(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 January 2017 (19.01.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/011519 A1

(51) International Patent Classification:
A01N 63/00 (2006.01) *A61K 48/00* (2006.01)
A61K 38/46 (2006.01)

(21) International Application Number:
PCT/US2016/042024

(22) International Filing Date:
13 July 2016 (13.07.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

62/191,918	13 July 2015 (13.07.2015)	US
62/247,469	28 October 2015 (28.10.2015)	US
62/315,438	30 March 2016 (30.03.2016)	US

(71) Applicant: **SANGAMO BIOSCIENCES, INC.** [US/US]; Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, California 94804 (US).

(72) Inventors: **LEE, Gary K.**; c/o Sangamo BioSciences, Inc., Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, California 94804 (US). **RILEY, Brigit E.**; c/o Sangamo BioSciences, Inc., Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, California 94804 (US). **ST. MARTIN, Susan J.**; c/o Sangamo BioSciences, Inc., Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, California 94804 (US). **WECHSLER, Thomas**; c/o Sangamo BioSciences, Inc., Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, California 94804 (US).

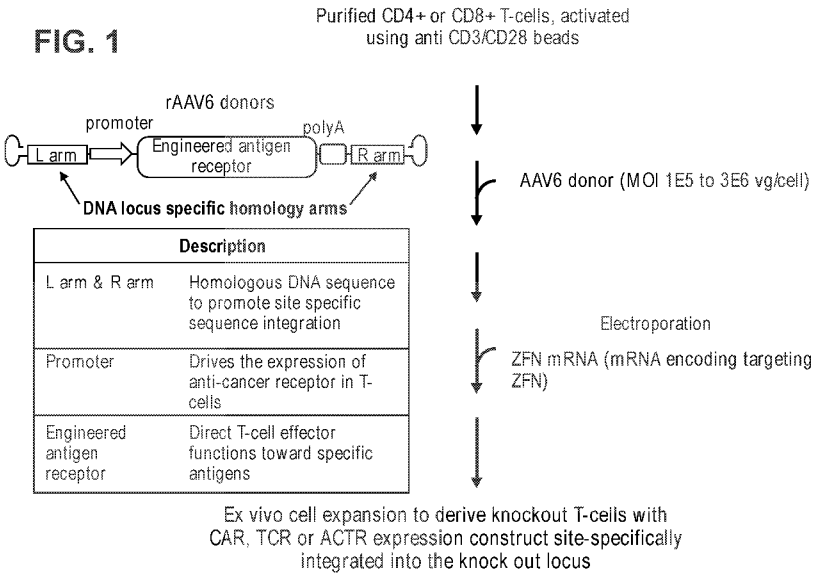
(74) Agent: **PASTERNAK, Dahna S.**; Pasternak Patent Law, 1900 Embarcadero Rd., Suite 211, Palo Alto, California 94303 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,

(54) Title: DELIVERY METHODS AND COMPOSITIONS FOR NUCLEASE-MEDIATED GENOME ENGINEERING

FIG. 1



Description	
L arm & R arm	Homologous DNA sequence to promote site specific sequence integration
Promoter	Drives the expression of anti-cancer receptor in T-cells
Engineered antigen receptor	Direct T-cell effector functions toward specific antigens

(57) Abstract: The present disclosure is in the field of genome engineering, particularly targeted modification of the genome of a cell.

WO 2017/011519 A1

TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

DELIVERY METHODS AND COMPOSITIONS FOR NUCLEASE-MEDIATED GENOME ENGINEERING

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the benefit of U.S. Provisional Application No. 62/191,918, filed July 13, 2015; U.S. Provisional Patent Application No. 62/247,469, filed October 28, 2015; and U.S. Provisional Patent Application No. 62/315,438, filed March 30, 2016, the disclosures of which are hereby incorporated by reference in their entireties.

TECHNICAL FIELD

[0002] The present disclosure is in the field of genome engineering, particularly targeted modification of the genome of a cell.

BACKGROUND

[0003] Various methods and compositions for targeted cleavage of genomic DNA have been described. Such targeted cleavage events can be used, for example, to induce targeted mutagenesis, induce targeted deletions of cellular DNA sequences, and facilitate targeted recombination at a predetermined chromosomal locus. *See, e.g.,* U.S. Patent Nos. ,255,250; 9,200,266; 9,045,763; 9,005,973; 9,150,847; 8,956,828; 8,945,868; 8,703,489; 8,586,526; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,067,317; 7,262,054; 7,888,121; 7,972,854; 7,914,796; 7,951,925; 8,110,379; 8,409,861; U.S. Patent Publications 20030232410; 20050208489; 20050026157; 20050064474; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983; 20130196373; 20140120622; 20150056705; 20150335708; 20160030477 and 20160024474, the disclosures of which are incorporated by reference in their entireties for all purposes.

[0004] These methods often involve the use of engineered cleavage systems to induce a double strand break (DSB) or a nick in a target DNA sequence such that repair of the break by an error born process such as non-homologous end joining (NHEJ) or repair using a repair template (homology directed repair or HDR) can

result in the knock out of a gene or the insertion of a sequence of interest (targeted integration). Cleavage can occur through the use of specific nucleases such as engineered zinc finger nucleases (ZFN), transcription-activator like effector nucleases (TALENs), using the CRISPR/Cas system (including Cas and/or Cfp1) with an engineered crRNA/tracr RNA ('single guide RNA') to guide specific cleavage and/or using nucleases based on the Argonaute system (*e.g.*, from *T. thermophilus*, known as 'TtAgo', (Swarts *et al* (2014) *Nature* 507(7491): 258-261).

[0005] Targeted cleavage using one of the above mentioned nuclease systems can be exploited to insert a nucleic acid into a specific target location using either HDR or NHEJ-mediated processes. However, delivering both the nuclease system and the donor to the cell can be problematic. For example, delivery of a donor or a nuclease via transduction of a plasmid into the cell can be toxic to the recipient cell, especially to a cell which is a primary cell and so not as robust as a cell from a cell line.

[0006] One method often utilized for delivery of nucleic acids to cells involves the use of viral nucleic acid delivery vectors. In particular, the adeno associated virus (AAV) is widely used to deliver nucleic acid because of its efficiency and relative non-toxicity. The AAV genome can be nearly depleted of viral nucleic acid and replaced with nucleic acids encoding donor transgenes or engineered nucleases to facilitate integration of the transgene into a recipient cell's DNA.

[0007] AAV transduction of mammalian cells depends on both primary and secondary co-receptors on the target cells. While the primary receptor is important for initial adhesion of the virus to the target cell (and its tropism), the secondary receptor mediates endocytosis of the AAV virus into the cell. For example, for serotype AAV6 the primary receptor has been identified as alpha 2,3 N-linked sialic acid (Wu *et al*, (2006) *J.Virol.* 80(18):9093), and the secondary receptor as EGFR. Furthermore, the use of additional secondary co-receptors has also been proposed (Weller *et al*, (2010) *Nat Med* 16(6): 662).

[0008] Delivery (transplantation) of cells and/or tissues *in vivo* can often be hampered by antibody-mediated responses. For example, some kidney transplant patients are prone to acute rejection mediated by the development of host antibodies against the transplant tissue. Accordingly, physicians routinely use steroid therapy to suppress the antibody response following transplantation (see for example Ku *et al* (1973) *Br. Med J* 4:702) and also can use rituximab (anti-CD20 antibody) for B cell

suppression (for example Becker *et al* (2004) *Am J Transpl* 4:996). Antibody-mediated responses are also challenges facing the use of AAV delivery due to prevalence of background anti-AAV antibodies in the human population and the de novo development of these antibodies following dosing with a AAV mediated delivery system (see Kotterman *et al* (2015) *Gene Ther* 22(2):116-126).

[0009] In the body, there are complex mechanisms that can regulate either the activation or the suppression of the cellular members of the immune system. For example, dendritic cells (DCs) have been established as central players in the balance between immune activation versus immune tolerance. They are the most potent antigen presenting cells in the immune system and specifically capture and present antigens to naïve T cells. Immature DCs interact with potential antigens through specific receptors such as Toll-like receptors where the antigen is brought into the cell by micropinocytosis. The antigen is then broken up into smaller peptides that are presented to T cells by the major histocompatibility complexes. In addition, mature DCs secrete inflammatory mediators such as IL-1β, IL-12, IL-6 and TNF which further serve to activate the T cells. On the other side, DCs also play a role in tolerizing the body to some antigens in order to maintain central and peripheral tolerance. Tolerogenic DCs (tolDC) have low amounts of co-stimulatory signals on the cell surfaces and have a reduced expression of the inflammatory mediators described above. However, these tolDCs express large amounts of anti-inflammatory cytokines like IL-10 and when these cells interact with naïve T cells, the T cells are driven to become anergic/regulatory T cells (CD8+ Tregs). In fact, it has been shown that this process is enhanced upon repeated stimulation of T cells with these immature/tolerogenic DCs. Several factors have also been identified that work in concert with tolDCs to induce different types of Tregs. For example, naïve T cells co-exposed with tolDCs and HGF, VIP peptide, TSLP or Vitamin D3 leads to the induction of CD4+CD25+Foxp3+ Tregs, co-exposure with TGF-β or IL-10 leads to Tr1 T regs and co-exposure with corticosteroids, rapamycin, retinoic acid can lead to CD4+/CD8+ Tregs (Raker *et al* (2015) *Front Immunol* 6:art 569 and Osorio *et al* (2015) *Front Immunol* 6: art 535).

[0010] CD34+ stem or progenitor cells are a heterogeneous set of cells characterized by their ability to self-renew and/or differentiate into the cells of the lymphoid lineage (*e.g.* T cells, B cells, NK cells) and myeloid lineage (*e.g.* monocytes, erythrocytes, eosinophiles, basophiles, and neutrophils). Their

heterogeneous nature arises from the fact that within the CD34+ stem cell population, there are multiple subgroups which often reflect the multipotency (whether lineage committed) of a specific group. For example, CD34+ cells that are CD38- are more primitive, immature CD34+ progenitor cell, (also referred to as long term hematopoietic progenitors), while those that are CD34+CD38+ (short term hematopoietic progenitors) are lineage committed (*see* Stella *et al* (1995) *Hematologica* 80:367-387). When this population then progresses further down the differentiation pathway, the CD34 marker is lost. CD34+ stem cells have enormous potential in clinical cell therapy. However, in part due to their heterogeneous nature, performing genetic manipulations such as gene knock out, transgene insertion and the like upon the cells can be difficult. Specifically, these cells are poorly transduced by conventional delivery vectors, the most primitive stem cells are sensitive to modification, there is limited HDR following induced DNA DSBs, and there is insufficient HSC maintenance in prolonged standard culture conditions. Additionally, other cells of interest (for non-limiting example only, cardiomyocytes, medium spiny neurons, primary hepatocytes, embryonic stem cells, induced pluripotent stem cells and muscle cells) can be less successfully transduced for genome editing than others.

[0011] Thus, there remains a need for additional compositions and methods for genome engineering to deliver nucleic acids efficiently to CD34+ cells and other cells of interest using viral vectors.

SUMMARY

[0012] The present invention describes compositions and methods for use in gene therapy and genome engineering. Specifically, the methods and compositions described relate to introducing nucleic acids into cells such as primary cells including hematopoietic stem cells/progenitor cells (HSC/PC) and T cells. In addition, the methods and compositions of the invention are useful for delivery of AAV particles (vectors) comprising donor DNAs of interest to such cells.

[0013] In some aspects, the invention comprises delivery of at least one nuclease to a cell (*e.g.,* an HSC/PC or other hematopoietic lineage cells such as T cell, B cell, or NK cell) for the purpose of genome engineering. In some embodiments, the nuclease is delivered as a peptide, while in others it is delivered as a nucleic acid encoding the nuclease. In some embodiments, more than one nuclease is used. In some preferred embodiments, the nucleic acid encoding the nuclease is an mRNA,

and in some instances, the mRNA is protected. The nuclease may comprise a zinc finger nuclease (ZFN), a TALE-nuclease (TALEN) or a CRISPR/Cas (Cas and/or Cfp1) or TtAgo nuclease system or a combination thereof. In a preferred embodiment, the nucleic acid encoding the nuclease(s) is delivered via electroporation.

[0014] In another aspect, described herein is a method of introducing a nucleic acid into an isolated cell (*e.g.,* a hematopoietic stem cell, a T-cell, a B-cell or an NK cell), the method comprising: administering to the cell at least one adeno-associated vector comprising a donor molecule (*e.g.,* a transgene that is expressed in the cell) in the presence of at least one inhibitor of growth factor receptor binding. In certain embodiments, the growth factor inhibitor inhibits binding to an epidermal growth factor receptor (EGFR), a fibroblast growth factor receptor (FGFR), a Met/hepatocyte growth factor receptor (HGFR), a lipoarabinomannan receptor (LamR), a αVβ5 integrin receptor, an Intercellular Adhesion Molecule 1 receptor (Icam-1) and/or a Platelet-derived growth factor receptor. The transgene may be episomal or may be integrated into the genome of the cell. In any of the methods described herein, the transgene may encode a chimeric antigen receptor (CAR). In addition, any of the methods as described herein may further comprise introducing at least one nuclease into the cell, wherein the transgene is integrated into one or more genes of the cell following cleavage of the one or more genes by the nuclease. Additional nucleases may be used for inactivation (knockout) of additional genes, with or without targeted integration. In certain embodiments, the nuclease cleaves a programmed cell death 1 (PD1) gene, a Cytotoxic T-Lymphocyte Antigen 4 (CTLA-4) gene, a beta 2-microglobulin (B2M) and/or a T-cell receptor alpha (TRAC) gene. In certain embodiments, the cell comprises at least one gene with an integrated transgene and which gene is inactivated (KO) and at least one second (different) gene in which is also inactivated (KO). The at least one second gene may be inactivated with or without integration of a transgene.

[0015] Thus, the invention provides methods and compositions for introducing a nucleic acid into a cell, including methods and composition for increasing the efficiency of delivery of a nucleic acid to a cell. In certain embodiments, at least 50% to 60% (or any value therebetween), more preferably at least 60 to 70% (or any value therebetween), even more preferably at least 70% to 80% (or any value therebetween) or even more preferably greater than 80% (any

value between 80-100%) of the cells are modified by introduction of the nucleic acid thereinto (*e.g.*, into the genome of cell). In some embodiments, the delivery encompasses use of a viral vector. In preferred embodiments, the vector is an AAV vector. In some aspects, increased efficiency of viral vector delivery is accomplished through selectively inhibiting the ability of one or more viral receptors on a cell to bind to the viral vector (*e.g.*, AAV), thereby increasing delivery of the nucleic acid carried by the viral vector to the cell through one or more alternate receptors. In some embodiments, binding of the viral vector (*e.g.*, AAV) to an epidermal growth factor receptor (EGFR) is blocked or inhibited, while in other embodiments, binding to a fibroblast growth factor receptor (FGFR) is blocked or inhibited. In other embodiments, receptors that may be blocked or inhibited from binding to the viral vector (*e.g.*, AAV) include, but are not limited to, a Met/hepatocyte growth factor receptor (HGFR), a lipoarabinomannan receptor (LamR), a αVβ5 integrin receptor, a Intercellular Adhesion Molecule 1 receptor (Icam-1) and/or a Platelet-derived growth factor receptor (PDGFR, including PDGFR beta and PDGFR alpha). In preferred embodiments, inhibition of a receptor increases the efficiency of viral delivery by 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold. The inhibitors may be given prior (*e.g.* 4-5 days, 2-3 days, 1 day, 12-24 hours, 6-11 hours, 1-5 hours or less than 1 hour (or any time therebetween)) to treatment of the cell with the AAV. The inhibitors may also be given simultaneously when the viral vector (*e.g.*, AAV) is delivered to the cell and/or after the viral vector (*e.g.*, AAV) is delivered to cell (any time up to one day or even longer). In some embodiments, the inhibitor(s) used is/are gefitinib, BGJ398, SU11274, CP-673451, and/or Crenolanib for inhibition of EGFR, FGFR, HGFR, and/or PDGFR. The PDGFR inhibitors CP-673451 (also referred to as PDGFR1) and Crenolanib (also referred to as PDGFR2), inhibit both PDGFR alpha and beta, albeit with different affinities. In any of the methods described herein, the viral vector may carry a nucleic acid encoding one or more nucleases and/or one or more donors (*e.g.*, sequences encoding therapeutic proteins).

[0016] In one aspect, a viral vector is used to deliver a transgene such that the transgene does not integrate into the genome. In some embodiments, the transgene comprises an inducible promoter. In other embodiments, the transgene comprises a constitutive promoter. In still further embodiments, the viral vector is an AAV or a lentiviral vector. The viral vector may be delivered before, during or after delivery of

one or more molecules that inhibit binding of the viral vector to one or more cell surface receptors.

[0017] In one aspect, provided herein is a method of integrating one or more transgenes into a genome of an isolated cell, the method comprising sequentially or concurrently introducing the transgene and at least one nuclease into the cell such that the nuclease mediates targeted integration of the transgene. The methods comprise administering to the cell at least one adeno-associated vector comprising a donor molecule in the presence of at least one inhibitor of growth factor receptor binding or a B-cell inhibitor and further comprise introducing at least one nuclease into the cell, wherein the transgene (*e.g.,* chimeric antigen receptor) is integrated into the genome following cleavage by the nuclease (*e.g.,* cleavage of PD1, CTLA-4 and/or TRAC by a targeted nuclease). Thus, in certain aspects, a method of integrating one or more transgenes into a genome of an isolated cell, the method comprising: introducing, into the cell, (a) a donor vector comprising the one or more transgenes (*e.g.,* encoding chimeric antigen receptors (CARs) and/or engineered TCR) and (b) at least one nuclease, wherein the at least one nuclease cleaves the genome of the cell such that the one or more transgenes are integrated into the genome of the cell, and further wherein (i) if the donor vector is introduced into the cell before the at least one nuclease, the at least one nuclease is introduced into the cell within 48 hours after donor vector is introduced and; (ii) if the at least one nuclease is introduced before the donor vector, the donor vector is introduced into the cell within 4 hours after the at least one nuclease is introduced. In certain embodiments, the methods can comprise (a) introducing a donor vector comprising the one or more transgenes into the cell; (b) culturing the cell for less than 48 hours (*e.g.,* seconds to 48 hours or any time therebetween); and (c) introducing at least one nuclease into the cell, wherein the at least one nuclease cleaves the genome of the cell such that the one or more transgenes are integrated into the genome of the cell. Alternatively, the methods can comprise: (a) introducing at least one nuclease into the cell; (b) culturing the cell for less than 24 hours (*e.g.,* seconds to 24 hours or any time therebetween); and (c) introducing a donor vector comprising the one or more transgenes into the cell, wherein the at least one nuclease cleaves the genome of the cell such that the one or more transgenes are integrated into the genome of the cell. The method steps may be repeated for integration of additional transgenes into the same and/or different loci. In certain embodiments, the cell is cultured (step (b)) for less than 24 hours (*e.g.,* seconds to 24

hours or any time therebetween). In still further embodiments, the cell is cultured for less than 4 hours, for example, when the nuclease(s) is introduced before introduction of the donor vector. Any of these methods may further comprise the step of administering a molecule that inhibits binding of a viral vector (*e.g.,* carrying the nuclease(s) and/or the donor vector) to a cell receptor prior to, simultaneously and/or after the step of introducing the nucleases and/or donor vector to the cell.

[0018] Any cell can be used, for example, a hematopoietic stem cell (*e.g.,* CD34+ cell) or T-cell (*e.g.,* CD4+ or CD8+ cell (including Treg cells)), or B-cell or NK cell. In some embodiments, the T cell is a tumor infiltrating T-lymphocyte (TIL). The donor vector may be introduced as a viral or non-viral vector, for example an AAV vector (*e.g.,* AAV6 or AAV6 chimeric vector such as AAV2/6, etc.). The nuclease (*e.g.,* ZFN, TALEN, TtAgo and/or CRISPR/Cas) may also be introduced using viral or non-viral vectors, for example in mRNA form. In certain embodiments, the nuclease targets a safe-harbor gene (*e.g.,* a CCR5 gene, an AAVS1 gene, a *Rosa* gene, an albumin gene, etc.). The transgene may encode a protein, for example a therapeutic protein that is lacking or deficient in a subject with a disorder (*e.g.,* lysosomal storage disease, hemoglobinopathy, hemophilia, etc.). In certain embodiments, a method of providing one or more proteins to a subject in need thereof is described, the method comprising: introducing one or more transgenes encoding the one or more proteins into an isolated cell according to any of the methods described herein and introducing the cell into the subject such that the one or more proteins are provided to the subject.

[0019] In other aspects, the invention comprises delivery of a donor nucleic acid to a target cell. The donor may be delivered prior to, after, or along with the nucleic acid encoding the nuclease(s) and/or optional viral receptor inhibitor. In certain embodiments, the donor is delivered simultaneously with the nuclease(s) and/or optional viral receptor inhibitor. In other embodiments, the donor is delivered prior to the nuclease(s), including any time before, for example, immediately before, 1 to 60 minutes before (or any time therebetween), 1 to 24 hours before (or any time therebetween), 1 to 48 hours (or any time therebetween) or more than 48 hours before. In certain embodiments, the donor is delivered after the nuclease, preferably within 4 hours. In certain embodiments, provided herein is a method of introducing a nucleic acid into a cell, the method comprising: administering a donor molecule comprising the nucleic acid into the cell; administering a nuclease to the cell, wherein the

nuclease is administered 1 to 48 hours after or within 4 hours before the donor molecule and further wherein the donor molecule is integrated into the genome of the cell following cleavage by the nuclease. In other embodiments, a method of introducing one or more transgenes into a genome of a cell, the method comprising: introducing at least one nuclease into the cell, wherein the at least one nuclease cleaves the genome of the cell such that the one or more transgenes are integrated into the genome of the cell, and further wherein (i) if the donor vector is introduced into the cell before the at least one nuclease, the at least one nuclease is introduced into the cell within 48 hours after donor vector is introduced and; (ii) if the at least one nuclease is introduced before the donor vector, the donor vector is introduced into the cell within 4 hours after the at least one nuclease is introduced. The donor nucleic acid comprises an exogenous sequence (transgene) to be integrated into the genome of the cell, for example, an endogenous locus. The transgene is preferably integrated at or near (*e.g.*, within 1-50 base pairs) of the site of cleavage by the nuclease(s). In some embodiments, the donor comprises a full length gene or fragment thereof flanked by regions of homology with the targeted cleavage site. In some embodiments, the donor lacks homologous regions and is integrated into a target locus through homology independent mechanism (*i.e.* NHEJ). In other embodiments, the donor comprises an smaller piece of nucleic acid flanked by homologous regions for use in the cell (*i.e.* for gene correction). In some embodiments, the donor comprises a gene encoding a functional or structural component such as a shRNA, RNAi, miRNA or the like. In other embodiments the donor comprises a gene encoding a regulatory element that binds to and/or modulates expression of a gene of interest.

