H1), 50-65 (H2) and 95-102 (H3) in the heavy chain variable domain; Kabat et al., *Sequences of Proteins of Immunological Interest*, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, Md. (1991)) and/or those residues from a “hypervariable loop” (i.e., residues 26-32 (L1), 50-52 (L2) and 91-96 (L3) in the light chain variable domain and 26-32 (H1), 53-55 (H2) and 96-101 (H3) in the heavy chain variable domain; Chothia and Lesk, 1987, *J. Mol. Biol.* 196:901-917).

**[0035]** As used herein, the term “Framework Region” or “FR” residues are those variable domain residues other than the hypervariable region residues as herein defined.

**[0036]** As used herein, the term “antibody” refers to natural or synthetic antibodies that bind a target antigen. The term includes polyclonal and monoclonal antibodies. In addition to intact immunoglobulin molecules, also included in the term “antibodies” are binding proteins, fragments, and polymers of those immunoglobulin molecules, and human or humanized versions of immunoglobulin molecules that bind the target antigen.

**[0037]** As used herein, the term “cell-penetrating antibody” refers to an immunoglobulin protein, fragment, variant thereof, or fusion protein based thereon that is transported into the cytoplasm and/or nucleus of living mammalian cells. The “cell-penetrating anti-DNA antibody” specifically binds DNA (e.g., single-stranded and/or double-stranded DNA). In some embodiments, the antibody is transported into the cytoplasm of the cells without the aid of a carrier or conjugate. In other embodiments, the antibody is conjugated to a cell-penetrating moiety, such as a cell penetrating peptide. In some embodiments, the cell-penetrating antibody is transported in the nucleus with or without a carrier or conjugate.

**[0038]** In addition to intact immunoglobulin molecules, also included in the term “antibodies” are fragments, binding proteins, and polymers of immunoglobulin molecules, chimeric antibodies containing sequences from more than one species, class, or subclass of immunoglobulin, such as human or humanized antibodies, and recombinant proteins containing a least the idiotype of an immunoglobulin that specifically binds DNA. The antibodies can be tested for their desired activity using the in vitro assays described herein, or by analogous methods, after which their in vivo therapeutic activities are tested according to known clinical testing methods.

**[0039]** As used herein, the term “variant” refers to a polypeptide or polynucleotide that differs from a reference polypeptide or polynucleotide, but retains essential proper-

ties. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more modifications (e.g., substitutions, additions, and/or deletions). A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polypeptide may be naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally.

**[0040]** Modifications and changes can be made in the structure of the polypeptides of in disclosure and still obtain a molecule having similar characteristics as the polypeptide (e.g., a conservative amino acid substitution). For example, certain amino acids can be substituted for other amino acids in a sequence without appreciable loss of activity. Because it is the interactive capacity and nature of a polypeptide that defines that polypeptide’s biological functional activity, certain amino acid sequence substitutions can be made in a polypeptide sequence and nevertheless obtain a polypeptide with like properties.

**[0041]** In making such changes, the hydropathic index of amino acids can be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a polypeptide is generally understood in the art. It is known that certain amino acids can be substituted for other amino acids having a similar hydropathic index or score and still result in a polypeptide with similar biological activity. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. Those indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (−0.4); threonine (−0.7); serine (−0.8); tryptophan (−0.9); tyrosine (−1.3); proline (−1.6); histidine (−3.2); glutamate (−3.5); glutamine (−3.5); aspartate (−3.5); asparagine (−3.5); lysine (−3.9); and arginine (−4.5).

**[0042]** It is believed that the relative hydropathic character of the amino acid determines the secondary structure of the resultant polypeptide, which in turn defines the interaction of the polypeptide with other molecules, such as enzymes, substrates, receptors, antibodies, antigens, and cofactors. It is known in the art that an amino acid can be substituted by another amino acid having a similar hydropathic index and still obtain a functionally equivalent polypeptide. In such changes, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

**[0043]** Substitution of like amino acids can also be made on the basis of hydrophilicity, particularly where the biological functional equivalent polypeptide or peptide thereby created is intended for use in immunological embodiments. The following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 $\pm$ 1); glutamate (+3.0 $\pm$ 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); proline (−0.5 $\pm$ 1); threonine (−0.4); alanine (−0.5); histidine (−0.5); cysteine (−1.0); methionine (−1.3); valine (−1.5); leucine (−1.8); isoleucine (−1.8); tyrosine (−2.3); phenylalanine (−2.5); tryptophan (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still

obtain a biologically equivalent, and in particular, an immunologically equivalent polypeptide. In such changes, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  is preferred, those within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

**[0044]** As outlined above, amino acid substitutions are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include (original residue: exemplary substitution): (Ala: Gly, Ser), (Arg: Lys), (Asn: Gln, His), (Asp: Glu, Cys, Ser), (Gln: Asn), (Glu: Asp), (Gly: Ala), (His: Asn, Gln), (Ile: Leu, Val), (Leu: Ile, Val), (Lys: Arg), (Met: Leu, Tyr), (Ser: Thr), (Thr: Ser), (Tyr: Trp, Phe), and (Val: Ile, Leu). Embodiments of this disclosure thus contemplate functional or biological equivalents of a polypeptide as set forth above. In particular, embodiments of the polypeptides can include variants having about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to the polypeptide of interest.

**[0045]** As used herein, the term “percent (%) sequence identity” is defined as the percentage of nucleotides or amino acids in a candidate sequence that are identical with the nucleotides or amino acids in a reference nucleic acid sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign (DNASTAR) software. Appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared can be determined by known methods.

**[0046]** For purposes herein, the % sequence identity of a given nucleotides or amino acids sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given sequence C that has or includes a certain % sequence identity to, with, or against a given sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z,$$

where W is the number of nucleotides or amino acids scored as identical matches by the sequence alignment program in that program's alignment of C and D, and where Z is the total number of nucleotides or amino acids in D. It will be appreciated that where the length of sequence C is not equal to the length of sequence D, the % sequence identity of C to D will not equal the % sequence identity of D to C.

**[0047]** As used herein, the term “specifically binds” refers to the binding of an antibody to its cognate antigen (for example, DNA) while not significantly binding to other antigens. Specific binding of an antibody to a target under such conditions requires the antibody be selected for its specificity to the target. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See, e.g., Harlow and Lane (1988) *Antibodies*, A Laboratory

Manual, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity. Preferably, an antibody “specifically binds” to an antigen with an affinity constant ( $K_a$ ) greater than about  $10^5 \text{ mol}^{-1}$  (e.g.,  $10^6 \text{ mol}^{-1}$ ,  $10^7 \text{ mol}^{-1}$ ,  $10^8 \text{ mol}^{-1}$ ,  $10^9 \text{ mol}^{-1}$ ,  $10^{10} \text{ mol}^{-1}$ ,  $10^{11} \text{ mol}^{-1}$ , and  $10^{12} \text{ mol}^{-1}$  or more) with that second molecule.

**[0048]** As used herein, the term “monoclonal antibody” or “MAb” refers to an antibody obtained from a substantially homogeneous population of antibodies, i.e., the individual antibodies within the population are identical except for possible naturally occurring mutations that may be present in a small subset of the antibody molecules.

**[0049]** As used herein a “gene editing potentiating factor” or “gene editing potentiating agent” or “potentiating factor” or “potentiating agent” refers to a compound that increases the efficacy of editing (e.g., mutation, including insertion, deletion, substitution, etc.) of a gene, genome, or other nucleic acid by a donor oligonucleotide relative to use of the donor oligonucleotide in the absence of the compound.

**[0050]** As used herein, the term “subject” means any individual who is the target of administration. The subject can be a vertebrate, for example, a mammal. Thus, the subject can be a human. The term does not denote a particular age or sex.

**[0051]** As used herein, the term “effective amount” means that the amount of the composition used is of sufficient quantity to ameliorate one or more causes or symptoms of a disease or disorder. Such amelioration only requires a reduction or alteration, not necessarily elimination. The precise dosage will vary according to a variety of factors such as subject-dependent variables (e.g., age, immune system health, etc.), the disease or disorder being treated, as well as the route of administration and the pharmacokinetics of the agent being administered.

**[0052]** As used herein, the term “pharmaceutically acceptable” refers to a material that is not biologically or otherwise undesirable, i.e., the material may be administered to a subject without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition in which it is contained.

**[0053]** As used herein, the term “carrier” or “excipient” refers to an organic or inorganic ingredient, natural or synthetic inactive ingredient in a formulation, with which one or more active ingredients are combined. The carrier or excipient would naturally be selected to minimize any degradation of the active ingredient and to minimize any adverse side effects in the subject, as would be well known to one of skill in the art.

**[0054]** As used herein, the term “treat” refers to the medical management of a patient with the intent to cure, ameliorate, stabilize, or prevent a disease, pathological condition, or disorder. This term includes active treatment, that is, treatment directed specifically toward the improvement of a disease, pathological condition, or disorder, and also includes causal treatment, that is, treatment directed toward removal of the cause of the associated disease, pathological condition, or disorder. In addition, this term includes palliative treatment, that is, treatment designed for the relief of symptoms rather than the curing of the disease, pathological condition, or disorder; preventative treatment, that is, treatment directed to minimizing or partially or completely inhibiting the development of the associated disease, patho-

logical condition, or disorder; and supportive treatment, that is, treatment employed to supplement another specific therapy directed toward the improvement of the associated disease, pathological condition, or disorder.

**[0055]** As used herein, “targeting moiety” is a substance which can direct a particle or molecule to a receptor site on a selected cell or tissue type, can serve as an attachment molecule, or serve to couple or attach another molecule. As used herein, “direct” refers to causing a molecule to preferentially attach to a selected cell or tissue type. This can be used to direct cellular materials, molecules, or drugs, as discussed below.

**[0056]** As used herein, the term “inhibit” or “reduce” means to decrease an activity, response, condition, disease, or other biological parameter. This can include, but is not limited to, the complete ablation of the activity, response, condition, or disease. This may also include, for example, a 10% reduction in the activity, response, condition, or disease as compared to the native or control level. Thus, the reduction can be a 10, 20, 30, 40, 50, 60, 70, 80, 90, 100%, or any amount of reduction in between as compared to native or control levels.

**[0057]** As used herein, a “fusion protein” refers to a polypeptide formed by the joining of two or more polypeptides through a peptide bond formed between the amino terminus of one polypeptide and the carboxyl terminus of another polypeptide. The fusion protein can be formed by the chemical coupling of the constituent polypeptides or it can be expressed as a single polypeptide from a nucleic acid sequence encoding the single contiguous fusion protein. A single chain fusion protein is a fusion protein having a single contiguous polypeptide backbone. Fusion proteins can be prepared using conventional techniques in molecular biology to join the two genes in frame into a single nucleic acid sequence, and then expressing the nucleic acid in an appropriate host cell under conditions in which the fusion protein is produced.

**[0058]** Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein.

**[0059]** Use of the term “about” is intended to describe values either above or below the stated value in a range of approx.  $\pm 10\%$ ; in other embodiments the values may range in value either above or below the stated value in a range of approx.  $\pm 5\%$ ; in other embodiments the values may range in value either above or below the stated value in a range of approx.  $\pm 2\%$ ; in other embodiments the values may range in value either above or below the stated value in a range of approx.  $\pm 1\%$ . The preceding ranges are intended to be made clear by context, and no further limitation is implied.

**[0060]** All methods described herein can be performed in any suitable order unless otherwise indicated or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the embodiments and does not pose a limitation on the scope of the embodiments unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

## II. Compositions

**[0061]** Compositions for enhancing targeted gene editing and methods of use thereof are disclosed. The disclosed methods typically include contacting cells with both a potentiating agent and a donor oligonucleotide. Exemplary potentiating agents and donor oligonucleotides are provided. The potentiating agent and donor oligonucleotide can be part of the same or different compositions.

**[0062]** In some embodiments, potentiating agents can engage one or more endogenous high fidelity DNA repair pathways or inhibit/modulate error prone (i.e. low fidelity) DNA repair pathways. Potentiating agents include, for example, modulators of DNA damage and/or DNA repair factors, modulators of homologous recombination factors, cell adhesion modulators, cell cycle modulators, cell proliferation modulators, and stem cell mobilizers. The potentiating factor may modulate (e.g., alter, inhibit, promote, compete with) one or more endogenous high fidelity DNA repair pathways or inhibit/modulate error prone (i.e. low fidelity) DNA repair pathways. In preferred embodiments, the potentiating factor may be an inhibitor of a DNA damage, DNA repair, or homologous recombination factor. In more preferred embodiments, the potentiating factor may be an inhibitor of RAD51.

**[0063]** For example, an inhibitor of a DNA damage and/or DNA repair factor may be used as a potentiating agent. An inhibitor of a homologous recombination factor may be used as a potentiating agent.

**[0064]** Cells repair DNA breaks mainly through endogenous non-homologous end joining (NHEJ) DNA-repair, the predominant but error-prone pathway that can introduce or delete nucleotides at the DNA-break region. NHEJ is therefore amenable to permanent silencing of target genes. Alternatively, cells can also repair double-strand breaks by homology-directed repair (HDR), a more accurate mechanism involving homologous recombination in the presence of a template DNA strand. Typically, targeted genome editing is directed to correction of a mutated sequence in a genome by replacing the mutated sequence with a corrective sequence provided by a template/donor DNA. As such, there is ongoing effort in the field to identify and utilize mechanisms that favor homologous recombination of a template/donor DNA to enhance efficiency of targeted genome editing. Modulating the expression and/or activity of factors involved in DNA repair is a promising approach to enhance precision genome engineering.

**[0065]** The term “DNA repair” refers to a collection of processes by which a cell identifies and corrects damage to DNA molecules. Single-strand defects are repaired by base excision repair (BER), nucleotide excision repair (NER), or mismatch repair (MMR). Double-strand breaks are repaired by non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ), or homologous recombination. After DNA damage, cell cycle checkpoints are activated, which pause the cell cycle to give the cell time to repair the damage before continuing to divide. Checkpoint mediator proteins include BRCA1, MDC1, 53BP1, p53, ATM, ATR, CHK1, CHK2, and p21. Accordingly, a factor involved in any of the above-mentioned processes, including BER, NER, MMR, NHEJ, MMEJ, homologous recombination, or DNA synthesis and the like, may be described as a DNA damage and/or DNA repair factor.

**[0066]** Non-limiting examples of DNA damage, DNA repair, DNA synthesis, or homologous recombination fac-

tors include XRCC1, ADPRT (PARP-1), ADPRTL2, (PARP-2), POLYMERASE BETA, CTPS, MLH1, MSH2, FANCD2, PMS2, p53, p21, PTEN, RPA, RPA1, RPA2, RPA3, XPD, ERCC1, XPF, MMS19, RAD51, RAD51B, RAD51C, RAD51D, DMC1, XRCCR, XRCC3, BRCA1, BRCA2, PALB2, RAD52, RAD54, RAD50, MREU, NB51, WRN, BLM, KU70, KU80, ATM, ATR, CIPK1, CHK2, FANCA, FANCB, FANCC, FANCD1, FANCD2, FANCE, FANCF, FANCG, FANCC, FANCD1, FANCD2, FANCE, FANCF, FANCG, RAD1, and RAD9. In a preferred embodiment, the DNA damage factor or DNA repair factor is RAD51.

**[0067]** RAD51 recombinase, an ortholog of *E. coli* RecA, is a key protein in homologous recombination in mammalian cells. RAD51 promotes the repair of double-strand breaks, the most harmful type of DNA lesion. Double-strand breaks can be induced by various chemical agents and ionizing radiation, and are also formed during the repair of inter-strand crosslinks. Once double-strand breaks are formed, they are processed first by exonucleases to generate extensive 3' single-stranded DNA (ssDNA) tails (Cejka et al., *Nature.*, 467(7311):112-16 (2010); Mimitou & Symington, *DNA Repair.*, 8(9):983-95 (2009)). These tracks of ssDNA rapidly become coated by single strand DNA-binding protein, RPA, which is ultimately displaced from the ssDNA by RAD51. RAD51 has ATP-dependent DNA binding activity, and so binds the ssDNA tails, and multimerizes to form helical nucleoprotein filaments that promote search for homologous dsDNA sequences (Kowalczykowski, *Nature.*, 453(7194):463-6 (2008)). The ability of RAD51 to displace RPA on ssDNA in cells requires several mediator proteins, which include BRCA2, RAD52, the RAD51 paralog complexes, and other proteins (Thompson & Schild, *Mutat Res.*, 477:131-53 (2001)). Once homologous dsDNA sequences are found, RAD51 promotes DNA strand exchange between the ssDNA that resides within the filament and homologous dsDNA, i.e., an invasion of ssDNA into homologous DNA duplex that results in the displacement of the identical ssDNA from the duplex and formation of a joint molecule. Joint molecules, key intermediates of DSB repair, provide both the template and the primer for DNA repair synthesis that is required for double-strand break repair (Paques & Haber, *Microbiol. Mol. Biol. Rev.*, 63(2):349-404 (1999)).

**[0068]** By promoting DNA strand exchange, RAD51 plays a key role in homologous recombination. The protein is evolutionarily conserved from bacteriophages to mammals. In all organisms, RAD51 orthologs play an important role in DNA repair and homologous recombination (Krough & Symington, *Annu. Rev. Genet.*, 38:233-71 (2004); Helleday et al., *DNA Repair.*, 6(7):923-35 (2007); Huang et al., *Proc. Natl. Acad. Sci. USA.*, 93(10):4827-32 (1996)).

**[0069]** In preferred embodiments, the potentiating agent is one that antagonizes or reduces expression and/or activity of RAD51, XRCC4, or a combination thereof. For example, in some embodiments, the potentiating agent is a RAD51 and/or XRCC4 inhibitor. Non-limiting examples of potentiating agents include, ribozymes, triplex forming molecules, siRNAs, shRNAs, miRNAs, aptamers, antisense oligonucleotides, small molecules, and antibodies.

**[0070]** Methods for designing and producing any of the foregoing factors are well-known in the art and can be used. For example, predesigned anti-RAD51 siRNAs are commercially available through Dharmacon (as described in the Examples) and may be used as potentiating agents. Like-

wise, anti-XRCC4 siRNAs, shRNAs and miRNAs are known in the art and are readily available. Further, small molecule inhibitors of XRCC4 and RAD51 are known in the art (e.g., Jekimovs, et al., *Front. Oncol.*, 4:86 (2014)) and can be used as potentiating agents in accordance with the disclosed methods.

#### **[0071] A. Cell-Penetrating Antibodies**

**[0072]** In some embodiments, the potentiating agent is a cell-penetrating antibody. Although the cell-penetrating molecules are generally referred to herein as "cell-penetrating antibodies," it will be appreciated that fragments and binding proteins, including antigen-binding fragments, variants, and fusion proteins such as scFv, di-scFv, tr-scFv, and other single chain variable fragments, and other cell-penetrating molecules disclosed herein are encompassed by the phrase also expressly provided for use in compositions and methods disclosed herein.

**[0073]** Cell-penetrating antibodies for use in the compositions and methods may be anti-DNA antibodies. The cell-penetrating antibody may bind single stranded DNA and/or double stranded DNA. The cell-penetrating antibody may be an anti-RNA antibody (e.g., the antibody specifically binds RNA).

**[0074]** Autoantibodies to double-stranded deoxyribonucleic acid (dsDNA) are frequently identified in the serum of patients with systemic lupus erythematosus (SLE) and are often implicated in disease pathogenesis. Therefore, in some embodiments, cell-penetrating antibodies (e.g., cell-penetrating anti-DNA antibodies) can be derived or isolated from patients with SLE or animal models of SLE.

**[0075]** In preferred embodiments, the anti-DNA antibodies are monoclonal antibodies, or antigen binding fragments or variants thereof. In some embodiments, the anti-DNA antibodies are conjugated to a cell-penetrating moiety, such as a cell penetrating peptide to facilitate entry into the cell and transport to the cytoplasm and/or nucleus. Examples of cell penetrating peptides include, but are not limited to, Polyarginine (e.g., R9), Antennapedia sequences, TAT, HIV-Tat, Penetratin, Antp-3A (Antp mutant), Buforin II, Transportan, MAP (model amphipathic peptide), K-FGF, Ku70, Prion, pVEC, Pep-1, SynB1, Pep-7, HN-1, BGSC (Bis-Guanidinium-Spermidine-Cholesterol, and BGTC (Bis-Guanidinium-Tren-Cholesterol). In other embodiments, the antibody is modified using TransMabs™ technology (In-Nexus Biotech., Inc., Vancouver, BC).

**[0076]** In preferred embodiments, the anti-DNA antibody is transported into the cytoplasm and/or nucleus of the cells without the aid of a carrier or conjugate. For example, the monoclonal antibody 3E10 and active fragments thereof that are transported in vivo to the nucleus of mammalian cells without cytotoxic effect are disclosed in U.S. Pat. Nos. 4,812,397 and 7,189,396 to Richard Weisbart. Briefly, the antibodies may be prepared by fusing spleen cells from a host having elevated serum levels of anti-DNA antibodies (e.g., MRL/lpr mice) with myeloma cells in accordance with known techniques or by transforming the spleen cells with an appropriate transforming vector to immortalize the cells. The cells may be cultured in a selective medium and screened to select antibodies that bind DNA.

**[0077]** In some embodiments, the cell-penetrating antibody may bind and/or inhibit Rad51. See for example, the cell-penetrating antibody described in Turchick, et al., *Nucleic Acids Res.*, 45(20): 11782-11799 (2017).

[0078] Antibodies that can be used in the compositions and methods include whole immunoglobulin (i.e., an intact antibody) of any class, fragments thereof, and synthetic proteins containing at least the antigen binding variable domain of an antibody. The variable domains differ in sequence among antibodies and are used in the binding and specificity of each particular antibody for its particular antigen. However, the variability is not usually evenly distributed through the variable domains of antibodies. It is typically concentrated in three segments called complementarity determining regions (CDRs) or hypervariable regions both in the light chain and the heavy chain variable domains. The more highly conserved portions of the variable domains are called the framework (FR). The variable domains of native heavy and light chains each comprise four FR regions, largely adopting a beta-sheet configuration, connected by three CDRs, which form loops connecting, and in some cases forming part of, the beta-sheet structure. The CDRs in each chain are held together in close proximity by the FR regions and, with the CDRs from the other chain, contribute to the formation of the antigen binding site of antibodies. Therefore, the antibodies typically contain at least the CDRs necessary to maintain DNA binding and/or interfere with DNA repair.

[0079] 1. 3E10 Sequences

[0080] In some embodiments, the cell-penetrating anti-DNA antibody is the monoclonal anti-DNA antibody 3E10, or a variant, derivative, fragment, or humanized form thereof that binds the same or different epitope(s) as 3E10. Thus, the cell-penetrating anti-DNA antibody may have the same or different epitope specificity as monoclonal antibody 3E10 produced by ATCC No. PTA 2439 hybridoma. The anti-DNA antibody can have the paratope of monoclonal antibody 3E10. The anti-DNA antibody can be a single chain variable fragment of an anti-DNA antibody, or conservative variant thereof. For example, the anti-DNA antibody can be a single chain variable fragment of 3E10 (3E10 Fv), or a variant thereof.

[0081] Amino acid sequences of monoclonal antibody 3E10 are known in the art. For example, sequences of the 3E10 heavy and light chains are provided below, where single underlining indicates the CDR regions identified according to the Kabat system, and in SEQ ID NOS:12-14 italics indicates the variable regions and double underlining indicates the signal peptide. CDRs according to the IMGT system are also provided.

[0082] a. 3E10 Heavy Chain

[0083] In some embodiments, a heavy chain variable region of 3E10 is: EVQLVESGGGLVKPGGSRKLS-CAASGFTFSDYGMHWVRQAPEKGLEWVA YISSGSSTIYYADTVKGRFTISRDNAKNTLFLQMTSLRSED

[0084] In some embodiments, a 3E10 heavy chain is expressed as

(3E10 WT Heavy Chain; SEQ ID NO: 12)  
MGWSCIIILFLVATATGVHSEVQLVESGGGLVKPGG  
SRKLSCAASGFTFSDYGMHWVRQAPEKGLEWVAI

-continued

SSGSSTIYYADTVKGRFTISRDNAKNTLFLQMTSL  
RSED

TAMYYCARRGLLLDYWGQGTTLTVSAASTKG

PSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVS

WNSGALTSGVHTFPAVLQSSGLYSLSSVTVPS

LGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP

PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCV

VVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYN

STYRVVSVLT

TVLHQDWLNGKEYKCKVSNKALPAPI

EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT

LVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDG

SFFLYSKLTVDKSRWQQGNVFC

CSVMHEALHNHYT

QKSLSLSPGK.

[0085] Variants of the 3E10 antibody which incorporate mutations into the wild type sequence are also known in the art, as disclosed for example, in Zack, et al., *J. Immunol.*, 157(5):2082-8 (1996). For example, amino acid position 31 of the heavy chain variable region of 3E10 has been determined to be influential in the ability of the antibody and fragments thereof to penetrate nuclei and bind to DNA (bolded in SEQ ID NOS:1, 2 and 13). A D31N mutation (bolded in SEQ ID NOS:2 and 13) in CDR1 penetrates nuclei and binds DNA with much greater efficiency than the original antibody (Zack, et al., *Immunology and Cell Biology*, 72:513-520 (1994), Weisbart, et al., *J. Autoimmun.*, 11, 539-546 (1998); *Weisbart, Int. J. Oncol.*, 25, 1867-1873 (2004)).

[0086] In some embodiments, an amino acid sequence for a preferred variant of a heavy chain variable region of 3E10 is:

(SEQ ID NO: 2)  
EVQLVESGGGLVKPGGSRKLSCAASGFTFSNYGMHW  
VRQAPEKGLEWVAI

ISSGSSTIYYADTVKGRFTISR

DNAKNTLFLQMTSLRSED

TAMYYCARRGLLLDYWGQ

GTTLTVSS

[0087] In some embodiments, a 3E10 heavy chain is expressed as

(3E10 D31N Variant Heavy Chain; SEQ ID NO: 13)  
MGWSCIIILFLVATATGVHSEVQLVESGGGLVKPGG

SRKLSCAASGFTFSNYGMHWVRQAPEKGLEWVAI

SSGSSTIYYADTVKGRFTISRDNAKNTLFLQMTSL

RSED

TAMYYCARRGLLLDYWGQGTTLTVSAASTKG

PSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVS

WNSGALTSGVHTFPAVLQSSGLYSLSSVTVPS

LGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP

-continued  
PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCV  
VVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYN  
STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI  
EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTC  
LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDG  
SFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYT  
QKSLSLSPGK.

[0088] In some embodiments, the C-terminal serine of SEQ ID NOS:1 or 2 is absent or substituted, with, for example, an alanine, in 3E10 heavy chain variable region.

[0089] The complementarity determining regions (CDRs) as identified by Kabat are shown with underlining above and include CDR H1.1 (original sequence): DYGMH (SEQ ID NO:15); CDR H1.2 (with D31N mutation): NYGMH (SEQ ID NO:16); CDR H2.1: YISSGSSTIYYADTVKG (SEQ ID NO:17); CDR H3.1: RGLLLDY (SEQ ID NO:18).

[0090] A variant of Kabat CDR H2.1 is YISSGSSTIYY-ADSVKG (SEQ ID NO:19).

[0091] Additionally, or alternatively, the heavy chain complementarity determining regions (CDRs) can be defined according to the IMGT system. The complementarity determining regions (CDRs) as identified by the IMGT system include CDR H1.3 (original sequence): GFTFSDYG (SEQ ID NO:20); CDR H1.4 (with D31N mutation): GFTFSNYG (SEQ ID NO:21); CDR H2.2: ISSGSSTI (SEQ ID NO:22); CDR H3.2: ARRGLLLDY (SEQ ID NO:23).

[0092] b. 3E10 Light Chain

[0093] In some embodiments, a light chain variable region of 3E10 is:

(SEQ ID NO: 7)  
DIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSYMHYQQKP  
GQPPKLLIKYASYLESQVPAFSGSGSGTDFTLNIHPVEEEDAA  
TYYCQHSREFPWTFGGGTKLEIK.

[0094] An amino acid sequence for the light chain variable region of 3E10 can also be:

(SEQ ID NO: 8)  
DIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSYMHYQQKP  
GQPPKLLIKYASYLESQVPAFSGSGSGTDFTLNIHPVEEEDAA  
TYYCQHSREFPWTFGGGTKLELK.

[0095] In some embodiments, a 3E10 light chain is expressed as

(3E10 WT Light Chain; SEQ ID NO: 14)  
MGWSCIIILFLVATATGVHSDIVLTQSPASLAVSLGQRATISCR  
RASKSVSTSSYSYMHYQQKPGQPPKLLIKYASYLESQVPA  
RFGSGSGSGTDFTLNIHPVEEEDAAATYYCQHSREFPWTFGGG

-continued  
TKLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPR  
EAKVQWKVDNALQSGNSQESVTEQDSKDSYSTLSSTLTLSK  
ADYEKHKVYACEVTHQGLSPVTKSFNRGEC

[0096] Other 3E10 light chain sequences are known in the art. See, for example, Zack, et al., *J. Immunol.*, 15; 154(4): 1987-94 (1995); GenBank: L16981.1—Mouse Ig rearranged L-chain gene, partial cds; GenBank: AAA65681.1—immunoglobulin light chain, partial [*Mus musculus*]).

[0097] The complementarity determining regions (CDRs) as identified by Kabat are shown with underlining, including

CDR L1.1:  
(SEQ ID NO: 24)  
RASKSVSTSSYSYMH;  
CDR L2.1:  
(SEQ ID NO: 25)  
YASYLES;  
CDR L3.1:  
(SEQ ID NO: 26)  
QHSREFPWT.

[0098] A variant of Kabat CDR L1.1 is RASKSVSTSSY-SYLA (SEQ ID NO:27).

[0099] A variant of Kabat CDR L2.1 is YASYLQS (SEQ ID NO:28).

[0100] Additionally, or alternatively, the heavy chain complementarity determining regions (CDRs) can be defined according to the IMGT system. The complementarity determining regions (CDRs) as identified by the IMGT system include CDR L1.2 KSVSTSSYSY (SEQ ID NO:29); CDR L2.2: YAS (SEQ ID NO:30); CDR L3.2: QHS-REFPWT (SEQ ID NO:26).

[0101] In some embodiments, the C-terminal end of sequence of SEQ ID NOS:7 or 8 further includes an arginine in the 3E10 light chain variable region.

[0102] 2. Humanized 3E10

[0103] In some embodiments, the antibody is a humanized antibody. Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. These non-human amino acid residues are often referred to as “import” residues, which are typically taken from an “import” variable domain. Antibody humanization techniques generally involve the use of recombinant DNA technology to manipulate the DNA sequence encoding one or more polypeptide chains of an antibody molecule.

[0104] Exemplary 3E10 humanized sequences are discussed in WO 2015/106290 and WO 2016/033324, and provided below.

**[0105]** a. Humanized 3E10 Heavy Chain Variable Regions**[0106]** In some embodiments, a humanized 3E10 heavy chain variable domain includes

(hVH1, SEQ ID NO: 3)  
EVQLVQSGGGLIQPGGSLRLSCAASGFTFSNYGMHWVRQA

PGKGLEWVSYSISGSSTIYYADSVKGRFTISRDNKNTLY

LQMNSLRAEDTAVYYCARRGLLLDYWGQGTTLTVSS,  
or

(hVH2, SEQ ID NO: 4)  
EVQLVESGGGLIQPGGSLRLSCAASGFTFSNYGMHWVRQA

PGKGLEWVSYSISGSSTIYYADSVKGRFTISRDNKNTLY

LQMTSLRAEDTAVYYCARRGLLLDYWGQGTTLTVSS,  
or

(hVH3, SEQ ID NO: 5)  
EVQLQESGGGVVQPGGSLRLSCAASGFTFSNYGMHWIRQA

PGKGLEWVSYSISGSSTIYYADSVKGRFTISRDNKNTLY

LQMNSLRSED TAVYYCARRGLLLDYWGQGTTLTVSS

(hVH4, SEQ ID NO: 6)  
EVQLVESGGGLVQPGGSLRLSCSASGFTFSNYGMHWVRQA

PGKGLEWVSYSISGSSTIYYADTVKGRFTISRDNKNTLY

LQMSSLRAEDTAVYYCVKRGGLLLDYWGQGTTLTVSS

**[0107]** b. Humanized 3E10 Light Chain Variable Regions**[0108]** In some embodiments, a humanized 3E10 light chain variable domain includes

(hVL1, SEQ ID NO: 9)  
DIQMTQSPSSLSASVGRVTITCRASKSVSTSSSYSLAWY

QQKPEKAPKLLIKYASYLQSGVPSRFSGSGSGTDFTLTIS

SLQPEDFATYYCQHSREFPWTFGAGTKLELK,  
or

(hVL2, SEQ ID NO: 10)  
DIQMTQSPSSLSASVGRVTISCRASKSVSTSSYSYMHY

QQKPEKAPKLLIKYASYLQSGVPSRFSGSGSGTDFTLTIS

SLQPEDVATYYCQHSREFPWTFGAGTKLELK,  
or

(hVL3, SEQ ID NO: 11)  
DIVLTQSPASLAVSPGQRATITCRASKSVSTSSYSYMHY

QQKPGQPPKLLIYYASYLESVGPARGSGSGSGTDFTLTIN

PVEANDTANYYCQHSREFPWTFGQGTKVEIK

**[0109]** 3. Fragments, Variants, and Fusion Proteins**[0110]** The anti-DNA antibody can be composed of an antibody fragment or fusion protein including an amino acid sequence of a variable heavy chain and/or variable light chain that is at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% identical to the amino acid sequence of the variable heavy chain and/or light chain of 3E10 or a humanized form thereof (e.g., any of SEQ ID NOS:1-11, or the heavy and/or light chains of any of SEQ ID NOS:12-14).**[0111]** The anti-DNA antibody can be composed of an antibody fragment or fusion protein that includes one or more CDR(s) that is at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% identical to the amino acid sequence of the CDR(s) of 3E10, or a variant or humanized form thereof (e.g., CDR(s) of any of SEQ ID NOS:1-11, or SEQ ID NOS:12-14, or SEQ ID NOS:15-30). The determination of percent identity of two amino acid sequences can be determined by BLAST protein comparison. In some embodiments, the antibody includes one, two, three, four, five, or all six of the CDRs of the above-described preferred variable domains.**[0112]** Preferably, the antibody include one of each of a heavy chain CDR1, CDR2, and CDR3 in combination with one of each of a light chain CDR1, CDR2, and CDR3.**[0113]** Predicted complementarity determining regions (CDRs) of the light chain variable sequence for 3E10 are provided above. See also GenBank: AAA65681.1—immunoglobulin light chain, partial [*Mus musculus*] and GenBank: L34051.1—Mouse Ig rearranged kappa-chain mRNA V-region. Predicted complementarity determining regions (CDRs) of the heavy chain variable sequence for 3E10 are provide above. See also, for example, Zack, et al., *Immunology and Cell Biology*, 72:513-520 (1994), GenBank Accession number AAA65679.1. Zach, et al., *J. Immunol.* 154 (4), 1987-1994 (1995) and GenBank: L16982.1—Mouse Ig reagrranged H-chain gene, partial cds.**[0114]** Thus, in some embodiments, the cell-penetrating antibody contains the CDRs, or the entire heavy and light chain variable regions, of SEQ ID NO:1 or 2, or the heavy chain region of SEQ ID NO:12 or 13; or a humanized form thereof in combination with SEQ ID NO:7 or 8, or the light chain region of SEQ ID NO:14; or a humanized form thereof. In some embodiments, the cell-penetrating antibody contains the CDRs, or the entire heavy and light chain variable regions, of SEQ ID NO:3, 4, 5, or 6 in combination with SEQ ID NO:9, 10, or 11.**[0115]** Also included are fragments of antibodies which have bioactivity. The fragments, whether attached to other sequences or not, include insertions, deletions, substitutions, or other selected modifications of particular regions or specific amino acids residues, provided the activity of the fragment is not significantly altered or impaired compared to the nonmodified antibody or antibody fragment.**[0116]** Techniques can also be adapted for the production of single-chain antibodies specific to an antigenic protein of the present disclosure. Methods for the production of single-chain antibodies are well known to those of skill in the art. A single chain antibody can be created by fusing together the variable domains of the heavy and light chains using a short peptide linker, thereby reconstituting an antigen binding site on a single molecule. Single-chain antibody variable fragments (scFvs) in which the C-terminus of one variable domain is tethered to the N-terminus of the other variable domain via a 15 to 25 amino acid peptide or linker have been developed without significantly disrupting antigen binding or specificity of the binding. The linker is chosen to permit the heavy chain and light chain to bind together in their proper conformational orientation.**[0117]** The anti-DNA antibodies can be modified to improve their therapeutic potential. For example, in some embodiments, the cell-penetrating anti-DNA antibody is conjugated to another antibody specific for a second thera-

peutic target in the cytoplasm and/or nucleus of a target cell. For example, the cell-penetrating anti-DNA antibody can be a fusion protein containing 3E10 Fv and a single chain variable fragment of a monoclonal antibody that specifically binds the second therapeutic target. In other embodiments, the cell-penetrating anti-DNA antibody is a bispecific antibody having a first heavy chain and a first light chain from 3E10 and a second heavy chain and a second light chain from a monoclonal antibody that specifically binds the second therapeutic target.

**[0118]** Divalent single-chain variable fragments (di-scFvs) can be engineered by linking two scFvs. This can be done by producing a single peptide chain with two VH and two VL regions, yielding tandem scFvs. ScFvs can also be designed with linker peptides that are too short for the two variable regions to fold together (about five amino acids), forcing scFvs to dimerize. This type is known as diabodies. Diabodies have been shown to have dissociation constants up to 40-fold lower than corresponding scFvs, meaning that they have a much higher affinity to their target. Still shorter linkers (one or two amino acids) lead to the formation of trimers (triabodies or tribodies). Tetraabodies have also been produced. They exhibit an even higher affinity to their targets than diabodies. In some embodiments, the anti-DNA antibody may contain two or more linked single chain variable fragments of 3E10 (e.g., 3E10 di-scFv, 3E10 tri-scFv), or conservative variants thereof. In some embodiments, the anti-DNA antibody is a diabody or triabody (e.g., 3E10 diabody, 3E10 triabody). Sequences for single and two or more linked single chain variable fragments of 3E10 are provided in WO 2017/218825 and WO 2016/033321.

**[0119]** The function of the antibody may be enhanced by coupling the antibody or a fragment thereof with a therapeutic agent. Such coupling of the antibody or fragment with the therapeutic agent can be achieved by making an immunoconjugate or by making a fusion protein, or by linking the antibody or fragment to a nucleic acid such as DNA or RNA (e.g., siRNA), comprising the antibody or antibody fragment and the therapeutic agent.

**[0120]** A recombinant fusion protein is a protein created through genetic engineering of a fusion gene. This typically involves removing the stop codon from a cDNA sequence coding for the first protein, then appending the cDNA sequence of the second protein in frame through ligation or overlap extension PCR. The DNA sequence will then be expressed by a cell as a single protein. The protein can be engineered to include the full sequence of both original proteins, or only a portion of either. If the two entities are proteins, often linker (or “spacer”) peptides are also added which make it more likely that the proteins fold independently and behave as expected.

**[0121]** In some embodiments, the cell-penetrating antibody is modified to alter its half-life. In some embodiments, it is desirable to increase the half-life of the antibody so that it is present in the circulation or at the site of treatment for longer periods of time. For example, it may be desirable to maintain titers of the antibody in the circulation or in the location to be treated for extended periods of time. In other embodiments, the half-life of the anti-DNA antibody is decreased to reduce potential side effects. Antibody fragments, such as 3E10Fv may have a shorter half-life than full size antibodies. Other methods of altering half-life are known and can be used in the described methods. For example, antibodies can be engineered with Fc variants that

extend half-life, e.g., using Xtend™ antibody half-life prolongation technology (Xencor, Monrovia, Calif.).

**[0122]** a. Linkers

**[0123]** The term “linker” as used herein includes, without limitation, peptide linkers. The peptide linker can be any size provided it does not interfere with the binding of the epitope by the variable regions. In some embodiments, the linker includes one or more glycine and/or serine amino acid residues. Monovalent single-chain antibody variable fragments (scFvs) in which the C-terminus of one variable domain are typically tethered to the N-terminus of the other variable domain via a 15 to 25 amino acid peptide or linker. The linker is chosen to permit the heavy chain and light chain to bind together in their proper conformational orientation. Linkers in diabodies, triabodies, etc., typically include a shorter linker than that of a monovalent scFv as discussed above. Di-, tri-, and other multivalent scFvs typically include three or more linkers. The linkers can be the same, or different, in length and/or amino acid composition. Therefore, the number of linkers, composition of the linker(s), and length of the linker(s) can be determined based on the desired valency of the scFv as is known in the art. The linker(s) can allow for or drive formation of a di-, tri-, and other multivalent scFv.

**[0124]** For example, a linker can include 4-8 amino acids. In a particular embodiment, a linker includes the amino acid sequence GQSSRSS (SEQ ID NO:31). In another embodiment, a linker includes 15-20 amino acids, for example, 18 amino acids. In a particular embodiment, the linker includes the amino acid sequence GQSSRSSGGGSSGGGGS (SEQ ID NO:32). Other flexible linkers include, but are not limited to, the amino acid sequences Gly-Ser, Gly-Ser-Gly-Ser (SEQ ID NO:33), Ala-Ser, Gly-Gly-Gly-Ser (SEQ ID NO:34), (Gly<sub>4</sub>-Ser)<sub>2</sub> (SEQ ID NO:35) and (Gly<sub>4</sub>-Ser)<sub>4</sub> (SEQ ID NO:36), and (Gly-Gly-Gly-Gly-Ser)<sub>3</sub> (SEQ ID NO:37).

**[0125]** b. Exemplary Anti-DNA scFv Sequences

**[0126]** Exemplary murine 3E10 scFv sequences, including mono-, di-, and tri-scFv are disclosed in WO 2016/033321 and WO 2017/218825 and provided below. Cell-penetrating antibodies for use in the disclosed compositions and methods include exemplary scFv, and fragments and variants thereof.

**[0127]** The amino acid sequence for scFv 3E10 (D31N) is:

(SEQ ID NO: 38)

```
AGIHDIVL TQSPASLAVSLGQRATISCRASKSVSTSSYSY
MHWYQQKPGQPPKLLIKYASYLESQVPAFSGSGSGTDFT
LNIHPVEEEDAATYYCQHSREFPWTFGGGTKLEIKRADAA
PGGGSGGGSGGGSGGGSEVQLVESGGGLVKPGGSRKLSCAA
SGFTFSNYGMHWVRQAPEKGLEWVAYISSGSSTIYYADTV
KGRFTISRDNKNTLFLQMTSLRSED TAMY CARRGLLLD
YWGQGTTLTVSSLEQKLI SEEDLNSAVDHHHHHH.
```

**[0128]** Annotation of scFv Protein Domains with Reference to SEQ ID NO:38

**[0129]** AGIH sequence increases solubility (amino acids 1-4 of SEQ ID NO:38)

**[0130]** Vk variable region (amino acids 5-115 of SEQ ID NO:38) Initial (6 aa) of light chain CH1 (amino acids 116-121 of SEQ ID NO:38)

[0131] (GGGGS)<sub>3</sub> (SEQ ID NO:37) linker (amino acids 122-136 of SEQ ID NO:38)  
[0132] VH variable region (amino acids 137-252 of SEQ ID NO:38)  
[0133] Myc tag (amino acids 253-268 SEQ ID NO:38)  
[0134] His 6 tag (amino acids 269-274 of SEQ ID NO:38)  
[0135] Amino Acid Sequence of 3E10 Di-scFv (D31N)  
[0136] Di-scFv 3E10 (D31N) is a di-single chain variable fragment including 2X the heavy chain and light chain variable regions of 3E10 and wherein the aspartic acid at position 31 of the heavy chain is mutated to an asparagine. The amino acid sequence for di-scFv 3E10 (D31N) is:

(SEQ ID NO: 39)  
AGIH DIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSY  
MHWYQQKPGQPPKLLIKYASYLESVGPARGSGSGTDFT  
LNIHPVEEEDAATYYCQHSREFPWTFGGGTKLEIKRADAA  
PGGGSGGGSGGGGSEVQLVESGGGLVKPGGSRKLSCAA  
SGFTFSNYGMHWVRQAPEKGLEWVAYISSGSSTIYYADTV  
KGRFTISRDNAKNTLFLQMTSLRSED TAMYYCARRGLLLD  
YWGQGTTLTVSSASTKGPSVFPLAPLESSGSDIVLTQSPA  
SLAVSLGQRATISCRASKSVSTSSYSYMHWYQQKPGQPPK  
LLIKYASYLESVGPARGSGSGTDFTLNIHPVEEEDAAT  
YYCQHSREFPWTFGGGTKLEIKRADAAPGGGSGGGGSGG  
GGSEVQLVESGGGLVKPGGSRKLSCAASGFTFSNYGMHW  
RQAPEKGLEWVAYISSGSSTIYYADTVKGRFTISRDNAKN  
TLFLQMTSLRSED TAMYYCARRGLLLDYWGQGTTLTVSSL  
EQKLISEEDLNSAVDHHHHHH.

[0137] Annotation of di-scFv Protein Domains with Reference to SEQ ID NO:39  
[0138] AGIH sequence increases solubility (amino acids 1-4 of SEQ ID NO:39)  
[0139] Vk variable region (amino acids 5-115 of SEQ ID NO:39)  
[0140] Initial (6 aa) of light chain CH1 (amino acids 116-121 of SEQ ID NO:39)  
[0141] (GGGGS)<sub>3</sub> (SEQ ID NO:37) linker (amino acids 122-136 of SEQ ID NO:39)  
[0142] VH variable region (amino acids 137-252 of SEQ ID NO:39)  
[0143] Linker between Fv fragments consisting of human IgG CH1 initial 13 amino acids (amino acids 253-265 of SEQ ID NO:39)  
[0144] Swivel sequence (amino acids 266-271 of SEQ ID NO:39)  
[0145] Vk variable region (amino acids 272-382 of SEQ ID NO:39)  
[0146] Initial (6 aa) of light chain CH1 (amino acids 383-388 of SEQ ID NO:39)  
[0147] (GGGGS)<sub>3</sub> (SEQ ID NO:37) linker (amino acids 389-403 of SEQ ID NO:39)  
[0148] VH variable region (amino acids 404-519 of SEQ ID NO:39)

[0149] Myc tag (amino acids 520-535 of SEQ ID NO:39)  
[0150] His 6 tag (amino acids 536-541 of SEQ ID NO:39)  
[0151] Amino Acid Sequence for Tri-scFv  
[0152] Tri-scFv 3E10 (D31N) is a tri-single chain variable fragment including 3X the heavy chain and light chain variable regions of 310E and wherein the aspartic acid at position 31 of the heavy chain is mutated to an asparagine. The amino acid sequence for tri-scFv 3E10 (D31N) is:

(SEQ ID NO: 40)  
AGIH DIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSY  
MHWYQQKPGQPPKLLIKYASYLESVGPARGSGSGTDFT  
LNIHPVEEEDAATYYCQHSREFPWTFGGGTKLEIKRADAA  
PGGGSGGGSGGGGSEVQLVESGGGLVKPGGSRKLSCAA  
SGFTFSNYGMHWVRQAPEKGLEWVAYISSGSSTIYYADTV  
KGRFTISRDNAKNTLFLQMTSLRSED TAMYYCARRGLLLD  
YWGQGTTLTVSSASTKGPSVFPLAPLESSGSDIVLTQSPA  
SLAVSLGQRATISCRASKSVSTSSYSYMHWYQQKPGQPPK  
LLIKYASYLESVGPARGSGSGTDFTLNIHPVEEEDAAT  
YYCQHSREFPWTFGGGTKLEIKRADAAPGGGSGGGGSGG  
GGSEVQLVESGGGLVKPGGSRKLSCAASGFTFSNYGMHW  
RQAPEKGLEWVAYISSGSSTIYYADTVKGRFTISRDNAKN  
TLFLQMTSLRSED TAMYYCARRGLLLDYWGQGTTLTVSSA  
STKGPSVFPLAPLESSGSDIVLTQSPASLAVSLGQRATIS  
CRASKSVSTSSYSYMHWYQQKPGQPPKLLIKYASYLESV  
PARFSGSGSGTDFTLNIHPVEEEDAATYYCQHSREFPWTF  
GGGTKLEIKRADAAPGGGSGGGGSGGGGSEVQLVESGGG  
LVKPGGSRKLSCAASGFTFSNYGMHWVRQAPEKGLEWVAY  
ISSGSSTIYYADTVKGRFTISRDNAKNTLFLQMTSLRSED  
TAMYYCARRGLLLDYWGQGTTLTVSSLEQKLISEEDLNSA  
VDHHHHHH.