[0020] In other aspects, the donor is delivered by viral and/or non-viral gene transfer methods. In preferred embodiments, the donor is delivered to the cell via an adeno associated virus (AAV). Any AAV vector can be used, including, but not limited to, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 and combinations thereof. In some instances, the AAV comprises LTRs that are of a heterologous serotype in comparison with the capsid serotype (*e.g.*, AAV2 ITRs with AAV5, AAV6, or AAV8 capsids). The donor may be delivered using the same gene transfer system as used to deliver the nuclease (including on the same vector) or may be delivered using a different delivery system that is used for the nuclease. In certain embodiments, the donor is delivered using a viral vector (*e.g.*, AAV) and the

nuclease(s) is(are) delivered in mRNA form. The cell may also be treated with one or more molecules that inhibit binding of the viral vector to a cell surface receptor as described herein prior to, simultaneously and/or after delivery of the viral vector (*e.g.*, carrying the nuclease(s) and/or donor).

[0021] The sequence of interest of the donor molecule may comprise one or more sequences encoding a functional polypeptide (*e.g.*, a cDNA), with or without a promoter. In some instances, the donor comprises a promoter for expression only in a specific cell type (*e.g.*, a T cell or B cell or NK cell specific promoter). In certain embodiments, the nucleic acid sequence comprises a sequence encoding an antibody, an antigen, an enzyme, a growth factor, a receptor (cell surface or nuclear), a hormone, a lymphokine, a cytokine, a reporter, functional fragments of any of the above and combinations of the above. The nucleic acid sequences may also encode one or more proteins that are lacking and/or aberrantly expressed in a subject with a disease or a disorder, including by way of example only a lysosomal storage disease or a hemophilia. The nucleic acid sequences may also encode a one or more proteins useful in cancer therapies, for example one or more chimeric antigen receptors (CARs) and/or an engineered T cell receptor (TCR). In embodiments in which the functional polypeptide encoding sequences are promoterless, expression of the integrated sequence is then ensured by transcription driven by an endogenous promoter or other control element in the region of interest. In other embodiments, a "tandem" cassette is integrated into the selected site in this manner, the first component of the cassette comprising a promoterless sequence as described above, followed by a transcription termination sequence, and a second sequence, encoding an autonomous expression cassette. Additional sequences (coding or non-coding sequences) may be included in the donor molecule between the homology arms, including but not limited to, sequences encoding a 2A peptide, SA site, IRES, *etc.*

[0022] In another aspect, described herein are methods of integrating a donor nucleic acid into the genome of a cell via homology-independent mechanisms. The methods comprise creating a double-stranded break (DSB) in the genome of a cell and cleaving the donor molecule using a nuclease, such that the donor nucleic acid is integrated at the site of the DSB. In certain embodiments, the donor nucleic acid is integrated via non-homology dependent methods (*e.g.*, NHEJ). As noted above, upon *in vivo* cleavage the donor sequences can be integrated in a targeted manner into the genome of a cell at the location of a DSB. The donor sequence can include one or

more of the same target sites for one or more of the nucleases used to create the DSB. Thus, the donor sequence may be cleaved by one or more of the same nucleases used to cleave the endogenous gene into which integration is desired. In certain embodiments, the donor sequence includes different nuclease target sites from the nucleases used to induce the DSB. DSBs in the genome of the target cell may be created by any mechanism. In certain embodiments, the DSB is created by one or more zinc-finger nucleases (ZFNs), fusion proteins comprising a zinc finger binding domain, which is engineered to bind a sequence within the region of interest, and a cleavage domain or a cleavage half-domain. In other embodiments, the DSB is created by one or more TALE DNA-binding domains (naturally occurring or non-naturally occurring) fused to a nuclease domain (TALEN). In yet further embodiments, the DSB is created using a CRISPR/Cas (Cas and/or Cfp1) or TtAgo nuclease system where an engineered single guide RNA or its functional equivalent is used as needed to guide the nuclease to a targeted site in a genome.

[0023] In other aspects, the nuclease(s) binds to and/or cleaves a safe-harbor gene, for example a CCR5 gene, a PPP1R12C (also known as AAVS1) gene, a *Rosa* gene or an albumin gene in mammalian cells. In addition, to aid in selection in mammalian systems, the HPRT locus may be used.

[0024] In some aspects, the nuclease(s) binds to and/or cleaves a check point inhibitor gene, for example PD-1, CTLA4, receptors for the B7 family of inhibitory ligands, or cleaves a receptor or ligand gene involved in signaling through LAG3, 2B4, BTLA, TIM3, A2aR, and killer inhibitor receptors (KIRs and C-type lectin receptors), *see* Pardoll (2012) *Nat Rev Cancer* 12(4):252.

[0025] In other aspects, the nuclease(s) binds to and/or cleaves a gene that encodes a factor involved in rejection, for example, genes encoding subunits of the HLA complex (class I: HLA-A, HLA-B, HLA-C, HLA-E, HLA-F, HLA-G, beta-2 microglobulin (B2M); class II: HLA-DMA, HLA-DOA, HLA-DPA1, HLA-DQA, HLA-DRA, HLA-DMB, HLA-DOB, HLA-DPB1, HLA-DQB, HLA-DRB) or TCR. In some embodiments, the nuclease(s) target a gene encoding a product involved in the peptide loading process and antigen processing for the HLA complexes (*e.g.* TAP, tapasin, calreticulin, calnexin, LMP2, LMP7 or Erp57).

[0026] In one aspect, the donor is a regulatory protein of interest (*e.g.* ZFP TFs, TALE TFs or a CRISPR/Cas TF) that binds to and/or modulates expression of a gene of interest. In one embodiment, the regulatory proteins bind to a DNA sequence

and prevent binding of other regulatory factors. In another embodiment, the binding of a the regulatory protein may modulate (*i.e.* induce or repress) expression of a target DNA.

[0027] In other aspects, the donor encodes a protein capable of redirecting a T cell. In some embodiments, the protein is an engineered antigen receptor. In further embodiments, the engineered receptor is a chimeric antigen receptor (CAR) or a T cell receptor (TCR), where the TCR in some embodiments is an affinity enhanced engineered TCR or a naturally occurring TCR. In other aspects, the engineered protein is an antibody coupled T-cell receptor (ACTR).

[0028] In other aspects, provided herein is a cell which has been genetically modified (*e.g.,* transgenic) as described herein, for example using a nuclease to introduce the genetic modification. One or more genes of the cell may be modified, for example via targeted integration into one or more genes and/or partial or full inactivation (KO) of one or more genes, including genetic modifications involving targeted integration (with or without inactivation of the target gene) of one or more genes and KO (with or without targeted integration of one or more genes (*e.g.,* targeted integration and inactivation (KO) of one gene and KO of a second, different gene). In certain embodiments, the cell is made by the methods described herein. In certain embodiments, the cell comprises a transgene that is integrated into a safe-harbor locus, such as *CCR5*, *AAVS1*, *ALB*, *Rosa26* and/or *HPRT*. The genetic modification to the cell (*e.g.,* integration of nucleic acid) may be for example within, at or near a site comprising 12 or more (*e.g.,* 12-35 or any value therebetween) contiguous nucleotides of the target site in the gene to which the DNA-binding molecule of a nuclease binds. The cells comprising the integrated transgene may express the transgene from an endogenous promoter or, alternatively, the transgene may include regulatory and control elements such as exogenous promoters that drive expression of the transgene (*e.g.,* when integrated into a safe harbor locus). In certain embodiments, the cells comprising the transgene do not include any viral vector sequences integrated into the genome. The cells may be any eukaryotic cell, for example CD34+ stem cells (*e.g.,* patient-derived stem cells mobilized in patients from the bone marrow into the peripheral blood via granulocyte colony-stimulating factor (GCSF) or other mobilizing agent administration or harvested directly from the bone marrow or umbilical cords). The cells can be harvested, purified, cultured, and the nucleases and/or donor introduced into the cell by any suitable method.

[0029] Compositions such as pharmaceutical compositions comprising the genetically modified cells as described herein are also provided. In some embodiments, the compositions comprise CD34+ HSC/PC or HSC/PC cell population. In other embodiments, the compositions comprise T cells (*e.g.* CD4+ and/or CD8+ T cells or TILs). In still further embodiments, the T cell compositions comprise only CD4+ or only CD8+ cells.

[0030] In another aspect, provided are methods of using the genetically modified cells as described herein. In certain embodiments, genetically modified blood cell precursors ("HSC/PC") are given in a bone marrow transplant and the HSC/PC differentiate and mature *in vivo*. In some embodiments, the HSC/PC are isolated following G-CSF-induced mobilization, and in others, the cells are isolated from human bone marrow or umbilical cords. In some aspects, the HSC/PC are edited by treatment with a nuclease designed to knock out a specific gene or regulatory sequence. In other aspects, the HSC/PC are modified with an engineered nuclease and a donor nucleic acid such that a wild type gene or other gene of interest is inserted and expressed and/or an endogenous aberrant gene is corrected. In some embodiments, the modified HSCs/PC are administered to the patient following mild myeloablative pre-conditioning. In other aspects, the HSC/PC are administered after full myeloablation such that following engraftment, 100% of the hematopoietic cells are derived from the modified HSC/PC.

[0031] The methods and compositions of the invention may also include additional treatment of the subject (*e.g.,* animal) to increase *in vivo* delivery efficiency of viral vectors (*e.g.,* AAV) to cells in target tissues. In some embodiments, treatments are provided before, during and/or after delivery of AAV. In some embodiments, treatments include the provision of steroids to the subject to inhibit the humoral antibody response. Non-limiting examples of suitable steroids include methylprednisolone (*e.g.* Medrol®, Solu Medrol®) and prednisolone. In other embodiments, treatments include use of inhibitors of the humoral response including B-cell inhibitors such as rituximab (e.g. Rituxan®). In still further embodiments, treatment methods combine regiments to increase delivery efficiency such as treating the animal with at least steroids and B-cell inhibitors. These treatment regiments can be used before, during or after treatment of the animal with AAV.

[0032] In some embodiments, the method and compositions of the invention as described herein can be used to induce tolerance in a mammal to a therapeutic

protein such that the levels of the therapeutic protein encoded by the transgene remain at therapeutically relevant levels following a transient rise in anti-therapeutic protein antibodies. Thus, provided herein is a method of inducing tolerance to a therapeutic protein in a subject, the method comprising genetically modifying a cell in a subject using the method as described herein (*e.g.*, so that the cell produces the therapeutic protein), optionally with treatment of the subject with additional compositions (*e.g.*, steroids and/or B-cell inhibitors) such that the animal becomes tolerized to the therapeutic protein. In some embodiments, insertion of the therapeutic protein into the recipient cells is done at the same time as treatment with an immune-inhibitory steroid or B-cell inhibitor, whereas in other instances, no immunomodulatory is used. In some instances, the immunomodulatory agent is administered only if anti-therapeutic protein antibodies are generated. In further instances, the immunomodulatory agent is discontinued after a period of time. Thus, a method of introducing a nucleic acid into a cell of a subject (*e.g,*. a subject with a disorder such as a hemophilia) is provided, the method comprising: administering to the subject at least one adeno-associated vector (AAV) comprising a donor molecule (*e.g.*, transgene encoding a therapeutic protein such as a clotting factor, optionally Factor VIII and/or Factor IX) and at least one steroid and/or at least one B-cell inhibitor. In certain embodiments, the subject is a mammal and the transgene encodes a therapeutic protein, and the mammal becomes tolerized to the therapeutic protein.

[0033] In any of the methods and compositions described herein, the therapeutic transgene used encodes a clotting factor. In preferred embodiments, the transgene encodes a FVIII protein or a F.IX protein. In especially preferred embodiment, the transgene encoding the FVIII protein encodes a FVIII that is lacking the B-domain (B-Domain Delete or BDD).

[0034] In other embodiments, provided are methods of using genetically modified T cells as described herein. In some instances, autologous T cells (derived from the patient) are used while in other embodiments, allogenic (derived from a donor) are used. In some instances, the T cells are isolated from a donor or patient by apheresis and then are treated *ex vivo* to achieve the desired genetic modifications. The modified T cells are then expanded to achieve greater numbers of cells for infusion into the patient. In some embodiments, the isolated T cells are expanded first, and then are modified to achieve the desired genetic modifications. In some aspects, the T cells are edited by treatment with a nuclease designed to knock out a

specific gene or regulatory sequence. In other aspects, the T cells are modified with an engineered nuclease and a donor nucleic acid such that a wild type gene or other gene of interest is inserted and expressed and/or an endogenous aberrant gene is corrected. In some embodiments, TILs are isolated from excised tumor tissues by known methods (for example only, *See, e.g.,* Ellebaek *et al* (2012) *J. Transl Med* 10:169), and in further embodiments, patients may be subjected to lymphodepletive therapy following excision of the tumor tissue but prior to infusion of the modified TILs. In still further embodiments, the T cell is a regulatory T cell (Treg). Furthermore, the cell may be arrested in the G2 phase of the cell cycle.

[0035] In some aspects, the present invention includes methods and compositions for treating or preventing a specific disease in a mammal. In some embodiments, the methods and compositions are used to treat a cancer, for example follicular lymphoma, neuroblastoma, non-Hodgkin lymphoma, lymphoma, glioblastoma, chronic lymphocytic leukemia or CLL and acute lymphocytic leukemia or ALL, ovarian cancer, prostate, colorectal, renal cell and carcinoma (*see, e.g.,* Kershaw *et al* (2014) *Clin Transl Immunol* 3, e16, doi:10.1038/cti.2014.7). In other embodiments, the methods and compositions are used to treat an infectious disease, for example HIV, HCV, HBV, (see Sautto *et al* (2015) *Gut* 0:1-12), Ebola, CMV, EBV and adenovirus. In still further aspects, the methods and compositions of the invention include treatment of autoimmune disease, for example rheumatoid arthritis, multiple sclerosis, inflammatory bowel disease, psoriasis, lupus and scleroderma.

[0036] In another aspect, a method of treating a cancer in subject is provided, the method comprising introducing a nucleic acid into an isolated cell by any of the methods described , wherein donor molecule comprises a sequence encoding a CAR such that the cell expresses the CAR and administering the cell expressing the CAR to the subject. In certain embodiments, the CAR-encoding sequence is integrated into a PD1, CTLA-4 and/or TRAC gene following cleavage of the gene by a nuclease. In certain embodiments, the CAR-encoding sequence is integrated into a first gene (*e.g.,* a PD1, CTLA-4 and/or TRAC gene) following cleavage of the gene by a nuclease following cleavage and a second (different) gene (*e.g.,* a PD1, CTLA-4 and/or TRAC gene) is inactivated by a second nuclease.

[0037] In some embodiments, the transgenic HSC/PC cell or T cell and/or animal includes a transgene that encodes a human gene. In some instances, the transgenic animal comprises a knock out at the endogenous locus corresponding to

exogenous transgene, thereby allowing the development of an *in vivo* system where the human protein may be studied in isolation. Such transgenic models may be used for screening purposes to identify small molecules or large biomolecules or other entities which may interact with or modify the human protein of interest. In some aspects, the transgene is integrated into the selected locus (*e.g.,* safe-harbor) into a stem cell (*e.g.,* an embryonic stem cell, an induced pluripotent stem cell, a hematopoietic stem or precursor cell, etc.), T cell (*e.g.* CD4+, CD8+ (including Treg) or TIL) or animal embryo obtained by any of the methods described herein, and then the embryo is implanted such that a live animal is born. The animal is then raised to sexual maturity and allowed to produce offspring wherein at least some of the offspring comprise edited endogenous gene sequence or the integrated transgene.

[0038] A kit, comprising the AAVs and nucleic acids of the invention, is also provided. The kit may comprise nucleic acids encoding the nucleases, (*e.g.* RNA molecules or ZFN, TALEN, TtAgo or CRISPR/Cas system encoding genes contained in a suitable expression vector), or aliquots of the nuclease proteins, donor molecules, suitable stemness modifiers, instructions for performing the methods of the invention, and the like. The kit may also comprise donor molecules of interest such as selection or screening markers.

[0039] These and other aspects will be readily apparent to the skilled artisan in light of disclosure as a whole.

BRIEF DESCRIPTION OF THE DRAWINGS

[0040] **Figure 1** depicts the design of a T-cell transgene donor vector. The donor comprises right (R) and left (L) homology arms which are homologous to the genomic sequence flanking the cleavage site. The donor also comprises a promoter sequence linked to a transgene of interest. The transgene encodes an engineered antigen receptor such as a CAR, TCR or ACTR. A poly A signaling sequence is also in the donor to achieve efficient expression of the transgene. Depicted also is the method used to transduce the T cells.

[0041] **Figure 2** is a graph depicting the effect that serum present in the culture media has on the targeted integration of a donor into a CD3+ T cell. Integration of the donor is measured by the percent of integrated RFLP detected at increasing doses of AAV2/6 comprising the RFLP donor. CCR5-specific ZFNs were delivered via mRNA electroporation.

[0042] **Figure 3** is a graph depicting nuclease activity in AAV 2/6 transduced cells. Hep3B cells were transduced with an MOI (multiplicity of infection) of 3e4 vg/cell of AAV2/6 comprising a pair of human albumin specific ZFNs in the presence of the indicated inhibitors (that inhibit binding of AAV to a viral receptor). Three days following transduction, genomic DNA was isolated from the cells and analyzed for nuclease activity in the form of small insertions and deletions around the ZFN cleavage site (% indels). Inhibitors were used in the concentrations shown and the data shows that inhibitors of PDGFR increase the amount of activity observed.

[0043] **Figure 4** is a graph depicting nuclease activity of nucleases delivered via mRNA transduction by Lipofectamine RNAiMAX. The data demonstrates that unlike the results shown in Figure 3, where AAV2/6 delivery was used increasing the concentration of PDGFR inhibitors has no effect on the number of indels detected. Therefore, the increase in indel frequency observed in the presence of the PDGFR inhibitors is not due to the inhibitors acting directly on the DNA break repair pathways.

[0044] **Figure 5** is a graph depicting nuclease activity following treatment of Hep3B cells with AAV comprising albumin-specific ZFNs and a combination of PDGFR and EGFR inhibitors where each condition was done in duplicate. The results demonstrate that there is no increase in indel formation in the presence of both inhibitor types.

[0045] **Figure 6** is a graph depicting nuclease activity when the inhibitors are used in HepG2 cells. The concentration of the different inhibitors is shown below the data and demonstrates that at this dose of ZFN-comprising AAV2/6 (MOI of 3e4 vg/cell), PDGFR inhibitors lead to increased detectable nuclease activity.

[0046] **Figure 7** is a graph depicting nuclease activity in Hep3B cells in the presence or absence of serum and the indicated viral receptor inhibitors. Each data set shows the results either plus or minus serum, where all the inhibitors were examined with the use of the ZFN-comprising AAV2/6. The bars on the left of each set correspond to the data measured in the absence of serum while the bars on the right of each set are the results for those experiments done with serum. The data demonstrate that in Hep3B the overall AAV2/6 transduction was down in the presence of serum by about 40-fold. However, in the presence of the PDGFR inhibitors, AAV transduction was robust with 20% indels detectable.

[0047] **Figure 8** is a graph showing the fold increase in indel detection in the presence of the PDGFR inhibitors in the presence or absence of serum. The data is plotted as the fold change in indel formation detected as compared with ZFN alone for the two conditions. This data demonstrates that indels detected increased 70 fold in the presence of the PDGFR inhibitors when the experiment was done in serum as compared to ZFNs alone in serum.

[0048] **Figure 9** is a graph depicting the effect of the PDGFR inhibitor 2 (Crenolanib) on AAV uptake in primary hepatocytes. The data demonstrates that 9 μM of the PDGFR inhibitor caused a large increase in indels detected as compared to no inhibitor.

[0049] **Figures 10A and 10B** show a dosing schematic for non-human primate studies. Figure 10A reflects Rituxan® and Solu-Medrol® dosing post- test article administration. Figure 10B reflects Rituxan® and Solu-Medrol® dosing pre-test article administration.

[0050] **Figures 11A through 11D** are graphs depicting peak human F.IX levels by total AAV-F.IX dose following treatment of NHP with AAV donors carrying Factor IX (F.IX). Figures 11A and 11B reflect dose curves for NHP study with Rituxan®/Solu-Medrol® post-test article administration. Figure 11A depicts dose curves for Group 3 (*see* Examples; 1.5e15 vg/kg; 1:1:8; high dose ZFNs+hF9 donor); Group 4 (9e13 vg/kg; 1:1:4; high dose ZFNs+hF9 donor); Group 5 (6e13 vg/kg; 1:1:2; high dose ZFNs+hF9 donor); Group 6 (5e13 vg/kg; 1:1:8; mid dose ZFNs+hF9 donor); Group 7 (mid dose ZFNs+hF9 donor; 1:1:2; 2e13 vg/kg) and Group 8 (low dose ZFNs+hF9 donor; 1:1:8; 1.5e13 vg/kg). Figure 11B summarizes the peak circulating hF.IX levels by ZFN+hF9 donor dose levels for the 1:1:8 ratio of ZFN:ZFN:hF9 donor. Figure 11C reflects dose curves for NHP study with Rituxan®/Solu-Medrol® pre- test article administration. Figure 11D depicts compiled data for both post- (grey symbols) and pre- (black symbols) Rituxan®/Solu-Medrol® dosing.

[0051] **Figures 12A through 12C** are graphs depicting levels of gene modification (% Indels) following treatment in NHP for the 1:1:8 dose ratio groups. Figure 12A shows % Indels at day 28 for the NHP study with Rituxan®/Solu-Medrol® post- test article administration. Group 3 (1.5e15 vg/kg; high dose ZFNs+hF9), Group 6 (5e13 vg/kg; 1:1:8; mid dose ZFNs+hF9 donor); and Group 8 (low dose ZFNs+hF9 donor; 1:1:8; 1.5e13 vg/kg). Figure 12B reflects % Indels at day

61 for the NHP study with Rituxan®/Solu-Medrol® pre- test article administration. Figure 12C shows compiled data for both post- (grey symbols) and pre- (black symbols) Rituxan®/Solu-Medrol® dosing.