[0153] Annotation of Tri-scFv Protein Domains with Reference to SEQ ID NO:40  
[0154] AGIH sequence increases solubility (amino acids 1-4 of SEQ ID NO:40)  
[0155] Vk variable region (amino acids 5-115 of SEQ ID NO:40)  
[0156] Initial (6 aa) of light chain CH1 (amino acids 116-121 of SEQ ID NO:40)  
[0157] (GGGGS)<sub>3</sub> (SEQ ID NO:37) linker (amino acids 122-136 of SEQ ID NO:40)  
[0158] VH variable region (amino acids 137-252 of SEQ ID NO:40)  
[0159] Linker between Fv fragments consisting of human IgG CH1 initial 13 amino acids (amino acids 253-265 of SEQ ID NO:40)  
[0160] Swivel sequence (amino acids 266-271 of SEQ ID NO:40)

[0161] Vk variable region (amino acids 272-382 of SEQ ID NO:40)

[0162] Initial (6 aa) of light chain CH1 (amino acids 383-388 of SEQ ID NO:40)

[0163] (GGGGS)<sub>3</sub> (SEQ ID NO:37) linker (amino acids 389-403 of SEQ ID NO:40)

[0164] VH variable region (amino acids 404-519 of SEQ ID NO:40)

[0165] Linker between Fv fragments consisting of human IgG CH1 initial 13 amino acids (amino acids 520-532 of SEQ ID NO:40)

[0166] Swivel sequence (amino acids 533-538 of SEQ ID NO:40)

[0167] Vk variable region (amino acids 539-649 of SEQ ID NO:40)

[0168] Initial (6 aa) of light chain CH1 (amino acids 650-655 of SEQ ID NO:40)

[0169] (GGGGS)<sub>3</sub> (SEQ ID NO:37) linker (amino acids 656-670 of SEQ ID NO:40)

[0170] VH variable region (amino acids 671-786 of SEQ ID NO:40)

[0171] Myc tag (amino acids 787-802 of SEQ ID NO:40)

[0172] His 6 tag (amino acids 803-808 of SEQ ID NO:40)

[0173] WO 2016/033321 and Noble, et al., *Cancer Research*, 75(11):2285-2291 (2015), show that di-scFv and tri-scFv have some improved and additional activities compared to their monovalent counterpart. The subsequences corresponding to the different domains of each of the exemplary fusion proteins are also provided above. One of skill in the art will appreciate that the exemplary fusion proteins, or domains thereof, can be utilized to construct fusion proteins discussed in more detail above. For example, in some embodiments, the di-scFv includes a first scFv including a Vk variable region (e.g., amino acids 5-115 of SEQ ID NO:39, or a functional variant or fragment thereof), linked to a VH variable domain (e.g., amino acids 137-252 of SEQ ID NO:39, or a functional variant or fragment thereof), linked to a second scFv including a Vk variable region (e.g., amino acids 272-382 of SEQ ID NO:39, or a functional variant or fragment thereof), linked to a VH variable domain (e.g., amino acids 404-519 of SEQ ID NO:39, or a functional variant or fragment thereof). In some embodiments, a tri-scFv includes a di-scFv linked to a third scFv domain including a Vk variable region (e.g., amino acids 539-649 of SEQ ID NO:40, or a functional variant or fragment thereof), linked to a VH variable domain (e.g., amino acids 671-786 of SEQ ID NO:40, or a functional variant or fragment thereof).

[0174] The Vk variable regions can be linked to VH variable domains by, for example, a linker (e.g., (GGGGS)<sub>3</sub> (SEQ ID NO:37), alone or in combination with a (6 aa) of light chain CH1 (amino acids 116-121 of SEQ ID NO:39). Other suitable linkers are discussed above and known in the art. scFv can be linked by a linker (e.g., human IgG CH1 initial 13 amino acids (253-265) of SEQ ID NO:39), alone or in combination with a swivel sequence (e.g., amino acids 266-271 of SEQ ID NO:39). Other suitable linkers are discussed above and known in the art.

[0175] Therefore, a di-scFv can include amino acids 5-519 of SEQ ID NO:39. A tri-scFv can include amino acids 5-786 of SEQ ID NO:40. In some embodiments, the fusion proteins include additional domains. For example, in some

embodiments, the fusion proteins include sequences that enhance solubility (e.g., amino acids 1-4 of SEQ ID NO:39). Therefore, in some embodiments, a di-scFv can include amino acids 1-519 of SEQ ID NO:39. A tri-scFv can include amino acids 1-786 of SEQ ID NO:40. In some embodiments that fusion proteins include one or more domains that enhance purification, isolation, capture, identification, separation, etc., of the fusion protein. Exemplary domains include, for example, Myc tag (e.g., amino acids 520-535 of SEQ ID NO:39) and/or a His tag (e.g., amino acids 536-541 of SEQ ID NO:39). Therefore, in some embodiments, a di-scFv can include the amino acid sequence of SEQ ID NO:39. A tri-scFv can include the amino acid sequence of SEQ ID NO:40. Other substitutable domains and additional domains are discussed in more detail above.

[0176] An exemplary 3E10 humanized Fv sequence is discussed in WO 2016/033324:

(SEQ ID NO: 41)

```

DIVLTQSPASLAVSPGQRATITCRASKSVSTSSYSYMHVY
QQKPGQPPKLLIYYASYLESVGPARGSGSGSGTDFTLTIN
PVEANDTANYYCQHSREFPWTFGQGTKVEIKGGGSGGGG
SGGGGSEVQLVESGGGLVQPGGSLRLSCSASGFTFSNYGM
HWVRQAPGKGLYVSYISSGSSTIYYADTVKGRFTISRDN
SKNTLYLQMSSLRAEDTAVYYCVKRGLLLDYWGQGLTVTV
SS.

```

[0177] B. Donor Oligonucleotides

[0178] Donor oligonucleotides are provided for use in the compositions and methods. In some embodiments, the composition includes or is administered in combination with a donor oligonucleotide. The donor oligonucleotide may or may not be covalently linked to the potentiating agent. For example, the donor oligonucleotide may form a non-covalent complex with the cell-penetrating antibody.

[0179] Donor oligonucleotide strategies include, but are not limited to, small fragment homologous replacement (e.g., polynucleotide small DNA fragments (SDFs)) and single-stranded oligodeoxynucleotide-mediated gene modification (e.g., ssODN/SSOs).

[0180] Generally, in the case of gene therapy, the donor oligonucleotide includes a sequence that can correct a mutation(s) in the host genome, though in some embodiments, the donor introduces a mutation that can, for example, reduce expression of an oncogene or a receptor that facilitates HIV infection. In addition to containing a sequence designed to introduce the desired correction or mutation, the donor oligonucleotide may also contain synonymous (silent) mutations (e.g., 7 to 10). The additional silent mutations can facilitate detection of the corrected target sequence using allele-specific PCR of genomic DNA isolated from treated cells.

[0181] The donor oligonucleotide can exist in single stranded (ss) or double stranded (ds) form (e.g., ssDNA, dsDNA). Thus, the oligonucleotide (e.g., DNA or RNA, or combination thereof) may be single stranded or double stranded. Preferably, the oligonucleotide is single stranded DNA.

[0182] The donor oligonucleotide can be of any length. For example, the size of the donor oligonucleotide may be

between 1 to 800 nucleotides. In one embodiment, the donor oligonucleotide is between 25 and 200 nucleotides. In some embodiments, the donor oligonucleotide is between 100 and 150 nucleotides. In a further embodiment, the donor nucleotide is about 40 to 80 nucleotides in length. The donor oligonucleotide may be about 60 nucleotides in length. ssDNAs of length 25-200 are active. Most studies have been with ssDNAs of length 60-70. Longer ones as 70-150 also work. The preferred length is 60.

**[0183]** The donor sequence can contain one or more nucleic acid sequence alterations compared to the sequence of the region targeted for recombination, for example, a substitution, a deletion, or an insertion of one or more nucleotides. Successful recombination of the donor sequence results in a change of the sequence of the target region. Donor oligonucleotides are also referred to herein as donor fragments, donor nucleic acids, donor DNA, or donor DNA fragments. It is understood in the art that a greater number of homologous positions within the donor fragment will increase the probability that the donor fragment will be recombined into the target sequence, target region, or target site.

**[0184]** The donor oligonucleotides may contain at least one mutated, inserted or deleted nucleotide relative to the target DNA sequence. Target sequences can be within the coding DNA sequence of the gene or within introns. Target sequences can also be within DNA sequences which regulate expression of the target gene, including promoter or enhancer sequences or sequences that regulate RNA splicing.

**[0185]** The donor oligonucleotides can contain a variety of mutations relative to the target sequence. Representative types of mutations include, but are not limited to, point mutations, deletions and insertions. Deletions and insertions can result in frameshift mutations or deletions. Point mutations can cause missense or nonsense mutations. These mutations may disrupt, reduce, stop, increase, improve, or otherwise alter the expression of the target gene.

**[0186]** The donor oligonucleotide may correspond to the wild type sequence of a gene (or a portion thereof), for example, a mutated gene involved with a disease or disorder (e.g., hemophilia, muscular dystrophy, globinopathies, cystic fibrosis, xeroderma pigmentosum, lysosomal storage diseases). Exemplary genes include genes encoding coagulation factor VIII, coagulation factor IX, dystrophin, beta-globin, CFTR, XPC, XPD, and DNA polymerase eta.

**[0187]** Compositions may include one or more (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more) different donor oligonucleotide sequences. Use of more than one donor oligonucleotide may be useful, for example, to create a heterozygous target gene where the two alleles contain different modifications.

**[0188]** Donor oligonucleotides are preferably DNA oligonucleotides, composed of the principal naturally-occurring nucleotides (uracil, thymine, cytosine, adenine and guanine) as the heterocyclic bases, deoxyribose as the sugar moiety, and phosphate ester linkages. Donor oligonucleotides may include modifications to nucleobases, sugar moieties, or backbone/linkages, depending on the desired structure of the replacement sequence at the site of recombination or to provide some resistance to degradation by nucleases. For example, the terminal three inter-nucleoside linkages at each end of a ssDNA oligonucleotide (both 5' and 3' ends) may be replaced with phosphorothioate linkages in lieu of the usual phosphodiester linkages, thereby providing increased resis-

tance to exonucleases. Modifications to the donor oligonucleotide should not prevent the donor oligonucleotide from successfully recombining at the recombination target sequence.

#### **[0189]** 1. Donor Oligonucleotide Design

**[0190]** A polynucleotide including a donor sequence to be inserted is provided to the cell to be edited. A “donor sequence”, “donor polynucleotide” or “donor oligonucleotide” means a nucleic acid sequence to be inserted at the target site. The donor polynucleotide typically contains sufficient homology to a genomic sequence at the target site, e.g., 70%, 80%, 85%, 90%, 95%, or 100% homology with the nucleotide sequences at the target site, to support homology-directed repair between it and the genomic sequence to which it bears homology.

**[0191]** The donor sequence may or may not be identical to the genomic sequence that it replaces. The donor sequence may correspond to the wild type sequence (or a portion thereof) of the target sequence (e.g., a gene). The donor sequence may contain at least one or more single base changes, insertions, deletions, inversions or rearrangements with respect to the genomic sequence, so long as sufficient homology is present to support homology-directed repair. In some embodiments, the donor sequence includes a non-homologous sequence flanked by two regions of homology, such that homology-directed repair between the target DNA region and the two flanking sequences results in insertion of the non-homologous sequence at the target region.

**[0192]** The donor oligonucleotide is thought to simply recombine with the site in the chromosome to which it has homology. An exogenous nuclease is not used, so endogenous DNA repair and/or replication factors are believed to be involved to support the homology-directed recombination of the donor DNA into its target site. It is believed that 3E10 promotes recombination by shifting the balance of DNA repair and recombination pathways from one that is RAD51 mediated to one that is RAD52 mediated.

**[0193]** When the genome editing compositions include a donor polynucleotide sequence that includes at least a segment with homology to the target DNA sequence, the methods can be used to site specifically add, i.e., insert or replace, nucleic acid material to a target DNA sequence (e.g., to “knock in” a nucleic acid that allows expression of a protein, an siRNA, an miRNA, etc.), to add a tag (e.g., 6×His, a fluorescent protein (e.g., a green fluorescent protein; a yellow fluorescent protein, etc.), hemagglutinin (HA), FLAG, etc.), to add a regulatory sequence to a gene (e.g., promoter, polyadenylation signal, internal ribosome entry sequence (IRES), 2A peptide, start codon, stop codon, splice signal, localization signal, etc.), or to modify a nucleic acid sequence (e.g., introduce a mutation).

#### **[0194]** 2. Oligonucleotide Compositions

**[0195]** Any of the donor oligonucleotides or other nucleic acids can include one or more modifications or substitutions to the nucleobases or linkages. Modifications should not prevent, and preferably enhance, the activity, persistence, or function of gene editing. For example, modifications to oligonucleotides should not prevent, and preferably enhance duplex invasion, and/or strand displacement. Modified bases and base analogues, modified sugars and sugar analogues and/or various suitable linkages known in the art are also suitable for use in the molecules herein.

**[0196]** a. Heterocyclic Bases

**[0197]** The principal naturally-occurring nucleotides include uracil, thymine, cytosine, adenine and guanine as the heterocyclic bases. Gene editing molecules can include chemical modifications to their nucleotide constituents. For example, target sequences with adjacent cytosines can be problematic. Chemical modification of nucleotides may be useful to increase binding affinity and/or stability under physiologic conditions.

**[0198]** Chemical modifications of heterocyclic bases or heterocyclic base analogs may be effective to increase the binding affinity of a nucleotide or its stability in a complex (e.g., duplex or triplex). Chemically-modified heterocyclic bases include, but are not limited to, inosine, 5-(1-propynyl) uracil (pU), 5-(1-propynyl) cytosine (pC), 5-methylcytosine, 8-oxo-adenine, pseudocytosine, pseudoisocytosine, 5 and 2-amino-5-(2'-deoxy- $\beta$ -D-ribofuranosyl)pyridine (2-amino-pyridine), and various pyrrolo- and pyrazolopyrimidine derivatives.

**[0199]** b. Backbone

**[0200]** The nucleotide subunits of the oligonucleotides may contain certain modifications. For example, the phosphate backbone of the oligonucleotide may be replaced in its entirety by repeating N-(2-aminoethyl)-glycine units and/or phosphodiester bonds may be replaced by peptide bonds or phosphorothioate linkages, either partial or complete. The heterocyclic bases may be linked to the backbone by methylene carbonyl bonds, which allow them to form duplexes with DNA via Watson-Crick base pairing with high affinity and sequence-specificity.

**[0201]** Other backbone modifications include peptide and amino acid variations and modifications. The backbone constituents of donor oligonucleotides may be peptide linkages, or alternatively, they may be non-peptide linkages. Examples include acetyl caps, amino spacers such as 8-amino-3,6-dioxaoctanoic acid (referred to herein as O-linkers), amino acids such as lysine are particularly useful if positive charges are desired in the oligonucleotide, and the like.

**[0202]** Backbone modifications of oligonucleotides should not prevent the molecules from binding with high specificity to the DNA target site and mediating information transfer.

**[0203]** c. Modified Nucleic Acids

**[0204]** Oligonucleotides are composed of a chain of nucleotides which are linked to one another. Canonical nucleotides typically include a heterocyclic base (nucleic acid base), a sugar moiety attached to the heterocyclic base, and a phosphate moiety which esterifies a hydroxyl function of the sugar moiety. The principal naturally-occurring nucleotides include uracil, thymine, cytosine, adenine and guanine as the heterocyclic bases, and ribose or deoxyribose sugar linked by phosphodiester bonds. As used herein "modified nucleotide" or "chemically modified nucleotide" defines a nucleotide that has a chemical modification of one or more of the heterocyclic base, sugar moiety or phosphate moiety constituents. The charge of the modified nucleotide may be reduced compared to DNA or RNA oligonucleotides of the same nucleobase sequence. The oligonucleotide may have low negative charge, no charge, or positive charge such that electrostatic repulsion with the nucleotide duplex at the target site is reduced compared to DNA or RNA oligonucleotides with the corresponding nucleobase sequence.

**[0205]** Examples of modified nucleotides with reduced charge include modified internucleotide linkages such as phosphate analogs having achiral and uncharged intersubunit linkages (e.g., Sterchak, E. P. et al., *Organic Chem.*, 52:4202, (1987)), and uncharged morpholino-based polymers having achiral intersubunit linkages (see, e.g., U.S. Pat. No. 5,034,506). Some internucleotide linkage analogs include morpholidate, acetal, and polyamide-linked heterocycles. Locked nucleic acids (LNA) are modified RNA nucleotides (see, for example, Braasch, et al., *Chem. Biol.*, 8(1):1-7 (2001)). LNAs form hybrids with DNA which are more stable than DNA/DNA hybrids. Therefore, LNA can be used. LNA binding efficiency can be increased in some embodiments by adding positive charges to it. Commercial nucleic acid synthesizers and standard phosphoramidite chemistry are used to make LNAs.

**[0206]** Molecules may also include nucleotides with modified heterocyclic bases, sugar moieties or sugar moiety analogs. Modified nucleotides may include modified heterocyclic bases or base analogs as described above. Sugar moiety modifications include, but are not limited to, 2'-O-aminoethoxy, 2'-O-aminoethyl (2'-OAE), 2'-O-methoxy, 2'-O-methyl, 2'-guanidoethyl (2'-OGE), 2'-O,4'-C-methylene (LNA), 2'-O-(methoxyethyl) (2'-OME) and 2'-O-(N-(methyl)acetamido) (2'-OMA).

**[0207]** In some embodiments, the donor oligonucleotide includes 1, 2, 3, 4, 5, 6, or more optional phosphorothioate internucleoside linkages. In some embodiments, the donor includes phosphorothioate internucleoside linkages between first 2, 3, 4 or 5 nucleotides, and/or the last 2, 3, 4, or 5 nucleotides in the donor oligonucleotide.

**[0208]** C. Pharmaceutical Compositions

**[0209]** The cell-penetrating anti-DNA antibody and donor oligonucleotide compositions can be used therapeutically in combination with a pharmaceutically acceptable carrier.

**[0210]** The compositions are preferably employed for therapeutic uses in combination with a suitable pharmaceutical carrier. Such compositions include an effective amount of the composition, and a pharmaceutically acceptable carrier or excipient.

**[0211]** It is understood by one of ordinary skill in the art that nucleotides administered in vivo are taken up and distributed to cells and tissues (Huang, et al., *FEBS Lett.*, 558(1-3):69-73 (2004)). For example, Nyce, et al. have shown that antisense oligodeoxynucleotides (ODNs) when inhaled bind to endogenous surfactant (a lipid produced by lung cells) and are taken up by lung cells without a need for additional carrier lipids (Nyce, et al., *Nature*, 385:721-725 (1997)). Small nucleic acids are readily taken up into T24 bladder carcinoma tissue culture cells (Ma, et al., *Antisense Nucleic Acid Drug Dev.*, 8:415-426 (1998)).

**[0212]** The compositions including a potentiating agent, such as a cell-penetrating antibody, and donor oligonucleotides may be in a formulation for administration topically, locally or systemically in a suitable pharmaceutical carrier. Remington's Pharmaceutical Sciences, 15th Edition by E. W. Martin (Mark Publishing Company, 1975), discloses typical carriers and methods of preparation. The compound may also be encapsulated in suitable biocompatible particles formed of biodegradable or non-biodegradable polymers or proteins or liposomes for targeting to cells. Such systems are well known to those skilled in the art and may be optimized

for use with the appropriate nucleic acid. In some embodiments, the donor oligonucleotide is encapsulated in nanoparticles.

**[0213]** Various methods for nucleic acid delivery are described, for example, in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York (1989); and Ausubel, et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York (1994). Such nucleic acid delivery systems include the desired nucleic acid, by way of example and not by limitation, in either “naked” form as a “naked” nucleic acid, or formulated in a vehicle suitable for delivery, such as in a complex with a cationic molecule or a liposome forming lipid, or as a component of a vector, or a component of a pharmaceutical composition. The nucleic acid delivery system can be provided to the cell either directly, such as by contacting it with the cell, or indirectly, such as through the action of any biological process. The nucleic acid delivery system can be provided to the cell by endocytosis, receptor targeting, coupling with native or synthetic cell membrane fragments, physical means such as electroporation, combining the nucleic acid delivery system with a polymeric carrier such as a controlled release film or nanoparticle or microparticle, using a vector, injecting the nucleic acid delivery system into a tissue or fluid surrounding the cell, simple diffusion of the nucleic acid delivery system across the cell membrane, or by any active or passive transport mechanism across the cell membrane. Additionally, the nucleic acid delivery system can be provided to the cell using techniques such as antibody-related targeting and antibody-mediated immobilization of a viral vector.

**[0214]** Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, optionally with an added preservative. The compositions may take such forms as sterile aqueous or nonaqueous solutions, suspensions and emulsions, which can be isotonic with the blood of the subject in certain embodiments. Examples of nonaqueous solvents are polypropylene glycol, polyethylene glycol, vegetable oil such as olive oil, sesame oil, coconut oil, arachis oil, peanut oil, mineral oil, injectable organic esters such as ethyl oleate, or fixed oils including synthetic mono or di-glycerides. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, 1,3-butandiol, Ringer’s dextrose, dextrose and sodium chloride, lactated Ringer’s or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, and electrolyte replenishers (such as those based on Ringer’s dextrose). The materials may be in solution, emulsions, or suspension (for example, incorporated into particles, liposomes, or cells). Typically, an appropriate amount of a pharmaceutically-acceptable salt is used in the formulation to render the formulation isotonic. Trehalose, typically in the amount of 1-5%, may be added to the pharmaceutical compositions. The pH of the solution is preferably from about 5 to about 8, and more preferably from about 7 to about 7.5.

**[0215]** Pharmaceutical compositions may include carriers, thickeners, diluents, buffers, preservatives, and surface-active agents. Carrier formulation can be found in Remington’s Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa. Those of skill in the art can readily determine the various parameters for preparing and formulating the compositions without resort to undue experimentation.

**[0216]** The compositions (cell-penetrating antibody and donor oligonucleotide) alone or in combination with other suitable components, can also be made into aerosol formulations (i.e., they can be “nebulized”) to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and air. For administration by inhalation, the compounds are delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant.

**[0217]** In some embodiments, the compositions (cell-penetrating antibody and donor oligonucleotide) include pharmaceutically acceptable carriers with formulation ingredients such as salts, carriers, buffering agents, emulsifiers, diluents, excipients, chelating agents, preservatives, solubilizers, or stabilizers. The donor oligonucleotides may be conjugated to lipophilic groups like cholesterol and lauric and lithocholic acid derivatives with C32 functionality to improve cellular uptake. For example, cholesterol has been demonstrated to enhance uptake and serum stability of siRNA in vitro (Lorenz, et al., *Bioorg. Med. Chem. Lett.*, 14(19):4975-4977 (2004)) and in vivo (Soutschek, et al., *Nature*, 432(7014):173-178 (2004)). In addition, it has been shown that binding of steroid conjugated oligonucleotides to different lipoproteins in the bloodstream, such as LDL, protect integrity and facilitate biodistribution (Rump, et al., *Biochem. Pharmacol.*, 59(11):1407-1416 (2000)). Other groups that can be attached or conjugated to the compound described above to increase cellular uptake, include acridine derivatives; cross-linkers such as psoralen derivatives, azidophenacyl, proflavin, and azidoproflavin; artificial endonucleases; metal complexes such as EDTA-Fe(II) and porphyrin-Fe(II); alkylating moieties; nucleases such as alkaline phosphatase; terminal transferases; abzymes; cholesteryl moieties; lipophilic carriers; peptide conjugates; long chain alcohols; phosphate esters; radioactive markers; non-radioactive markers; carbohydrates; and polylysine or other polyamines U.S. Pat. No. 6,919,208 to Levy, et al., also describes methods for enhanced delivery. These pharmaceutical formulations may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

**[0218]** Further carriers include sustained release preparations such as semi-permeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped particles, e.g., films, liposomes or microparticles. Implantation includes inserting implantable drug delivery systems, e.g., microspheres, hydrogels, polymeric reservoirs, cholesterol matrixes, polymeric systems, e.g., matrix erosion and/or diffusion systems and non-polymeric systems. Inhalation includes administering the composition (cell-penetrating antibody and donor oligonucleotide) with an aerosol in an inhaler, either alone or attached to a carrier that can be absorbed. For systemic administration, it may be preferred that the composition is encapsulated in liposomes.

**[0219]** The compositions may be delivered in a manner which enables tissue-specific uptake of the agent and/or nucleotide delivery system, using invasive devices such as vascular or urinary catheters, and using interventional devices such as stents having drug delivery capability and configured as expansive devices or stent grafts.

**[0220]** The formulations (containing the cell-penetrating antibody and donor oligonucleotide) may be delivered using

a bioerodible implant by way of diffusion or by degradation of the polymeric matrix. In certain embodiments, the administration of the formulation may be designed to result in sequential exposures to the composition, over a certain time period, for example, hours, days, weeks, months or years. This may be accomplished, for example, by repeated administrations of a formulation or by a sustained or controlled release delivery system in which the compositions are delivered over a prolonged period without repeated administrations.

**[0221]** Other delivery systems suitable include time-release, delayed release, sustained release, or controlled release delivery systems. Such systems may avoid repeated administrations in many cases, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include, for example, polymer-based systems such as polylactic and/or polyglycolic acids, polyanhydrides, polycaprolactones, copolyoxalates, polyesteramides, polyorthoesters, polyhydroxybutyric acid, and/or combinations of these. Microcapsules of the foregoing polymers containing nucleic acids are described in, for example, U.S. Pat. No. 5,075,109. Other examples include non-polymer systems that are lipid-based including sterols such as cholesterol, cholesterol esters, and fatty acids or neutral fats such as mono-, di- and triglycerides; hydrogel release systems; liposome-based systems; phospholipid based-systems; silastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; or partially fused implants. The formulation may be as, for example, microspheres, hydrogels, polymeric reservoirs, cholesterol matrices, or polymeric systems. In some embodiments, the system may allow sustained or controlled release of the composition to occur, for example, through control of the diffusion or erosion/degradation rate of the formulations containing the cell-penetrating antibody and/or donor oligonucleotide.

**[0222]** Active agent(s) (cell-penetrating antibody and donor oligonucleotide) and compositions thereof can be formulated for pulmonary or mucosal administration. The administration can include delivery of the composition to the lungs, nasal, oral (sublingual, buccal), vaginal, or rectal mucosa. The term aerosol as used herein refers to any preparation of a fine mist of particles, which can be in solution or a suspension, whether or not it is produced using a propellant. Aerosols can be produced using standard techniques, such as ultrasonication or high-pressure treatment.

**[0223]** For administration via the upper respiratory tract, the formulation can be formulated into a solution, e.g., water or isotonic saline, buffered or un-buffered, or as a suspension, for intranasal administration as drops or as a spray. Preferably, such solutions or suspensions are isotonic relative to nasal secretions and of about the same pH, ranging e.g., from about pH 4.0 to about pH 7.4 or, from pH 6.0 to pH 7.0. Buffers should be physiologically compatible and include, simply by way of example, phosphate buffers.

**[0224]** The potentiating agent, such as a cell-penetrating antibody, and donor oligonucleotides can be delivered to the target cells using a nanoparticle delivery vehicle. In some embodiments, some of the compositions are packaged in nanoparticles and some are not. For example, in some embodiments, the donor oligonucleotide is incorporated into nanoparticles while the cell-penetrating antibody is not.

Nanoparticles generally refers to particles in the range of between 500 nm to less than 0.5 nm, preferably having a diameter that is between 50 and 500 nm, more preferably having a diameter that is between 50 and 300 nm. Cellular internalization of polymeric particles is highly dependent upon their size, with nanoparticulate polymeric particles being internalized by cells with much higher efficiency than microparticulate polymeric particles. For example, Desai, et al. have demonstrated that about 2.5 times more nanoparticles that are 100 nm in diameter are taken up by cultured Caco-2 cells as compared to microparticles having a diameter on 1  $\mu$ M (Desai, et al., *Pharm. Res.*, 14:1568-73 (1997)). Nanoparticles also have a greater ability to diffuse deeper into tissues in vivo.

**[0225]** Examples of preferred biodegradable polymers include synthetic polymers that degrade by hydrolysis such as poly(hydroxy acids), such as polymers and copolymers of lactic acid and glycolic acid, other degradable polyesters, polyanhydrides, poly(ortho)esters, polyesters, polyurethanes, poly(butyric acid), poly(valeric acid), poly(caprolactone), poly(hydroxyalkanoates), poly(lactide-co-caprolactone), and poly(amine-co-ester) polymers, such as those described in Zhou, et al., *Nature Materials*, 11:82-90 (2012) and WO 2013/082529, U.S. Published Application No. 2014/0342003, and PCT/US2015/061375.

**[0226]** The nucleic acids can be complexed to polycations to increase the encapsulation efficiency of the nucleic acids into the nanoparticles. The term "polycation" refers to a compound having a positive charge, preferably at least 2 positive charges, at a selected pH, preferably physiological pH. Polycationic moieties have between about 2 to about 15 positive charges, preferably between about 2 to about 12 positive charges, and more preferably between about 2 to about 8 positive charges at selected pH values.

**[0227]** Many polycations are known in the art. Suitable constituents of polycations include basic amino acids and their derivatives such as arginine, asparagine, glutamine, lysine and histidine; cationic dendrimers; and amino polysaccharides. Suitable polycations can be linear, such as linear tetralysine, branched or dendrimeric in structure.

**[0228]** Exemplary polycations include, but are not limited to, synthetic polycations based on acrylamide and 2-acrylamido-2-methylpropanetrimethylamine, poly(N-ethyl-4-vinylpyridine) or similar quaternized polypyridine, diethylaminoethyl polymers and dextran conjugates, polymyxin B sulfate, lipopolyamines, poly(allylamines) such as the strong polycation poly(dimethyldiallylammonium chloride), polyethyleneimine, polybrene, and polypeptides such as pro-tamine, the histone polypeptides, polylysine, polyarginine and polyornithine. Suitable naturally occurring polyamines include, but are not limited to, spermine, spermidine, cadaverine and putrescine. In some embodiments, the particles themselves are a polycation (e.g., a blend of PLGA and poly(beta amino ester).

**[0229]** Functional molecules can be associated with, linked, conjugated, or otherwise attached directly or indirectly to the carriers utilized for delivery. Targeting moieties can be associated with, linked, conjugated, or otherwise attached directly or indirectly to the donor oligonucleotide, or to a nanoparticle or other delivery vehicle thereof. Targeting molecules can be proteins, peptides, nucleic acid molecules, saccharides or polysaccharides that bind to a receptor or other molecule on the surface of a targeted cell.

The degree of specificity and the avidity of binding to the graft can be modulated through the selection of the targeting molecule.

**[0230]** Examples of moieties include, for example, targeting moieties which provide for the delivery of molecules to specific cells, e.g., antibodies to hematopoietic stem cells, CD34<sup>+</sup> cells, T cells or any other preferred cell type, as well as receptor and ligands expressed on the preferred cell type. Preferably, the moieties target hematopoietic stem cells. Examples of molecules targeting extracellular matrix (“ECM”) include glycosaminoglycan (“GAG”) and collagen. In one embodiment, the external surface of polymer particles may be modified to enhance the ability of the particles to interact with selected cells or tissue. The method described above wherein an adaptor element conjugated to a targeting molecule is inserted into the particle is preferred. However, in another embodiment, the outer surface of a polymer micro- or nanoparticle having a carboxy terminus may be linked to targeting molecules that have a free amine terminus.

**[0231]** Other useful ligands attached to polymeric micro- and nanoparticles include pathogen-associated molecular patterns (PAMPs). PAMPs target Toll-like Receptors (TLRs) on the surface of the cells or tissue, or signal the cells or tissue internally, thereby potentially increasing uptake. PAMPs conjugated to the particle surface or co-encapsulated may include: unmethylated CpG DNA (bacterial), double-stranded RNA (viral), lipopolysaccharide (bacterial), peptidoglycan (bacterial), lipoarabinomannin (bacterial), zymosan (yeast), mycoplasmal lipoproteins such as MALP-2 (bacterial), flagellin (bacterial) poly(inosinic-cytidylic) acid (bacterial), lipoteichoic acid (bacterial) or imidazoquinolines (synthetic).

**[0232]** In another embodiment, the outer surface of the particle may be treated using a mannose amine, thereby mannosylating the outer surface of the particle. This treatment may cause the particle to bind to the target cell or tissue at a mannose receptor on the antigen presenting cell surface. Alternatively, surface conjugation with an immunoglobulin molecule containing an Fc portion (targeting Fc receptor), heat shock protein moiety (HSP receptor), phosphatidylserine (scavenger receptors), and lipopolysaccharide (LPS) are additional receptor targets on cells or tissue.

**[0233]** Lectins that can be covalently attached to micro- and nanoparticles to render them target specific to the mucin and mucosal cell layer.

**[0234]** The choice of targeting moiety will depend on the method of administration of the nanoparticle composition and the cells or tissues to be targeted. The targeting molecule may generally increase the binding affinity of the particles for cell or tissues or may target the nanoparticle to a particular tissue in an organ or a particular cell type in a tissue. The covalent attachment of any of the natural components of mucin in either pure or partially purified form to the particles would decrease the surface tension of the bead-gut interface and increase the solubility of the bead in the mucin layer. The attachment of polyamino acids containing extra pendant carboxylic acid side groups, e.g., polyaspartic acid and polyglutamic acid, increases bioadhesiveness. Using polyamino acids in the 15,000 to 50,000 kDa molecular weight range yields chains of 120 to 425 amino acid residues attached to the surface of the particles.

The polyamino chains increase bioadhesion by means of chain entanglement in mucin strands as well as by increased carboxylic charge.

### III. Methods

**[0235]** Methods of modifying the genome of a cell include contacting the cell with an effective amount of (i) a cell-penetrating antibody, and (ii) a donor oligonucleotide containing a sequence that can correct a mutation in a cell's genome. Genomic modification may occur at a higher frequency when cells are contacted with both (i) and (ii), than when contacted with (ii) in the absence of (i). Preferably, the method does not involve contacting the cell with a nuclease (e.g., ZFN, Cas9) or peptide nucleic acid (PNA). The methods can be used to perform in vitro, ex vivo or in vivo gene editing.

**[0236]** Potentiating agent and donor oligonucleotide can be contacted with the cells together in the same or different admixtures, or potentiating agent and donor oligonucleotide can be contacted with cells separately. For example, cells can be first contacted with potentiating agent, followed by donor oligonucleotide. Alternatively, cells can be first contacted with donor oligonucleotide, followed by potentiating agent. In some embodiments, donor oligonucleotide and potentiating agent are mixed in solution and contacted with cells simultaneously. In a preferred embodiment, donor DNA is mixed with potentiating agent in solution and the combination is added to the cells in culture or injected into an animal to be treated.

**[0237]** The effective amount or therapeutically effective amount of the potentiating agent, and/or donor oligonucleotide can be a dosage sufficient to treat, inhibit, or alleviate one or more symptoms of a disease or disorder, or to otherwise provide a desired pharmacologic and/or physiologic effect, for example, reducing, inhibiting, or reversing one or more of the pathophysiological mechanisms underlying a disease or disorder.

**[0238]** An effective amount may also be an amount effective to increase the rate of recombination of a donor fragment relative to administration of the donor fragment in the absence of the potentiating agent. The formulation of the potentiating agent, and/or donor oligonucleotide is made to suit the mode of administration. Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions containing the potentiating agent and/or donor oligonucleotide. The precise dosage will vary according to a variety of factors such as subject-dependent variables (e.g., age, immune system health, clinical symptoms etc.).

**[0239]** The potentiating agent and donor oligonucleotide can be administered or otherwise contacted with target cells once, twice, or three times daily; one, two, three, four, five, six, seven times a week; one, two, three, four, five, six, seven or eight times a month. For example, in some embodiments, the potentiating agent and donor oligonucleotide is administered every two or three days, or on average about 2 to about 4 times about week.

**[0240]** The potentiating agent and donor oligonucleotide may or may not be administered at the same time. In some embodiments, the potentiating agent is contacted with a cell prior to the donor oligonucleotide. The potentiating agent

can be administered to the subject, for example, 1, 2, 3, 4, 5, 6, 8, 10, 12, 18, or 24 hours, or 1, 2, 3, 4, 5, 6, or 7 days, or any combination thereof prior to administration of the donor oligonucleotide to the subject.

**[0241]** In preferred embodiments, the potentiating agent and donor oligonucleotide are administered in an amount effective to induce gene modification in at least one target allele to occur at frequency of at least 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25% of target cells. In some embodiments, particularly ex vivo applications, gene modification occurs in at least one target allele at a frequency of about 0.1-25%, or 0.5-25%, or 1-25%, 2-25%, or 3-25%, or 4-25% or 5-25% or 6-25%, or 7-25%, or 8-25%, or 9-25%, or 10-25%, 11-25%, or 12-25%, or 13%-25% or 14%-25% or 15-25%, or 2-20%, or 3-20%, or 4-20% or 5-20% or 6-20%, or 7-20%, or 8-20%, or 9-20%, or 10-20%, 11-20%, or 12-20%, or 13%-20% or 14%-20% or 15-20%, 2-15%, or 3-15%, or 4-15% or 5-15% or 6-15%, or 7-15%, or 8-15%, or 9-15%, or 10-15%, 11-15%, or 12-15%, or 13%-15% or 14%-15%.

**[0242]** In some embodiments, particularly in vivo applications, gene modification occurs in at least one target allele at a frequency of about 0.1% to about 15%, or about 0.2% to about 15%, or about 0.3% to about 15%, or about 0.4% to about 15%, or about 0.5% to about 15%, or about 0.6% to about 15%, or about 0.7% to about 15%, or about 0.8% to about 15%, or about 0.9% to about 15%, or about 1.0% to about 15%, or about 1.1% to about 15%, or about 1.1% to about 15%, 1.2% to about 15%, or about 1.3% to about 15%, or about 1.4% to about 15%, or about 1.5% to about 15%, or about 1.6% to about 15%, or about 1.7% to about 15%, or about 1.8% to about 15%, or about 1.9% to about 15%, or about 2.0% to about 15%, or about 2.5% to about 15%, or about 3.0% to about 15%, or about 3.5% to about 15%, or about 4.0% to about 15%, or about 4.5% to about 15%, or about 5.0% to about 15%, or about 1% to about 15%, about 1.5% to about 15%, about 2.0% to about 15%, or about 2.5% to about 15%, or about 3.0% to about 15%, or about 3.5% to about 15%, or about 4.0% to about 15%, or about 4.5% to about 15%.

**[0243]** In some embodiments, gene modification occurs with low off-target effects. In some embodiments, off-target modification is undetectable using routine analysis such as, but not limited to, those described in the Examples. In some embodiments, off-target incidents occur at a frequency of 0-1%, or 0-0.1%, or 0-0.01%, or 0-0.001%, or 0-0.0001%, or 0-0000.1%, or 0-0.000001%. In some embodiments, off-target modification occurs at a frequency that is about  $10^2$ ,  $10^3$ ,  $10^4$ , or  $10^5$ -fold lower than at the target site.

**[0244]** A. Ex vivo Gene Therapy In some embodiments, ex vivo gene therapy of cells is used for the treatment of a genetic disorder in a subject. For ex vivo gene therapy, cells are isolated from a subject and contacted ex vivo with the compositions (potentiating agent and donor oligonucleotide) to produce cells containing altered sequences in or adjacent to genes. In a preferred embodiment, the cells are isolated from the subject to be treated or from a syngenic host.

**[0245]** Target cells are removed from a subject prior to contacting with a potentiating agent and donor oligonucleotide. The cells can be hematopoietic progenitor or stem cells. In a preferred embodiment, the target cells are CD34<sup>+</sup> hematopoietic stem cells. Hematopoietic stem cells (HSCs),

such as CD34<sup>+</sup> cells are multipotent stem cells that give rise to all the blood cell types including erythrocytes. Therefore, CD34<sup>+</sup> cells can be isolated from a patient with, for example, thalassemia, sickle cell disease, or a lysosomal storage disease, the mutant gene altered or repaired ex-vivo using the compositions and methods, and the cells reintroduced back into the patient as a treatment or a cure.

**[0246]** Stem cells can be isolated and enriched by one of skill in the art. Methods for such isolation and enrichment of CD34<sup>+</sup> and other cells are known in the art and disclosed for example in U.S. Pat. Nos. 4,965,204; 4,714,680; 5,061,620; 5,643,741; 5,677,136; 5,716,827; 5,750,397 and 5,759,793. As used herein in the context of compositions enriched in hematopoietic progenitor and stem cells, “enriched” indicates a proportion of a desirable element (e.g. hematopoietic progenitor and stem cells) which is higher than that found in the natural source of the cells. A composition of cells may be enriched over a natural source of the cells by at least one order of magnitude, preferably two or three orders, and more preferably 10, 100, 200 or 1000 orders of magnitude.

**[0247]** In humans, CD34<sup>+</sup> cells can be recovered from cord blood, bone marrow or from blood after cytokine mobilization effected by injecting the donor with hematopoietic growth factors such as granulocyte colony stimulating factor (G-CSF), granulocyte-monocyte colony stimulating factor (GM-CSF), stem cell factor (SCF) subcutaneously or intravenously in amounts sufficient to cause movement of hematopoietic stem cells from the bone marrow space into the peripheral circulation. Initially, bone marrow cells may be obtained from any suitable source of bone marrow, e.g. tibiae, femora, spine, and other bone cavities. For isolation of bone marrow, an appropriate solution may be used to flush the bone, which solution will be a balanced salt solution, conveniently supplemented with fetal calf serum or other naturally occurring factors, in conjunction with an acceptable buffer at low concentration, generally from about 5 to 25 mM. Convenient buffers include Hepes, phosphate buffers, lactate buffers, etc.

**[0248]** Cells can be selected by positive and negative selection techniques. Cells can be selected using commercially available antibodies which bind to hematopoietic progenitor or stem cell surface antigens, e.g. CD34, using methods known to those of skill in the art. For example, the antibodies may be conjugated to magnetic beads and immunogenic procedures utilized to recover the desired cell type. Other techniques involve the use of fluorescence activated cell sorting (FACS). The CD34 antigen, which is found on progenitor cells within the hematopoietic system of non-leukemic individuals, is expressed on a population of cells recognized by the monoclonal antibody My-10 (i.e., express the CD34 antigen) and can be used to isolate stem cell for bone marrow transplantation. My-10 deposited with the American Type Culture Collection (Rockville, Md.) as HB-8483 is commercially available as anti-HPCA 1. Additionally, negative selection of differentiated and “dedicated” cells from human bone marrow can be utilized, to select against substantially any desired cell marker. For example, progenitor or stem cells, most preferably CD34<sup>+</sup> cells, can be characterized as being any of CD3<sup>-</sup>, CD7<sup>-</sup>, CD8<sup>-</sup>, CD10<sup>-</sup>, CD14<sup>-</sup>, CD15<sup>-</sup>, CD19<sup>-</sup>, CD20<sup>-</sup>, CD33<sup>-</sup>, Class II HLA<sup>+</sup> and Thy-1<sup>+</sup>.

**[0249]** Once progenitor or stem cells have been isolated, they may be propagated by growing in any suitable medium. For example, progenitor or stem cells can be grown in

conditioned medium from stromal cells, such as those that can be obtained from bone marrow or liver associated with the secretion of factors, or in medium including cell surface factors supporting the proliferation of stem cells. Stromal cells may be freed of hematopoietic cells employing appropriate monoclonal antibodies for removal of the undesired cells.

**[0250]** The isolated cells are contacted ex vivo with a combination of a potentiating agent, such as a cell-penetrating antibody or anti-RAD51 factor, and donor oligonucleotides in amounts effective to cause the desired alterations in or adjacent to genes in need of repair or alteration, for example the human beta-globin or  $\alpha$ -L-iduronidase gene. These cells are referred to herein as modified cells. A solution of the cell-penetrating antibody and donor oligonucleotide may simply be added to the cells in culture. Alternatively, transfection techniques may be used. Methods for transfection of cells with oligonucleotides are well known in the art (Koppelhus, et al., *Adv. Drug Deliv. Rev.*, 55(2): 267-280 (2003)). It may be desirable to synchronize the cells in S-phase to further increase the frequency of gene correction. Methods for synchronizing cultured cells, for example, by double thymidine block, are known in the art (Zielke, et al., *Methods Cell Biol.*, 8:107-121 (1974)).

**[0251]** The modified cells can be maintained or expanded in culture prior to administration to a subject. Culture conditions are generally known in the art depending on the cell type. Conditions for the maintenance of CD34<sup>+</sup> in particular have been well studied, and several suitable methods are available. A common approach to ex vivo multi-potential hematopoietic cell expansion is to culture purified progenitor or stem cells in the presence of early-acting cytokines such as interleukin-3. It has also been shown that inclusion, in a nutritive medium for maintaining hematopoietic progenitor cells ex vivo, of a combination of thrombopoietin (TPO), stem cell factor (SCF), and flt3 ligand (Flt-3L; i.e., the ligand of the flt3 gene product) was useful for expanding primitive (i.e., relatively non-differentiated) human hematopoietic progenitor cells in vitro, and that those cells were capable of engraftment in SCID-hu mice (Luens et al., 1998, *Blood* 91:1206-1215). In other known methods, cells can be maintained ex vivo in a nutritive medium (e.g., for minutes, hours, or 3, 6, 9, 13, or more days) including murine prolactin-like protein E (mPLP-E) or murine prolactin-like protein F (mPIP-F; collectively mPLP-E/IF) (U.S. Pat. No. 6,261,841). It will be appreciated that other suitable cell culture and expansion methods can be used as well. Cells can also be grown in serum-free medium, as described in U.S. Pat. No. 5,945,337.

**[0252]** In another embodiment, the modified hematopoietic stem cells are differentiated ex vivo into CD4<sup>+</sup> cells culture using specific combinations of interleukins and growth factors prior to administration to a subject using methods well known in the art. The cells may be expanded ex vivo in large numbers, preferably at least a 5-fold, more preferably at least a 10-fold and even more preferably at least a 20-fold expansion of cells compared to the original population of isolated hematopoietic stem cells.

**[0253]** In another embodiment cells, for ex vivo gene therapy can be dedifferentiated somatic cells. Somatic cells can be reprogrammed to become pluripotent stem-like cells that can be induced to become hematopoietic progenitor cells. The hematopoietic progenitor cells can then be treated with a potentiating agent, such as a cell-penetrating anti-

body, and donor oligonucleotides as described above with respect to CD34<sup>+</sup> cells to produce recombinant cells having one or more modified genes. Representative somatic cells that can be reprogrammed include, but are not limited to fibroblasts, adipocytes, and muscles cells. Hematopoietic progenitor cells from induced stem-like cells have been successfully developed in the mouse (Hanna, J. et al. *Science*, 318:1920-1923 (2007)).

**[0254]** To produce hematopoietic progenitor cells from induced stem-like cells, somatic cells are harvested from a host. In a preferred embodiment, the somatic cells are autologous fibroblasts. The cells are cultured and transduced with vectors encoding Oct4, Sox2, Klf4, and c-Myc transcription factors. The transduced cells are cultured and screened for embryonic stem cell (ES) morphology and ES cell markers including, but not limited to AP, SSEA1, and Nanog. The transduced ES cells are cultured and induced to produce induced stem-like cells. Cells are then screened for CD41 and c-kit markers (early hematopoietic progenitor markers) as well as markers for myeloid and erythroid differentiation.

**[0255]** The modified hematopoietic stem cells or modified induced hematopoietic progenitor cells are then introduced into a subject. Delivery of the cells may be affected using various methods and includes most preferably intravenous administration by infusion as well as direct depot injection into periosteal, bone marrow and/or subcutaneous sites.

**[0256]** The subject receiving the modified cells may be treated for bone marrow conditioning to enhance engraftment of the cells. The recipient may be treated to enhance engraftment, using a radiation or chemotherapeutic treatment prior to the administration of the cells. Upon administration, the cells will generally require a period of time to engraft. Achieving significant engraftment of hematopoietic stem or progenitor cells typically takes weeks to months.

**[0257]** A high percentage of engraftment of modified hematopoietic stem cells is not envisioned to be necessary to achieve significant prophylactic or therapeutic effect. It is believed that the engrafted cells will expand over time following engraftment to increase the percentage of modified cells. For example, in some embodiments, the modified cells have a corrected  $\alpha$ -L-iduronidase gene. Therefore, in a subject with Hurler syndrome, the modified cells can improve or cure the condition. It is believed that engraftment of only a small number or small percentage of modified hematopoietic stem cells will be required to provide a prophylactic or therapeutic effect.