[0052] **Figures 13A through 13D** are graphs depict a summary of human FVIII plasma levels for the NHP study using AAV donors carrying Factor VIII (F8) B-Domain Deleted (FVIII-BDD) proteins. Figure 13A shows results for Group 2 animals (AAV2/6, 2E+12 vg/kg); Figure 13B shows results for Group 3 animals (AAV2/6, 6E+12 vg/kg); Figure 13C shows results for Group 4 animals (AAV2/8, 6E+12 vg/kg); and Figure 13D shows results for Group 5 animals (AAV2/8, 6E+12 vg/kg). 1U/mL of human factor VIII is considered physiological normal, and thus equals 100 % of normal physiological circulating human factor VIII.

[0053] **Figure 14** shows the dosing scheme with for the human FVIII-BDD non-human primate (NHP) studies including removal of all immunosuppression at Day 103. Overview of Rituxan and Solu-Medrol dosing. Rituxan (10 mg/kg; IV) dosing was pre-test article administration while methylprednisolone (Solu-Medrol) (10 mg/kg; IM) dosing was daily up until Day 103.

[0054] **Figure 15** is a graph depicting the peak human FVIII antigen levels over the study following treatment in non-human primates (NHP). At dose levels of 2E+12 vg/kg (n =3), peak values of 111.0%, 493.9% and 840.0% (overall mean 481.6% as measured by hFVIII ELISA) of normal hFVIII plasma levels in humans were achieved. At a higher dose representing 6E+12 vg/kg (n =3), peak values of 450.0%, 625.6% and 886.7% [overall mean 654.1%] of hFVIII plasma levels were achieved.

[0055] **Figures 16A through 16C** are graphs depicting the results from individual cynomolgus monkeys (n=3) dosed with the low dose (2E+12 vg/kg, Group 2) of AAV2/6-FVIII-BDD cDNA over a time period of 168 days post dosing. In all three graphs, concentrations of FVIII-BDD in the plasma, as measured through ELISA, are shown in black. Additionally, concentrations of neutralizing anti-FVIII antibody (shown as Bethesda Units) in plasma are shown in grey. The dotted horizontal line represents the Bethesda Unit cutoff point, below which values would not be considered positive for anti-FVIII neutralizing antibodies. The Solu-Medrol was stopped at day 103- indicated by the vertical dashed line. Each graph shows the results for a single monkey (animals 2101, 2102 and 2103).

[0056] **Figures 17A through 17C** are graphs depicting the results from individual cynomolgus monkeys (n=3) dosed with the high dose (6E+12 vg/kg, Group 3) of AAV2/6-FVIII-BDD cDNA over a time period of 180 days post dosing. In all three graphs, concentrations of FVIII-BDD in the plasma, as measured through ELISA, are shown in black. Additionally, concentrations of neutralizing anti-FVIII antibody (shown as Bethesda Units) in plasma are shown in grey. The dotted horizontal line represents the Bethesda Unit cutoff point, below which values would not be considered positive for anti-FVIII neutralizing antibodies. The Solu-Medrol was stopped at day 103- indicated by the vertical dashed line. Each graph shows the results for a single monkey (animals 3101, 3102 and 3103).

[0057] **Figures 18A through 18D** are graphs depicting the results from individual cynomolgus monkeys (n=3) dosed with the high dose (6E+12 vg/kg, Group 4) of AAV2/8-FVIII-BDD cDNA over a time period of 168 days post dosing. In graphs 18A- 18C, concentrations of FVIII-BDD in the plasma, as measured through ELISA, are shown in black. Additionally, concentrations of neutralizing anti-FVIII antibody (shown as Bethesda Units) in plasma are shown in grey. Figure 18D is a 'blow up' of the lower values in the graph for animal 4103 (note that the y axis in 18A-C goes from 0-5 U/mL of FVIII antigen while 18D goes from 0-1 U/mL of FVIII antigen. The dotted horizontal line represents the Bethesda Unit cutoff point, below which values would not be considered positive for anti-FVIII neutralizing antibodies. The Solu-Medrol was stopped at day 103- indicated by the vertical dashed line. Each graph shows the results for a single monkey (animals 4101, 4102 and 4103).

[0058] **Figures 19A through 19E** are graphs depicting the results from individual cynomolgus monkeys (n=3) dosed with the high dose (6E+12 vg/kg, Group 5) of AAV2/8-FVIII-BDD cDNA over a time period of 168 days post dosing. In graphs 19A- 19C, concentrations of FVIII-BDD in the plasma, as measured through ELISA, are shown in black. Additionally, concentrations of neutralizing anti-FVIII antibody (shown as Bethesda Units) in plasma are shown in grey. Figures 19D and 19E are 'blow ups' of the lower values in the graph (note that the y axis in 19A-C goes from 0-5 U/mL of FVIII antigen while the axis for 19D and 19E goes from 0-1 U/mL of FVIII antigen. The dotted horizontal line represents the Bethesda Unit cutoff point, below which values would not be considered positive for anti-FVIII neutralizing antibodies. The Solu-Medrol was stopped at day 103- indicated by the

vertical dashed line. Each graph shows the results for a single monkey (animals 5101, 5102 and 5103).

[0059] **Figures 20A through 20D** show FACs analysis of TRAC and B2M single and double knockouts (KOs) and targeted integration (TI). Figure 20A shows that 88% of TRAC-nuclease/AAV GFP donor-treated cells had inactivated TRAC genes (left panel) and that 71% of cells had the AAV-delivered GFP donor integrated into the TRAC gene (right panel). Figure 20B shows that 93% of B2M-nuclease/AAV GFP donor-treated cells had inactivated B2M genes (left panel) and that 72% of cells had the AAV-delivered GFP donor integrated into the B2M gene (right panel). Figures 20C and 20D show analysis of cells in which both TRAC and B2M were targeted in the same cell (double KO/modified cells).

DETAILED DESCRIPTION

[0060] Disclosed herein are compositions and methods for transduction of a cell for use in gene therapy or genome engineering. In particular, nuclease-mediated (*i.e.* ZFN, TALEN, TtAgo or CRISPR/Cas (Cas and/or Cfp1) system) targeted integration of an exogenous sequence or genome alteration by targeted cleavage followed by non-homologous end joining, is efficiently achieved in a cell. Particularly useful for transduction and engineering of HSC/PC and primary T cells, however, the methods and compositions can also be used for other cell types.

[0061] Delivery of ZFNs and donor template DNA was optimized as detailed herein using viral vectors and/or molecules that inhibit binding of the viral vector to cell surface receptors. The methods and compositions described herein can be used in anycell type, including any hematopoietic stem cell or precursor cell, such as CD34+ cells. CD34+ cells can include primitive (CD133+CD90+, or CD90-), early (CD34+, CD133+) and committed (CD34+CD133-) CD34+ subsets as well as T cells. T cells can comprise CD4+ or CD8+ cells or TILs. The methods and compositions contained in the instant application can also relate to use *in vivo* for delivery of nucleic acids to primary cells via AAV. The methods described herein result in long-term multilineage engraftment in animals treated with the modified cells.

General

[0062] Practice of the methods, as well as preparation and use of the compositions disclosed herein employ, unless otherwise indicated, conventional

techniques in molecular biology, biochemistry, chromatin structure and analysis, computational chemistry, cell culture, recombinant DNA and related fields as are within the skill of the art. These techniques are fully explained in the literature. *See*, for example, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, Second edition, Cold Spring Harbor Laboratory Press, 1989 and Third edition, 2001; Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 and periodic updates; the series METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, Third edition, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wassarman and A. P. Wolffe, eds.), Academic Press, San Diego, 1999; and METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

Definitions

[0063] The terms "nucleic acid," "polynucleotide," and "oligonucleotide" are used interchangeably and refer to a deoxyribonucleotide or ribonucleotide polymer, in linear or circular conformation, and in either single- or double-stranded form. For the purposes of the present disclosure, these terms are not to be construed as limiting with respect to the length of a polymer. The terms can encompass known analogues of natural nucleotides, as well as nucleotides that are modified in the base, sugar and/or phosphate moieties (*e.g.*, phosphorothioate backbones). In general, an analogue of a particular nucleotide has the same base-pairing specificity; *i.e.*, an analogue of A will base-pair with T.

[0064] The terms "polypeptide," "peptide" and "protein" are used interchangeably to refer to a polymer of amino acid residues. The term also applies to amino acid polymers in which one or more amino acids are chemical analogues or modified derivatives of a corresponding naturally-occurring amino acids.

[0065] "Binding" refers to a sequence-specific, non-covalent interaction between macromolecules (*e.g.*, between a protein and a nucleic acid). Not all components of a binding interaction need be sequence-specific (*e.g.*, contacts with phosphate residues in a DNA backbone), as long as the interaction as a whole is sequence-specific. Such interactions are generally characterized by a dissociation

constant (K_d) of 10^{-6} M^{-1} or lower. "Affinity" refers to the strength of binding: increased binding affinity being correlated with a lower K_d.

[0066] A "binding protein" is a protein that is able to bind to another molecule. A binding protein can bind to, for example, a DNA molecule (a DNA-binding protein), an RNA molecule (an RNA-binding protein) and/or a protein molecule (a protein-binding protein). In the case of a protein-binding protein, it can bind to itself (to form homodimers, homotrimers, *etc.*) and/or it can bind to one or more molecules of a different protein or proteins. A binding protein can have more than one type of binding activity. For example, zinc finger proteins have DNA-binding, RNA-binding and protein-binding activity.

[0067] A "zinc finger DNA binding protein" (or binding domain) is a protein, or a domain within a larger protein, that binds DNA in a sequence-specific manner through one or more zinc fingers, which are regions of amino acid sequence within the binding domain whose structure is stabilized through coordination of a zinc ion. The term zinc finger DNA binding protein is often abbreviated as zinc finger protein or ZFP.

[0068] A "TALE DNA binding domain" or "TALE" is a polypeptide comprising one or more TALE repeat domains/units. The repeat domains are involved in binding of the TALE to its cognate target DNA sequence. A single "repeat unit" (also referred to as a "repeat") is typically 33-35 amino acids in length and exhibits at least some sequence homology with other TALE repeat sequences within a naturally occurring TALE protein.

[0069] Zinc finger and TALE binding domains can be "engineered" to bind to a predetermined nucleotide sequence, for example via engineering (altering one or more amino acids) of the recognition helix region of a naturally occurring zinc finger or TALE protein. Therefore, engineered DNA binding proteins (zinc fingers or TALEs) are proteins that are non-naturally occurring. Non-limiting examples of methods for engineering DNA-binding proteins are design and selection. A designed DNA binding protein is a protein not occurring in nature whose design/composition results principally from rational criteria. Rational criteria for design include application of substitution rules and computerized algorithms for processing information in a database storing information of existing ZFP and/or TALE designs and binding data. See, for example, U.S. Patents 8,586,526; 6,140,081; 6,453,242; and 6,534,261; see also WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 and WO 03/016496.

[0070] A "selected" zinc finger protein or TALE is a protein not found in nature whose production results primarily from an empirical process such as phage display, interaction trap or hybrid selection. See *e.g.*, U.S. Patent Nos. 8,586,526; 5,789,538; US 5,925,523; US 6,007,988; US 6,013,453; US 6,200,759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197, WO 02/099084.

[0071] "TtAgo" is a prokaryotic Argonaute protein thought to be involved in gene silencing. TtAgo is derived from the bacteria *Thermus thermophilus*. (See, *e.g.*, Swarts *et al, ibid*, G. Sheng *et al.*, (2013) *Proc. Natl. Acad. Sci. U.S.A.* 111, 652). A "TtAgo system" is all the components required including, for example, guide DNAs for cleavage by a TtAgo enzyme.

[0072] "Recombination" refers to a process of exchange of genetic information between two polynucleotides, including but not limited to, donor capture by non-homologous end joining (NHEJ) and homologous recombination. For the purposes of this disclosure, "homologous recombination (HR)" refers to the specialized form of such exchange that takes place, for example, during repair of double-strand breaks in cells via homology-directed repair mechanisms. This process requires nucleotide sequence homology, uses a "donor" molecule to template repair of a "target" molecule (*i.e.*, the one that experienced the double-strand break), and is variously known as "non-crossover gene conversion" or "short tract gene conversion," because it leads to the transfer of genetic information from the donor to the target. Without wishing to be bound by any particular theory, such transfer can involve mismatch correction of heteroduplex DNA that forms between the broken target and the donor, and/or "synthesis-dependent strand annealing," in which the donor is used to resynthesize genetic information that will become part of the target, and/or related processes. Such specialized HR often results in an alteration of the sequence of the target molecule such that part or all of the sequence of the donor polynucleotide is incorporated into the target polynucleotide.

[0073] In the methods of the disclosure, one or more targeted nucleases as described herein create a double-stranded break in the target sequence (*e.g.*, cellular chromatin) at a predetermined site, and a "donor" polynucleotide, having homology to the nucleotide sequence in the region of the break, can be introduced into the cell. The presence of the double-stranded break has been shown to facilitate integration of the donor sequence. The donor sequence may be physically integrated or,

alternatively, the donor polynucleotide is used as a template for repair of the break via homologous recombination, resulting in the introduction of all or part of the nucleotide sequence as in the donor into the cellular chromatin. Thus, a first sequence in cellular chromatin can be altered and, in certain embodiments, can be converted into a sequence present in a donor polynucleotide. Thus, the use of the terms "replace" or "replacement" can be understood to represent replacement of one nucleotide sequence by another, (*i.e.*, replacement of a sequence in the informational sense), and does not necessarily require physical or chemical replacement of one polynucleotide by another.

[0074] In any of the methods described herein, additional pairs of zinc-finger proteins or TALEN can be used for additional double-stranded cleavage of additional target sites within the cell.

[0075] Any of the methods described herein can be used for insertion of a donor of any size and/or partial or complete inactivation of one or more target sequences in a cell by targeted integration of donor sequence that disrupts expression of the gene(s) of interest. Cell lines with partially or completely inactivated genes are also provided.

[0076] Furthermore, the methods of targeted integration as described herein can also be used to integrate one or more exogenous sequences. The exogenous nucleic acid sequence can comprise, for example, one or more genes or cDNA molecules, or any type of coding or noncoding sequence, as well as one or more control elements (*e.g.,* promoters). In addition, the exogenous nucleic acid sequence may produce one or more RNA molecules (*e.g.*, small hairpin RNAs (shRNAs), inhibitory RNAs (RNAis), microRNAs (miRNAs), *etc.*).

[0077] In certain embodiments of methods for targeted recombination and/or replacement and/or alteration of a sequence in a region of interest in cellular chromatin, a chromosomal sequence is altered by homologous recombination with an exogenous "donor" nucleotide sequence. Such homologous recombination is stimulated by the presence of a double-stranded break in cellular chromatin, if sequences homologous to the region of the break are present.

[0078] In any of the methods described herein, the exogenous nucleotide sequence (the "donor sequence" or "transgene") can contain sequences that are homologous, but not identical, to genomic sequences in the region of interest, thereby stimulating homologous recombination to insert a non-identical sequence in the

region of interest. Thus, in certain embodiments, portions of the donor sequence that are homologous to sequences in the region of interest exhibit between about 80 to 99% (or any integer therebetween) sequence identity to the genomic sequence that is replaced. In other embodiments, the homology between the donor and genomic sequence is higher than 99%, for example if only 1 nucleotide differs as between donor and genomic sequences of over 100 contiguous base pairs. In certain cases, a non-homologous portion of the donor sequence can contain sequences not present in the region of interest, such that new sequences are introduced into the region of interest. In these instances, the non-homologous sequence is generally flanked by sequences of 50-1,000 base pairs (or any integral value therebetween) or any number of base pairs greater than 1,000, that are homologous or identical to sequences in the region of interest. In other embodiments, the donor sequence is non-homologous to the first sequence, and is inserted into the genome by non-homologous recombination mechanisms.

[0079] "Cleavage" refers to the breakage of the covalent backbone of a DNA molecule. Cleavage can be initiated by a variety of methods including, but not limited to, enzymatic or chemical hydrolysis of a phosphodiester bond. Both single-stranded cleavage and double-stranded cleavage are possible, and double-stranded cleavage can occur as a result of two distinct single-stranded cleavage events. DNA cleavage can result in the production of either blunt ends or staggered ends. In certain embodiments, fusion polypeptides are used for targeted double-stranded DNA cleavage.

[0080] A "cleavage half-domain" is a polypeptide sequence which, in conjunction with a second polypeptide (either identical or different) forms a complex having cleavage activity (preferably double-strand cleavage activity). The terms “first and second cleavage half-domains;” “+ and – cleavage half-domains” and “right and left cleavage half-domains” are used interchangeably to refer to pairs of cleavage half-domains that dimerize.

[0081] An “engineered cleavage half-domain” is a cleavage half-domain that has been modified so as to form obligate heterodimers with another cleavage half-domain (*e.g.*, another engineered cleavage half-domain). *See, also,* U.S. Patent Publication Nos. 2005/0064474, 20070218528, 2008/0131962 and 2011/0201055, incorporated herein by reference in their entireties.

[0082] The term "sequence" refers to a nucleotide sequence of any length, which can be DNA or RNA; can be linear, circular or branched and can be either single-stranded or double stranded. The term "donor sequence" refers to a nucleotide sequence that is inserted into a genome. A donor sequence can be of any length, for example between 2 and 100,000,000 nucleotides in length (or any integer value therebetween or thereabove), preferably between about 100 and 100,000 nucleotides in length (or any integer therebetween), more preferably between about 100 and 5,000 nucleotides in length (or any value therebetween) and even more preferable, between about 100 and 2,000 base pairs (or any value therebetween).

[0083] A "homologous, non-identical sequence" refers to a first sequence which shares a degree of sequence identity with a second sequence, but whose sequence is not identical to that of the second sequence. For example, a polynucleotide comprising the wild-type sequence of a mutant gene is homologous and non-identical to the sequence of the mutant gene. In certain embodiments, the degree of homology between the two sequences is sufficient to allow homologous recombination therebetween, utilizing normal cellular mechanisms. Two homologous non-identical sequences can be any length and their degree of non-homology can be as small as a single nucleotide (*e.g.*, for correction of a genomic point mutation by targeted homologous recombination) or as large as 10 or more kilobases (*e.g.*, for insertion of a gene at a predetermined ectopic site in a chromosome). Two polynucleotides comprising the homologous non-identical sequences need not be the same length. For example, an exogenous polynucleotide (*i.e.*, donor polynucleotide) of between 20 and 10,000 nucleotides or nucleotide pairs can be used.

[0084] Techniques for determining nucleic acid and amino acid sequence identity are known in the art. Typically, such techniques include determining the nucleotide sequence of the mRNA for a gene and/or determining the amino acid sequence encoded thereby, and comparing these sequences to a second nucleotide or amino acid sequence. Genomic sequences can also be determined and compared in this fashion. In general, identity refers to an exact nucleotide-to-nucleotide or amino acid-to-amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. Two or more sequences (polynucleotide or amino acid) can be compared by determining their percent identity using standard techniques. Typically the percent identities between sequences are at least 70-75%, preferably 80-82%,

more preferably 85-90%, even more preferably 92%, still more preferably 95%, and most preferably 98% sequence identity.

[0085] Alternatively, the degree of sequence similarity between polynucleotides can be determined by hybridization of polynucleotides under conditions that allow formation of stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. Two nucleic acid, or two polypeptide sequences are substantially homologous to each other when the sequences exhibit at least about 70%-75%, preferably 80%-82%, more preferably 85%-90%, even more preferably 92%, still more preferably 95%, and most preferably 98% sequence identity over a defined length of the molecules, as determined using the methods known in the art. Conditions for hybridization are well-known to those of skill in the art. Hybridization stringency refers to the degree to which hybridization conditions disfavor the formation of hybrids containing mismatched nucleotides, with higher stringency correlated with a lower tolerance for mismatched hybrids. Factors that affect the stringency of hybridization are well-known to those of skill in the art and include, but are not limited to, temperature, pH, ionic strength, and concentration of organic solvents such as, for example, formamide and dimethylsulfoxide. As is known to those of skill in the art, hybridization stringency is increased by higher temperatures, lower ionic strength and lower solvent concentrations.

[0086] "Chromatin" is the nucleoprotein structure comprising the cellular genome. Cellular chromatin comprises nucleic acid, primarily DNA, and protein, including histones and non-histone chromosomal proteins. The majority of eukaryotic cellular chromatin exists in the form of nucleosomes, wherein a nucleosome core comprises approximately 150 base pairs of DNA associated with an octamer comprising two each of histones H2A, H2B, H3 and H4; and linker DNA (of variable length depending on the organism) extends between nucleosome cores. A molecule of histone H1 is generally associated with the linker DNA. For the purposes of the present disclosure, the term “chromatin” is meant to encompass all types of cellular nucleoprotein, both prokaryotic and eukaryotic. Cellular chromatin includes both chromosomal and episomal chromatin.

[0087] A "chromosome," is a chromatin complex comprising all or a portion of the genome of a cell. The genome of a cell is often characterized by its karyotype,

which is the collection of all the chromosomes that comprise the genome of the cell. The genome of a cell can comprise one or more chromosomes.

[0088] An "episome" is a replicating nucleic acid, nucleoprotein complex or other structure comprising a nucleic acid that is not part of the chromosomal karyotype of a cell. Examples of episomes include plasmids and certain viral genomes.

[0089] An "accessible region" is a site in cellular chromatin in which a target site present in the nucleic acid can be bound by an exogenous molecule which recognizes the target site. Without wishing to be bound by any particular theory, it is believed that an accessible region is one that is not packaged into a nucleosomal structure. The distinct structure of an accessible region can often be detected by its sensitivity to chemical and enzymatic probes, for example, nucleases.

[0090] A "target site" or "target sequence" is a nucleic acid sequence that defines a portion of a nucleic acid to which a binding molecule will bind, provided sufficient conditions for binding exist.

[0091] An "exogenous" molecule is a molecule that is not normally present in a cell, but can be introduced into a cell by one or more genetic, biochemical or other methods. "Normal presence in the cell" is determined with respect to the particular developmental stage and environmental conditions of the cell. Thus, for example, a molecule that is present only during embryonic development of muscle is an exogenous molecule with respect to an adult muscle cell. Similarly, a molecule induced by heat shock is an exogenous molecule with respect to a non-heat-shocked cell. An exogenous molecule can comprise, for example, a functioning version of a malfunctioning endogenous molecule or a malfunctioning version of a normally-functioning endogenous molecule.