**[0258]** In preferred embodiments, the cells to be administered to a subject will be autologous, e.g. derived from the subject, or syngenic.

**[0259]** In some embodiments, the compositions and methods can be used to edit embryonic genomes in vitro. The methods typically include contacting an embryo in vitro with an effective amount of potentiating agent and donor oligonucleotide to induce at least one alteration in the genome of the embryo. Most preferably the embryo is a single cell zygote, however, treatment of male and female gametes prior to and during fertilization, and embryos having 2, 4, 8, or 16 cells and including not only zygotes, but also morulas and blastocytes, are also provided. Typically, the embryo is contacted with the compositions on culture days 0-6 during or following in vitro fertilization.

**[0260]** The contacting can be adding the compositions to liquid media bathing the embryo. For example, the compo-

sitions can be pipetted directly into the embryo culture media, whereupon they are taken up by the embryo.

**[0261]** B. In Vivo Gene Therapy

**[0262]** In some embodiments, in vivo gene therapy of cells is used for the treatment of a genetic disorder in a subject. The compositions (cell-penetrating antibody and donor oligonucleotide) can be administered directly to a subject for in vivo gene therapy.

**[0263]** In general, methods of administering compounds, including antibodies, oligonucleotides and related molecules, are well known in the art. In particular, the routes of administration already in use for nucleic acid therapeutics, along with formulations in current use, provide preferred routes of administration and formulation for the donor oligonucleotides described above. Preferably the compositions (potentiating agent and donor oligonucleotide) are injected or infused into the organism undergoing genetic manipulation, such as an animal requiring gene therapy.

**[0264]** The compositions (e.g., potentiating agent such as a cell-penetrating antibody and donor oligonucleotide) can be administered by a number of routes including, but not limited to, intravenous, intraperitoneal, intraamniotic, intramuscular, subcutaneous, or topical (sublingual, rectal, intranasal, pulmonary, rectal mucosa, and vaginal), and oral (sublingual, buccal).

**[0265]** In some embodiments, the compounds are formulated for pulmonary delivery, such as intranasal administration or oral inhalation.

**[0266]** Administration of the formulations may be accomplished by any acceptable method that allows the potentiating agent and donor oligonucleotide to reach their targets. The administration may be localized (i.e., to a particular region, physiological system, tissue, organ, or cell type) or systemic, depending on the condition being treated. Compositions and methods for in vivo delivery are also discussed in WO 2017/143042.

**[0267]** The methods can also include administering an effective amount of potentiating agent and donor oligonucleotide to an embryo or fetus, or the pregnant mother thereof, in vivo. In some methods, compositions are delivered in utero by injecting and/or infusing the compositions into a vein or artery, such as the vitelline vein or the umbilical vein, or into the amniotic sac of an embryo or fetus. See, e.g., Ricciardi, et al., *Nat Commun.* 2018 Jun. 26; 9(1):2481. doi: 10.1038/s41467-018-04894-2, and WO 2018/187493.

**[0268]** C. Diseases to be Treated

**[0269]** Gene therapy is apparent when studied in the context of human genetic diseases, for example, cystic fibrosis, hemophilia, globinopathies such as sickle cell anemia and beta-thalassemia, xeroderma pigmentosum, and lysosomal storage diseases, though the strategies are also useful for treating non-genetic disease such as HIV, in the context of ex vivo-based cell modification and also for in vivo cell modification. The methods using a potentiating agent, such as a cell-penetrating antibody, and donor oligonucleotides, are especially useful to treat genetic deficiencies, disorders and diseases caused by mutations in single genes, for example, to correct genetic deficiencies, disorders and diseases caused by point mutations. If the target gene contains a mutation that is the cause of a genetic disorder, then the methods can be used for mutagenic repair that may restore the DNA sequence of the target gene to normal. The target sequence can be within the coding DNA sequence of the gene or within an intron. The target sequence can also be

within DNA sequences that regulate expression of the target gene, including promoter or enhancer sequences.

**[0270]** In the methods herein, cells that have been contacted with the potentiating agent and donor oligonucleotide may be administered to a subject. The subject may have a disease or disorder such as hemophilia, muscular dystrophy, globinopathies, cystic fibrosis, xeroderma pigmentosum, or lysosomal storage diseases. In such embodiments, gene modification may occur in an effective amount to reduce one or more symptoms of the disease or disorder in the subject. Exemplary sequences for donor oligonucleotides designed to correct mutations in globinopathies, cystic fibrosis, HIV, and lysosomal storage diseases are known in the art and disclosed in, for example, published international applications WO 2017/143042, WO 2017/143061, WO 2018/187493, and published U.S. Application No. 2017/0283830, each of which is specifically incorporated by reference in its entirety.

**[0271]** Donor oligonucleotides can be either single stranded or double stranded, and can target one or both strands of the genomic sequence at a target locus. Exemplary donor oligonucleotides are provided below. The donors are typically presented as single stranded DNA sequences targeting one strand of the target genomic locus. However, even where not expressly provided below, the reverse complement of each donor, and double stranded DNA sequences, are also disclosed based on the provided sequences. In some embodiments, the donor oligonucleotide is a functional fragment of the disclosed sequence, or the reverse complement, or double stranded DNA thereof.

**[0272]** 1. Globinopathies

**[0273]** Worldwide, globinopathies account for significant morbidity and mortality. Over 1,200 different known genetic mutations affect the DNA sequence of the human alpha-like (HBZ, HBA2, HBA1, and HBQ1) and beta-like (HBE1, HBG1, HBD, and HBB) globin genes. Two of the more prevalent and well-studied globinopathies are sickle cell anemia and (3-thalassemia. Substitution of valine for glutamic acid at position 6 of the (3-globin chain in patients with sickle cell anemia predisposes to hemoglobin polymerization, leading to sickle cell rigidity and vasoocclusion with resulting tissue and organ damage. In patients with  $\beta$ -thalassemia, a variety of mutational mechanisms results in reduced synthesis of  $\beta$ -globin leading to accumulation of aggregates of unpaired, insoluble  $\alpha$ -chains that cause ineffective erythropoiesis, accelerated red cell destruction, and severe anemia.

**[0274]** Together, globinopathies represent the most common single-gene disorders in man. Donor oligonucleotides herein are effective at binding to the human  $\beta$ -globin both in vitro and in living cells, both ex vivo and in vivo in animals. Experimental results also demonstrate correction of a thalassemia-associated mutation in vivo in a transgenic mouse carrying a human beta globin gene with the IVS2-654 thalassemia mutation (in place of the endogenous mouse beta globin).

**[0275]**  $\beta$ -thalassemia is an unstable hemoglobinopathy leading to the precipitation of  $\alpha$ -hemoglobin within RBCs resulting in a severe hemolytic anemia. Patients experience jaundice and splenomegaly, with substantially decreased blood hemoglobin concentrations necessitating repeated transfusions, typically resulting in severe iron overload with time. Cardiac failure due to myocardial siderosis is a major cause of death from (3-thalassemia by the end of the third

decade. Reduction of repeated blood transfusions in these patients is therefore of primary importance to improve patient outcomes.

[0276] Exemplary  $\beta$ -globin Gene Mutations A portion of the GenBank sequence of the chromosome-11 human-native hemoglobin-gene cluster (GenBank: UO1317.1—Human beta globin region on chromosome 11—LOCUS HUMHBB, 73308 bp ds-DNA) from base 60001 to base 66060 is presented below. The start of the gene coding sequence at position 62187-62189 is indicated by wave underlining. This portion of the GenBank sequence contains the native (3 globin gene sequence. In sickle cell hemoglobin the adenine base at position 62206 is mutated to a thymine. Other common point mutations occur in intron 2 (IVS2), which is highlighted in the sequence below by italics. Mutations include IVS2-1, IVS2-566, IVS2-654, IVS2-705, and IVS2-745, which are also shown in bold and heavy underlining; numbering relative to the start of intron 2. Gene editing molecules can be designed based on the guidance provided herein and otherwise known in the art. Exemplary donor oligonucleotide sequences, are provided in, for example, WO 1996/040271, WO/2010/123983, and U.S. Pat. No. 8,658,608, and can be altered to include one or more of the modifications herein. Target regions can be reference based on the coding strand of genomic DNA, or the complementary non-coding sequence thereto (e.g., the Watson or Crick stand). Exemplary sites targeted by the donor oligonucleotide can be any region encompassing any disease-related mutation (shown below) in the beta globin gene.

(SEQ ID NO: 42)

CAGAATACTATAAAATATAACCTATATTATA

ATTTCATAAAGTCTGTGCATTTTCTTTGAC

CCAGGATATTTGCAAAAGACATATTCAAAC

TTCCGCAGAACACTTTATTTTACATATACA

TGCCTCTTATATCAGGGATGTGAAACAGGG

TCTTGAAACTGTCTAAATCTAAACAATG

CTAATGCAGGTTTAAATTTAATAAAATAAA

ATCCAAAATCTAACAGCCAAGTCAAATCTG

TATGTTTTAACATTTAAATATTTTAAAGA

CGTCTTTTCCCAGGATTCAACATGTGAAAT

CTTTTCTCAGGGATACACGTGTGCCTAGAT

CCTCATTGCTTTAGTTTTTTACAGAGGAAT

GAATATAAAAAGAAAATACTTAAATTTTAT

CCCTCTTACCTCTATAATCATACATAGGCA

TAATTTTTTAACCTAGGCTCCAGATAGCCA

TAGAAGAACCAAACACTTTCTGCGTGTGTG

AGAATAATCAGAGTGAGATTTTTTACAAG

TACCTGATGAGGGTTGAGACAGGTAGAAAA

AGTGAGAGATCTCTATTTATTTAGCAATAA

-continued

TAGAGAAAGCATTTAAGAGAATAAAGCAAT

GGAAATAAGAAATTTGTAAATTTCTTCTG

ATAACTAGAAATAGAGGATCCAGTTTCTTT

TGGTTAACCTAAATTTTATTTTATTTTATT

GTTTTATTTTATTTTATTTTATTTTATTTT

GTGTAATCGTAGTTTCAGAGTGTTAGAGCT

GAAAGGAAGAAGTAGGAGAAACATGCAAAG

TAAAAGTATAACACTTTCCTTACTAAACCG

ACTGGGTTTCCAGGTAGGGGCAGGATTCAG

GATGACTGACAGGGCCCTTAGGGAACACTG

AGACCCTACGCTGACCTCATAAATGCTTGC

TACCTTTGCTGTTTTAATTACATCTTTTAA

TAGCAGGAAGCAGAACTCTGCACCTTCAAAA

GTTTTTCCTCACCTGAGGAGTTAATTTAGT

ACAAGGGGAAAAAGTACAGGGGGATGGGAG

AAAGGCGATCACGTTGGGAAGCTATAGAGA

AAGAAGAGTAAATTTTAGTAAAGGAGGTTT

AAACAAACAAAATATAAAGAGAAATAGGAA

CTTGAATCAAGGAAATGATTTTAAACGCA

GTATTCTTAGTGACTAGAGGAAAAAATA

ATCTGAGCCAAGTAGAAGACCTTTTCCCCT

CCTACCCCTACTTTCTAAGTCACAGAGGCT

TTTTGTTCCTCCAGACACTCTTGACAGATTA

GTCCAGGCAGAAACAGTTAGATGTCCCCAG

TTAACCTCCTATTTGACACCACTGATTACC

CCATTGATAGTCACACTTTGGGTTGTAAGT

GACTTTTTTATTTATTTGTATTTTGACTGC

ATTAAGAGGTCTCTAGTTTTTTATCTCTTG

TTTCCCAAACCTAATAAGTAACTAATGCA

CAGAGCACATTGATTTGTATTTATTTCTATT

TTTAGACATAATTTATTAGCATGCATGAGC

AAATTAAGAAAAACAACAACAAATGAATGC

CTTACCAGAAGGTTTTAATCCAAATAAGGA

GAAGATATGCTTAGAACTGAGGTAGAGTTT

TCATCCATTCTGTCCTGTAAGTATTTTGCA

TATTCTGGAGACGCAGGAAGAGATCCATCT

ACATATCCCAAAGCTGAATTATGGTAGACA

AAGCTCTTCCACTTTTAGTGATCAATTTT

CCAAATATTACGTAAATACACTTGCAAAGG

-continued  
AGGATGTTTTTAGTAGCAATTTGTACTGAT  
GGTATGGGGCCAAGAGATATATCTTAGAGG  
GAGGGCTGAGGGTTTGAAGTCCAACCTCCTA  
AGCCAGTGCCAGAAGAGCCAAGGACAGGTA  
CGGCTGTCATCACTTAGACCTCACCTGTG  
GAGCCACACCCTAGGGTTGGCCAATCTACT  
CCCAGGAGCAGGGAGGGCAGGAGCCAGGGC  
TGGGCATAAAAGTCAGGGCAGAGCCATCTA  
TTGCTTACATTTGCTTCTGACACAACCTGTG  
TTCAC TAGCAACCTCAAACAGACACC**ATGG**  
TGCACCTGACTCCTG**AGG**GAGAAGTCTGCCG  
TTACTGCCCTGTGGGGCAAGGTGAACGTGG  
ATGAAGTTGGTGGTGAGGCCCTGGGCAGGT  
TGGTATCAAGGTTACAAGACAGGTTTAAGG  
AGACCAATAGAACTGGGCATGTGGAGACA  
GAGAAGACTCTTGGGTTTCTGATAGGCACT  
GACTCTCTCTGCCTATTGGTCTATTTTCCC  
ACCCTTAGGCTGCTGGTGGTCTACCCTTGG  
ACCCAGAGGTTCTTTGAGTCCTTTGGGGAT  
CTGTCCACTCCTGATGCTGTTATGGGCAAC  
CCTAAGGTGAAGGCTCATGGCAAGAAAGTG  
CTCGGTGCCTTTAGTGATGGCCTGGCTCAC  
CTGGACAACCTCAAGGGCACCTTTGCCACA  
CTGAGTGAGCTGCACTGTGACAAGCTGCAC  
GTGGATCCTGAGAACTTCAGG**GTG**AGTCTA  
TGGGACCCTTGATGTTTTCTTTCCCTTCT  
TTTCTATGGTTAAGTTCATGTCATAGGAAG  
GGGAGAAGTAACAGGGTACAGTTTAGAATG  
GGAAACAGACGAATGATTGCATCAGTGTGG  
AAGTCTCAGGATCGTTTTAGTTTCTTTTAT  
TTGCTGTTCATAACAATTGTTTTCTTTTGT  
TTAATTCTTG**CTTCTTTTTTTTTCTTCTC**  
**CGCAATTTT**ACTATTATACTTAATGCCTT  
AACATTGTGTATAACAAAAGGAAATATCTC  
TGAGATACATTAAGTAACTTAAAAAAAAC  
TTTACACAGTCTGCCTAGTACATTACTATT  
TGGAATATATGTGTGCTTATTTGCATATTC  
ATAATCTCCCTACTTTATTTTCTTTTATTT  
TTAATTGATACATAATCATTATACATATTT

-continued  
ATGGGTTAAAGTGTAATGTTTTAATATGTG  
TACACATATTGACCAAATCAGGGTAATTTT  
GCATTTGTAATTTTAAAAATGCTTTCTTC  
TTTTAATATACTTTTTTGTATTCTTATTT  
CTAATACTTTCCCTAATCTCTTTCTTTCAG  
GGCAATAATGATACAATGTATCATGCCTCT  
**TTGC**ACCATTCTAAAGAATAACAGTGATAA  
TTTCTGGGTTAAGG**CA**ATAGCAATATTTCT  
GCATATAAAATATTTCTGCATATAAAATTGTA  
ACTG**A**T**G**TAAGAGGTTTCATATTGCTAATA  
GCAGCTACAATCCAG**CT**ACCATTCTGCTTT  
TATTTTATGGTTGGGATAAGGCTGGATTAT  
TCTGAGTCCAAGCTAGG**CCCTTTT**GCTAAT  
CATGTTCATACCTCTTAT**CTTCCTCC**CACA  
GCTCCTGGGCAACGTGCTGGTCTGTGTGCT  
GGCCCATCACTTTGGCAAAGAATTCACCCC  
ACCAGTGCAGGCTGCCTATCAGAAAGTGGT  
GGCTGGTGTGGCTAATGCCCTGGCCACAA  
GTATCACTAAGCTCGCTTCTTGCTGTCCA  
ATTTCTATTAAAGGTTCCCTTTGTTCCCTAA  
GTCCAAC TACTAACTGGGGGATATTATGA  
AGGGCCTTGAGCATCTGGATTCTGCCTAAT  
AAAAAACATTTATTTTCATTGCAATGATGT  
ATTTAAATTATTTCTGAATATTTTACTAAA  
AAGGGAATGTGGGAGGTCAGTGCATTTAAA  
ACATAAAGAAATGAAGAGCTAGTTCAAACC  
TTGGGAAAATACACTATATCTTAAACTCCA  
TGAAAGAAGGTGAGGCTGCAAACAGCTAAT  
GCACATTGGCAACAGCCCTGATGCCTATGC  
CTTATTCATCCCTCAGAAAAGGATTCAAGT  
AGAGGCTTGATTTGGAGGTTAAAGTTTTC  
TATGCTGTATTTTACATTACTTATTGTTTT  
ACCTGTCCTCATGAATGTCTTTTCACTACC  
CATTTGCTTATCCTGCATCTCTCAGCCTTG  
ACTCCACTCAGTTCTCTTGCTTAGAGATAC  
CACCTTTCCCTGAAGTGTTCCTTCCATGT  
TTTACGGCGAGATGGTTTCTCCTCGCCTGG  
CCACTCAGCCTTAGTTGTCTCTGTTGTCTT  
ATAGAGGTCTACTTGAAGAAGGAAAAACAG

-continued  
GGGGCATGGTTTGA CTGTCTGTGAGCCCT  
TCTTCCCTGCCTCCCCACTCACAGTGACC  
CGGAATCTGCAGTGCTAGTCTCCCGGAACT  
ATCACTCTTTCACAGTCTGCTTTGGAAGGA  
CTGGGCTTAGTATGAAAAGTTAGGACTGAG  
AAGAATTTGAAAGGGGGCTTTTTGTAGCTT  
GATATTCACTACTGTCTTATTACCCTATCA  
TAGGCCCACCCCAAATGGAAGTCCCATCT  
TCCTCAGGATGTTTAAGATTAGCATTGAG  
AAGAGATCAGAGGTCTGCTGGCTCCCTTAT  
CATGTCCCTTATGGTGCTTCTGGCTCTGCA  
GTTATTAGCATAGTGTTACCATCAACCACC  
TTAACTTCATTTTTCTTATTCAATACCTAG  
GTAGGTAGATGCTAGATTCTGGAAATAAAA  
TATGAGTCTCAAGTGGTCCTTGTCTCTCT  
CCCAGTCAAATCTGAATCTAGTTGGCAAG  
ATTCTGAAATCAAGGCATATAATCAGTAAT  
AAGTGATGATAGAAGGGTATATAGAAGAAT  
TTTATTATATGAGAGGGTGAAACCTAAAT  
GAAATGAAATCAGACCCTTGTCTTACACCA  
TAAACAAAAATAAATTTGAATGGGTAAAG  
AATTAACTAAGACCTAAACCATAAAAAT  
TTTTAAAGAAATCAAAGAAGAAAATTCTA  
ATATTCATGTTGCAGCCGTTTTTTGAATTT  
GATATGAGAAGCAAAGGCAACAAAAGGAAA  
AATAAAGAAGTGAGGCTACATCAAATAAA  
AAATTTCCACACAAAAAAGAAAACAATGAA  
CAAATGAAAGGTGAACCATGAAATGGCATA  
TTTGCAAACCAAATATTTCTTAAATATTTT  
GGTTAATATCCAAAATATATAAGAAACACA  
GATGATTCAATAACAAACAAAAATTAATA  
ATAGGAAAAATAAAAAATTAAGAAGAA  
AATCCTGCCATTTATGCGAGAATTGATGAA  
CCTGGAGGATGTAAACTAAGAAAAATAAG  
CCTGACACAAAAAGACAAATACTACACAAC  
CTTGCTCATATGTGAAACATAAAAAAGTCA  
CTCTCATGGAAACAGACAGTAGAGGTATGG  
TTTCCAGGGGTGGGGGTGGGAGAATCAGG  
AAACTATTACTCAAAGGGTATAAAATTTCA

-continued  
GTTATGTGGGATGAATAAATTCTAGATATC  
TAATGTACAGCATCGTGACTGTAGTTAATT  
GTACTGTAAGTATATTTAAAATTTGCAAAG  
AGAGTAGATTTTTTTGTTTTTTTAGATGGA  
GTTTTGCTCTTGTTGTCCAGGCTGGAGTGC  
AATGGCAAGATCTTGCTCACTGCAACCTC  
CGCCTCCTGGGTTCAAGCAAATCTCTGCC  
TCAGCCTCCCGAGTAGCTGGGATTACAGGC  
ATGCGACACCATGCCAGCTAATTTTGTAT  
TTTTAGTAGAGACGGGGTTTCTCCATGTTG  
GTCAGGCTGATCCGCCTCCTCGGCCACCAA  
AGGGCTGGGATTACAGGCGTGACCACCGGG  
CCTGGCCGAGAGTAGATCTTAAAGCATTT  
ACCACAAGAAAAAGGTAAGTATGTGAGATA  
ATGGGTATGTTAATTAGCTTGATTGTGGTA  
ATCATTTACAAGGTATACATATATTAATA  
CATCATGTTGTACACCTTAAATATATACAA  
TTTTTATTTGTGAATGATACCTCAATAAAG  
TTGAAGAATAATAAAAAAGAATAGACATCA  
CATGAATTAAAAAACTAAAAATAAAAAA  
TGCATCTTGATGATTAGAATTGCATTCTTG  
ATTTTTCAGATACAAATATCCATTTGACTG.

Exemplary β-Globin Donors

[0277] In some embodiments, the a potentiating agent, such as a cell-penetrating antibody, are used in combination with a donor oligonucleotide for correction of IVS2-654 mutation that includes the sequence

[0278] 5' AAAGAATAACAGTGATAATTTCTGGGT-  
TAAGGCAATAGCAATAT CTCTGCATATAAATAT 3'  
(SEQ ID NO:43) with the correcting IVS2-654 nucleotide underlined, or a functional fragment thereof that is suitable and sufficient to correct the IVS2-654 mutation.

[0279] Other exemplary donor sequences include, but are not limited to,

DonorGFP-IVS2-1 (Sense)

[0280]

(SEQ ID NO: 44)  
5' GTTCAGCGTGTCCGCGAGGGCGAG  
GTGAGTCTATGGGACCCTTGATGTTT 3' ,

DonorGFP-IVS2-1 (Antisense)

[0281] 5' AAACATCAAGGGTCCCATAGACT-  
CACCTCGCCCTCGCCGGACAC GCTGAAC 3' (SEQ ID

NO:45), and, or a functional fragment thereof that is suitable and sufficient to correct a mutation.

[0282] In some embodiments, a Sickle Cells Disease mutation can be corrected using a donor having the sequence

[0283] 5' CTTGCCCCACAGGGCAGTAACGGCAGAT-

TTTTCTTC CGGCGTTA AATGCAC-CATGGTGTCTGTTTGAGGT 3' (SEQ ID NO:46), or a functional fragment thereof that is suitable and sufficient to correct a mutation, wherein the three boxed nucleotides represent the corrected codon 6 which reverts the mutant Valine (associated with human sickle cell disease) back to the wildtype Glutamic acid and nucleotides in bold font (without underlining) represent changes to the genomic DNA but not to the encoded amino acid; or

[0284] 5'ACAGACACCATGGTGCACCTGACTCCTG AGGAGAAGTCTGCCGT TACTGCC 3' (SEQ ID NO:47), or a functional fragment thereof that is suitable and sufficient to correct a mutation, wherein the bolded and underlined residue is the correction, or

[0285] 5' T(s)T (s)G(s) CCC-CACAGGGCAGTAACGGCAGACTTCTCCTC AGG AGTCAGGTGCACCATGGTGTCTGT (s)T (s)T (s) G3' (SEQ ID NO:48), or a functional fragment thereof that is suitable and sufficient to correct a mutation, wherein the bolded and underlined residue is the correction and “(s)” indicates an optional phosphorothioate internucleoside linkage.

[0286] 2. Cystic Fibrosis

[0287] The compositions and methods can be used to treat cystic fibrosis. Cystic fibrosis (CF) is a lethal autosomal recessive disease caused by defects in the cystic fibrosis transmembrane conductance regulator (CFTR), an ion channel that mediates Cl-transport. Lack of CFTR function results in chronic obstructive lung disease and premature death due to respiratory failure, intestinal obstruction syndromes, exocrine and endocrine pancreatic dysfunction, and infertility (Davis, et al., *Pediatr Rev.*, 22(8):257-64 (2001)). The most common mutation in CF is a three base-pair deletion (F508del) resulting in the loss of a phenylalanine residue, causing intracellular degradation of the CFTR protein and lack of cell surface expression (Davis, et al., *Am J Respir Crit Care Med.*, 173(5):475-82 (2006)). In addition to this common mutation there are many other mutations that occur and lead to disease including a class of mutations due to premature stop codons, nonsense mutations. In fact nonsense mutations account for approximately 10% of disease causing mutations. Of the nonsense mutations G542X and W1282X are the most common with frequencies of 2.6% and 1.6% respectively.

[0288] Although CF is one of the most rigorously characterized genetic diseases, current treatment of patients with CF focuses on symptomatic management rather than primary correction of the genetic defect. Gene therapy has remained an elusive target in CF, because of challenges of in vivo delivery to the lung and other organ systems (Armstrong, et al., *Archives of disease in childhood* (2014) doi: 10.1136/archdischild-2012-302158. PubMed PMID: 24464978). In recent years, there have been many advances in gene therapy for treatment of diseases involving the hematolymphoid system, where harvest and ex vivo manipulation of cells for autologous transplantation is possible: some examples include the use of zinc finger nucleases targeting CCR5 to produce HIV-1 resistant cells (Holt,

et al., *Nature biotechnology*, 28(8):839-47 (2010)) correction of the ABCD1 gene by lentiviral vectors for treatment of adrenoleukodystrophy (Cartier, et al., *Science*, 326(5954): 818-23 (2009)) and correction of SCID due to ADA deficiency using retroviral gene transfer (Aiuti, et al., *The New England Journal Of Medicine*, 360(5):447-58 (2009).

[0289] Unfortunately, harvest and autologous transplant is not an option in CF, due to the involvement of the lung and other internal organs. As one approach, the UK Cystic Fibrosis Gene Therapy Consortium has tested liposomes to deliver plasmids containing cDNA encoding CFTR to the lung (Alton, et al., *Thorax*, 68(11):1075-7 (2013)), Alton, et al., *The Lancet Respiratory Medicine*, (2015). doi: 10.1016/S2213-2600(15)00245-3. PubMed PMID: 26149841.) other clinical trials have used viral vectors for delivery of the CFTR gene or CFTR expression plasmids that are compacted by polyethylene glycol-substituted lysine 30-mer peptides with limited success (Konstan, et al., *Human Gene Therapy*, 15(12):1255-69 (2004)). Delivery of plasmid DNA for gene addition without targeted insertion does not result in correction of the endogenous gene and is not subject to normal CFTR gene regulation, and virus-mediated integration of the CFTR cDNA could introduce the risk of non-specific integration into important genomic sites.

[0290] A potentiating agent, such as a cell-penetrating antibody, and donor DNA oligonucleotides can be used to correct mutations leading to cystic fibrosis. In preferred embodiments, the compositions are administered by intranasal or pulmonary delivery. In some embodiments, systemic administration such as IV injection or infusion is utilized to treat CF, due to the need to correct the mutation in multiple tissues—e.g., gut. The compositions can be administered in an effective amount to induce or enhance gene correction in an amount effective to reduce one or more symptoms of cystic fibrosis. For example, in some embodiments, the gene correction occurs at an amount effective to improve impaired response to cyclic AMP stimulation, improve hyperpolarization in response to forskolin, reduction in the large lumen negative nasal potential, reduction in inflammatory cells in the bronchoalveolar lavage (BAL), improve lung histology, or a combination thereof. In some embodiments, the target cells are cells, particularly epithelial cells, that make up the sweat glands in the skin, that line passageways inside the lungs, liver, pancreas, or digestive or reproductive systems. In particular embodiments, the target cells are bronchial epithelial cells. While permanent genomic change using the gene-editing methods herein is less transient than plasmid-based approaches and the changes will be passed on to daughter cells, some modified cells may be lost over time with regular turnover of the respiratory epithelium. In some embodiments, the target cells are lung epithelial progenitor cells. Modification of lung epithelial progenitors can induce more long-term correction of phenotype.

[0291] Sequences for the human cystic fibrosis transmembrane conductance regulator (CFTR) are known in the art, see, for example, GenBank Accession number: AH006034.1, and compositions and methods of targeted correction of CFTR are described in McNeer, et al., *Nature Communications*, 6:6952, (DOI 10.1038/ncomms7952), 11 pages.

## Exemplary CFTR Donors

[0292] In some embodiments, a donor that can be used for CFTR gene correction includes the sequence

[0293] 5'TTCTGTATCTATATTCATCATAGGAAACAC-CAAAGATAATGTTC TCCTTAATGGTGCCAGG3' (SEQ ID NO:49), or a functional fragment thereof that is suitable and sufficient to correct the F508del mutation in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. The complementary sequence matching the other strand of the CFTR gene at this location, or functional fragments thereof, is also suitable.

[0294] In some embodiments, a donor that can be used for CFTR gene correction includes the sequence

[0295] T (s)C (s)T (s) TGGGATTCAATAACCTTGCA GACAGTGGAGGAAGGCCTTTGG CGTGATAACCACAG (s)G(s)T(s)G (SEQ ID NO:50) or a functional fragment thereof that is suitable and sufficient to correct a mutation in CFTR, wherein the bolded and underlined nucleotides are inserted mutations, and "(s)" indicates an optional phosphorothioate internucleoside linkage.

[0296] In some embodiments, a donor that can be used for CFTR gene correction includes the sequence

[0297] T (s)C(s)C(s) AAGTTTGCAGAGAAAGA TAATATAGTCCTTGAG AAGGAGGAATCAC CCTGAGTGGA (s) G (s) G (s) T (SEQ ID NO:51), or a functional fragment thereof that is suitable and sufficient to correct a mutation in CFTR, wherein the bolded and underlined nucleotides are inserted mutations, and "(s)" indicates an optional phosphorothioate internucleoside linkage.

[0298] In some embodiments, in addition to containing sequence designed to correct a mutation in the CFTR gene, the donor oligonucleotides may also contain additional, synonymous (silent) mutations. The additional silent mutations can facilitate detection of the corrected target sequence using allele-specific PCR of genomic DNA isolated from treated cells.

## [0299] 3. Lysosomal Storage Diseases

[0300] The disclosed compositions and methods can also be used to treat lysosomal storage diseases. Lysosomal storage diseases (LSDs) are a group of more than 50 clinically-recognized, rare inherited metabolic disorders that result from defects in lysosomal function (Walkley, *J. Inherit. Metab. Dis.*, 32(2):181-9 (2009)). Lysosomal storage disorders are caused by dysfunction of the cell's lysosome organelle, which is part of the larger endosomal/lysosomal system. Together with the ubiquitin-proteasomal and autophagosomal systems, the lysosome is essential to substrate degradation and recycling, homeostatic control, and signaling within the cell. Lysosomal dysfunction is usually the result of a deficiency of a single enzyme necessary for the metabolism of lipids, glycoproteins (sugar containing proteins) or mucopolysaccharides (long unbranched polysaccharides consisting of a repeating disaccharide unit; also known as glycosaminoglycans, or GAGs) which are fated for breakdown or recycling. Enzyme deficiency reduces or prevents break down or recycling of the unwanted lipids, glycoproteins, and GAGs, and results in buildup or "storage" of these materials within the cell. Most lysosomal diseases show widespread tissue and organ involvement, with brain, viscera, bone and connective tissues often being affected. More than two-thirds of lysosomal diseases affect the brain. Neurons appear particularly vul-

nerable to lysosomal dysfunction, exhibiting a range of defects from specific axonal and dendritic abnormalities to neuron death.

[0301] Individually, LSDs occur with incidences of less than 1:100,000, however, as a group the incidence is as high as 1 in 1,500 to 7,000 live births (Staretz-Chacham, et al., *Pediatrics*, 123(4):1191-207 (2009)). LSDs are typically the result of inborn genetic errors. Most of these disorders are autosomal recessively inherited, however a few are X-linked recessively inherited, such as Fabry disease and Hunter syndrome (MPS II). Affected individuals generally appear normal at birth, however the diseases are progressive. Development of clinical disease may not occur until years or decades later, but is typically fatal. Lysosomal storage diseases affect mostly children and they often die at a young and unpredictable age, many within a few months or years of birth. Many other children die of this disease following years of suffering from various symptoms of their particular disorder. Clinical disease may be manifest as mental retardation and/or dementia, sensory loss including blindness or deafness, motor system dysfunction, seizures, sleep and behavioral disturbances, and so forth. Some people with Lysosomal storage disease have enlarged livers (hepatomegaly) and enlarged spleens (splenomegaly), pulmonary and cardiac problems, and bones that grow abnormally.

[0302] Treatment for many LSDs is enzyme replacement therapy (ERT) and/or substrate reduction therapy (SRT), as well as treatment or management of symptoms. The average annual cost of ERT in the United States ranges from 90,000 to \$565,000. While ERT has significant systemic clinical efficacy for a variety of LSDs, little or no effects are seen on central nervous system (CNS) disease symptoms, because the recombinant proteins cannot penetrate the blood-brain barrier. Allogeneic hematopoietic stem cell transplantation (HSCT) represents a highly effective treatment for selected LSDs. It is currently the only means to prevent the progression of associated neurologic sequelae. However, HSCT is expensive, requires an HLA-matched donor and is associated with significant morbidity and mortality. Recent gene therapy studies suggest that LSDs are good targets for this type of treatment.

[0303] WO 2011/133802 provides examples of donor oligonucleotides that can be utilized in the compositions and methods.

[0304] For example, the compositions and methods can be employed to treat Gaucher's disease (GD). Gaucher's disease, also known as Gaucher syndrome, is the most common lysosomal storage disease. Gaucher's disease is an inherited genetic disease in which lipid accumulates in cells and certain organs due to deficiency of the enzyme glucocerebrosidase (also known as acid  $\beta$ -glucosidase) in lysosomes. Glucocerebrosidase enzyme contributes to the degradation of the fatty substance glucocerebroside (also known as glucosylceramide) by cleaving b-glycoside into b-glucose and ceramide subunits (Scriver C R, Beaudet A L, Valle D, Sly W S. *The metabolic and molecular basis of inherited disease*. 8th ed. New York: McGraw-Hill Pub, 2001: 3635-3668). When the enzyme is defective, the substance accumulates, particularly in cells of the mononuclear cell lineage, and organs and tissues including the spleen, liver, kidneys, lungs, brain and bone marrow.

[0305] There are two major forms: non-neuropathic (type 1, most commonly observed type in adulthood) and neuropathic (type 2 and 3). GBA (GBA glucosidase, beta, acid),

the only known human gene responsible for glucosidase-mediated GD, is located on chromosome 1, location 1q21. More than 200 mutations have been defined within the known genomic sequence of this single gene (NCBI Reference Sequence: NG\_009783.1). The most commonly observed mutations are N370S, L444P, RecNcil, 84GG, R463C, and recTL. 84 GG is a null mutation in which there is no capacity to synthesize enzyme. However, N370S mutation is almost always related with type 1 disease and milder forms of disease. Very rarely, deficiency of sphingolipid activator protein (Gaucher factor, SAP-2, saposin C) may result in GD. In some embodiments, donor oligonucleotides are designed and used to correct mutations in GBA.

**[0306]** In another embodiment, compositions and the methods are used to treat Fabry disease (also known as Fabry's disease, Anderson-Fabry disease, angiokeratoma corporis diffusum and alpha-galactosidase A deficiency), a rare X-linked recessive disorder, resulting from a deficiency of the enzyme alpha galactosidase A (a-GAL A, encoded by GLA). The human gene encoding GLA has a known genomic sequence (NCBI Reference Sequence: NG\_007119.1) and is located at Xp22 of the X chromosome. Mutations in GLA result in accumulation of the glycolipid globotriaosylceramide (abbreviated as Gb3, GL-3, or ceramide trihexoside) within the blood vessels, other tissues, and organs, resulting in impairment of their proper function (Karen, et al., *Dermatol. Online J.*, 11 (4): 8 (2005)). The condition affects hemizygous males (i.e. all males), as well as homozygous, and potentially heterozygous (carrier), females. Males typically experience severe symptoms, while women can range from being asymptomatic to having severe symptoms. This variability is thought to be due to X-inactivation patterns during embryonic development of the female. In some embodiments, donor oligonucleotides are designed and used to correct mutations in GLA.

**[0307]** The compositions and methods can be used to treat Hurler syndrome (HS). Hurler syndrome, also known as mucopolysaccharidosis type I (MPS I),  $\alpha$ -L-iduronidase deficiency, and Hurler's disease, is a genetic disorder that results in the buildup of mucopolysaccharides due to a deficiency of  $\alpha$ -L iduronidase, an enzyme responsible for the degradation of mucopolysaccharides in lysosomes (Dib and Pastories, *Genet. Mol. Res.*, 6(3):667-74 (2007)). MPS I is divided into three subtypes based on severity of symptoms. All three types result from an absence of, or insufficient levels of, the enzyme  $\alpha$ -L-iduronidase. MPS I H or Hurler syndrome is the most severe of the MPS I subtypes. The other two types are MPS I S or Scheie syndrome and MPS I H-S or Hurler-Scheie syndrome. Without  $\alpha$ -L-iduronidase, heparan sulfate and dermatan sulfate, the main components of connective tissues, build-up in the body. Excessive amounts of glycosaminoglycans (GAGs) pass into the blood circulation and are stored throughout the body, with some excreted in the urine. Symptoms appear during childhood, and can include developmental delay as early as the first year of age. Patients usually reach a plateau in their development between the ages of two and four years, followed by progressive mental decline and loss of physical skills (Scott et al., *Hum. Mutat.* 6: 288-302 (1995)). Language may be limited due to hearing loss and an enlarged tongue, and eventually site impairment can result from clouding of cornea and retinal degeneration. Carpal tunnel syndrome (or similar compression of nerves elsewhere in the body) and restricted joint movement are also common.

**[0308]** Exemplary Donors

**[0309]** In some embodiments, the donor oligonucleotide with the sequence

**[0310]** 5' AGGACGGTCCCGGCCTGCGACACTTCCGCCCAT-AATTGTTCTTCA TCTGCGGGGCGGGGGGGGG3' (SEQ ID NO:52), or a functional fragment thereof that is suitable and sufficient to correct the W402X mutation (which is a common mutation for mucopolysaccharidosis type I associated with a severe Hurler syndrome) is administered with a potentiating agent, such as a cell-penetrating antibody, to correct the W402X mutation in cells.

**[0311]** An exemplary donor oligonucleotide can have the sequence

**[0312]** 5' GGGACGGCGCCCCACATAGGCCAAATT-CAATTGCTGATCCCAGCT TAA-GACGTACTGGTCAGCCTGGC 3' (SEQ ID NO:53), or a functional fragment thereof that is suitable to correct the Q70X mutation in the alpha-L-iduronidase gene associated with Hurler's syndrome is administered with a potentiating agent, such as a cell-penetrating antibody, to correct the of Q70X mutation in cells.

**[0313]** In some embodiments, in addition to containing sequence designed to correct the point mutation at Q70X or W402X mutation, the donor oligonucleotides may also contain 7 to 10 additional, synonymous (silent) mutations. The additional silent mutations can facilitate detection of the corrected target sequence using allele-specific PCR of genomic DNA isolated from treated cells.

**[0314]** D. Combination Therapies

**[0315]** Each of the different components for gene editing here can be administered alone or in any combination and further in combination with one or more additional active agents. In all cases, the combination of agents can be part of the same admixture, or administered as separate compositions. In some embodiments, the separate compositions are administered through the same route of administration. In other embodiments, the separate compositions are administered through different routes of administration.

**[0316]** Examples of preferred additional active agents include other conventional therapies known in the art for treating the desired disease or condition. For example, in the treatment of sickle cell disease, the additional therapy may be hydroxyurea.

**[0317]** In the treatment of cystic fibrosis, the additional therapy may include mucolytics, antibiotics, nutritional agents, etc. Specific drugs are outlined in the Cystic Fibrosis Foundation drug pipeline and include, but are not limited to, CFTR modulators such as KALYDECO® (ivacaftor), ORKAMBI™ (lumacaftor+ivacaftor), ataluren (PTC124), VX-661+ivacaftor, riociguat, QBW251, N91115, and QR-010; agents that improve airway surface liquid such as hypertonic saline, bronchitol, and P-1037; mucus alteration agents such as PULMOZYME® (dornase alfa); anti-inflammatories such as ibuprofen, alpha 1 anti-trypsin, CTX-4430, and JBT-101; anti-infective such as inhaled tobramycin, azithromycin, CAYSTON® (aztreonam for inhalation solution), TOBI inhaled powder, levofloxacin, ARIKACE® (nebulized liposomal amikacin), AEROVANC® (vancomycin hydrochloride inhalation powder), and gallium; and nutritional supplements such as aquADEKs, pancrelipase enzyme products, lipotamase, and burlulipase.

**[0318]** In the treatment of HIV, the additional therapy maybe an antiretroviral agents including, but not limited to,

a non-nucleoside reverse transcriptase inhibitor (NNRTIs), a nucleoside reverse transcriptase inhibitor (NRTIs), a protease inhibitors (PIs), a fusion inhibitors, a CCR5 antagonists (CCR5s) (also called entry inhibitors), an integrase strand transfer inhibitors (INSTIs), or a combination thereof.

[0319] In the treatment of lysosomal storage disease, the additional therapy could include, for example, enzyme replacement therapy, bone marrow transplantation, or a combination thereof.

[0320] E. Determining Gene Modification

[0321] Sequencing and allele-specific PCR are preferred methods for determining if gene modification has occurred. PCR primers are designed to distinguish between the original allele, and the new predicted sequence following recombination. Other methods of determining if a recombination event has occurred are known in the art and may be selected based on the type of modification made. Methods include, but are not limited to, analysis of genomic DNA, for example by sequencing, allele-specific PCR, droplet digital PCR, or restriction endonuclease selective PCR (REMS-PCR); analysis of mRNA transcribed from the target gene for example by Northern blot, in situ hybridization, real-time or quantitative reverse transcriptase (RT) PCT; and analysis of the polypeptide encoded by the target gene, for example, by immunostaining, ELISA, or FACS. In some cases, modified cells will be compared to parental controls. Other methods may include testing for changes in the function of the RNA transcribed by, or the polypeptide encoded by the target gene. For example, if the target gene encodes an enzyme, an assay designed to test enzyme function may be used.

[0322] The present invention will be further understood by reference to the following non-limiting examples.

## EXAMPLES

### Example 1: 3E10 Enhances Gene Editing by NPs Containing Only Donor DNA

#### Materials and Methods

[0323] Donor DNA and Nanoparticle Synthesis

[0324] The single-stranded donor DNA oligomer was prepared by standard DNA synthesis and 5' and 3'-end protected by inclusion of three phosphorothioate internucleoside linkages at each end. The sequence of the donor DNA matches positions 624 to 684 in  $\beta$ -globin intron 2 and is as follows, with the correcting IVS2-654 nucleotide underlined:

(SEQ ID NO: 43)  
5' AAAGAATAACAGTGATAATTTCTGGGTTAAG  
GCAATAGCAATATCTCTGCATATAAATAT3'.

[0325] Nanoparticles containing donor DNAs or donor DNAs and PNAs were synthesized by a double-emulsion solvent evaporation protocol as previously described (Bahal, et al., *Nat. Commun.*, 7:13304 (2016)). Briefly, poly(lactic-co-glycolic) acid (PLGA) was dissolved in dichloromethane at a concentration of 40 mg/ml. Prior to encapsulation, PNA and donor DNA were mixed at a 2:1 molar ratio and added dropwise to the PLGA solution under vortex. For NPs containing donor DNA alone, DNA was added dropwise at a molar ratio of 2 nmoles/mg of polymer. The resulting mixture was sonicated three times for 10 seconds using an

amplitude of 38%. The water-in-oil emulsion was subsequently added dropwise to a surfactant solution containing polyvinyl alcohol (5% w/v). Following the second emulsion, the sonication step was repeated as described. The resulting nanoparticle solution was added to 25 ml of a 0.3% PVA solution and allowed to stir for 3 hours at room temperature. After stirring and particle 'hardening,' the nanoparticles were washed 3 times via centrifugation (16,100 g, 15 min, 4° C.) before being flash frozen and lyophilized in cryoprotectant (trehalose, mg:mg). Dry nanoparticles were stored at -20° C. until later use.

[0326] Cell Isolation, Culture and Treatment

[0327] Mouse embryonic fibroblasts (MEFs) were isolated from a transgenic mouse model carrying a  $\beta$ -globin/GFP fusion transgene consisting of human  $\beta$ -globin intron 2 with a thalassemia-associated IVS2-654 (C→T) mutation inserted within the GFP coding sequence (Sazani, et al., *Nat. Biotechnol.*, 20:1228-1233 (2002); Bahal, et al., *Nat. Commun.*, 7:13304 (2016)). This results in incorrect splicing of  $\beta$ -globin/GFP mRNA and lack of GFP expression.

[0328] By simple addition to the cell culture, MEFs were treated with 2 mg of nanoparticles containing donor DNA alone or PNA in combination with donor DNA, at a seeding density of 5,000 cells/well in DMEM media, containing 10% FBS.

[0329] In some samples, prior to treatment with donor DNA containing nanoparticles, MEFs were treated with 3E10 WT (containing wild type 3E10 heavy and light chain sequences) at a final concentration of 7.5  $\mu$ M. 3E10 WT was prepared as a full-length recombinant antibody by expression from a recombinant expression construct in 293 cells based on the publicly available sequence and purified by standard techniques. The cells were analyzed for gene editing 72 h later by fluorescence via flow cytometry.

#### Results

[0330] As shown in FIG. 1, treatment with donor DNA NPs alone resulted in some detectable editing slightly above background (untreated control). The addition of 3E10 significantly increased this editing (FIG. 1). Treatment with donor DNA NPs+3E10 achieved higher percentage editing compared to cells treated with NPs containing donor DNA/PNAs (an approach that has previously been established for gene editing (Bahal, et al., *Nat. Commun.*, 7:13304 (2016))).

### Example 2: 3E11 Enhances Editing of the Beta Globin Gene by Naked Donor DNA k Bone Marrow Cells Ex Vivo

#### Materials and Methods

[0331] Donor DNA

The single-stranded donor DNA oligomer was prepared by standard DNA synthesis and 5' and 3'-end protected by inclusion of three phosphorothioate internucleoside linkages at each end. The donor DNA sequence is a 60-nucleotide single stranded DNA matching the wild type human beta globin sequence centered on codon 6, the position of the sickle cell mutation. The sequence of the donor DNA is as follows:

(SEQ ID NO: 48)

5' TTGCCCCACAGGGCAGTAACGGCAGACTTCT

CCTCAGGAGTCAGGTGCACCATGGTGTCTGTTTG3'.

**[0332]** Mouse Model for Sickle Cells Disease

**[0333]** In sickle cell disease (SCD), the mutation (GAG->GTG) at codon 6 results in glutamic acid changed to valine. For correction (editing) of this SCD mutation site, studies were performed in the Townes mouse model.

**[0334]** The Townes mouse model was developed by Ryan T M, Ciavatta D J, Townes T M., "Knockout-transgenic mouse model of sickle cell disease." *Science*. 1997 Oct. 31; 278(5339):873-6. PMID: 9346487.

**[0335]** Townes mice exclusively express human sickle hemoglobin (HbS). They were produced by generating transgenic mice expressing human  $\alpha$ -,  $\gamma$ -, and  $\beta^S$ -globin that were then bred with knockout mice that had deletions of the murine  $\alpha$ - and  $\beta$ -globin genes. Thus, the resulting progeny no longer express mouse  $\alpha$ - and  $\beta$ -globin. Instead, they express exclusively human  $\alpha$ - and  $\beta^S$ -globin. Hence, the mice express human sickle hemoglobin and possess many of the major hematologic and histopathologic features of individuals with SCD.

**[0336]** Cell Isolation, Culture and Treatment

**[0337]** To further evaluate the effects of 3E10 on donor DNA directed gene editing, a transgenic mouse model of sickle cell disease, designed with human hemoglobin alpha (ha) and human sickle hemoglobin beta ( $\beta^S$ ), as described above, was used (Townes model, Jackson Laboratory). Bone marrow cells were harvested from these mice and cultured in 12-well plates at a density of 200,000 cells per well in RPMI with 20% FBS and 1% pen-strep.