[0092] An exogenous molecule can be, among other things, a small molecule, such as is generated by a combinatorial chemistry process, or a macromolecule such as a protein, nucleic acid, carbohydrate, lipid, glycoprotein, lipoprotein, polysaccharide, any modified derivative of the above molecules, or any complex comprising one or more of the above molecules. Nucleic acids include DNA and RNA, can be single- or double-stranded; can be linear, branched or circular; and can be of any length. Nucleic acids include those capable of forming duplexes, as well as triplex-forming nucleic acids. See, for example, U.S. Patent Nos. 5,176,996 and 5,422,251. Proteins include, but are not limited to, DNA-binding proteins,

transcription factors, chromatin remodeling factors, methylated DNA binding proteins, polymerases, methylases, demethylases, acetylases, deacetylases, kinases, phosphatases, integrases, recombinases, ligases, topoisomerases, gyrases and helicases.

[0093] An exogenous molecule can be the same type of molecule as an endogenous molecule, *e.g.*, an exogenous protein or nucleic acid. For example, an exogenous nucleic acid can comprise an infecting viral genome, a plasmid or episome introduced into a cell, or a chromosome that is not normally present in the cell. Methods for the introduction of exogenous molecules into cells are known to those of skill in the art and include, but are not limited to, lipid-mediated transfer (*i.e.*, liposomes, including neutral and cationic lipids), electroporation, direct injection, cell fusion, particle bombardment, biopolymer nanoparticle delivery (see Nitta and Numata (2013) *Int J Mol Sci* 14:1629), calcium phosphate co-precipitation, DEAE-dextran-mediated transfer and viral vector-mediated transfer. An exogeneous molecule can also be the same type of molecule as an endogenous molecule but derived from a different species than the cell is derived from. For example, a human nucleic acid sequence may be introduced into a cell line originally derived from a mouse or hamster..

[0094] By contrast, an "endogenous" molecule is one that is normally present in a particular cell at a particular developmental stage under particular environmental conditions. For example, an endogenous nucleic acid can comprise a chromosome, the genome of a mitochondrion, or other organelle, or a naturally-occurring episomal nucleic acid. Additional endogenous molecules can include proteins, for example, transcription factors and enzymes.

[0095] As used herein, the term “product of an exogenous nucleic acid” includes both polynucleotide and polypeptide products, for example, transcription products (polynucleotides such as RNA) and translation products (polypeptides).

[0096] A "fusion" molecule is a molecule in which two or more subunit molecules are linked, preferably covalently. The subunit molecules can be the same chemical type of molecule, or can be different chemical types of molecules. Examples of the first type of fusion molecule include, but are not limited to, fusion proteins (for example, a fusion between a ZFP or TALE DNA-binding domain and one or more activation domains) and fusion nucleic acids (for example, a nucleic acid encoding the fusion protein described *supra*). Examples of the second type of fusion

molecule include, but are not limited to, a fusion between a triplex-forming nucleic acid and a polypeptide, and a fusion between a minor groove binder and a nucleic acid.

[0097] Expression of a fusion protein in a cell can result from delivery of the fusion protein to the cell or by delivery of a polynucleotide encoding the fusion protein to a cell, wherein the polynucleotide is transcribed, and the transcript is translated, to generate the fusion protein. Trans-splicing, polypeptide cleavage and polypeptide ligation can also be involved in expression of a protein in a cell. Methods for polynucleotide and polypeptide delivery to cells are presented elsewhere in this disclosure.

[0098] A "gene," for the purposes of the present disclosure, includes a DNA region encoding a gene product (see *infra*), as well as all DNA regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. Accordingly, a gene includes, but is not necessarily limited to, promoter sequences, terminators, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites and locus control regions.

[0099] "Gene expression" refers to the conversion of the information, contained in a gene, into a gene product. A gene product can be the direct transcriptional product of a gene (*e.g.*, mRNA, tRNA, rRNA, antisense RNA, ribozyme, structural RNA or any other type of RNA) or a protein produced by translation of an mRNA. Gene products also include RNAs which are modified, by processes such as capping, polyadenylation, methylation, and editing, and proteins modified by, for example, methylation, acetylation, phosphorylation, ubiquitination, ADP-ribosylation, myristilation, and glycosylation.

[0100] "Modulation" of gene expression refers to a change in the activity of a gene. Modulation of expression can include, but is not limited to, gene activation and gene repression. Genome editing (*e.g.,* cleavage, alteration, inactivation, random mutation) can be used to modulate expression. Gene inactivation refers to any reduction in gene expression as compared to a cell that does not include a ZFP, TALE or CRISPR/Cas system as described herein. Thus, gene inactivation may be partial or complete.

[0101] A "region of interest" is any region of cellular chromatin, such as, for example, a gene or a non-coding sequence within or adjacent to a gene, in which it is desirable to bind an exogenous molecule. Binding can be for the purposes of targeted DNA cleavage and/or targeted recombination. A region of interest can be present in a chromosome, an episome, an organellar genome (*e.g.*, mitochondrial, chloroplast), or an infecting viral genome, for example. A region of interest can be within the coding region of a gene, within transcribed non-coding regions such as, for example, leader sequences, trailer sequences or introns, or within non-transcribed regions, either upstream or downstream of the coding region. A region of interest can be as small as a single nucleotide pair or up to 2,000 nucleotide pairs in length, or any integral value of nucleotide pairs.

[0102] "Eukaryotic" cells include, but are not limited to, fungal cells (such as yeast), plant cells, animal cells, mammalian cells and human cells (*e.g.*, T-cells).

[0103] "Secretory tissues" are those tissues in an animal that secrete products out of the individual cell into a lumen of some type which are typically derived from epithelium. Examples of secretory tissues that are localized to the gastrointestinal tract include the cells that line the gut, the pancreas, and the gallbladder. Other secretory tissues include the liver, tissues associated with the eye and mucous membranes such as salivary glands, mammary glands, the prostate gland, the pituitary gland and other members of the endocrine system. Additionally, secretory tissues include individual cells of a tissue type which are capable of secretion.

[0104] The terms "operative linkage" and "operatively linked" (or "operably linked") are used interchangeably with reference to a juxtaposition of two or more components (such as sequence elements), in which the components are arranged such that both components function normally and allow the possibility that at least one of the components can mediate a function that is exerted upon at least one of the other components. By way of illustration, a transcriptional regulatory sequence, such as a promoter, is operatively linked to a coding sequence if the transcriptional regulatory sequence controls the level of transcription of the coding sequence in response to the presence or absence of one or more transcriptional regulatory factors. A transcriptional regulatory sequence is generally operatively linked in *cis* with a coding sequence, but need not be directly adjacent to it. For example, an enhancer is a transcriptional regulatory sequence that is operatively linked to a coding sequence, even though they are not contiguous.

[0105] With respect to fusion polypeptides, the term "operatively linked" can refer to the fact that each of the components performs the same function in linkage to the other component as it would if it were not so linked. For example, with respect to a fusion polypeptide in which a ZFP, TALE or Cas DNA-binding domain is fused to an activation domain, the ZFP, TALE or Cas DNA-binding domain and the activation domain are in operative linkage if, in the fusion polypeptide, the ZFP, TALE of Cas DNA-binding domain portion is able to bind its target site and/or its binding site, while the activation domain is able to upregulate gene expression. When a fusion polypeptide in which a ZFP, TALE or Cas DNA-binding domain is fused to a cleavage domain, the ZFP, TALE or Cas DNA-binding domain and the cleavage domain are in operative linkage if, in the fusion polypeptide, the ZFP, TALE or Cas DNA-binding domain portion is able to bind its target site and/or its binding site, while the cleavage domain is able to cleave DNA in the vicinity of the target site (*e.g.*, 1 to 500 base pairs or any value therebetween on either side of the target site).

[0106] A "functional fragment" of a protein, polypeptide or nucleic acid is a protein, polypeptide or nucleic acid whose sequence is not identical to the full-length protein, polypeptide or nucleic acid, yet retains the same function as the full-length protein, polypeptide or nucleic acid. A functional fragment can possess more, fewer, or the same number of residues as the corresponding native molecule, and/or can contain one or more amino acid or nucleotide substitutions. Methods for determining the function of a nucleic acid (*e.g.*, coding function, ability to hybridize to another nucleic acid) are well-known in the art. Similarly, methods for determining protein function are well-known. For example, the DNA-binding function of a polypeptide can be determined, for example, by filter-binding, electrophoretic mobility-shift, or immunoprecipitation assays. DNA cleavage can be assayed by gel electrophoresis. See Ausubel *et al., supra*. The ability of a protein to interact with another protein can be determined, for example, by co-immunoprecipitation, two-hybrid assays or complementation, both genetic and biochemical. See, for example, Fields *et al.* (1989) *Nature* **340:**245-246; U.S. Patent No. 5,585,245 and PCT WO 98/44350.

[0107] A "vector" is capable of transferring gene sequences to target cells. Typically, "vector construct," "expression vector," and "gene transfer vector," mean any nucleic acid construct capable of directing the expression of a gene of interest and which can transfer gene sequences to target cells. Thus, the term includes cloning, and expression vehicles, as well as integrating vectors.

[0108] A "reporter gene" or "reporter sequence" refers to any sequence that produces a protein product that is easily measured, preferably although not necessarily in a routine assay. Suitable reporter genes include, but are not limited to, sequences encoding proteins that mediate antibiotic resistance (*e.g.*, ampicillin resistance, neomycin resistance, G418 resistance, puromycin resistance), sequences encoding colored or fluorescent or luminescent proteins (*e.g.*, green fluorescent protein, enhanced green fluorescent protein, red fluorescent protein, luciferase), and proteins which mediate enhanced cell growth and/or gene amplification (*e.g.*, dihydrofolate reductase). Epitope tags include, for example, one or more copies of FLAG, His, myc, Tap, HA or any detectable amino acid sequence. "Expression tags" include sequences that encode reporters that may be operably linked to a desired gene sequence in order to monitor expression of the gene of interest.

[0109] A "safe harbor" locus is a locus within the genome wherein a gene may be inserted without any deleterious effects on the host cell. Most beneficial is a safe harbor locus in which expression of the inserted gene sequence is not perturbed by any read-through expression from neighboring genes. Non-limiting examples of safe harbor loci in mammalian cells are the AAVS1 gene (U.S. Patent No. 8,110,379), the CCR5 gene (U.S. Publication No. 20080159996), the Rosa locus (WO 2010/065123) and/or the albumin locus (U.S. Publication Nos. 20130177960 and 20130177983). A safe harbor in a plant cell is the ZP15 locus (U.S. patent publication 20100199389).

[0110] The terms "subject" and "patient" are used interchangeably and refer to mammals such as human patients and non-human primates, as well as experimental animals such as rabbits, dogs, cats, rats, mice, and other animals. Accordingly, the term "subject" or "patient" as used herein means any mammalian patient or subject to which the or stem cells of the invention can be administered. Subjects of the present invention include those that have been exposed to one or more chemical toxins, including, for example, a nerve toxin.

[0111] "Stemness" refers to the relative ability of any cell to act in a stem cell-like manner, *i.e.*, the degree of toti-, pluri-, or oligopotentcy and expanded or indefinite self-renewal that any particular stem cell may have.

Nucleases

[0112] Described herein are compositions, particularly nucleases, such as ZFNs, TALEs, homing endonucleases, Ttago and/or CRISPR/Cas systems, that are useful for *in vivo* cleavage of a donor molecule carrying a transgene and nucleases for cleavage of the genome of a cell such that the transgene is integrated into the genome in a targeted manner. In certain embodiments, one or more of the nucleases are naturally occurring. In other embodiments, one or more of the nucleases are non-naturally occurring, *i.e.*, engineered in the DNA-binding domain and/or cleavage domain. For example, the DNA-binding domain of a naturally-occurring nuclease may be altered to bind to a selected target site (*e.g.*, a meganuclease that has been engineered to bind to site different than the cognate binding site). In other embodiments, the nuclease comprises heterologous DNA-binding and cleavage domains (*e.g.*, zinc finger nucleases; TAL-effector domain DNA binding proteins; meganuclease DNA-binding domains with heterologous cleavage domains). In other embodiments, the nuclease comprises a system such as the CRISPR/Cas or Ttago system.

A. DNA-binding domains

[0113] In certain embodiments, the composition and methods described herein employ a meganuclease (homing endonuclease) DNA-binding domain for binding to the donor molecule and/or binding to the region of interest in the genome of the cell. Naturally-occurring meganucleases recognize 15-40 base-pair cleavage sites and are commonly grouped into four families: the LAGLIDADG family, the GIY-YIG family, the His-Cyst box family and the HNH family. Exemplary homing endonucleases include I-*Sce*I, I-*Ceu*I, PI-*Psp*I, PI-*Sce*, I-*Sce*IV, I-*Csm*I, I-*Pan*I, I-*Sce*II, I-*Ppo*I, I-*Sce*III, I-*Cre*I, I-*Tev*I, I-*Tev*II and I-*Tev*III. Their recognition sequences are known. *See* also U.S. Patent No. 5,420,032; U.S. Patent No. 6,833,252; Belfort *et al.* (1997) *Nucleic Acids Res.* **25:**3379–3388; Dujon *et al.* (1989) *Gene* **82:**115–118; Perler *et al.* (1994) Nucleic Acids Res. 22, 1125–1127; Jasin (1996) *Trends Genet.* **12:**224–228; Gimble *et al.* (1996) *J. Mol. Biol.* **263:**163–180; Argast *et al.* (1998) *J. Mol. Biol.* **280:**345–353 and the New England Biolabs catalogue.

[0114] In certain embodiments, the methods and compositions described herein make use of a nuclease that comprises an engineered (non-naturally occurring)

homing endonuclease (meganuclease). The recognition sequences of homing endonucleases and meganucleases such as I-*Sce*I, I-*Ceu*I, PI-*Psp*I, PI-*Sce*, I-*Sce*IV, I-*Csm*I, I-*Pan*I, I-*Sce*II, I-*Ppo*I, I-*Sce*III, I-*Cre*I, I-*Tev*I, I-*Tev*II and I-*Tev*III are known. *See* also U.S. Patent No. 5,420,032; U.S. Patent No. 6,833,252; Belfort *et al.* (1997) *Nucleic Acids Res.* **25:**3379–3388; Dujon *et al.* (1989) *Gene* **82:**115–118; Perler *et al.* (1994) Nucleic Acids Res. 22, 1125–1127; Jasin (1996) *Trends Genet.* **12:**224–228; Gimble *et al.* (1996) *J. Mol. Biol.* **263:**163–180; Argast *et al.* (1998) *J. Mol. Biol.* **280:**345–353 and the New England Biolabs catalogue. In addition, the DNA-binding specificity of homing endonucleases and meganucleases can be engineered to bind non-natural target sites. *See*, for example, Chevalier *et al.* (2002) *Molec. Cell* **10:**895-905; Epinat *et al.* (2003) *Nucleic Acids Res.* **31:**2952-2962; Ashworth *et al.* (2006) *Nature* **441:**656-659; Paques *et al.* (2007) *Current Gene Therapy* **7:**49-66; U.S. Patent Publication No. 20070117128. The DNA-binding domains of the homing endonucleases and meganucleases may be altered in the context of the nuclease as a whole (*i.e.,* such that the nuclease includes the cognate cleavage domain) or may be fused to a heterologous cleavage domain.

[0115] In other embodiments, the DNA-binding domain of one or more of the nucleases used in the methods and compositions described herein comprises a naturally occurring or engineered (non-naturally occurring) TAL effector DNA binding domain. *See, e.g.,* U.S. Patent No. 8,586,526, incorporated by reference in its entirety herein. The plant pathogenic bacteria of the genus *Xanthomonas* are known to cause many diseases in important crop plants. Pathogenicity of *Xanthomonas* depends on a conserved type III secretion (T3S) system which injects more than 25 different effector proteins into the plant cell. Among these injected proteins are transcription activator-like (TAL) effectors which mimic plant transcriptional activators and manipulate the plant transcriptome (see Kay *et al* (2007) *Science* 318:648-651). These proteins contain a DNA binding domain and a transcriptional activation domain. One of the most well characterized TAL-effectors is AvrBs3 from *Xanthomonas campestgris* pv. *Vesicatoria* (see Bonas *et al* (1989) *Mol Gen Genet* 218: 127-136 and WO2010079430). TAL-effectors contain a centralized domain of tandem repeats, each repeat containing approximately 34 amino acids, which are key to the DNA binding specificity of these proteins. In addition, they contain a nuclear localization sequence and an acidic transcriptional activation domain (for a review see Schornack S, *et al* (2006) *J Plant Physiol* 163(3): 256-272). In addition, in the

phytopathogenic bacteria *Ralstonia solanacearum* two genes, designated brg11 and hpx17 have been found that are homologous to the AvrBs3 family of *Xanthomonas* in the *R. solanacearum* biovar 1 strain GMI1000 and in the biovar 4 strain RS1000 (See Heuer *et al* (2007) *Appl and Envir Micro* 73(13): 4379-4384). These genes are 98.9% identical in nucleotide sequence to each other but differ by a deletion of 1,575 bp in the repeat domain of hpx17. However, both gene products have less than 40% sequence identity with AvrBs3 family proteins of *Xanthomonas*. *See, e.g.,* U.S. Patent Publication No. 8,586,526, incorporated by reference in its entirety herein.

[0116] Specificity of these TAL effectors depends on the sequences found in the tandem repeats. The repeated sequence comprises approximately 102 bp and the repeats are typically 91-100% homologous with each other (Bonas *et al, ibid*). Polymorphism of the repeats is usually located at positions 12 and 13 and there appears to be a one-to-one correspondence between the identity of the hypervariable diresidues at positions 12 and 13 with the identity of the contiguous nucleotides in the TAL-effector's target sequence (see Moscou and Bogdanove, (2009) *Science* 326:1501 and Boch *et al* (2009) *Science* 326:1509-1512). Experimentally, the natural code for DNA recognition of these TAL-effectors has been determined such that an HD sequence at positions 12 and 13 leads to a binding to cytosine (C), NG binds to T, NI to A, C, G or T, NN binds to A or G, and ING binds to T. These DNA binding repeats have been assembled into proteins with new combinations and numbers of repeats, to make artificial transcription factors that are able to interact with new sequences and activate the expression of a non-endogenous reporter gene in plant cells (Boch *et al, ibid*). Engineered TAL proteins have been linked to a *Fok*I cleavage half domain to yield a TAL effector domain nuclease fusion (TALEN) exhibiting activity in a yeast reporter assay (plasmid based target). *See, e.g.,* U.S. Patent No. 8,586,526; Christian *et al* ((2010)< *Genetics* epub 10.1534/genetics.110.120717).

[0117] In certain embodiments, the DNA binding domain of one or more of the nucleases used for *in vivo* cleavage and/or targeted cleavage of the genome of a cell comprises a zinc finger protein. Preferably, the zinc finger protein is non-naturally occurring in that it is engineered to bind to a target site of choice. *See*, for example, *See*, for example, Beerli *et al.* (2002) *Nature Biotechnol.* **20:**135-141; Pabo *et al.* (2001) *Ann. Rev. Biochem.* **70:**313-340; Isalan *et al.* (2001) *Nature Biotechnol.* **19:**656-660; Segal *et al.* (2001) *Curr. Opin. Biotechnol.* **12:**632-637; Choo *et al.* (2000) *Curr. Opin. Struct. Biol.* **10:**411-416; U.S. Patent Nos. 6,453,242; 6,534,261;

6,599,692; 6,503,717; 6,689,558; 7,030,215; 6,794,136; 7,067,317; 7,262,054; 7,070,934; 7,361,635; 7,253,273; and U.S. Patent Publication Nos. 2005/0064474; 2007/0218528; 2005/0267061, all incorporated herein by reference in their entireties.

[0118] An engineered zinc finger binding domain can have a novel binding specificity, compared to a naturally-occurring zinc finger protein. Engineering methods include, but are not limited to, rational design and various types of selection. Rational design includes, for example, using databases comprising triplet (or quadruplet) nucleotide sequences and individual zinc finger amino acid sequences, in which each triplet or quadruplet nucleotide sequence is associated with one or more amino acid sequences of zinc fingers which bind the particular triplet or quadruplet sequence. See, for example, co-owned U.S. Patents 6,453,242 and 6,534,261, incorporated by reference herein in their entireties.

[0119] Exemplary selection methods, including phage display and two-hybrid systems, are disclosed in US Patents 5,789,538; 5,925,523; 6,007,988; 6,013,453; 6,410,248; 6,140,466; 6,200,759; and 6,242,568; as well as WO 98/37186; WO 98/53057; WO 00/27878; and WO 01/88197. In addition, enhancement of binding specificity for zinc finger binding domains has been described, for example, in co-owned WO 02/077227.

[0120] In addition, as disclosed in these and other references, zinc finger domains and/or multi-fingered zinc finger proteins may be linked together using any suitable linker sequences, including for example, linkers of 5 or more amino acids in length. See, also, U.S. Patent Nos. 8,772,453; 6,479,626; 6,903,185; and 7,153,949 for exemplary linker sequences. The proteins described herein may include any combination of suitable linkers between the individual zinc fingers of the protein.

[0121] Selection of target sites; ZFPs and methods for design and construction of fusion proteins (and polynucleotides encoding same) are known to those of skill in the art and described in detail in U.S. Patent Nos. 6,140,815; 5,789,538; 6,453,242; 6,534,261; 5,925,523; 6,007,988; 6,013,453; 6,200,759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 and WO 03/016496.

[0122] In addition, as disclosed in these and other references, zinc finger domains and/or multi-fingered zinc finger proteins may be linked together using any suitable linker sequences, including for example, linkers of 5 or more amino acids in

length. See, also, U.S. Patent Nos. 6,479,626; 6,903,185; and 7,153,949 for exemplary linker sequences 6 or more amino acids in length. The proteins described herein may include any combination of suitable linkers between the individual zinc fingers of the protein.

[0123] In certain embodiments, the DNA-binding domain is part of a CRISPR/Cas nuclease system. *See, e.g.,* U.S. Patent No. 8,697,359 and U.S. Patent Application No. 14/278,903. The CRISPR (clustered regularly interspaced short palindromic repeats) locus, which encodes RNA components of the system, and the cas (CRISPR-associated) locus, which encodes proteins (Jansen *et al.*, 2002. *Mol. Microbiol.* 43: 1565-1575; Makarova *et al.*, 2002. *Nucleic Acids Res.* 30: 482-496; Makarova *et al.*, 2006. *Biol. Direct* 1: 7; Haft *et al.*, 2005. *PLoSComput. Biol.* 1: e60) make up the gene sequences of the CRISPR/Cas nuclease system. CRISPR loci in microbial hosts contain a combination of CRISPR-associated (Cas) genes as well as non-coding RNA elements capable of programming the specificity of the CRISPR-mediated nucleic acid cleavage.

[0124] The Type II CRISPR is one of the most well characterized systems and carries out targeted DNA double-strand break in four sequential steps. First, two non-coding RNA, the pre-crRNA array and tracrRNA, are transcribed from the CRISPR locus. Second, tracrRNA hybridizes to the repeat regions of the pre-crRNA and mediates the processing of pre-crRNA into mature crRNAs containing individual spacer sequences. Third, the mature crRNA:tracrRNA complex directs Cas9 to the target DNA via Watson-Crick base-pairing between the spacer on the crRNA and the protospacer on the target DNA next to the protospacer adjacent motif (PAM), an additional requirement for target recognition. Finally, Cas9 mediates cleavage of target DNA to create a double-stranded break within the protospacer. Activity of the CRISPR/Cas system comprises of three steps: (i) insertion of alien DNA sequences into the CRISPR array to prevent future attacks, in a process called 'adaptation', (ii) expression of the relevant proteins, as well as expression and processing of the array, followed by (iii) RNA-mediated interference with the alien nucleic acid. Thus, in the bacterial cell, several of the so-called 'Cas' proteins are involved with the natural function of the CRISPR/Cas system and serve roles in functions such as insertion of the alien DNA etc.

[0125] In some embodiments, the CRISPR-Cpf1 system is used. The CRISPR-Cpf1 system, identified in *Francisella spp*, is a class 2 CRISPR-Cas system

that mediates robust DNA interference in human cells. Although functionally conserved, Cpf1 and Cas9 differ in many aspects including in their guide RNAs and substrate specificity (see Fagerlund *et al*, (2015) *Genom Bio* 16:251). A major difference between Cas9 and Cpf1 proteins is that Cpf1 does not utilize tracrRNA, and thus requires only a crRNA. The FnCpf1 crRNAs are 42–44 nucleotides long (19-nucleotide repeat and 23–25-nucleotide spacer) and contain a single stem-loop, which tolerates sequence changes that retain secondary structure. In addition, the Cpf1 crRNAs are significantly shorter than the ~100-nucleotide engineered sgRNAs required by Cas9, and the PAM requirements for FnCpfl are 5′-TTN-3′ and 5′-CTA-3′ on the displaced strand. Although both Cas9 and Cpf1 make double strand breaks in the target DNA, Cas9 uses its RuvC- and HNH-like domains to make blunt-ended cuts within the seed sequence of the guide RNA, whereas Cpf1 uses a RuvC-like domain to produce staggered cuts outside of the seed. Because Cpf1 makes staggered cuts away from the critical seed region, NHEJ will not disrupt the target site, therefore ensuring that Cpf1 can continue to cut the same site until the desired HDR recombination event has taken place. Thus, in the methods and compositions described herein, it is understood that the term "Cas" includes both Cas9 and Cfp1 proteins. Thus, as used herein, a "CRISPR/Cas system" refers both CRISPR/Cas and/or CRISPR/Cfp1 systems, including both nuclease and/or transcription factor systems.

[0126] In certain embodiments, Cas protein may be a "functional derivative" of a naturally occurring Cas protein. A "functional derivative" of a native sequence polypeptide is a compound having a qualitative biological property in common with a native sequence polypeptide. "Functional derivatives" include, but are not limited to, fragments of a native sequence and derivatives of a native sequence polypeptide and its fragments, provided that they have a biological activity in common with a corresponding native sequence polypeptide. A biological activity contemplated herein is the ability of the functional derivative to hydrolyze a DNA substrate into fragments. The term "derivative" encompasses both amino acid sequence variants of polypeptide, covalent modifications, and fusions thereof. Suitable derivatives of a Cas polypeptide or a fragment thereof include but are not limited to mutants, fusions, covalent modifications of Cas protein or a fragment thereof. Cas protein, which includes Cas protein or a fragment thereof, as well as derivatives of Cas protein or a fragment thereof, may be obtainable from a cell or synthesized chemically or by a combination

of these two procedures. The cell may be a cell that naturally produces Cas protein, or a cell that naturally produces Cas protein and is genetically engineered to produce the endogenous Cas protein at a higher expression level or to produce a Cas protein from an exogenously introduced nucleic acid, which nucleic acid encodes a Cas that is same or different from the endogenous Cas. In some case, the cell does not naturally produce Cas protein and is genetically engineered to produce a Cas protein.

[0127] In some embodiments, the DNA binding domain is part of a TtAgo system (see Swarts *et al*, *ibid*; Sheng *et al*, *ibid*). In eukaryotes, gene silencing is mediated by the Argonaute (Ago) family of proteins. In this paradigm, Ago is bound to small (19-31 nt) RNAs. This protein-RNA silencing complex recognizes target RNAs via Watson-Crick base pairing between the small RNA and the target and endonucleolytically cleaves the target RNA (Vogel (2014) *Science* 344:972-973). In contrast, prokaryotic Ago proteins bind to small single-stranded DNA fragments and likely function to detect and remove foreign (often viral) DNA (Yuan *et al*., (2005) *Mol. Cell* 19, 405; Olovnikov, *et al*. (2013) *Mol. Cell* 51, 594; Swarts *et al., ibid*). Exemplary prokaryotic Ago proteins include those from *Aquifex aeolicus*, *Rhodobacter sphaeroides*, and *Thermus thermophilus*.

[0128] One of the most well-characterized prokaryotic Ago protein is the one from *T. thermophilus* (TtAgo; Swarts *et al. ibid*). TtAgo associates with either 15 nt or 13-25 nt single-stranded DNA fragments with 5' phosphate groups. This "guide DNA" bound by TtAgo serves to direct the protein-DNA complex to bind a Watson-Crick complementary DNA sequence in a third-party molecule of DNA. Once the sequence information in these guide DNAs has allowed identification of the target DNA, the TtAgo-guide DNA complex cleaves the target DNA. Such a mechanism is also supported by the structure of the TtAgo-guide DNA complex while bound to its target DNA (G. Sheng *et al., ibid*). Ago from *Rhodobacter sphaeroides* (RsAgo) has similar properties (Olivnikov *et al. ibid*).

[0129] Exogenous guide DNAs of arbitrary DNA sequence can be loaded onto the TtAgo protein (Swarts *et al. ibid.*). Since the specificity of TtAgo cleavage is directed by the guide DNA, a TtAgo-DNA complex formed with an exogenous, investigator-specified guide DNA will therefore direct TtAgo target DNA cleavage to a complementary investigator-specified target DNA. In this way, one may create a targeted double-strand break in DNA. Use of the TtAgo-guide DNA system (or

orthologous Ago-guide DNA systems from other organisms) allows for targeted cleavage of genomic DNA within cells. Such cleavage can be either single- or double-stranded. For cleavage of mammalian genomic DNA, it would be preferable to use of a version of TtAgo codon optimized for expression in mammalian cells. Further, it might be preferable to treat cells with a TtAgo-DNA complex formed *in vitro* where the TtAgo protein is fused to a cell-penetrating peptide. Further, it might be preferable to use a version of the TtAgo protein that has been altered via mutagenesis to have improved activity at 37 degrees Celcius. Ago-RNA-mediated DNA cleavage could be used to effect a panopoly of outcomes including gene knock-out, targeted gene addition, gene correction, targeted gene deletion using techniques standard in the art for exploitation of DNA breaks.

[0130] Thus, the nuclease comprises a DNA-binding domain in that specifically binds to a target site in any gene into which it is desired to insert a donor (transgene). The DNA-binding domains described herein typically bind to a target site comprising 12 to 35 nucleotides (or any valuetherebetween). The nucleotides within the target sites that are bound by the DNA-binding domain may be contiguous or non-contiguous (*e.g.*, the DNA-binding domain may bind to less than all base pairs making up the target site).

B. Cleavage Domains

[0131] Any suitable cleavage domain can be operatively linked to a DNA-binding domain to form a nuclease. For example, ZFP DNA-binding domains have been fused to nuclease domains to create ZFNs – a functional entity that is able to recognize its intended nucleic acid target through its engineered (ZFP) DNA binding domain and cause the DNA to be cut near the ZFP binding site via the nuclease activity. See, *e.g.*, Kim *et al.* (1996) *Proc Natl Acad Sci USA* 93(3):1156-1160. More recently, ZFNs have been used for genome modification in a variety of organisms. *See*, for example, United States Patent Publications 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231; and International Publication WO 07/014275. Likewise, TALE DNA-binding domains have been fused to nuclease domains to create TALENs. *See, e.g.,* U.S. Patent No. 8,586,526.

[0132] As noted above, the cleavage domain may be heterologous to the DNA-binding domain, for example a zinc finger DNA-binding domain and a cleavage domain from a nuclease or a TALEN DNA-binding domain and a cleavage domain,

or meganuclease DNA-binding domain and cleavage domain from a different nuclease. Heterologous cleavage domains can be obtained from any endonuclease or exonuclease. Exemplary endonucleases from which a cleavage domain can be derived include, but are not limited to, restriction endonucleases and homing endonucleases. *See*, for example, 2002-2003 Catalogue, New England Biolabs, Beverly, MA; and Belfort *et al.* (1997) *Nucleic Acids Res.* **25:**3379-3388. Additional enzymes which cleave DNA are known (*e.g.*, S1 Nuclease; mung bean nuclease; pancreatic DNase I; micrococcal nuclease; yeast HO endonuclease; *see* also Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press,1993). One or more of these enzymes (or functional fragments thereof) can be used as a source of cleavage domains and cleavage half-domains.

[0133] Similarly, a cleavage half-domain can be derived from any nuclease or portion thereof, as set forth above, that requires dimerization for cleavage activity. In general, two fusion proteins are required for cleavage if the fusion proteins comprise cleavage half-domains. Alternatively, a single protein comprising two cleavage half-domains can be used. The two cleavage half-domains can be derived from the same endonuclease (or functional fragments thereof), or each cleavage half-domain can be derived from a different endonuclease (or functional fragments thereof). In addition, the target sites for the two fusion proteins are preferably disposed, with respect to each other, such that binding of the two fusion proteins to their respective target sites places the cleavage half-domains in a spatial orientation to each other that allows the cleavage half-domains to form a functional cleavage domain, *e.g.*, by dimerizing. Thus, in certain embodiments, the near edges of the target sites are separated by 5-8 nucleotides or by 15-18 nucleotides. However any integral number of nucleotides or nucleotide pairs can intervene between two target sites (*e.g.*, from 2 to 50 nucleotide pairs or more). In general, the site of cleavage lies between the target sites.

[0134] Restriction endonucleases (restriction enzymes) are present in many species and are capable of sequence-specific binding to DNA (at a recognition site), and cleaving DNA at or near the site of binding. Certain restriction enzymes (*e.g.*, Type IIS) cleave DNA at sites removed from the recognition site and have separable binding and cleavage domains. For example, the Type IIS enzyme *Fok* I catalyzes double-stranded cleavage of DNA, at 9 nucleotides from its recognition site on one strand and 13 nucleotides from its recognition site on the other. *See*, for example, US Patents 5,356,802; 5,436,150 and 5,487,994; as well as Li *et al.* (1992) *Proc. Natl.*

Acad. Sci. USA **89:**4275-4279; Li *et al.* (1993) *Proc. Natl. Acad. Sci. USA* **90:**2764-2768; Kim *et al.* (1994a) *Proc. Natl. Acad. Sci. USA* **91:**883-887; Kim *et al.* (1994b) *J. Biol. Chem.* **269:**31,978-31,982. Thus, in one embodiment, fusion proteins comprise the cleavage domain (or cleavage half-domain) from at least one Type IIS restriction enzyme and one or more zinc finger binding domains, which may or may not be engineered.

[0135] An exemplary Type IIS restriction enzyme, whose cleavage domain is separable from the binding domain, is *Fok* I. This particular enzyme is active as a dimer. Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* **95:** 10,570-10,575. Accordingly, for the purposes of the present disclosure, the portion of the *Fok* I enzyme used in the disclosed fusion proteins is considered a cleavage half-domain. Thus, for targeted double-stranded cleavage and/or targeted replacement of cellular sequences using zinc finger-*Fok* I fusions, two fusion proteins, each comprising a *Fok*I cleavage half-domain, can be used to reconstitute a catalytically active cleavage domain. Alternatively, a single polypeptide molecule containing a zinc finger binding domain and two *Fok* I cleavage half-domains can also be used. Parameters for targeted cleavage and targeted sequence alteration using zinc finger-*Fok* I fusions are provided elsewhere in this disclosure.

[0136] A cleavage domain or cleavage half-domain can be any portion of a protein that retains cleavage activity, or that retains the ability to multimerize (*e.g.*, dimerize) to form a functional cleavage domain.

[0137] Exemplary Type IIS restriction enzymes are described in U.S. Patent 7,888,121, incorporated herein in its entirety. Additional restriction enzymes also contain separable binding and cleavage domains, and these are contemplated by the present disclosure. *See*, for example, Roberts *et al.* (2003) *Nucleic Acids Res.* **31:**418-420.

[0138] In certain embodiments, the cleavage domain comprises one or more engineered cleavage half-domain (also referred to as dimerization domain mutants) that minimize or prevent homodimerization, as described, for example, in U.S. Patent Nos. 8,772,453; 8,623,618; 8,409,861; 8,034,598; 7,914,796; and 7,888,121, the disclosures of all of which are incorporated by reference in their entireties herein. Amino acid residues at positions 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, and 538 of *Fok* I are all targets for influencing dimerization of the *Fok* I cleavage half-domains.

[0139] Exemplary engineered cleavage half-domains of *Fok* I that form obligate heterodimers include a pair in which a first cleavage half-domain includes mutations at amino acid residues at positions 490 and 538 of *Fok* I and a second cleavage half-domain includes mutations at amino acid residues 486 and 499.

[0140] Thus, in one embodiment, a mutation at 490 replaces Glu (E) with Lys (K); the mutation at 538 replaces Iso (I) with Lys (K); the mutation at 486 replaced Gln (Q) with Glu (E); and the mutation at position 499 replaces Iso (I) with Lys (K). Specifically, the engineered cleavage half-domains described herein were prepared by mutating positions 490 (E→K) and 538 (I→K) in one cleavage half-domain to produce an engineered cleavage half-domain designated "E490K:I538K" and by mutating positions 486 (Q→E) and 499 (I→L) in another cleavage half-domain to produce an engineered cleavage half-domain designated "Q486E:I499L". The engineered cleavage half-domains described herein are obligate heterodimer mutants in which aberrant cleavage is minimized or abolished. *See, e.g.,* U.S. Patent Nos. 7,914,796 and 8,034,598, the disclosures of which are incorporated by reference in their entireties for all purposes. In certain embodiments, the engineered cleavage half-domain comprises mutations at positions 486, 499 and 496 (numbered relative to wild-type FokI), for instance mutations that replace the wild type Gln (Q) residue at position 486 with a Glu (E) residue, the wild type Iso (I) residue at position 499 with a Leu (L) residue and the wild-type Asn (N) residue at position 496 with an Asp (D) or Glu (E) residue (also referred to as a "ELD" and "ELE" domains, respectively). In other embodiments, the engineered cleavage half-domain comprises mutations at positions 490, 538 and 537 (numbered relative to wild-type FokI), for instance mutations that replace the wild type Glu (E) residue at position 490 with a Lys (K) residue, the wild type Iso (I) residue at position 538 with a Lys (K) residue, and the wild-type His (H) residue at position 537 with a Lys (K) residue or a Arg (R) residue (also referred to as "KKK" and "KKR" domains, respectively). In other embodiments, the engineered cleavage half-domain comprises mutations at positions 490 and 537 (numbered relative to wild-type FokI), for instance mutations that replace the wild type Glu (E) residue at position 490 with a Lys (K) residue and the wild-type His (H) residue at position 537 with a Lys (K) residue or a Arg (R) residue (also referred to as "KIK" and "KIR" domains, respectively). *See, e.g.,* U.S. Patent No. 8,772,453. In other embodiments, the engineered cleavage half domain

comprises the "Sharkey" and/or "Sharkey' " mutations (see Guo *et al*, (2010) *J. Mol. Biol.* 400(1):96-107).

[0141] Engineered cleavage half-domains described herein can be prepared using any suitable method, for example, by site-directed mutagenesis of wild-type cleavage half-domains (Fok I) as described in U.S. Patent Nos. 8,772,453; 8,623,618; 8,409,861; 8,034,598; 7,914,796; and 7,888,121.

[0142] Alternatively, nucleases may be assembled *in vivo* at the nucleic acid target site using so-called "split-enzyme" technology (*see e.g.* U.S. Patent Publication No. 20090068164). Components of such split enzymes may be expressed either on separate expression constructs, or can be linked in one open reading frame where the individual components are separated, for example, by a self-cleaving 2A peptide or IRES sequence. Components may be individual zinc finger binding domains or domains of a meganuclease nucleic acid binding domain.

[0143] Nucleases can be screened for activity prior to use, for example in a yeast-based chromosomal system as described in U.S. Patent No. 8,563,314.

[0144] Expression of the nuclease may be under the control of a constitutive promoter or an inducible promoter, for example the galactokinase promoter which is activated (de-repressed) in the presence of raffinose and/or galactose and repressed in presence of glucose.

[0145] The Cas9 related CRISPR/Cas system comprises two RNA non-coding components: tracrRNA and a pre-crRNA array containing nuclease guide sequences (spacers) interspaced by identical direct repeats (DRs). To use a CRISPR/Cas system to accomplish genome engineering, both functions of these RNAs must be present (see Cong *et al*, (2013) *Sciencexpress* 1/10.1126/science 1231143). In some embodiments, the tracrRNA and pre-crRNAs are supplied via separate expression constructs or as separate RNAs. In other embodiments, a chimeric RNA is constructed where an engineered mature crRNA (conferring target specificity) is fused to a tracrRNA (supplying interaction with the Cas9) to create a chimeric cr-RNA-tracrRNA hybrid (also termed a single guide RNA). (see Jinek *ibid* and Cong, *ibid*).

Target Sites

[0146] As described in detail above, DNA domains can be engineered to bind to any sequence of choice. An engineered DNA-binding domain can have a novel

binding specificity, compared to a naturally-occurring DNA-binding domain. Engineering methods include, but are not limited to, rational design and various types of selection. Rational design includes, for example, using databases comprising triplet (or quadruplet) nucleotide sequences and individual zinc finger amino acid sequences, in which each triplet or quadruplet nucleotide sequence is associated with one or more amino acid sequences of zinc fingers which bind the particular triplet or quadruplet sequence. *See*, for example, co-owned U.S. Patents 6,453,242 and 6,534,261, incorporated by reference herein in their entireties. Rational design of TAL-effector domains can also be performed. *See, e.g.,* U.S. Patent No. 8,586,526.

[0147] Exemplary selection methods applicable to DNA-binding domains, including phage display and two-hybrid systems, are disclosed in U.S. Patents 5,789,538; 5,925,523; 6,007,988; 6,013,453; 6,410,248; 6,140,466; 6,200,759; and 6,242,568; as well as WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 and GB 2,338,237. In addition, enhancement of binding specificity for zinc finger binding domains has been described, for example, in co-owned WO 02/077227.

[0148] Selection of target sites; nucleases and methods for design and construction of fusion proteins (and polynucleotides encoding same) are known to those of skill in the art and described in detail in U.S. Patent Application Publication Nos. 20050064474 and 20060188987, incorporated by reference in their entireties herein.

[0149] In addition, as disclosed in these and other references, DNA-binding domains (*e.g.,* multi-fingered zinc finger proteins) may be linked together using any suitable linker sequences, including for example, linkers of 5 or more amino acids. *See*, *e.g.*, U.S. Patent Nos. 6,479,626; 6,903,185; and 7,153,949 for exemplary linker sequences 6 or more amino acids in length. The proteins described herein may include any combination of suitable linkers between the individual DNA-binding domains of the protein. *See, also*, U.S. Patent No. 8,586,526.

[0150] Non-limiting examples of suitable target genes include a beta (β) globin gene (HBB), a gamma (δ) globin gene (HBG1), a B-cell lymphoma/leukemia 11A (BCL11A) gene, a Kruppel-like factor 1 (KLF1) gene, a CCR5 gene, a CXCR4 gene, a PPP1R12C (AAVS1) gene, an hypoxanthine phosphoribosyltransferase (HPRT) gene, an albumin gene, a Factor VIII gene, a Factor IX gene, a Leucine-rich repeat kinase 2 (LRRK2) gene, a Hungtingin (Htt) gene, a rhodopsin (RHO) gene, a Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) gene, a surfactant

protein B gene (SFTPB), a T-cell receptor alpha (TRAC) gene, a T-cell receptor beta (TRBC) gene, a programmed cell death 1 (PD1) gene, a Cytotoxic T-Lymphocyte Antigen 4 (CTLA-4) gene, an human leukocyte antigen (HLA) A gene, an HLA B gene, an HLA C gene, an HLA-DPA gene, an HLA-DQ gene, an HLA-DRA gene, a LMP7 gene, a Transporter associated with Antigen Processing (TAP) 1 gene, a TAP2 gene, a tapasin gene (TAPBP), a class II major histocompatibility complex transactivator (CIITA) gene, a dystrophin gene (DMD), a glucocorticoid receptor gene (GR), an IL2RG gene, a Rag-1 gene, an RFX5 gene, a FAD2 gene, a FAD3 gene, a ZP15 gene, a KASII gene, a MDH gene, and/or an EPSPS gene. In some aspects, the nuclease(s) binds to and/or cleaves a check point inhibitor gene, for example PD-1, CTLA4, receptors for the B7 family of inhibitory ligands, or cleaves a receptor or ligand gene involved in signaling through LAG3, 2B4, BTLA, TIM3, A2aR, and killer inhibitor receptors (KIRs and C-type lectin receptors), see Pardoll (2012) *Nat Rev Cancer* 12(4):252. *See, also,* U.S. Patent Nos. 8,956,828 and 8,945,868 and U.S. Patent Publication No. 20140120622 and 20150056705.

[0151] In other aspects, the nuclease(s) binds to and/or cleaves a gene that encodes a factor involved in rejection, for example, genes encoding subunits of the HLA complex (class I: HLA-A, HLA-B, HLA-C, HLA-E, HLA-F, HLA-G, B2M; class II: HLA-DMA, HLA-DOA, HLA-DPA1, HLA-DQA, HLA-DRA, HLA-DMB, HLA-DOB, HLA-DPB1, HLA-DQB, HLA-DRB) or TCR. In some embodiments, the nuclease(s) target a gene encoding a product involved in the peptide loading process and antigen processing for the HLA complexes (e.g. TAP, tapasin, calreticulin, calnexin, LMP2, LMP7 or Erp57). *See, e.g.,* U.S. Patent No. 8,956,828 and 8,945,868.