**[0338]** Prior to treatment, 3E10 D31N (containing 3E10 D31N heavy chain and wild type light chain sequences) was co-incubated with naked donor DNA for 5 minutes (to allow non-covalent antibody-DNA complexes to form) in sodium acetate buffer at room temperature. Bone marrow cells were then treated with the antibody-DNA mixture at a final concentration of 7.5  $\mu$ M 3E10 and 0.5  $\mu$ g/ $\mu$ l of donor DNA.

**[0339]** After 72 hours, the cells were washed 3 times prior to genomic DNA (gDNA) isolation (SV Wizard, Promega). Freshly isolated gDNA was analyzed by droplet digital PCR (ddPCR) for quantitative determination of gene editing frequencies.

## Results

**[0340]** Experiments were designed to determine if 3E10's ability to bind to DNA and to penetrate cells could obviate the need for nanoparticle delivery (Weisbart et al., *Sci. Rep.*, 5:12022 (2015)). Therefore, rather than encapsulate donor DNA into nanoparticles (as described in Example 1), the ability of 3E10 in combination with naked donor DNA to mediate gene editing in bone marrow cells from a transgenic model of sickle cell disease was evaluated.

**[0341]** As shown in FIG. 2, substantial gene editing frequencies in the range of 8% were observed when cells were treated with pre-incubated mixture of 3E10 and donor DNA. To ensure that editing was not an artifact, ddPCR controls consisting of donor alone and donor spiked into genomic DNA from untreated cells were included as samples. No background signal was detected in the controls, ruling out artefacts (FIG. 2).

## Example 3: 3E10 Enhances Editing of the Beta Globin Gene by Naked Donor DNA in MEFs

## Materials and Methods

**[0342]** Mouse embryonic fibroblasts (MEFs) were isolated from Townes mouse embryos (the same sickle cell transgenic mouse model used in Example 2). MEFs were subsequently seeded in a 12-well plate at a seeding density of 200,000 cells per well in DMEM media containing 20% FBS and 1% pen-strep.

**[0343]** Prior to treatment, 3E10 D31N (containing 3E10 D31N heavy chain and wild type light chain sequences) and donor DNA were incubated for 5 minutes. MEFs were subsequently treated at a final concentration of 7.5  $\mu$ M 3E10 and 0.5  $\mu$ g/ $\mu$ l of donor DNA.

**[0344]** After 72 hours, the cells were washed 3 times prior to gDNA isolation (SV Wizard, Promega). Freshly isolated genomic DNA was analyzed by digital droplet PCR (ddPCR) for editing frequencies.

## Results

**[0345]** To extend the findings with 3E10 plus donor DNA in treated bone marrow cells (described above in Example 2) to another cell type, the efficiency of gene editing in MEFs by 3E10 in combination with naked donor DNA was evaluated. Compared to the very low levels of gene editing observed in MEFs treated with donor DNA alone, the combination of 3E10 and donor DNA achieved high levels of gene editing in the range of 13% (FIG. 3).

## Example 4: 3E10 and Donor DNA Alone Achieves In Vivo Editing in Townes Mice

## Materials and Methods

**[0346]** To test whether 3E10 could promote donor DNA directed gene editing in vivo in a living animal, the Townes model (the same sickle cell transgenic mouse model used in Examples 2 and 3) was used. Prior to treatment of mice, 1 mg of 3E10 D31N (containing 3E10 D31N heavy chain and wild type light chain sequences) was mixed with 330  $\mu$ g of donor DNA for 5 minutes in sodium acetate buffer at room temperature. Two mice per experimental condition were subsequently injected intraperitoneally (i.p.) with a total of 4 doses of 3E10/donor DNA given at 2 day intervals. Each dose consisted of 1 mg of 3E10 and 330  $\mu$ g of donor DNA. After 2 months, bone marrow cells were harvested and analyzed via digital droplet PCR (ddPCR) for gene editing in the beta globin gene at the sickle cell mutation

## Results

**[0347]** Compared to control mice treated with blank PLGA nanoparticles, mice treated with 3E10 plus donor DNA demonstrated significantly higher levels of gene editing, with a frequency of 1.5% (FIG. 4).

**[0348]** Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of skill in the art to which the invention belongs. Publications cited herein and the materials for which they are cited are specifically incorporated by reference.

**[0349]** Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

## SEQUENCE LISTING

Sequence total quantity: 53

SEQ ID NO: 1 moltype = AA length = 116

FEATURE Location/Qualifiers

source 1..116

mol\_type = protein

note = synthetic polynucleotide

organism = synthetic construct

SEQUENCE: 1

EVQLVESGGG LVKPGGSRKL SCAASGFTFS DYGMHWVRQA PEKGLEWVAY ISSGSSTIYY 60

ADTVKGRFTI SRDNAKNTLF LQMTSLRSED TAMYVCARRG LLLDYWGQGT TLTVSS 116

SEQ ID NO: 2 moltype = AA length = 116

FEATURE Location/Qualifiers

source 1..116

mol\_type = protein

note = synthetic polynucleotide

organism = synthetic construct

SEQUENCE: 2

EVQLVESGGG LVKPGGSRKL SCAASGFTFS NYGMHWVRQA PEKGLEWVAY ISSGSSTIYY 60

ADTVKGRFTI SRDNAKNTLF LQMTSLRSED TAMYVCARRG LLLDYWGQGT TLTVSS 116

SEQ ID NO: 3 moltype = AA length = 116

FEATURE Location/Qualifiers

source 1..116

mol\_type = protein

note = synthetic polynucleotide

organism = synthetic construct

SEQUENCE: 3

EVQLVQSGGG LIQPGGSLRL SCAASGFTFS NYGMHWVRQA PGKGLEWVSY ISSGSSTIYY 60

ADSVKGRFTI SRDNSKNTLY LQMNSLRAED TAVYVCARRG LLLDYWGQGT TVTVSS 116

SEQ ID NO: 4 moltype = AA length = 116

FEATURE Location/Qualifiers

source 1..116

mol\_type = protein

note = synthetic polynucleotide

organism = synthetic construct

SEQUENCE: 4

EVQLVESGGG LIQPGGSLRL SCAASGFTFS NYGMHWVRQA PGKGLEWVSY ISSGSSTIYY 60

ADSVKGRFTI SRDNSKNTLY LQMTSLRAED TAVYVCARRG LLLDYWGQGT TLTVSS 116

SEQ ID NO: 5 moltype = AA length = 116

FEATURE Location/Qualifiers

source 1..116

mol\_type = protein

note = synthetic polynucleotide

organism = synthetic construct

SEQUENCE: 5

EVQLQESGGG VVQPGGSLRL SCAASGFTFS NYGMHWIRQA PGKGLEWVSY ISSGSSTIYY 60

ADSVKGRFTI SRDNSKNTLY LQMNSLRSED TAVYVCARRG LLLDYWGQGT LVTVSS 116

SEQ ID NO: 6 moltype = AA length = 116

FEATURE Location/Qualifiers

source 1..116

mol\_type = protein

note = synthetic polynucleotide

organism = synthetic construct

SEQUENCE: 6

EVQLVESGGG LVQPGGSLRL SCSASGFTFS NYGMHWVRQA PGKGLEVSY ISSGSSTIYY 60

ADTVKGRFTI SRDNSKNTLY LQMSSLRAED TAVYCVKRG LLLDYWGQGT LVTVSS 116

SEQ ID NO: 7 moltype = AA length = 111

FEATURE Location/Qualifiers

source 1..111

mol\_type = protein

note = synthetic polynucleotide

organism = synthetic construct

SEQUENCE: 7

DIVLTQSPAS LAVSLGQRAT ISCRASKSVS TSSYSYMHY QQKPGQPPKL LIKYASYLES 60

GVPARFSGSG SGTDFTLNIH PVEEEDAATY YCQHSREFPW TFGGGTKLEI K 111

SEQ ID NO: 8 moltype = AA length = 111

FEATURE Location/Qualifiers

source 1..111

mol\_type = protein

-continued

	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 8		
DIVLTQSPAS	LAVSLGQRAT ISCRASKSVS TSSYSYMHWY QOKPGQPPKL LIKYASYLES	60
GVPARFSGSG	SGTDFHLNIH PVEEEDAATY YCQHSREFPW TFGGGTKLEL K	111
SEQ ID NO: 9	moltype = AA length = 111	
FEATURE	Location/Qualifiers	
source	1..111	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 9		
DIQMTQSPSS	LSASVGDRVT ITCRASKSVS TSSYSYLAHY QOKPEKAPKL LIKYASYLQS	60
GVPSRFSGSG	SGTDFTLTIS SLQPEDFATY YCQHSREFPW TFGAGTKLEL K	111
SEQ ID NO: 10	moltype = AA length = 111	
FEATURE	Location/Qualifiers	
source	1..111	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 10		
DIQMTQSPSS	LSASVGDRVT ISCRASKSVS TSSYSYMHWY QOKPEKAPKL LIKYASYLQS	60
GVPSRFSGSG	SGTDFTLTIS SLQPEDVATY YCQHSREFPW TFGAGTKLEL K	111
SEQ ID NO: 11	moltype = AA length = 111	
FEATURE	Location/Qualifiers	
source	1..111	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 11		
DIVLTQSPAS	LAVSPGQRAT ITCRASKSVS TSSYSYMHWY QOKPGQPPKL LIYYASYLES	60
GVPARFSGSG	SGTDFTLTIN PVEANDTANY YCQHSREFPW TFGQGTKVEI K	111
SEQ ID NO: 12	moltype = AA length = 465	
FEATURE	Location/Qualifiers	
source	1..465	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 12		
MGWSCIIILFL	VATATGVHSE VQLVESGGGL VKPGGSRKLS CAASGFTFSD YGMHWVRQAP	60
EKGLEWVAYI	SSGSSTIYYA DTVKGRFTIS RDNAKNTLFL QMTSLRSEDY AMYYCARRGL	120
LLDYWGQGT	LTVSAASTKG PSVFPLAPSS KSTSGGTAAL GCLVKDYFPE PVTVSWNSGA	180
LTSGVHTFPA	VLQSSGLYSL SSVVTVPSST LGTQTYICNV NHKPSNTKVD KKVEPKSCDK	240
THTCPPCPAP	ELLGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV	300
EVHNAKTKPR	EEQYNSTYRV VSVLTVLHQD WLNKEYKCK VSNKALPAPI EKTISKAKGQ	360
PREPQVYTL	PSRDELTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG	420
SFFLYSKLTV	DKSRWQQGNV FSCSVMEAL HNHYTQKSL LSPGK	465
SEQ ID NO: 13	moltype = AA length = 465	
FEATURE	Location/Qualifiers	
source	1..465	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 13		
MGWSCIIILFL	VATATGVHSE VQLVESGGGL VKPGGSRKLS CAASGFTFSN YGMHWVRQAP	60
EKGLEWVAYI	SSGSSTIYYA DTVKGRFTIS RDNAKNTLFL QMTSLRSEDY AMYYCARRGL	120
LLDYWGQGT	LTVSAASTKG PSVFPLAPSS KSTSGGTAAL GCLVKDYFPE PVTVSWNSGA	180
LTSGVHTFPA	VLQSSGLYSL SSVVTVPSST LGTQTYICNV NHKPSNTKVD KKVEPKSCDK	240
THTCPPCPAP	ELLGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV	300
EVHNAKTKPR	EEQYNSTYRV VSVLTVLHQD WLNKEYKCK VSNKALPAPI EKTISKAKGQ	360
PREPQVYTL	PSRDELTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG	420
SFFLYSKLTV	DKSRWQQGNV FSCSVMEAL HNHYTQKSL LSPGK	465
SEQ ID NO: 14	moltype = AA length = 237	
FEATURE	Location/Qualifiers	
source	1..237	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 14		
MGWSCIIILFL	VATATGVHSD IVLTQSPASL AVSLGQRATI SCRASKSVST SSYSYMHWYQ	60

QKPGQPPKLL IKYASYLESG VPARFSGSGS GTDFTLNIHP VEEEDAATYY CQHSREFPWT	120
FGGGTKLEIK RTVAAPSVFI FPPSDEQLKS GTASVVCLLN NFYPREAKVQ WKVDNALQSG	180
NSQESVTEQD SKDSTYSLSS TLTLISKADYE KHKVYACEVT HQGLSSPVTK SFNRGEC	237
SEQ ID NO: 15	moltype = AA length = 5
FEATURE	Location/Qualifiers
source	1..5
	mol_type = protein
	note = synthetic polynucleotide
	organism = synthetic construct
SEQUENCE: 15	
DYGMH	5
SEQ ID NO: 16	moltype = AA length = 5
FEATURE	Location/Qualifiers
source	1..5
	mol_type = protein
	note = synthetic polynucleotide
	organism = synthetic construct
SEQUENCE: 16	
NYGMH	5
SEQ ID NO: 17	moltype = AA length = 17
FEATURE	Location/Qualifiers
source	1..17
	mol_type = protein
	note = synthetic polynucleotide
	organism = synthetic construct
SEQUENCE: 17	
YISSGSSTIY YADTVKG	17
SEQ ID NO: 18	moltype = AA length = 7
FEATURE	Location/Qualifiers
source	1..7
	mol_type = protein
	note = synthetic polynucleotide
	organism = synthetic construct
SEQUENCE: 18	
RGLLLDY	7
SEQ ID NO: 19	moltype = AA length = 17
FEATURE	Location/Qualifiers
source	1..17
	mol_type = protein
	note = synthetic polynucleotide
	organism = synthetic construct
SEQUENCE: 19	
YISSGSSTIY YADSVKG	17
SEQ ID NO: 20	moltype = AA length = 8
FEATURE	Location/Qualifiers
source	1..8
	mol_type = protein
	note = synthetic polynucleotide
	organism = synthetic construct
SEQUENCE: 20	
GFTFSDYG	8
SEQ ID NO: 21	moltype = AA length = 8
FEATURE	Location/Qualifiers
source	1..8
	mol_type = protein
	note = synthetic polynucleotide
	organism = synthetic construct
SEQUENCE: 21	
GFTFSNYG	8
SEQ ID NO: 22	moltype = AA length = 8
FEATURE	Location/Qualifiers
source	1..8
	mol_type = protein
	note = synthetic polynucleotide
	organism = synthetic construct
SEQUENCE: 22	
ISSGSSTI	8

-continued

---

SEQ ID NO: 23	moltype = AA length = 9	
FEATURE	Location/Qualifiers	
source	1..9	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 23		
ARRGLLLDY		9
SEQ ID NO: 24	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
source	1..15	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 24		
RASKSVSTSS YSYMH		15
SEQ ID NO: 25	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
source	1..7	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 25		
YASYLES		7
SEQ ID NO: 26	moltype = AA length = 9	
FEATURE	Location/Qualifiers	
source	1..9	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 26		
QHSREFPWT		9
SEQ ID NO: 27	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
source	1..15	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 27		
RASKSVSTSS YSYLA		15
SEQ ID NO: 28	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
source	1..7	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 28		
YASYLQS		7
SEQ ID NO: 29	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
source	1..10	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 29		
KSVSTSSYSY		10
SEQ ID NO: 30	moltype = length =	
SEQUENCE: 30		
000		
SEQ ID NO: 31	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
source	1..7	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 31		
GQSSRSS		7

-continued

SEQ ID NO: 32	moltype = AA length = 18	
FEATURE	Location/Qualifiers	
source	1..18	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 32		
GQSSRSSSGG GSSGGGGS		18
SEQ ID NO: 33	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
source	1..4	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 33		
GSGS		4
SEQ ID NO: 34	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
source	1..4	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 34		
GGGS		4
SEQ ID NO: 35	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
source	1..10	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 35		
GGGSGGGGS		10
SEQ ID NO: 36	moltype = AA length = 20	
FEATURE	Location/Qualifiers	
source	1..20	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 36		
GGGSGGGGS GGGSGGGGS		20
SEQ ID NO: 37	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
source	1..15	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 37		
GGGSGGGGS GGGGS		15
SEQ ID NO: 38	moltype = AA length = 274	
FEATURE	Location/Qualifiers	
source	1..274	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 38		
AGIHDIVLTQ SPASLAVSLG QRATISCRAS KSVSTSSYSY MHWYQQKPGQ PPKLLIKYAS		60
YLESGVPARF SGSGSGTDFT LNIHPVEEED AATYYCQHSR EFPWTFGGGT KLEIKRADAA		120
PGGGGSGGGG SGGGGSEVQL VESGGGLVKP GGSRLSCAA SGFTFSNYGM HWVRQAPEKG		180
LEWVAYISSG SSTIYYADTV KGRFTISRDN AKNTLFLQMT SLRSEDAMY YCARRGLLLD		240
YWQGTTTLTV SSLEQKLISE EDLNSAVDHH HHHH		274
SEQ ID NO: 39	moltype = AA length = 541	
FEATURE	Location/Qualifiers	
source	1..541	
	mol_type = protein	
	note = synthetic polynucleotide	
	organism = synthetic construct	
SEQUENCE: 39		
AGIHDIVLTQ SPASLAVSLG QRATISCRAS KSVSTSSYSY MHWYQQKPGQ PPKLLIKYAS		60
YLESGVPARF SGSGSGTDFT LNIHPVEEED AATYYCQHSR EFPWTFGGGT KLEIKRADAA		120

-continued

PGGGGSGGGG	SGGGGSEVQL	VESGGGLVKP	GGSRKLSCAA	SGFTFSNYGM	HWVRQAPEKG	180
LEWVAYISSG	SSTIYYADTV	KGRFTISRDN	AKNTLFLQMT	SLRSEDAMY	YCARRGLLLD	240
YWQGTTTLTV	SSASTKGPSV	FPLAPLESSG	SDIVLTQSPA	SLAVSLGQRA	TISCRASKSV	300
STSSYSYMHW	YQKPGQPPK	LLIKYASYLE	SGVPARFSGS	SGTDFTLNI	HPVEEEDAAT	360
YYCQHSREFP	WTFGGGTKLE	IKRADAAPGG	GGSGGGGSGG	GGSEVQLVES	GGGLVKPGGS	420
RKLSCAASGF	TFSNYGMHWV	RQAPEKGLEW	VAYISSGSST	IYYADTVKGR	FTISRDNAKN	480
TLFLQMTSLR	SEDAMYCA	RRGLLLDYWG	QGTTLTVSSL	EQKLISEEDL	NSAVDHHHHH	540
H						541
SEQ ID NO: 40	moltype = AA length = 808					
FEATURE	Location/Qualifiers					
source	1..808					
	mol_type = protein					
	note = synthetic polynucleotide					
	organism = synthetic construct					
SEQUENCE: 40						
AGIHDIVLTQ	SPASLAVSLG	QRATISCRAS	KSVSTSSYSY	MHWYQKPGQ	PPKLLIKYAS	60
YLESGVPARF	SGSGSGTDFT	LNIHPVEEED	AATYYCQHSR	EFPWTFGGGT	KLEIKRADAA	120
PGGGGSGGGG	SGGGGSEVQL	VESGGGLVKP	GGSRKLSCAA	SGFTFSNYGM	HWVRQAPEKG	180
LEWVAYISSG	SSTIYYADTV	KGRFTISRDN	AKNTLFLQMT	SLRSEDAMY	YCARRGLLLD	240
YWQGTTTLTV	SSASTKGPSV	FPLAPLESSG	SDIVLTQSPA	SLAVSLGQRA	TISCRASKSV	300
STSSYSYMHW	YQKPGQPPK	LLIKYASYLE	SGVPARFSGS	SGTDFTLNI	HPVEEEDAAT	360
YYCQHSREFP	WTFGGGTKLE	IKRADAAPGG	GGSGGGGSGG	GGSEVQLVES	GGGLVKPGGS	420
RKLSCAASGF	TFSNYGMHWV	RQAPEKGLEW	VAYISSGSST	IYYADTVKGR	FTISRDNAKN	480
TLFLQMTSLR	SEDAMYCA	RRGLLLDYWG	QGTTLTVSSA	STKGPSVFPL	APLESSGSDI	540
VLTSQPASLA	VSLGQRATIS	CRASKSVSTS	SYSYMHWYQQ	KPGQPPKLLI	KYASYLESGV	600
PARFSGSGSG	TDFTLNIHPV	EEEDAATYYC	QHSREFPWTF	GGGTKLEIKR	ADAAPGGGGS	660
GGGGSGGGGS	EVQLVESGGG	LVKPGGSRKL	SCAASGFTFS	NYGMHWVRQA	PEKGLEWVAY	720
ISSGSSTIYY	ADTVKGRFTI	SRDNAKNTLF	LQMTSLRSED	TAMYYCARRG	LLLDYWGQGT	780
TLTVSSLEQK	LISEEDLNSA	VDHHHHHH				808
SEQ ID NO: 41	moltype = AA length = 242					
FEATURE	Location/Qualifiers					
source	1..242					
	mol_type = protein					
	note = synthetic polynucleotide					
	organism = synthetic construct					
SEQUENCE: 41						
DIVLTQSPAS	LAVSPGQRAT	ITCRASKSVS	TSSYSYMHWY	QKPGQPPKL	LIYYASYLES	60
GVPARFSGSG	SGTDFTLTIN	PVEANDTANY	YCQHSREFPW	TFGQGTKVEI	KGGGGSGGGG	120
SGGGGSEVQL	VESGGGLVQP	GGSLRLSCSA	SGFTFSNYGM	HWVRQAPGKG	LEYVSYISSG	180
SSTIYYADTV	KGRFTISRDN	SKNTLYLQMS	SLRAEDTAVY	YCVKRGLLLD	YWQGTLVTV	240
SS						242
SEQ ID NO: 42	moltype = DNA length = 6060					
FEATURE	Location/Qualifiers					
source	1..6060					
	mol_type = other DNA					
	note = synthetic polynucleotide					
	organism = synthetic construct					
SEQUENCE: 42						
aaagctcttg	ctttgacaat	tttggtcttt	cagaatacta	taaatataac	ctatatattata	60
atttcataaa	gtctgtgcat	tttctttgac	ccaggatatt	tgcaaaagac	atattcaaac	120
ttccgcagaa	cactttat	cacatataca	tgctcttat	atcagggatg	tgaaacaggg	180
tcttgaaaac	tgtctaaatc	taaaacaatg	ctaatagcagg	tttaaattta	ataaaaataaa	240
atccaaaatc	taacagccaa	gtcaaactctg	tatgttttaa	catttaaaat	attttaaaga	300
cgtcttttcc	caggattcaa	catgtgaaat	cttttctcag	ggatacacgt	gtgcctagat	360
cctcattgct	ttagtttttt	acagaggaat	gaatataaaa	agaaaatact	taaattttat	420
ccctcttacc	tctataatca	tacataggca	taatttttta	acctaggctc	cagatagcca	480
tagaagaacc	aaacactttc	tgcgtgtgtg	agaataatca	gagtgaagatt	ttttcacaag	540
tacctgatga	gggttgagac	aggtagaaaa	agtgaagat	ctctatttat	ttagcaataa	600
tagagaaagc	atttaagaga	ataaagcaat	ggaaataaga	aatttgtaaa	tttccttctg	660
ataactagaa	atagaggatc	cagtttcttt	tggttaacct	aaattttatt	tcattttatt	720
gttttat	attttatttt	attttatttt	gtgtaatcgt	agtttcagag	tgtagagct	780
gaaaggaaga	agtaggagaa	acatgcaaag	taaaagtata	acactttcct	tactaaaccg	840
actgggtttc	caggtagggg	caggattcag	gatgactgac	agggccctta	gggaacactg	900
agaccctacg	ctgacctcat	aaatgcttgc	tacctttgct	gttttaatta	catcttttaa	960
tagcaggaag	cagaactctg	cacttcaaaa	gttttctctc	acctgaggag	ttaatttagt	1020
acaaggggaa	aaagtacagg	gggatgggag	aaaggcgatc	acgttgggaa	gctatagaga	1080
aagaagagta	aatttttagta	aaggaggttt	aaacaaacaa	aatataaaga	gaaataggaa	1140
cttgaatcaa	ggaaatgatt	ttaaaacgca	gtattcttag	tggaactagag	gaaaaaaata	1200
atctgagcca	agtagaagac	cttttcccct	cctaccctta	ctttctaagt	cacagaggct	1260
ttttgttccc	ccagacactc	ttgcagatta	gtccaggcag	aaacagttag	atgtccccag	1320
ttaacctcct	atttgacacc	actgattacc	ccattgatag	tcacactttg	ggttgtaagt	1380
gactttttat	ttatttgtat	ttttgactgc	attaagaggt	ctctagtttt	ttatctcttg	1440
tttcccaaaa	cctaataagt	aactaatgca	cagagcacat	tgatttgtat	ttattctatt	1500