[0152] In certain embodiments, the nuclease targets a "safe harbor" loci such as the AAVS1, HPRT, albumin and CCR5 genes in human cells, and Rosa26 in murine cells (*see, e.g.,* U.S. Patent Nos. 7,888,121; 7,972,854; 7,914,796; 7,951,925; 8,110,379; 8,409,861; 8,586,526; U.S. Patent Publications 20030232410; 20050208489; 20050026157; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 and 20130177960) and the Zp15 locus in plants (*see* U.S. Patent No. 8,329,986).

Donors

[0153] The present disclosure relates to nuclease-mediated targeted integration of an exogenous sequence into the genome of an HSC/PC. As noted above, insertion of an exogenous sequence (also called a "donor sequence" or "donor" or "transgene"), for example for correction of a mutant gene or for increased expression of a wild-type gene or for expression of a transgene. It will be readily apparent that the donor sequence is typically not identical to the genomic sequence where it is placed. A donor sequence can contain a non-homologous sequence flanked by two regions of homology to allow for efficient HDR at the location of interest. Additionally, donor sequences can comprise a vector molecule containing sequences that are not homologous to the region of interest in cellular chromatin. A donor molecule can contain several, discontinuous regions of homology to cellular chromatin. For example, for targeted insertion of sequences not normally present in a region of interest, said sequences can be present in a donor nucleic acid molecule and flanked by regions of homology to sequence in the region of interest.

[0154] Described herein are methods of targeted insertion of any polynucleotides for insertion into a chosen location. Polynucleotides for insertion can also be referred to as "exogenous" polynucleotides, "donor" polynucleotides or molecules or "transgenes." The donor polynucleotide can be DNA, single-stranded and/or double-stranded and can be introduced into a cell in linear or circular form. *See, e.g.*, U.S. Patent Publication Nos. 20100047805 and 20110207221. The donor sequence(s) can also be introduced in DNA MC form, which may be introduced into the cell in circular or linear form. If introduced in linear form, the ends of the donor sequence can be protected (*e.g.*, from exonucleolytic degradation) by methods known to those of skill in the art. For example, one or more dideoxynucleotide residues are added to the 3' terminus of a linear molecule and/or self-complementary oligonucleotides are ligated to one or both ends. *See*, for example, Chang *et al.* (1987) *Proc. Natl. Acad. Sci. USA* **84:**4959-4963; Nehls *et al.* (1996) *Science* **272:**886-889. Additional methods for protecting exogenous polynucleotides from degradation include, but are not limited to, addition of terminal amino group(s) and the use of modified internucleotide linkages such as, for example, phosphorothioates, phosphoramidates, and O-methyl ribose or deoxyribose residues. If introduced in double-stranded form, the donor may include one or more nuclease target sites, for

example, nuclease target sites flanking the transgene to be integrated into the cell's genome. *See, e.g.,* U.S. Patent Publication No. 20130326645.

[0155] A polynucleotide can be introduced into a cell as part of a vector molecule having additional sequences such as, for example, replication origins, promoters and genes encoding antibiotic resistance. Moreover, donor polynucleotides can be introduced as naked nucleic acid, as nucleic acid complexed with an agent such as a liposome, nanoparticle or poloxamer, or can be delivered by viruses (*e.g.*, adenovirus, AAV, herpesvirus, retrovirus, lentivirus and integrase defective lentivirus (IDLV)).

[0156] In certain embodiments, the double-stranded donor includes sequences (*e.g.,* coding sequences, also referred to as transgenes) greater than 1 kb in length, for example between 2 and 200 kb, between 2 and 10 kb (or any value therebetween). The double-stranded donor also includes at least one nuclease target site, for example. In certain embodiments, the donor includes at least 2 target sites, for example for a pair of ZFNs or TALENs. Typically, the nuclease target sites are outside the transgene sequences, for example, 5' and/or 3' to the transgene sequences, for cleavage of the transgene. The nuclease cleavage site(s) may be for any nuclease(s). In certain embodiments, the nuclease target site(s) contained in the double-stranded donor are for the same nuclease(s) used to cleave the endogenous target into which the cleaved donor is integrated via homology-independent methods.

[0157] The donor is generally inserted so that its expression is driven by the endogenous promoter at the integration site, namely the promoter that drives expression of the endogenous gene into which the donor is inserted (*e.g.,* globin, AAVS1, etc.). However, it will be apparent that the donor may comprise a promoter and/or enhancer, for example a constitutive promoter or an inducible or tissue specific promoter.

[0158] The donor molecule may be inserted into an endogenous gene such that all, some or none of the endogenous gene is expressed. In other embodiments, the transgene (*e.g.,* with or without peptide- encoding sequences) is integrated into any endogenous locus, for example a safe-harbor locus. *See, e.g.,* US patent publications 20080299580; 20080159996 and 201000218264.

[0159] Furthermore, although not required for expression, exogenous sequences may also include transcriptional or translational regulatory sequences, for example, promoters, enhancers, insulators, internal ribosome entry sites, sequences

encoding 2A peptides and/or polyadenylation signals. Additionally, splice acceptor sequences may be included. Exemplary splice acceptor site sequences are known to those of skill in the art and include, by way of example only, CTGACCTCTTCTCTTCCTCCCACAG, (SEQ ID NO:1) (from the human *HBB* gene) and TTTCTCTCCACAG (SEQ ID NO:2) (from the human Immunoglobulin-gamma gene).

[0160] The transgenes carried on the donor sequences described herein may be isolated from plasmids, cells or other sources using standard techniques known in the art such as PCR. Donors for use can include varying types of topology, including circular supercoiled, circular relaxed, linear and the like. Alternatively, they may be chemically synthesized using standard oligonucleotide synthesis techniques. In addition, donors may be methylated or lack methylation. Donors may be in the form of bacterial or yeast artificial chromosomes (BACs or YACs).

[0161] The double-stranded donor polynucleotides described herein may include one or more non-natural bases and/or backbones. In particular, insertion of a donor molecule with methylated cytosines may be carried out using the methods described herein to achieve a state of transcriptional quiescence in a region of interest.

[0162] The exogenous (donor) polynucleotide may comprise any sequence of interest (exogenous sequence). Exemplary exogenous sequences include, but are not limited to any polypeptide coding sequence (*e.g.*, cDNAs or fragments thereof), promoter sequences, enhancer sequences, epitope tags, marker genes, cleavage enzyme recognition sites and various types of expression constructs. Marker genes include, but are not limited to, sequences encoding proteins that mediate antibiotic resistance (*e.g.*, ampicillin resistance, neomycin resistance, G418 resistance, puromycin resistance), sequences encoding colored or fluorescent or luminescent proteins (*e.g.*, green fluorescent protein, enhanced green fluorescent protein, red fluorescent protein, luciferase), and proteins which mediate enhanced cell growth and/or gene amplification (*e.g.*, dihydrofolate reductase). Epitope tags include, for example, one or more copies of FLAG, His, myc, Tap, HA or any detectable amino acid sequence.

[0163] In a preferred embodiment, the exogenous sequence (transgene) comprises a polynucleotide encoding any polypeptide of which expression in the cell is desired, including, but not limited to antibodies, antigens, enzymes, receptors (cell surface or nuclear), hormones, lymphokines, cytokines, reporter polypeptides, growth

factors, and functional fragments of any of the above. The coding sequences may be, for example, cDNAs. The exogenous sequences may also be a fragment of a transgene for linking with an endogenous gene sequence of interest. For example, a fragment of a transgene comprising sequence at the 3' end of a gene of interest may be utilized to correct, via insertion or replacement, of a sequence encoding a mutation in the 3' end of an endogenous gene sequence. Similarly, the fragment may comprise sequences similar to the 5' end of the endogenous gene for insertion/replacement of the endogenous sequences to correct or modify such endogenous sequence. Additionally the fragment may encode a functional domain of interest (catalytic, secretory or the like) for linking *in situ* to an endogenous gene sequence to produce a fusion protein.

[0164] For example, the exogenous sequence may comprise a sequence encoding a polypeptide that is lacking or non-functional in the subject having a genetic disease, including but not limited to any of the following genetic diseases: achondroplasia, achromatopsia, acid maltase deficiency, adenosine deaminase deficiency (OMIM No.102700), adrenoleukodystrophy, aicardi syndrome, alpha-1 antitrypsin deficiency, alpha-thalassemia, androgen insensitivity syndrome, apert syndrome, arrhythmogenic right ventricular, dysplasia, ataxia telangictasia, barth syndrome, beta-thalassemia, blue rubber bleb nevus syndrome, canavan disease, chronic granulomatous diseases (CGD), cri du chat syndrome, cystic fibrosis, dercum's disease, ectodermal dysplasia, fanconi anemia, fibrodysplasia ossificans progressive, fragile X syndrome, galactosemis, Gaucher's disease, generalized gangliosidoses (*e.g.*, GM1), hemochromatosis, the hemoglobin C mutation in the 6th codon of beta-globin (HbC), hemophilia, Huntington's disease, Hurler Syndrome, hypophosphatasia, Klinefleter syndrome, Krabbes Disease, Langer-Giedion Syndrome, leukocyte adhesion deficiency (LAD, OMIM No. 116920), leukodystrophy, long QT syndrome, Marfan syndrome, Moebius syndrome, mucopolysaccharidosis (MPS), nail patella syndrome, nephrogenic diabetes insipdius, neurofibromatosis, Neimann-Pick disease, osteogenesis imperfecta, porphyria, Prader-Willi syndrome, progeria, Proteus syndrome, retinoblastoma, Rett syndrome, Rubinstein-Taybi syndrome, Sanfilippo syndrome, severe combined immunodeficiency (SCID), Shwachman syndrome, sickle cell disease (sickle cell anemia), Smith-Magenis syndrome, Stickler syndrome, Tay-Sachs disease, Thrombocytopenia Absent Radius (TAR) syndrome, Treacher Collins syndrome,

trisomy, tuberous sclerosis, Turner's syndrome, urea cycle disorder, von Hippel-Landau disease, Waardenburg syndrome, Williams syndrome, Wilson's disease, Wiskott-Aldrich syndrome, X-linked lymphoproliferative syndrome (XLP, OMIM No. 308240).

[0165] Additional exemplary diseases that can be treated by targeted integration include acquired immunodeficiencies, lysosomal storage diseases (*e.g.*, Gaucher's disease, GM1, Fabry disease and Tay-Sachs disease), mucopolysaccahidosis (*e.g.* Hunter's disease, Hurler's disease), hemoglobinopathies (*e.g.*, sickle cell diseases, HbC, α-thalassemia, β-thalassemia) and hemophilias.

[0166] In certain embodiments, the exogenous sequences can comprise a marker gene (described above), allowing selection of cells that have undergone targeted integration, and a linked sequence encoding an additional functionality. Non-limiting examples of marker genes include GFP, drug selection marker(s) and the like.

[0167] Additional gene sequences that can be inserted may include, for example, wild-type genes to replace mutated sequences. For example, a wild-type Factor IX gene sequence may be inserted into the genome of a stem cell in which the endogenous copy of the gene is mutated. The wild-type copy may be inserted at the endogenous locus, or may alternatively be targeted to a safe harbor locus.

[0168] In some embodiments, the donor sequence encodes a receptor that serves to direct the function of a T cell. Chimeric Antigen Receptors (CARs) are molecules designed to target immune cells to specific molecular targets expressed on cell surfaces. In their most basic form, they are receptors introduced to a cell that couple a specificity domain expressed on the outside of the cell to signaling pathways on the inside of the cell such that when the specificity domain interacts with its target, the cell becomes activated. Often CARs are made from variants of T-cell receptors (TCRs) where a specificity domain such as a scFv or some type of receptor is fused to the signaling domain of a TCR. These constructs are then introduced into a T cell allowing the T cell to become activated in the presence of a cell expressing the target antigen, resulting in the attack on the targeted cell by the activated T cell in a non-MHC dependent manner (see Chicaybam *et al* (2011) *Int Rev Immunol* 30:294-311). Alternatively, CAR expression cassettes can be introduced into an HSC/PC for later engraftment such that the CAR cassette is under the control of a T cell specific

promoter (*e.g.*, the FOXP3 promoter, see Mantel *et al* (2006) *J. Immunol* 176: 3593-3602).

[0169] Currently, tumor specific CARs targeting a variety of tumor antigens are being tested in the clinic for treatment of a variety of different cancers. Examples of these cancers and their antigens that are being targeted includes follicular lymphoma (CD20 or GD2), neuroblastoma (CD171), non-Hodgkin lymphoma (CD20), lymphoma (CD19), glioblastoma (IL13Rα2), chronic lymphocytic leukemia or CLL and acute lymphocytic leukemia or ALL (both CD19).

[0170] Virus specific CARs have also been developed to attack cells harboring virus such as HIV. For example, a clinical trial was initiated using a CAR specific for Gp100 for treatment of HIV (Chicaybam, ibid). Other virus specific CARs could be developed to target Ebola to knock out cells harboring the Ebola virus, or CAR containing T cells can be used post-transplant to target CMV, adenovirus and/or EBV using engineered T cells comprising CARs specific for these viruses.

[0171] CARs are also being developed for the treatment of autoimmune disease. Regulatory T cells (Tregs) are a subset of CD4+ T cells that constiutively express the IL-2 receptor alpha chain and the transcription factor FoxP3. Tregs are believed to suppress colitis, for example, by inhibiting effector T cell proliferation and the production of proinflammatory cytokines, as well as hindering components of the innate immune system. Due to the scarcity of T regs to a specific antigen, researchers are exploring the potential of adoptive transfer of engineered antigen specific Tregs where the T cells have been modified to express a CAR against an antigen associated with autoimmunity (*e.g*, CEA for colitis, *see* Blat *et al* (2014) *Mol Ther* 22(5): 1018).

[0172] The T cell receptor (TCR) is an essential part of the selective activation of T cells and is typically made from two chains, α and β, which co-assemble to form a heterodimer. The genomic loci that encode the TCR chains resemble antibody encoding loci in that the TCR α gene comprises V and J segments, while the β chain locus comprises D segments in addition to V and J segments. Additionally, the TCR complex makes up part of the CD3 antigen complex on T cells. During T cell activation, the TCR interacts with antigens displayed on the major histocompatability complex (MHC) of an antigen presenting cell. Recognition of the antigen-MHC complex by the TCR leads to T cell stimulation, which in turn leads to differentiation of both T helper cells (CD4+) and cytotoxic T lymphocytes (CD8+) in memory and

effector lymphocytes. Thus, use of an engineered TCR can also lead to alter the direction of T cell activity (see U.S. Patent No. 8,956,828), and the donor sequence can comprise engineered sequences encoding a TCR. In some embodiments, the endogenous TCR is also disrupted through use an engineered nuclease designed to cleave a gene encoding an endogenous TCR subunit (*i.e.* TRAC or TRBC). *See, e.g.*, U.S. Patent No. 8,956,828 and 8,945,868.

[0173] Antibody coupled-T cell receptor (ACTR) technology is the use of a single species of T cell comprising the ACTR molecule that is combined with a variety of different antibodies to direct the T cells. An ACTR is a membrane spanning protein composed of a CD3ζ signaling domain, a 4-1BB co-stimulatory domain, a CD8 membrane spanning and hinge domain, and a CD16 Fc receptor domain. This protein is expressed in T cells and then is combined *in vivo* with an antibody such that the antibody associates with the T cell expressing the ACTR, and the antibody-ACTR complex serves to target the TCR to a specific cell or tissue that expresses the specific antigen that the antibody binds to. In some embodiments, the T cell to be modified with the ACTR is derived from the patient. In other embodiments, the T cell is a 'universal' T cell where specific endogenous T cell receptors (*e.g.* TCR, MHC) have been inactivated. In some cases, the receptors are inactivated with an engineered nuclease, creating a bank of T cells that will not react with host antigens in the absence of the ACTR-antibody complex.

[0174] Construction of such expression cassettes, following the teachings of the present specification, utilizes methodologies well known in the art of molecular biology (see, for example, Ausubel or Maniatis). Before use of the expression cassette to generate a transgenic animal, the responsiveness of the expression cassette to the stress-inducer associated with selected control elements can be tested by introducing the expression cassette into a suitable cell line (*e.g.*, primary cells, transformed cells, or immortalized cell lines).

[0175] Furthermore, although not required for expression, exogenous sequences may also transcriptional or translational regulatory sequences, for example, promoters, enhancers, insulators, internal ribosome entry sites, sequences encoding 2A peptides and/or polyadenylation signals. Further, the control elements of the genes of interest can be operably linked to reporter genes to create chimeric genes (*e.g.*, reporter expression cassettes).

[0176] Targeted insertion of non-coding nucleic acid sequence may also be achieved. Sequences encoding antisense RNAs, RNAi, shRNAs and micro RNAs (miRNAs) may also be used for targeted insertions.

[0177] In additional embodiments, the donor nucleic acid may comprise non-coding sequences that are specific target sites for additional nuclease designs. Subsequently, additional nucleases may be expressed in cells such that the original donor molecule is cleaved and modified by insertion of another donor molecule of interest. In this way, reiterative integrations of donor molecules may be generated allowing for trait stacking at a particular locus of interest or at a safe harbor locus.

[0178] The donor(s) may be delivered prior to, simultaneously or after the nuclease(s) is(are) introduced into a cell. In certain embodiments, the donor(s) are delivered simultaneously with the nuclease(s). In other embodiments, the donors are delivered prior to the nuclease(s), for example, seconds to hours to days before the donors, including, but not limited to, 1 to 60 minutes (or any time therebetween) before the nuclease(s), 1 to 24 hours (or any time therebetween) before the nuclease(s) or more than 24 hours before the nuclease(s). In certain embodiments, the donor is delivered after the nuclease, preferably within 4 hours.

[0179] The donors may be delivered using the same delivery systems as the nuclease(s). When delivered simultaneously, the donors and nucleases may be on the same vector, for example an AAV vector (*e.g.*, AAV6). In certain embodiments, the donors are delivered using an AAV vector and the nuclease(s) are delivered in mRNA form.

Cells

[0180] Thus, provided herein are genetically modified cells, for example primary HSC/PC or T cells comprising a transgene, including a transgene that expresses a functional protein in the cell. Cells produced by the methods described herein are also provided. The transgene is integrated in a targeted manner into the cell's genome using one or more nucleases. In certain embodiments, the transgene is integrated into a safe harbor gene.

[0181] Unlike random integration, targeted integration ensures that the transgene is integrated into a specified gene or locus. The transgene may be integrated anywhere in the target gene. In certain embodiments, the transgene is integrated at or near the nuclease cleavage site, for example, within 1-300 (or any

value therebetwen) base pairs upstream or downstream of the site of cleavage, more preferably within 1-100 base pairs (or any value therebetween) of either side of the cleavage site, even more preferably within 1 to 50 base pairs (or any value therebetween) of either side of the cleavage site. In certain embodiments, the integrated sequence comprising the transgene does not include any vector sequences (*e.g.*, viral vector sequences).

[0182] Any cell type can be genetically modified as described herein, including but not limited to cells and cell lines. Other non-limiting examples of cells as described herein include T-cells (*e.g.*, CD4+, CD3+, CD8+ (including Tregs), etc.); dendritic cells; B-cells; autologous (*e.g.*, patient-derived) or heterologous pluripotent, totipotent or multipotent stem cells (*e.g.*, CD34+ cells, induced pluripotent stem cells (iPSCs), embryonic stem cells or the like). In certain embodiments, the cells as described herein are CD34+ cells derived from a patient with a disorder it is desired to treat.

[0183] The cells as described herein are useful in treating and/or preventing a disorder, for example, by *ex vivo* therapies. The nuclease-modified cells can be expanded and then reintroduced into the patient using standard techniques. *See, e.g.*, Tebas *et al* (2014) *New Eng J Med* 370(10):901. In the case of stem cells, after infusion into the subject, *in vivo* differentiation of these precursors into cells expressing the functional transgene also occurs. Pharmaceutical compositions comprising the cells as described herein are also provided. In addition, the cells may be cryopreserved prior to administration to a patient.

Delivery

[0184] The nucleases, polynucleotides encoding these nucleases, donor polynucleotides and compositions comprising the proteins and/or polynucleotides described herein may be delivered *in vivo* or *ex vivo* by any suitable means into any cell type.

[0185] Suitable cells include eukaryotic (*e.g.*, animal) and prokaryotic cells and/or cell lines. Non-limiting examples of such cells or cell lines generated from such cells include COS, CHO (*e.g.*, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (*e.g.*, HEK293-F, HEK293-H, HEK293-T), and

perC6 cells as well as insect cells such as *Spodoptera fugiperda* (Sf), or fungal cells such as *Saccharomyces*, *Pichia* and *Schizosaccharomyces*. In certain embodiments, the cell line is a CHO, MDCK or HEK293 cell line. Suitable cells also include stem cells such as, by way of example, embryonic stem cells, induced pluripotent stem cells, hematopoietic stem cells, neuronal stem cells and mesenchymal stem cells.

[0186] Methods of delivering nucleases as described herein are described, for example, in U.S. Patent Nos. 6,453,242; 6,503,717; 6,534,261; 6,599,692; 6,607,882; 6,689,558; 6,824,978; 6,933,113; 6,979,539; 7,013,219; and 7,163,824, the disclosures of all of which are incorporated by reference herein in their entireties.

[0187] Nucleases and/or donor constructs as described herein may also be delivered using vectors containing sequences encoding one or more of the ZFN(s), TALEN(s) or CRIPSR/Cas systems. Any vector systems may be used including, but not limited to, plasmid vectors, retroviral vectors, lentiviral vectors, adenovirus vectors, poxvirus vectors; herpesvirus vectors and adeno-associated virus vectors, etc. *See, also,* U.S. Patent Nos. 6,534,261; 6,607,882; 6,824,978; 6,933,113; 6,979,539; 7,013,219; and 7,163,824, incorporated by reference herein in their entireties. Furthermore, it will be apparent that any of these vectors may comprise one or more of the sequences needed for treatment. Thus, when one or more nucleases and a donor construct are introduced into the cell, the nucleases and/or donor polynucleotide may be carried on the same vector or on different vectors (DNA MC(s)). When multiple vectors are used, each vector may comprise a sequence encoding one or multiple nucleases and/or donor constructs.