-continued

tttagacata	atattattagc	atgcatgagc	aaattaagaa	aaacaacaac	aatgaatgc	1560
atatatatgt	atatgtatgt	gtgtatatat	acacatatat	atatatatatt	tttttctttt	1620
cttaccagaa	ggttttaatc	caaataagga	gaagatatgc	ttagaactga	ggtagagttt	1680
tcatccattc	tgtcctgtaa	gtattttgca	tattctggag	acgcaggaag	agatccatct	1740
acatatccca	aagctgaatt	atggtagaca	aagctcttcc	acttttagtg	catcaatttc	1800
ttatttgtgt	aataagaaaa	ttgggaaaac	gatcttcaat	atgcttacca	agctgtgatt	1860
ccaaatatta	cgtaaatata	cttgcaaagg	aggatgtttt	tagtagcaat	ttgtactgat	1920
ggatatggggc	caagagatat	atcttagagg	gagggctgag	ggtttgaagt	ccaactccta	1980
agccagtgcc	agaagagcca	aggacaggta	cggctgtcat	cacttagacc	tcaccctgtg	2040
gagccacacc	ctagggttgg	ccaatctact	cccaggagca	gggagggcag	gagccagggc	2100
tgggcataaa	agtcagggca	gagccatcta	ttgcttacat	ttgcttctga	cacaactgtg	2160
ttacttagca	acctcaaaac	gacaccatgg	tgcacctgac	tcctgaggag	aagtctgccg	2220
ttactgccct	gtggggcaag	gtgaacgtgg	atgaagtggg	tggtgaggcc	ctgggcaggt	2280
tggtatcaag	gttacaagac	aggtttaagg	agaccaatag	aaactgggca	tgtggagaca	2340
gagaagactc	ttgggtttct	gataggcact	gactctctct	gcctattggg	ctattttccc	2400
acccttaggc	tgctgggtgg	ctacccttgg	accagagggt	tctttgagtc	ctttggggat	2460
ctgtccactc	ctgatgctgt	tatgggcaac	cctaagggtga	aggctcatgg	caagaaagtg	2520
ctcgggtgct	ttagtgtatg	cctggctcac	ctggacaacc	tcaagggcac	ctttgccaca	2580
ctgagtgage	tgcactgtga	caagctgcac	gtggatcctg	agaacttcag	ggtgagtcta	2640
tgggaccctt	gatgttttct	ttccccttct	tttctatggg	taagttcatg	tcataggaag	2700
gggagaagta	acaggttaca	gtttagaatg	ggaaacagac	gaatgattgc	atcagtgtgg	2760
aagtctcagg	atcgttttag	tttcttttat	ttgctgttca	taacaattgt	tttcttttgt	2820
ttaattcttg	ctttcttttt	ttttcttctc	cgcaattttt	actattatac	ttaatgcctt	2880
aacatttgtg	ataacaaaag	gaaatatctc	tgagatacat	taagtaactt	aaaaaaaaac	2940
tttacacagt	ctgcctagta	cattactatt	tggaaatata	gtgtgcttat	ttgcatattc	3000
ataatctccc	tactttattt	tcttttattt	ttaattgata	cataatcatt	atacatattt	3060
atgggttaaa	gtgtaatgtt	ttaatatgtg	tacacatatt	gaccaaatac	gggtaatttt	3120
gcatttgtta	ttttaaaaaa	tgttttcttc	ttttaatata	cttttttgtt	tatcttattt	3180
ctaatacttt	ccctaactct	tttctttcag	ggcaataatg	atacaatgta	tcatgcctct	3240
ttgcaccatt	ctaaagaata	acagtgataa	tttctgggtt	aaggcaatag	caatatttct	3300
gcataataat	atcttctgcat	ataaattgta	actgatgtaa	gaggtttcat	attgctaata	3360
gcagctacaa	tccagctacc	attctgcttt	tattttatgg	ttgggataag	gctggattat	3420
tctgagtcca	agctaggccc	ttttgctaata	catgttcata	cctcttatct	tcctcccaca	3480
gctcctgggc	aacgtgctgg	tctgtgtgct	ggcccatcac	tttgccaaag	aattcacccc	3540
accagtgcag	gctgcctatc	agaaagtggg	ggctgggtg	gctaattgcc	tggccacaaa	3600
gtatcactaa	gctcgctttc	ttgctgtcca	atttctatta	aaggttcctt	tgttccttaa	3660
gtccaactac	taaactgggg	gatattatga	agggccttga	gcatctggat	tctgcctaata	3720
aaaaaacatt	tattttcatt	gcaatgatgt	atttaaatta	tttctgaata	ttttactaaa	3780
aagggaatgt	gggaggtcag	tgcattttaa	acataaagaa	atgaagagct	agttcaaacc	3840
ttgggaaaat	acactatata	ttaaactcca	tgaagaagg	tgaggctgca	aacagctaata	3900
gcacattggc	aacagccctg	atgcctatgc	cttatctcat	cctcagaaaa	ggattcaagt	3960
agaggtctga	tttgagggtt	aaagttttgc	tatgctgtat	tttacattac	ttattgtttt	4020
agctgtcctc	atgaatgtct	tttctactac	catttgccta	tcctgcatct	ctcagccttg	4080
actccactca	gttctcttgc	ttagagatac	cactcttccc	ctgaagtgtt	ccttccatgt	4140
tttacggcga	gatggtttct	cctcgccctg	ccactcagcc	ttagttgtct	ctgttgtctt	4200
atagaggtct	acttgaagaa	ggaaaaacag	ggggcatggg	ttgactgtcc	tgtgagccct	4260
tcttccctgc	ctccccact	cacagtgacc	cggaaatctg	agtgctagtc	tcctcggaact	4320
atcactcttt	cacagtctgc	tttggaagga	ctgggcttag	tatgaaaagt	taggactgag	4380
aagaattttga	aagggggctt	tttgtagctt	gatattcact	actgtcttat	taccctatca	4440
tagggcccacc	ccaaatggaa	gtcccattct	tcctcaggat	gtttaagatt	agcattcagg	4500
aagagatcag	aggtctgctg	gtccccttat	catgtccctt	atgggtgctt	tggctctgca	4560
gtatttagca	tagtgttacc	atcaaccacc	ttaacttcat	ttttcttatt	caataacctag	4620
gtaggtagat	gctagattct	ggaaataaaa	ttaggtctct	aagtggctct	tgtcctctct	4680
cccagtcaaa	ttctgaatct	agttggcaag	attctgaaat	caaggcataat	aatcagtaata	4740
aagtgatgat	agaagggtat	atagaagaat	tttattatat	gagaggggtga	aacctaaaat	4800
gaaatgaaat	cagacccttg	tcttacacca	taaacaaaaa	taaatttgaa	tgggttaaag	4860
aattaaacta	agacctaaaa	ccataaaaaa	ttttaaagaa	atcaaaaagaa	gaaaattcta	4920
atattcatgt	tgcagccgtt	ttttgaattt	gatattgagaa	gcaaaggcaa	caaaaggaaa	4980
aataaagaag	tgaggctaca	tcaaactaaa	aaatttccac	acaaaaaaga	aaacaatgaa	5040
caaataaag	gtgaaccatg	aaatggcata	tttgcaaac	aaatatttct	taaatatttt	5100
ggttaatatc	caaaatatat	aagaaacaca	gatgattcaa	taacaaacaa	aaaattaaaa	5160
ataggaaaaat	aaaaaaatta	aaaagaagaa	aatcctgcca	tttatgctgag	aattgatgaa	5220
cctggaggat	gtaaaaactaa	gaaaaataag	cctgacacaa	aaagacaaat	actacacaa	5280
cttgtctata	tgtgaaacat	aaaaaagtca	ctctcatgga	aacagacagt	agagggtatg	5340
tttccagggg	ttgggggtgg	gagaatcagg	aaactattac	tcaaagggtga	taaaatttca	5400
gttatgtggg	atgaataaat	tctagatatc	taatgtacag	catcgtgact	gtagttaatt	5460
gtactgtaag	tatatttaaa	atttgcaaag	agagttagatt	tttttgtttt	tttagatgga	5520
gttttgctct	tgttgctccag	gctggagtgc	aatggcaaga	tcttggtcca	ctgcaacctc	5580
cgctcctggg	gttcaagcaa	atctcctgcc	tcagcctccc	gagtagctgg	gattacaggc	5640
atgcgacacc	atgccagct	aattttgtat	ttttagtaga	gacgggggtt	ctccatgttg	5700
gtcaggctga	tccgcctcct	cggccaccaa	agggtctggga	ttacaggcgt	gaccaccggg	5760
cctggccgag	agtagatctt	aaaagcattt	accacaagaa	aaaggtaact	atgtgagata	5820
atgggtatgt	taatttagctt	gattgtggta	atcatttcac	aaggatataca	tatattaaaa	5880
catcatgttg	tacaccttaa	atatatacaa	tttttatattg	tgaatgatac	ctcaataaag	5940
ttgaagaata	ataaaaaaga	atagacatca	catgaattaa	aaaactaaaa	aataaaaaaa	6000
tgcattctga	tgattagaat	tgcattcttg	atttttcaga	tacaaatatc	catttgactg	6060

-continued

---

SEQ ID NO: 43           moltype = DNA   length = 60  
FEATURE               Location/Qualifiers  
source                1..60  
                      mol\_type = other DNA  
                      note = synthetic polynucleotide  
                      organism = synthetic construct  
modified\_base        1..2  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        57..58  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        2..3  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        3..4  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        58..59  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        59..60  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
SEQUENCE: 43  
aaagaataac agtgataatt tctgggttaa ggcaatagca atatctctgc atataaatat   60

SEQ ID NO: 44           moltype = DNA   length = 51  
FEATURE               Location/Qualifiers  
source                1..51  
                      mol\_type = other DNA  
                      note = synthetic polynucleotide  
                      organism = synthetic construct  
modified\_base        1..2  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        48..49  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        2..3  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        3..4  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        49..50  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        50..51  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
SEQUENCE: 44  
gttcagcgtg tccggcgagg gcgagggtgag tctatgggac ccttgatgtt t           51

SEQ ID NO: 45           moltype = DNA   length = 51  
FEATURE               Location/Qualifiers  
source                1..51  
                      mol\_type = other DNA  
                      note = synthetic polynucleotide  
                      organism = synthetic construct  
modified\_base        1..2  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        48..49  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        2..3  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        3..4  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage  
modified\_base        49..50  
                      mod\_base = OTHER  
                      note = optional phosphorothioate internucleoside linkage

-continued

---

modified\_base 50..51  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

SEQUENCE: 45  
aaacatcaag ggtcccatag actcacctcg ccctcgccgg acacgctgaa c 51

SEQ ID NO: 46 moltype = DNA length = 70  
FEATURE Location/Qualifiers  
source 1..70  
mol\_type = other DNA  
note = synthetic polynucleotide  
organism = synthetic construct

modified\_base 1..2  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 67..68  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 3..4  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 2..3  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 68..69  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 69..70  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

SEQUENCE: 46  
cttgccccac agggcagtaa cggcagattt ttcttcgggc gttaaattgca ccatgggtgtc 60  
tgtttgaggt 70

SEQ ID NO: 47 moltype = DNA length = 51  
FEATURE Location/Qualifiers  
source 1..51  
mol\_type = other DNA  
note = synthetic polynucleotide  
organism = synthetic construct

modified\_base 1..2  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 48..49  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 2..3  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 3..4  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 49..50  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 50..51  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

SEQUENCE: 47  
acagacacca tgggtcacct gactcctgag gagaagtctg ccgttactgc c 51

SEQ ID NO: 48 moltype = DNA length = 65  
FEATURE Location/Qualifiers  
source 1..65  
mol\_type = other DNA  
note = synthetic polynucleotide  
organism = synthetic construct

modified\_base 1..2  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 62..63  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 2..3  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

-continued

---

modified\_base 3..4  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 63..64  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 64..65  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

SEQUENCE: 48  
ttgccccaca gggcagtaac ggcagacttc tcctcaggag tcaggtgcac catggtgtct 60  
gtttg 65

SEQ ID NO: 49 moltype = DNA length = 61  
FEATURE Location/Qualifiers  
source 1..61  
mol\_type = other DNA  
note = synthetic polynucleotide  
organism = synthetic construct

modified\_base 1..2  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 58..59  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 3..4  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 2..3  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 59..60  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 60..61  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

SEQUENCE: 49  
ttctgtatct atattcatca taggaaacac caaagataat gttctcctta atgggtgccag 60  
g 61

SEQ ID NO: 50 moltype = DNA length = 61  
FEATURE Location/Qualifiers  
source 1..61  
mol\_type = other DNA  
note = synthetic polynucleotide  
organism = synthetic construct

modified\_base 1..2  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 58..59  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 2..3  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 3..4  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 59..60  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

modified\_base 60..61  
mod\_base = OTHER  
note = optional phosphorothioate internucleoside linkage

SEQUENCE: 50  
tcttgggatt caataacctt gcagacagtg gaggaaggcc tttggcgtga taccacaggt 60  
g 61

SEQ ID NO: 51 moltype = DNA length = 63  
FEATURE Location/Qualifiers  
source 1..63  
mol\_type = other DNA  
note = synthetic polynucleotide  
organism = synthetic construct

modified\_base 1..2

-continued

---

	mod_base = OTHER
modified_base	note = optional phosphorothioate internucleoside linkage 60..61
	mod_base = OTHER
modified_base	note = optional phosphorothioate internucleoside linkage 61..62
	mod_base = OTHER
modified_base	note = optional phosphorothioate internucleoside linkage 62..63
	mod_base = OTHER
modified_base	note = optional phosphorothioate internucleoside linkage 2..3
	mod_base = OTHER
modified_base	note = optional phosphorothioate internucleoside linkage 3..4
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
SEQUENCE: 51	
tccaagtttg cagagaaaga taatatagtc cttggagaag gaggaatcac cctgagtgga	60
ggt	63
SEQ ID NO: 52	moltype = DNA length = 64
FEATURE	Location/Qualifiers
source	1..64
	mol_type = other DNA
	note = synthetic polynucleotide
	organism = synthetic construct
modified_base	1..2
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	61..62
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	2..3
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	3..4
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	62..63
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	63..64
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
SEQUENCE: 52	
aggacgggtcc cggcctgcga cacttccgcc cataattggt cttcatctgc ggggcggggg	60
gggg	64
SEQ ID NO: 53	moltype = DNA length = 67
FEATURE	Location/Qualifiers
source	1..67
	mol_type = other DNA
	note = synthetic polynucleotide
	organism = synthetic construct
modified_base	64..65
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	1..2
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	3..4
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	2..3
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage
modified_base	65..66
	mod_base = OTHER
	note = optional phosphorothioate internucleoside linkage

-continued

---

```

modified_base      66..67
                   mod_base = OTHER
                   note = optional phosphorothioate internucleoside linkage

SEQUENCE: 53
gggacggcgc ccacataggc caaattcaat tgctgatccc agcttaagac gtactggcca 60
gcctggc                                         67

```

---

**1-48.** (canceled)

**49.** A composition comprising a non-covalent complex of  
(i) a donor oligonucleotide comprising a sequence that can correct at least one mutation in a cell's genome; and,  
(ii) a 3E10 antibody or a cell-penetrating fragment thereof.

**50.** The composition of claim **49**, wherein the 3E10 antibody or cell-penetrating fragment thereof is selected from the group consisting of a cell-penetrating monovalent, divalent, or multivalent single chain variable fragment (scFv) of 3E10, a cell-penetrating diabody of 3E10, and a cell-penetrating humanized form or variant of 3E10.

**51.** The composition of claim **50**, wherein the 3E10 antibody or cell-penetrating fragment thereof incorporates an aspartic acid (Asp) to asparagine (Asn) substitution in a position corresponding to the Asp31 position in the heavy chain of 3E10.

**52.** The composition of claim **49**, wherein the 3E10 antibody or a cell-penetrating fragment thereof comprises:

- (a) a heavy chain variable region (VH) complementarity determining region (CDR) 1 comprising the amino acid sequence of SEQ ID NO: 15;
- (b) a VH CDR2 comprising the amino acid sequence of SEQ ID NO: 17;
- (c) a VH CDR3 comprising the amino acid sequence of SEQ ID NO: 18;
- (d) a light chain variable region (VL) CDR1 comprising the amino acid sequence of SEQ ID NO: 24;
- (e) a VL CDR2 comprising the amino acid sequence of SEQ ID NO: 25; and,
- (f) a VL CDR3 comprising the amino acid sequence of SEQ ID NO: 26;

or

- (a) a VH CDR1 comprising the amino acid sequence of SEQ ID NO: 16;
- (b) a VH CDR2 comprising the amino acid sequence of SEQ ID NO: 17;
- (c) a VH CDR3 comprising the amino acid sequence of SEQ ID NO: 18;
- (d) a VL CDR1 comprising the amino acid sequence of SEQ ID NO: 24;
- (e) a VL CDR2 comprising the amino acid sequence of SEQ ID NO: 25; and,
- (f) a VL CDR3 comprising the amino acid sequence of SEQ ID NO: 26.

**53.** The composition of claim **49**, wherein the 3E10 antibody or a cell-penetrating fragment thereof comprises a VH sequence set forth in SEQ ID NOS: 1, 2, 3, 4, 5, or 6.

**54.** The composition of claim **49**, wherein the 3E10 antibody or a cell-penetrating fragment thereof comprises a VL sequence set forth in SEQ ID NOS: 7, 8, 9, 10, or 11.

**55.** The composition of claim **49**, wherein the 3E10 antibody or a cell-penetrating fragment thereof comprises a

VH sequence set forth in SEQ ID NOS: 1, 2, 3, 4, 5, or 6, and a VL sequence set forth in SEQ ID NOS: 7, 8, 9, 10, or 11.

**56.** The composition of claim **49**, wherein the donor oligonucleotide comprises single stranded or double stranded DNA.

**57.** The composition of claim **49**, wherein the donor oligonucleotide comprises one or more inserted, substituted, or deleted nucleotides relative to a target gene in the cell's genome.

**58.** The composition of claim **49**, wherein the cell's genome has a mutation underlying a disease or disorder selected from the group consisting of hemophilia, muscular dystrophy, globinopathies, cystic fibrosis, xeroderma pigmentosum, lysosomal storage diseases, immune deficiency syndromes, tyrosinemia, Fanconi anemia, spherocytosis, alpha-1-anti-trypsin deficiency, Wilson's disease, Leber's hereditary optic neuropathy, and chronic granulomatous disorder.

**59.** The composition of claim **58**, wherein the immune deficiency syndromes is X-linked severe combined immunodeficiency or ADA deficiency.

**60.** The composition of claim **58**, wherein the mutation is in a gene encoding coagulation factor VIII, a gene encoding coagulation factor IX, a gene encoding dystrophin, a gene encoding beta-globin, a CFTR gene, an XPC gene, an XPD gene, a gene encoding DNA polymerase eta, a FANCA gene, a FANCB gene, a FANCC gene, a FANCD1 gene, a FANCD2 gene, a FANCE gene, a FANCF gene, a FANCG gene, a FANCI gene, a FANCI gene, a FANCI gene, a FANCI gene, a SPTA1 gene or other spectrin genes, an ANK1 gene, a SERPINA1 gene, an ATP7B gene, an IL2RG gene, an ADA gene, an FAH gene, a CYBA gene, a CYBB gene, an NCF1 gene, an NCF2 gene, or an NCF4 gene.

**61.** The composition of claim **49**, wherein the composition does not comprise a nuclease, triplex-forming peptide nucleic acid oligomer, or nanoparticles.

**62.** The composition of claim **49**, wherein the 3E10 antibody or a cell-penetrating fragment thereof can be transported into the nucleus of the cell without the aid of a carrier or conjugate.

**63.** A pharmaceutical composition comprising the composition of claim **49** and a pharmaceutically acceptable excipient.

**64.** The pharmaceutical composition of claim **63**, wherein the composition is packaged in a polymeric nanoparticle.

**65.** The pharmaceutical composition of claim **64**, wherein the polymeric nanoparticle comprises a polyhydroxy acid polymer.

**66.** The pharmaceutical composition of claim **65**, wherein the polyhydroxy acid polymer is as poly(lactic-co-glycolic acid) (PLGA).

**67.** The pharmaceutical composition of claim **64**, wherein the polymeric nanoparticle comprises a targeting moiety, a cell penetrating peptide, or a combination thereof, which is

associated with, linked, conjugated, or otherwise attached directly or indirectly to the polymeric nanoparticle.

**68.** A method of modifying the genome of a cell comprising contacting the cell with an effective amount of a composition comprising a non-covalent complex of

- (i) a donor oligonucleotide comprising a sequence that can correct at least one mutation in a cell's genome; and,
- (ii) a 3E10 antibody or a cell-penetrating fragment thereof,

wherein the 3E10 antibody or a cell-penetrating fragment thereof comprises

- (a) a heavy chain variable region ( $V_H$ ) complementarity determining region (CDR) 1 comprising the amino acid sequence of SEQ ID NO: 15,
- (b) a  $V_H$  CDR2 comprising the amino acid sequence of SEQ ID NO: 17,
- (c) a  $V_H$  CDR3 comprising the amino acid sequence of SEQ ID NO: 18,
- (d) a light chain variable region ( $V_L$ ) CDR1 comprising the amino acid sequence of SEQ ID NO: 24,
- (e) a  $V_L$  CDR2 comprising the amino acid sequence of SEQ ID NO: 25, and
- (f) a  $V_L$  CDR3 comprising the amino acid sequence of SEQ ID NO: 26,

or

- (a) a  $V_H$  CDR1 comprising the amino acid sequence of SEQ ID NO: 16;
- (b) a  $V_H$  CDR2 comprising the amino acid sequence of SEQ ID NO: 17;
- (c) a  $V_H$  CDR3 comprising the amino acid sequence of SEQ ID NO: 18;
- (d) a  $V_L$  CDR1 comprising the amino acid sequence of SEQ ID NO: 24;
- (e) a  $V_L$  CDR2 comprising the amino acid sequence of SEQ ID NO: 25; and,
- (f) a  $V_L$  CDR3 comprising the amino acid sequence of SEQ ID NO: 26.

**69.** The method of claim **68**, wherein the method is conducted in vivo in a subject.

**70.** The method of claim **68**, wherein the method is conducted ex vivo.

**71.** The method of claim **68**, wherein the cell is a hematopoietic stem cell.

**72.** A modified cell obtained according to the method of claim **68**.

**73.** The method of claim **68**, wherein the donor oligonucleotide comprises single stranded or double stranded DNA.

**74.** The method of claim **68**, wherein the donor oligonucleotide comprises one or more inserted, substituted, or deleted nucleotides relative to a target gene in the cell's genome.

75. A method to treat a disease or disorder comprising administering a therapeutically effective amount of a modified cell obtained according to the method of claim 68 to a subject in need thereof.

**76.** The method of claim **75**, wherein the disease or disorder is selected from the group consisting of hemophilia, muscular dystrophy, globinopathies, cystic fibrosis, xeroderma pigmentosum, lysosomal storage diseases, immune deficiency syndromes, tyrosinemia, Fanconi anemia, spherocytosis, alpha-1-anti-trypsin deficiency, Wilson's disease, Leber's hereditary optic neuropathy, and chronic granulomatous disorder.

77. The method of claim 81, wherein the immune deficiency syndromes is X-linked severe combined immunodeficiency or ADA deficiency.

**78.** The method of claim **68**, wherein the at least one mutation is in a gene encoding coagulation factor VIII, a gene encoding coagulation factor IX, a gene encoding dystrophin, a gene encoding beta-globin, a CFTR gene, an XPC gene, an XPD gene, a gene encoding DNA polymerase eta, a FANCA gene, a FANCB gene, a FANCC gene, a FANCD1 gene, a FANCD2 gene, a FANCE gene, a FANCF gene, a FANCG gene, a FANCI gene, a FANCJ gene, a FANCL gene, a SPTA1 gene or other spectrin genes, an ANK1 gene, a SERPINA1 gene, an ATP7B gene, an IL2RG gene, an ADA gene, an FAH gene, a CYBA gene, a CYBB gene, an NCF1 gene, an NCF2 gene, or an NCF4 gene.

\* \* \* \* \*