[0188] Conventional viral and non-viral based gene transfer methods can be used to introduce nucleic acids encoding nucleases and donor constructs in cells (*e.g.*, mammalian cells) and target tissues. Non-viral vector delivery systems include DNA or RNA plasmids, DNA MCs, naked nucleic acid, and nucleic acid complexed with a delivery vehicle such as a liposome, nanoparticle or poloxamer. Viral vector delivery systems include DNA and RNA viruses, which have either episomal or integrated genomes after delivery to the cell. For a review of *in vivo* delivery of engineered DNA-binding proteins and fusion proteins comprising these binding proteins, see, *e.g.*, Rebar (2004) *Expert Opinion Invest. Drugs* 13(7):829-839; Rossi *et al.* (2007) *Nature Biotech.* 25(12):1444-1454 as well as general gene delivery references such as Anderson, *Science* 256:808-813 (1992); Nabel & Felgner, *TIBTECH* 11:211-217 (1993); Mitani & Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-

175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10):1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada *et al.*, in *Current Topics in Microbiology and Immunology* Doerfler and Böhm (eds.) (1995); and Yu *et al., Gene Therapy* 1:13-26 (1994).

[0189] Methods of non-viral delivery of nucleic acids include electroporation, lipofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, other nanoparticle, polycation or lipid:nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Sonoporation using, *e.g.*, the Sonitron 2000 system (Rich-Mar) can also be used for delivery of nucleic acids.

[0190] Additional exemplary nucleic acid delivery systems include those provided by Amaxa Biosystems (Cologne, Germany), Maxcyte, Inc. (Rockville, Maryland), BTX Molecular Delivery Systems (Holliston, MA) and Copernicus Therapeutics Inc, (*see* for example US6008336). Lipofection is described in *e.g.*, U.S. Patent Nos. 5,049,386; 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (*e.g.*, Transfectam™ and Lipofectin™). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those of Felgner, WO 91/17424, WO 91/16024.

[0191] The preparation of lipid:nucleic acid complexes, including targeted liposomes such as immunolipid complexes, is well known to one of skill in the art (*see, e.g.*, Crystal, *Science* 270:404-410 (1995); Blaese *et al., Cancer Gene Ther.* 2:291-297 (1995); Behr *et al., Bioconjugate Chem.* 5:382-389 (1994); Remy *et al., Bioconjugate Chem.* 5:647-654 (1994); Gao *et al., Gene Therapy* 2:710-722 (1995); Ahmad *et al., Cancer Res.* 52:4817-4820 (1992); U.S. Pat. Nos. 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, and 4,946,787).

[0192] Additional methods of delivery include the use of packaging the nucleic acids to be delivered into EnGeneIC delivery vehicles (EDVs). These EDVs are specifically delivered to target tissues using bispecific antibodies where one arm of the antibody has specificity for the target tissue and the other has specificity for the EDV. The antibody brings the EDVs to the target cell surface and then the EDV is brought into the cell by endocytosis. Once in the cell, the contents are released (*see* MacDiarmid *et al* (2009) *Nature Biotechnology* 27(7):643).

[0193] The use of RNA or DNA viral based systems for the delivery of nucleic acids encoding engineered ZFPs, TALEs and/or CRISPR/Cas systems take

advantage of highly evolved processes for targeting a virus to specific cells in the body and trafficking the viral payload to the nucleus. Viral vectors can be administered directly to patients (*in vivo*) or they can be used to treat cells *in vitro* and the modified cells are administered to patients (*ex vivo*). Conventional viral based systems for the delivery of ZFPs include, but are not limited to, retroviral, lentivirus, adenoviral, adeno-associated, vaccinia and herpes simplex virus vectors for gene transfer. Integration in the host genome is possible with the retrovirus, lentivirus, and adeno-associated virus gene transfer methods, often resulting in long term expression of the inserted transgene. Additionally, high transduction efficiencies have been observed in many different cell types and target tissues.

[0194] The tropism of a retrovirus can be altered by incorporating foreign envelope proteins, expanding the potential target population of target cells. Lentiviral vectors are retroviral vectors that are able to transduce or infect non-dividing cells and typically produce high viral titers. Selection of a retroviral gene transfer system depends on the target tissue. Retroviral vectors are comprised of *cis*-acting long terminal repeats with packaging capacity for up to 6-10 kb of foreign sequence. The minimum *cis*-acting LTRs are sufficient for replication and packaging of the vectors, which are then used to integrate the therapeutic gene into the target cell to provide permanent transgene expression. Widely used retroviral vectors include those based upon murine leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immunodeficiency virus (SIV), human immunodeficiency virus (HIV), and combinations thereof (*see, e.g.,* Buchscher *et al.*, *J. Virol.* 66:2731-2739 (1992); Johann *et al., J. Virol.* 66:1635-1640 (1992); Sommerfelt *et al.*, *Virol.* 176:58-59 (1990); Wilson *et al.*, *J. Virol.* 63:2374-2378 (1989); Miller *et al.*, *J. Virol.* 65:2220-2224 (1991); PCT/US94/05700).

[0195] In applications in which transient expression is preferred, adenoviral based systems can be used. Adenoviral based vectors are capable of very high transduction efficiency in many cell types and do not require cell division. With such vectors, high titer and high levels of expression have been obtained. This vector can be produced in large quantities in a relatively simple system. Adeno-associated virus ("AAV") vectors are also used to transduce cells with target nucleic acids, *e.g.*, in the *in vitro* production of nucleic acids and peptides, and for *in vivo* and *ex vivo* gene therapy procedures (*see, e.g.*, West *et al.*, *Virology* 160:38-47 (1987); U.S. Patent No. 4,797,368; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994);

Muzyczka, *J. Clin. Invest.* 94:1351 (1994). Construction of recombinant AAV vectors are described in a number of publications, including U.S. Pat. No. 5,173,414; Tratschin *et al.*, *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, *et al.*, *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984); and Samulski *et al.*, *J. Virol.* 63:03822-3828 (1989).

[0196] At least six viral vector approaches are currently available for gene transfer in clinical trials, which utilize approaches that involve complementation of defective vectors by genes inserted into helper cell lines to generate the transducing agent.

[0197] pLASN and MFG-S are examples of retroviral vectors that have been used in clinical trials (Dunbar *et al., Blood* 85:3048-305 (1995); Kohn *et al., Nat. Med.* 1:1017-102 (1995); Malech *et al., PNAS* 94:22 12133-12138 (1997)). PA317/pLASN was the first therapeutic vector used in a gene therapy trial. (Blaese *et al., Science* 270:475-480 (1995)). Transduction efficiencies of 50% or greater have been observed for MFG-S packaged vectors. (Ellem *et al., Immunol Immunother.* 44(1):10-20 (1997); Dranoff *et al., Hum. Gene Ther.* 1:111-2 (1997).

[0198] Recombinant adeno-associated virus vectors (rAAV) are a promising alternative gene delivery systems based on the defective and nonpathogenic parvovirus adeno-associated type 2 virus. All vectors are derived from a plasmid that retains only the AAV 145 bp inverted terminal repeats flanking the transgene expression cassette. Efficient gene transfer and stable transgene delivery due to integration into the genomes of the transduced cell are key features for this vector system. (Wagner *et al., Lancet* 351:9117 1702-3 (1998), Kearns *et al., Gene Ther.* 9:748-55 (1996)). Other AAV serotypes, including AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 and AAVrh.10 and any novel AAV serotype can also be used in accordance with the present invention. In some embodiments, chimeric AAV is used where the viral origins of the LTR sequences of the viral nucleic acid are heterologous to the viral origin of the capsid sequences. Examples include chimeric virus with LTRs derived from AAV2 and capsids derived from AAV5, AAV6, AAV8 or AAV9 (i.e. AAV2/5, AAV2/6, AAV2/8 and AAV2/9, respectively).

[0199] Replication-deficient recombinant adenoviral vectors (Ad) can be produced at high titer and readily infect a number of different cell types. Most adenovirus vectors are engineered such that a transgene replaces the Ad E1a, E1b,

and/or E3 genes; subsequently the replication defective vector is propagated in human 293 cells that supply deleted gene function in *trans*. Ad vectors can transduce multiple types of tissues *in vivo*, including nondividing, differentiated cells such as those found in liver, kidney and muscle. Conventional Ad vectors have a large carrying capacity. An example of the use of an Ad vector in a clinical trial involved polynucleotide therapy for antitumor immunization with intramuscular injection (Sterman *et al., Hum. Gene Ther.* 7:1083-9 (1998)). Additional examples of the use of adenovirus vectors for gene transfer in clinical trials include Rosenecker *et al., Infection* 24:1 5-10 (1996); Sterman *et al., Hum. Gene Ther.* 9:7 1083-1089 (1998); Welsh *et al., Hum. Gene Ther.* 2:205-18 (1995); Alvarez *et al., Hum. Gene Ther.* 5:597-613 (1997); Topf *et al., Gene Ther.* 5:507-513 (1998); Sterman *et al., Hum. Gene Ther.* 7:1083-1089 (1998).

[0200] Packaging cells are used to form virus particles that are capable of infecting a host cell. Such cells include 293 cells, which package AAV and adenovirus, and ψ2 cells or PA317 cells, which package retrovirus. Viral vectors used in gene therapy are usually generated by a producer cell line that packages a nucleic acid vector into a viral particle. The vectors typically contain the minimal viral sequences required for packaging and subsequent integration into a host (if applicable), other viral sequences being replaced by an expression cassette encoding the protein to be expressed. The missing viral functions are supplied in *trans* by the packaging cell line. For example, AAV vectors used in gene therapy typically only possess inverted terminal repeat (ITR) sequences from the AAV genome which are required for packaging and integration into the host genome. Viral DNA is packaged in a cell line, which contains a helper plasmid encoding the other AAV genes, namely *rep* and *cap*, but lacking ITR sequences. The cell line is also infected with adenovirus as a helper. The helper virus promotes replication of the AAV vector and expression of AAV genes from the helper plasmid. The helper plasmid is not packaged in significant amounts due to a lack of ITR sequences. Contamination with adenovirus can be reduced by, *e.g.*, heat treatment to which adenovirus is more sensitive than AAV. In some embodiments, AAV is produced using a baculovirus expression system.

[0201] In many gene therapy applications, it is desirable that the gene therapy vector be delivered with a high degree of specificity to a particular tissue type. Accordingly, a viral vector can be modified to have specificity for a given cell type by

expressing a ligand as a fusion protein with a viral coat protein on the outer surface of the virus. The ligand is chosen to have affinity for a receptor known to be present on the cell type of interest. For example, Han *et al.*, *Proc. Natl. Acad. Sci. USA* 92:9747-9751 (1995), reported that Moloney murine leukemia virus can be modified to express human heregulin fused to gp70, and the recombinant virus infects certain human breast cancer cells expressing human epidermal growth factor receptor. This principle can be extended to other virus-target cell pairs, in which the target cell expresses a receptor and the virus expresses a fusion protein comprising a ligand for the cell-surface receptor. For example, filamentous phage can be engineered to display antibody fragments (*e.g.*, FAB or Fv) having specific binding affinity for virtually any chosen cellular receptor. Although the above description applies primarily to viral vectors, the same principles can be applied to nonviral vectors. Such vectors can be engineered to contain specific uptake sequences which favor uptake by specific target cells.

[0202] Gene therapy vectors can be delivered *in vivo* by administration to an individual patient, typically by systemic administration (*e.g.*, intravenous, intraperitoneal, intramuscular, subdermal, or intracranial infusion) or topical application, as described below. Alternatively, vectors can be delivered to cells *ex vivo*, such as cells explanted from an individual patient (*e.g.*, lymphocytes, bone marrow aspirates, tissue biopsy) or universal donor hematopoietic stem cells, followed by reimplantation of the cells into a patient, usually after selection for cells which have incorporated the vector.

[0203] Vectors (*e.g.*, retroviruses, adenoviruses, liposomes, etc.) containing nucleases and/or donor constructs can also be administered directly to an organism for transduction of cells *in vivo*. Alternatively, naked DNA can be administered. Administration is by any of the routes normally used for introducing a molecule into ultimate contact with blood or tissue cells including, but not limited to, injection, infusion, topical application and electroporation. Suitable methods of administering such nucleic acids are available and well known to those of skill in the art, and, although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

[0204] Vectors suitable for introduction of polynucleotides (*e.g.* nuclease-encoding and/or double-stranded donors) described herein include non-integrating

lentivirus vectors (IDLV). *See*, for example, Ory *et al.* (1996) *Proc. Natl. Acad. Sci. USA* **93:**11382-11388; Dull *et al.* (1998) *J. Virol.* **72:**8463-8471; Zuffery *et al.* (1998) *J. Virol.* **72:**9873-9880; Follenzi *et al.* (2000) *Nature Genetics* **25:**217-222; U.S. Patent Publication No 2009/054985.

[0205] 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 available, as described below (*see, e.g., Remington's Pharmaceutical Sciences*, 17th ed., 1989).

[0206] It will be apparent that the nuclease-encoding sequences and donor constructs can be delivered using the same or different systems. For example, the nucleases and donors can be carried by the same DNA MC. Alternatively, a donor polynucleotide can be carried by a MC, while the one or more nucleases can be carried by a standard plasmid or AAV vector. Furthermore, the different vectors can be administered by the same or different routes (intramuscular injection, tail vein injection, other intravenous injection, intraperitoneal administration and/or intramuscular injection. The vectors can be delivered simultaneously or in any sequential order.

[0207] Thus, the instant disclosure includes *in vivo* or *ex vivo* treatment of diseases and conditions that are amenable to insertion of a transgenes encoding a therapeutic protein, for example treatment of hemophilias via nuclease-mediated integration of clotting factors such as Factor VIII (F8). The compositions are administered to a human patient in an amount effective to obtain the desired concentration of the therapeutic polypeptide in the serum or the target organ or cells. Administration can be by any means in which the polynucleotides are delivered to the desired target cells. For example, both *in vivo* and *ex vivo* methods are contemplated. Intravenous injection to the portal vein is a preferred method of administration. Other *in vivo* administration modes include, for example, direct injection into the lobes of the liver or the biliary duct and intravenous injection distal to the liver, including through the hepatic artery, direct injection in to the liver parenchyma, injection via the hepatic artery, and/or retrograde injection through the biliary tree. *Ex vivo* modes of administration include transduction *in vitro* of resected hepatocytes or other cells of the liver, followed by infusion of the transduced, resected hepatocytes back into the

portal vasculature, liver parenchyma or biliary tree of the human patient, see *e.g.*, Grossman *et al.*, (1994) *Nature Genetics*, 6:335-341.

[0208] The effective amount of nuclease(s) and donor to be administered will vary from patient to patient and according to the therapeutic polypeptide of interest. Accordingly, effective amounts are best determined by the physician administering the compositions and appropriate dosages can be determined readily by one of ordinary skill in the art. After allowing sufficient time for integration and expression (typically 4-15 days, for example), analysis of the serum or other tissue levels of the therapeutic polypeptide and comparison to the initial level prior to administration will determine whether the amount being administered is too low, within the right range or too high. Suitable regimes for initial and subsequent administrations are also variable, but are typified by an initial administration followed by subsequent administrations if necessary. Subsequent administrations may be administered at variable intervals, ranging from daily to annually to every several years. One of skill in the art will appreciate that appropriate immunosuppressive techniques may be recommended to avoid inhibition or blockage of transduction by immunosuppression of the delivery vectors, see *e.g.*, Vilquin *et al.*, (1995)*Human Gene Ther.*, 6:1391-1401.

[0209] Formulations for both *ex vivo* and *in vivo* administrations include suspensions in liquid or emulsified liquids. The active ingredients often are mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients include, for example, water, saline, dextrose, glycerol, ethanol or the like, and combinations thereof. In addition, the composition may contain minor amounts of auxiliary substances, such as, wetting or emulsifying agents, pH buffering agents, stabilizing agents or other reagents that enhance the effectiveness of the pharmaceutical composition.

[0210] The following Examples relate to exemplary embodiments of the present disclosure in which the nuclease comprises a zinc finger nuclease (ZFN), a TALEN or a CRISPR/Cas nuclease system. It will be appreciated that this is for purposes of exemplification only and that other nucleases can be used, for instance Ttago systems, homing endonucleases (meganucleases) with engineered DNA-binding domains and/or fusions of naturally occurring of engineered homing endonucleases (meganucleases) DNA-binding domains and heterologous cleavage domains and/or fusions of meganucleases and TALE proteins.

EXAMPLES

Example 1: **Assembly of Zinc Finger Nucleases**

[0211] ZFNs were assembled against the human PD1 genes and were tested for activity by ELISA and CEL1 assays as described in Miller *et al.* (2007) *Nat. Biotechnol.* 25:778-785. For TCR (*e.g.*, TRAC)-, B2M-, CTLA-4- and PD1-specific nucleases, see U.S. Patent Nos. 8,956,828; 8,945,868; 8,563,314 and U.S. Patent Publication Nos. 20140120622 and 20150056705, incorporated by reference herein.

Example 2: AAV transduction and gene modification TI

A. CD4+ or CD8+ T-cells

[0212] Primary CD4+ or CD8+ T cells are transduced with AAV6 vectors containing a promoter (*e.g.* EF1a or PGK) driving and engineered antigen receptor transgene (*e.g.* an CAR, TCR or ACTR as depicted in Figure 1) and mRNAs encoding PD1-specific ZFN or TRAC-specific ZFN in the presence of IL2 (20 ng/ml) and Dynabeads® Human T-Activator CD3/CD28 (Life Technology). Cells are then collected at 5 days post-infection (dpi) and analyzed for expression of the engineered antigen receptor. The results demonstrate that use of an AAV6 vector comprising an engineered antigen receptor transgene in combination with ZFNs specific for the PD1 checkpoint inhibitor or TCR alpha chain TRAC results in integration of the transgene into the PD1 or TRAC locus.

B. CD3+ T-cells

[0213] To transduce CD3+ T cells with AAV2/6 vectors carrying the CCR5-RFLP donor, CD3+ T-cells were exposed to AAV2/6-RFLP donor vectors for four hours under standard conditions (see PCT application PCT/US2015/041807). Standard conditions include the use of serum in the media. mRNA encoding CCR5-specific ZFN as described previously (DeKelver *et al* (2010) *Gen. Res.* 20:1133-1142) were introduced into the cells via electroporation at 60 μg/mL. To investigate the effect of serum has on the transduction efficiency, a comparison experiment was performed with no serum in the media during the four hour transduction. Following transduction and electroporation, the cells were expanded for five days under serum containing conditions, and then collected. Genomic DNA was isolated by Illumina deep sequencing by standard procedures to measure the efficiency of targeted integration of the RFLP containing donor ("% RFLP").

[0214] The results are depicted in Figure 2, and demonstrate that at low concentrations of AAV2/6 donor, targeted integration of the donor is inhibited in the presence of serum, which is not observed while serum is not present.

Example 3: AAV transduction and gene modification

[0215] AAV donor expression was also studied in cells in which one or more endogenous genes were modified (*e.g.,* inactivated with or without TI) using the compositions and methods and described herein. In particular, TRAC and/or B2M genes were modified essentially as described above in primary T-cells (CD3+). Briefly, nucleases targeted TRAC and/or B2M were administered to the cells along with AAV donors comprising a transgene encoding a GFP (250 ug/mL and AAV MOI of 1e5vg/cell for TRAC1 nucleases alone; 120 ug/mL and AAV MOI of 1e5vg/cell for B2M nucleases alone; 120 ug/mL for B2M and 60 ug/mL and AAV MOI of 1e5vg/cell for TRAC1 and B2M nucleases). T-cells were activated using anti-CD3/ CD28 beads, and cultured in media with serum replacement and IL-2. Two days post activation, activated cells were transduced with AAV6 GFP donor vectors (comprising homology arms specific to either the TRAC or B2M ZFN cut site) at 1E5 vg/cell. The next day, cells were transfected with mRNA encoding for ZFNs targeting either TRAC or B2M by electroporation with mRNA concentration ranged from 60 - 250 ug/mL. T-cells were then diluted with standard T-cell culture media and incubated at 30°C overnight. Cultures were subsequently expanded under standard T-cell expansion condition for 7-11 more days.

[0216] As shown in Figure 20, a large percentage of the cells (greater than 70% in all cases) exhibited targeted integration (TI) of the AAV donor into the TRAC or B2M locus. In addition, a similar percentage of cells receiving both TRAC and B2M targeted nucleases showed inactivation (KO) of both TRAC and of B2M as well as targeted integration of the AAV GFP donor (into TRAC when using donor with TRAC homology arms).

Example 4: *Ex vivo* methods

[0217] The genetically modified cells, including CD34+ HSPCs (*e.g.,* patient-derived CD34+ cells and/or modified CD4+, CD3+ and/or CD8+ T cells) as previously described (Aiuti *et al.* (2013) *Science* 341, 1233151), expressing one or more CARs as described herein are administered to subjects as previously described

(Aiuti *et al. ibid*), resulting in long-term multilineage engraftment in subjects treated with the modified cells.

Example 5: AAV transduction using PDGFr inhibitors

[0218] In order to test the relative contribution of co-receptors to AAV transduction, inhibitors against epidermal growth factor receptor (EGFR), hepatocyte growth factor receptor (HGFR), fibroblast growth factor receptor (FGFR) and platelet-derived growth factor receptor (PDGFR) were used prior to *in vitro* AAV2/6 transduction experiments in Hep3B (human hepatoma cell line) cells using zinc-finger nucleases (ZFNs) targeting the human albumin locus. Hep3B cells were plated at a density of 1 x 105 cells per well in a 48-well tissue culture plate in 300 μL complete growth media the day before transduction with AAV2/6. On the morning of transduction, the cells were washed three times with serum-free media and incubated in the serum-free media for three hours. Growth factor receptor inhibitors were added to the wells and incubated for one hour. AAV2/6 particles to deliver ZFNs targeted to the albumin locus were then added to the wells. After three hours, serum was added to the wells to a final concentration of 10%. The cells were harvested on day 4 post-transduction. Genomic DNA was extracted and analyzed by deep sequencing (MiSeq) at the albumin locus. The growth factor receptor inhibitors used are shown below in Table 1, and they were used in the experiment in the concentrations indicated in Figure 3, where each inhibitor is labeled by its target. For example, Gefitinib, an inhibitor of EGFR, is labeled as "EGFRi".

[0219] As shown in Figure 3, compared to control, only EGFRi showed mild inhibition (as expected due to its role as AAV6 co-receptor) while HGFRi and FGFRi showed mild stimulation of AAV transduction. In contrast, both PDGFR inhibitors used, CP-673451 and Crenolanib, showed significant dose-dependent stimulation of indel formation up to 8 fold (at conc. of 9μM).

Table 1: Growth factor receptor inhibitors

Name of inhibitor	Target
Gefitinib	EGFR
BGJ398	FGFR
SU11274	HGFR

CP-673451	PDGFR (1)
Crenolanib	PDGFR (2)

[0220] To ensure that the inhibitors were not somehow inadvertently stimulating error-prone DSB repair on their own (and thus having no effect on AAV transduction), delivery of the ZFNs by AAV and by mRNA delivery were compared. Hep3B cells were plated at a density of 1 x 10^5 cells per well in a 48-well tissue culture plate in 300 µL complete growth media the day before transduction with AAV2/6 or transfection with mRNA. On the morning of transduction and transfection, the cells were washed three times with serum-free media and incubated in the serum-free media for three hours. Growth factor receptor inhibitors were added to the wells and incubated for one hour. For AAV2/6 transduction, AAV2/6 particles to deliver ZFNs targeted to the albumin locus were then added to the wells. For mRNA transfection, 5 ng of mRNA per ZFN was delivered with Opti-MEM and RNAiMax Lipofectamine reagent to the cells. After three hours, serum was added to the wells to a final concentration of 10%. The cells were harvested on day 4 post-transduction. Genomic DNA was extracted and analyzed by deep sequencing (MiSeq) at the albumin locus.

[0221] As shown in Figure 4, there is no appreciable difference in the nuclease activity (% indels) detected in the cells treated with increasing doses of inhibitors when the ZFNs are delivered as RNA, indicating that the inhibitors effect is on transduction of AAV.

[0222] To investigate the mechanism of PDGFRi stimulation of AAV2/6 transduction, Hep3B cells were treated with both PDGFRi and EGFRi at the same time. Hep3B cells were plated at a density of 1 x 105 cells per well in a 48-well tissue culture plate in 300 µL complete growth media the day before transduction with AAV2/6. On the morning of transduction, the cells were washed three times with serum-free media and incubated in the serum-free media for three hours. Growth factor receptor inhibitors were added to the wells and incubated for one hour. AAV2/6 particles to deliver ZFNs targeted to the albumin locus were then added to the wells. After three hours, serum was added to the wells to a final concentration of 10%. The cells were harvested on day 4 post-transduction. Genomic DNA was extracted and analyzed by deep sequencing (MiSeq) at the albumin locus.

[0223] As shown in Figure 5, the mild decrease of overall AAV transduction (shown as % indels) when EGFRi is present suggests that EGFR is at least partly being used during increased AAV2/6 transduction. Contribution by other receptors is likely.

[0224] While Hep3B has been described to express high amounts of surface EGFR, other cell lines like HepG2 do not express EGFR at high levels. In order to demonstrate that the methods and compositions described herein work independently of EGFR expression levels, we tested AAV2/6 transduction in the presence of PDGFRi in HepG2 cells. HepG2/C3a cells were plated at a density of 1 x 105 cells per well in a 48-well tissue culture plate in 300 µL complete growth media the day before transduction with AAV2/6. On the morning of transduction, the cells were washed three times with serum-free media and incubated in the serum-free media for three hours. Growth factor receptor inhibitors were added to the wells and incubated for one hour. AAV2/6 particles to deliver ZFNs targeted to the albumin locus were then added to the wells. After three hours, serum was added to the wells to a final concentration of 10%. The cells were harvested on day 4 post-transduction. Genomic DNA was extracted and analyzed by deep sequencing (MiSeq) at the albumin locus.

[0225] As shown in Figure 6, as in Hep3B cells, the AAV2/6 transduction rates were up to 7- fold higher in combination with PDGFRi compared to controls in HepG2 cells. This demonstrates that PDGFR inhibition is a technique which can be used in a variety of cell types independent of EGFR expression status.

[0226] While all previous experiments were carried out in the absence of serum to prevent interference of EGFR present in serum, we also tested if PDGFR inhibition also stimulates AAV2/6 transduction in the presence of serum. Hep3B cells were plated at a density of 1 x 105 cells per well in a 48-well tissue culture plate in 300 µL complete growth media the day before transduction with AAV2/6. On the morning of transduction for the no-serum wells, the cells were washed three times with serum-free media and incubated in the serum-free media for three hours. Growth factor receptor inhibitors were added to the wells and incubated for one hour. AAV2/6 particles to deliver ZFNs targeted to the albumin locus were then added to the wells and incubated for three hours. Serum was added to the wells to a final concentration of 10%. For the full-serum wells, the cells were washed with complete growth media, and AAV2/6 particles were added to the wells. The cells were

harvested on day 4 post-transduction. Genomic DNA was extracted and analyzed by deep sequencing (MiSeq) at the albumin locus.

[0227] As shown in Figure 7, in HepG2 cells, we found that overall AAV2/6 transduction was down in the presence of serum by about 40-fold. However, when PDGRi was added AAV2/6 transduction was very robust even in presence of serum (up to 20% indels). When calculated as fold-change over untreated AAV2/6 transduction (see Figure 8), the PDGRi increased AAV transduction in a dose dependent-manner 4-fold in absence of serum (up to 50% indels) and up to 70-fold in the presence of serum (up to 20% indels). This demonstrates that PDGFR inhibition can be used *in vivo* in presence of serum proteins to enhance delivery of nucleic acids using viral vectors such as AAV.

[0228] We next tested the effect of PDGFR inhibitors human patient-derived primary hepatocytes. For human hepatocytes, 48-well cell culture dishes were purchased pre-coated with collagen (Life Technologies). Plates were incubated for 1 hour at 37 °C. Thawing/plating media was prepared by combining 18 mL InVitroGRO CP medium (BioreclamationIVT) and 400 µL Torpedo antibiotic mix (Celsis In Vitro Technologies). Once the plates were prepared, the female plate-able human hepatocytes (Lot# AKB,) were transferred from the liquid nitrogen vapor phase directly into the 37 °C water bath and thawed under gently stirring. The cells were transferred directly into a 50 ml conical tube containing 5 mL of pre-warmed thawing/plating medium. To retrieve all the cells, the vial was washed with 1 mL of thawing/plating medium, which was added to the cells. Following resuspension of the cells, a small aliquot (20 µL) was removed to perform a cell count and to determine cell viability using trypan blue solution (1:5; Corning, Cat# 25-900-C1). The cells were then centrifuged at 75 x g for 5 minutes, the supernatant was removed and the cells were resuspended at 1e6 cells/mL. Cells were seeded at 3e5 cells/well in the coated 48-well plates. Cells were then incubated in a 37°C/5% CO2 incubator.

[0229] One day after the cells were plated, cells were switched to HCM maintenance medium (Lonza, HBM and HCM SingleQuots). The following day, cells were fed with fresh HCM medium. Growth factor receptor inhibitors were added to the wells and incubated for one hour. AAV2/6 particles were mixed at the appropriate MOI with HCM medium (in 300 µL per well) and added to the cells. After 24 hrs, the medium was replaced with fresh HCM medium to ensure maximal

health of the primary hepatocyte cultures. The cells were harvested on day 4 post-transduction. Genomic DNA was extracted and analyzed by deep sequencing (MiSeq) at the albumin locus.

[0230] As shown in Figure 9, compared to DMSO treated controls, increased AAV2/6 transduction was observed with a high dose of PDGFRi (9 μM) treatment but not with low dose (3 μM). This 5-fold increase indicates that the effectiveness of the methods described herein in primary cells and therefore *in vivo*. Importantly, these cells were grown in absence of serum.

Example 6: Delivery of AAV to non-human primates *in vivo*

[0231] To investigate the effect of combining steroid and anti-B cell regimes on efficiency of AAV delivery as measured by nuclease activity, the following study was performed. In these studies, AAV2/6 was the predominant serotype used (unless indicated elsewhere) where the terminology 2/6 means that the AAV comprises the AAV2 ITRs but has the AAV6 capsid. The studies utilized zinc finger nucleases designed to cleave the *M. fascicularis* albumin locus (SBS#37804 "Left ZFN"/ SBS#43083 "Right ZFN", see PCT publication WO2015/127439). Donors used were designed to introduce human F.IX or human FVIII transgene via nuclease mediated targeted integration of the transgene. Donor constructs are described in WO2015127439 (human F.IX) and WO2015089077 (human FVIII).

[0232] Cynomolgus monkeys (*M. fascicularis*) were housed in stainless steel cages equipped with a stainless mesh floor and an automatic watering valve. The study complied with all applicable sections of the Final Rules of the Animal Welfare Act regulations (Code of Federal Regulations, Title 9).

[0233] Control Article (Formulation Buffer, PBS, 35 mM NaCl, 1% sucrose, 0.05% pluronic 188, pH 7.1) and test article were thawed and dispensed on Day 1 of the studies where test and control articles were administered via intravenous infusion into a peripheral vein at 1 mL/minute. For the rituximab dosing, animals received doses of 10 mg/kg/dose at a concentration of 10 mg/mL in a volume of 1 mL/kg by intravenous dosing. For the methylprednisolone (Solu-Medrol®), the agent was administered at a dose of 10 mg/kg/dose at a concentration of 20 mg/mL, in a dose volume of 0.5 mL/kg, administered intramuscularly.

[0234] The dosing scheme is depicted in Figure 10 where Figure 10A shows a combination delayed dosing of Solu-Medrol® and Rituxan®. The AAV6 test article

was administered to the animals on day 1 and Solu-Medrol® given daily starting on day 5 through day 16, and Rituxan® given on days 7 and 14. The second scheme is shown in Figure 10B where Rituxan® was given on day -8 and on the day before the test article, and the Solu-Medrol® was given on the same day as the test article, and then given daily afterwards.

[0235] The dosing groups are shown below in Tables 2 and 3. Table 2 depicts the groups used in the scheme depicted in Figure 10A while Table 3 shows the groups used in the scheme for Figure 10B. The test articles shown in Tables 2 and 3 are characterized by the detail in the tables.

Table 2: Delayed treatment of steroid + B cell inhibitor, test articles used

Group No.	No. of Males [a]	Test Article Identification	AAV Transgene Description	Ratio (ZFN:ZFN: hF9 Donor)	Component Dose Level (vg/kg)	Total AAV Vector Dose Level (vg/kg)
1	1	SGMO.01	Formulation Buffer	NA	NA	NA
2	2	SGMO.02	Left ZFN Right ZFN	NA	1.5E+13 each ZFN	3E+13
3	3	SGMO.03	Left ZFN Right ZFN hF9 donor	1:1:8 High Dose	1.5E+13 each ZFN 1.2E+14	1.5E+14
4	3	SGMO.04	Left ZFN Right ZFN hF9 donor	1:1:4 High Dose	1.5E+13 each ZFN 6E+13	9E+13
5	3	SGMO.05	Left ZFN Right ZFN hF9 donor	1:1:2 High Dose	1.5E+13 each ZFN 3E+13	6E+13
6	3	SGMO.06	Left ZFN Right ZFN hF9 donor	1:1:8 Mid Dose	5E+12 each ZFN 4E+13	5E+13
7	3	SGMO.07	Left ZFN Right ZFN hF9 donor	1:1:2 Mid Dose	5E+12 each ZFN 1E+13	2E+13
8	2	SGMO.08	Left ZFN Right ZFN hF9 donor	1:1:8 Low Dose	1.5E+12 each ZFN 1.2E+13	1.5E+13
9	1	SGMO.09	hF9 donor	NA	3E+13	3E+13

Table 3: Pre-treament with B cell inhibitor, combined concurrent test article and steroid treatment, test articles

Group No.	No. of Males [a]	Test Article Identification	AAV Transgene Description	Ratio (ZFN:ZFN: hF9 Donor)	Component Dose Level (vg/kg)	Total AAV Vector Dose Level (vg/kg)
1	3	SGMO.01	Formulation Buffer	NA	NA	NA
2	4	SGMO.02	Left ZFN	1:1:8 Mid Dose	2.25E+12 each ZFN	2.25E+13
			Right ZFN			
			hF9 donor		1.8E+13	
3	5	SGMO.03	Left ZFN	1:1:8 Low Dose	1.2E+12 each ZFN	1.2E+13
			Right ZFN			
			hF9 donor		9E+12	

[0236] Samples were taken to evaluate standard hematology, coagulation, clinical chemistry parameters, for detection of human FIX and anti-FIX antibodies, and for urinalysis. Coagulation was analyzed by measuring activated partial thromboplastin time (aPTT), and prothrombin time. In addition, biopsies were taken to analyze AAV6 vector genome copy number in genomic DNA, and to measure nuclease activity.

[0237] The results for the delayed (post-test article, Figure 10A) treatment regime are shown in Figure 11. All dose groups showed an increase in hFIX concentration, with the highest levels seen in the highest total AAV2/6 dose group (Figure 11A). In the 1:1:8 dose groups (AAV-ZFN1:AAV-ZFN2:AAV-donor), the highest hFIX concentration was again seen with the highest total AAV2/6 dose (Figure 11B). For the pre-treatment regime (Figure 10B), FIX was detectable at both dosing groups (Figure 11C). When the data from both treatment regimens are compiled (Figure 11D), an increase in detectable hF.IX as the dose of total AAV2/6 increased was observed.

[0238] The animals were also analyzed for the levels of gene modification (% of genes comprising an insertion or a deletion at the cleavage site, "indels"). Levels of gene modification detected at Day 28 (Figure 12) indicated that in the post-test article administration of the immunomodulators (Figure 10A), gene modification increased as total AAV2/6 dose increased. When the rituximab was given prior to test article, and then the steroid was given concurrently with the test article and then daily thereafter, gene modification increased as total AAV2/6 dose increase (Figure 12A). When the gene modification levels from both regimes was compiled (Figure 12B), the general trend was that increased AAV2/6 dose related to increased gene modification

(Figure 12C). Importantly, the use of the steroid treatment prior to treatment with the test article allowed the use of a lower dose of test article to achieve a given level of gene modification as compared to treatment without the pre-dosing steroid treatment, while still achieving therapeutic levels of protein expression.

[0239] A similar experiment was also carried in NHP using a FVIII-BDD transgene rather than the F.IX donor. In this experiment, AAV6 and AAV8 serotypes were evaluated. Table 4 below shows the identity of the dosing groups. The difference between the F8 transgene expression cassette in Groups 2-4 ("cDNA 1") and Group 5 ("cDNA2") is that the Group 5 donor had a slightly different promoter module (hybrid liver promoter, see McIntosh *et al* (2013) *Blood* 121(17):3335), but the rest of the F8-BDD transgene expression cassette (including the coding region) was the same. In this experiment, ZFN were not used to cause targeted integration into the genome.

Table 4: NHP groups with FVIII transgene cassette

Group No.	No. of Males [a]	Test Article Identification	AAV Transgene Description	Ratio (ZFN:ZFN: hF9 Donor)	Serotype	Component Dose Level (vg/kg)	Total AAV Vector Dose Level (vg/kg)
1	2	SGMO.01	Formulation Buffer	NA	NA	NA	NA
2	3	SGMO.02	hF8-BDD cDNA 1	NA	AAV2/6	2E+12	2E+12
3	3	SGMO.03	hF8-BDD cDNA 1	NA	AAV2/6	6E+12	6E+12
4	3	SGMO.04	hF8-BDD cDNA 1	NA	AAV2/8	6E+12	6E+12
5	3	SGMO.05	hF8-BDD cDNA 2	NA	AAV2/8	6E+12	6E+12

For this experiment, the dosing regimen outlined in Figure 10B was used (where the rituximab was administered pre-test article and the steroid administered concurrently with the test article and daily thereafter), and the experiment was carried out for 14 days.

[0240] The data is presented in Figures 13-19, were each monkey in each group is given a data point. The data indicates that the higher doses (compare Figure 13A with Figure 13B) of test article in the AAV6 serotype background gave expression of FVIII-BDD at nearly 10X the level found in normal human plasma. The

data for test article in the AAV2/8 serotype showed an increase in the FVIII activity, but not to the same extent as was observed for AAV2/6.

[0241] Subsequent to the initial 14 day period described above, the experiment was continued up to 168 days post the single dose of the AAV-FVIII-BDD. The co-dosing of the steroid was stopped at day 103 (Figure 14). Determination of the hFVIII-BDD levels in the plasma of the monkeys was determined using a custom ELISA as follows. 96-well half-area HB (high binding) polystyrene microplates (Corning) were coated overnight at 4 °C with mouse monoclonal anti-hFVIII antibody (Green Mountain, Burlington, VT) in 0.2 M carbonate bicarbonate buffer pH 9.4 (Thermo Fisher Scientific, Waltham MA). The following day the plates were washed four times using 1X TBST (Thermo Fisher Scientific, Waltham MA). 96-well plates were then blocked two hours at room temperature using 3% BSA/TBS blocking buffer, followed by washing four times with 1X TBST. Plasma was added to the plate and incubated with rocking at room temperature for two hours, followed by washing four times with 1X TBST. Detection antibody, biotinylated monoclonal mouse anti-FVIII antibody (Green Mountain, Burlington, VT) was added and incubated for one hour at room temperature, followed by washing four times with 1X TBST. Streptavidin HRP (Jackson ImmunoResearch, West Grove PA) was added and incubated for one hour at room temperature followed by washing four times with 1X TBST. TMB Ultra (Thermo Fisher Scientific, Waltham MA) was added and allowed to develop for ten minutes, reaction was stopped with stop solution and absorbance read at 450 nM using a plate reader. Background absorbance readings were negligible (typically 0). The presence of inhibitory anti-FVIII antibodies were determined using a Bethesda assay (for example, see Kasper *et al* (1975) *Thromb Diath Haemorrh* 34:869-72). Figure 15 shows the peak concentration of FVIII-BDD delivered via AAV2/6 detected in the plasma over the course of the study.

[0242] For the low dose animals (n=3), comprising the FVIII-BDD cDNA donor in AAV2/6, dosed at 2E+12 vg/kg, following detection of robust hFVIII antigen levels (Ag), hFVIII-BDD levels decreased with a concomitant increase in Bethesda Units (BU). BU decreased over time and the hFVIII Ag increased (Figure 16). The results demonstrated that following the cessation of the immunosuppressive therapy, the levels of human FVIII antigen dropped.

[0243] For the high dose animals (n=3), comprising the FVIII-BDD cDNA donor in AAV2/6, dosed at 6E+12 vg/kg, a similar pattern was observed (see Figure 17). However in one animal, 3101, following the removal of Solu-medrol, anti-FVIII antibodies were not detected despite a detectable and persistent level of FVIII antigen (representing 200% of normal hFVIII levels), which could be indicative of a tolerance of animal to the human FVIII antigen.

[0244] When the experiment was done with delivery of the high dose in the AAV2/8 vector, similar results were seen except that the amount of FVIII antigen detectable in the plasma was less that seen using the AAV2/6 vector (Figure 18). Similarly, when a different FVIII-BDD cDNA promoter module was tested (Group 5, described above) in the AAV2/8 vector, FVIII-BDD plasma levels were similar to those seen in Group 4 (Figure 19). However, as above, there were two individuals that maintained a detectable amount of FVIII-BDD expression (5101 and 5102), and an individual in Group 4 (4103, Figure 18D) without a marked antibody response following the removal of the Solu-medrol, again suggestive of tolerization to the antigen following the robust response levels seen in the initial days of the experiment.

[0245] As shown in Figure 19D, Animal No. 5101 appears tolerized to hFVIII-BDD as after removal of methylprednisolone, hFVIII-BDD levels remain stable for 8 weeks at approximately 0.1U/mL (representing 10% of normal hFVIII levels). As shown in Figure 19E, Animal No. 5102 appears tolerized to hFVIII-BDD as after removal of methylprednisolone hFVIII-BDD levels remain stable for 8 weeks at approximately 0.6U/mL (representing 60% of normal hFVIII levels). It is worth noting that normal levels of hFVIII in human plasma is approximately 1U/mL or 200 ng/mL, and that expression of even 1%-5% of normal (>0.001 U/mL) has therapeutic efficacy (Llung RC (1999) *Thromb Haemost* 82(2):525-530).

[0246] All patents, patent applications and publications mentioned herein are hereby incorporated by reference in their entirety.

[0247] Although disclosure has been provided in some detail by way of illustration and example for the purposes of clarity of understanding, it will be apparent to those skilled in the art that various changes and modifications can be practiced without departing from the spirit or scope of the disclosure. Accordingly, the foregoing descriptions and examples should not be construed as limiting.

CLAIMS

What is claimed is:

1. A method of introducing a nucleic acid into an isolated cell, the method comprising:

administering to the cell at least one adeno-associated vector (AAV) comprising a donor molecule in the presence of at least one inhibitor of growth factor receptor binding.

2. The method of claim 1, wherein the growth factor inhibitor inhibits binding to an epidermal growth factor receptor (EGFR), a fibroblast growth factor receptor (FGFR), a Met/hepatocyte growth factor receptor (HGFR), a lipoarabinomannan receptor (LamR), a αVβ5 integrin receptor, a Intercellular Adhesion Molecule 1 receptor (Icam-1) and/or a Platelet-derived growth factor receptor.

3. The method of claim 1 or 2, wherein the donor molecule comprises a transgene that is expressed in the cell.

4. The method of any one of claims 1 to 3, wherein the transgene is integrated into the genome of the cell.

5. The method of claim 4, wherein the transgene encodes a chimeric antigen receptor (CAR).

6. The method of claim 4, further comprising introducing at least one nuclease into the cell, wherein the transgene is integrated into one or more genes of the cell following cleavage of the one or more genes by the nuclease, the method further comprising introducing a second nuclease that inactivates one or more additional genes of the cell.

7. The method of claim 6, wherein the nuclease cleaves a programmed cell death 1 (PD1) gene, a Cytotoxic T-Lymphocyte Antigen 4 (CTLA-4) gene, a beta 2-microglobulin (B2M) and/or a T-cell receptor alpha (TRAC) gene.

8. The method of any one of claims 1 to 7, wherein the cell is a hematopoietic stem cell, a T-cell, a B-cell or an NK cell.

9. A method of treating a cancer in subject,

the method comprising introducing a nucleic acid into a cell according to the method of any one of claims 1 to 8, wherein donor molecule comprises a sequence encoding a CAR such that the cell expresses the CAR;

administering the cell to the subject.

10. A method of introducing a nucleic acid into a cell of a subject, the method comprising:

administering to the subject at least one adeno-associated vector (AAV) comprising a donor molecule and at least one steroid and/or at least one B-cell inhibitor.

11. The method of claim 10, wherein the donor molecule comprising a sequence encoding a transgene.

12. The method of claim 10 or claim 11, wherein the subject is a mammal and the transgene encodes a therapeutic protein, and further wherein the mammal becomes tolerized to the therapeutic protein.

13. The method of claim 12, wherein the subject has a hemophilia and the donor molecule encodes a clotting factor such as Factor VIII or Factor IX.

14. The method of introducing a nucleic acid into a cell, the method comprising:

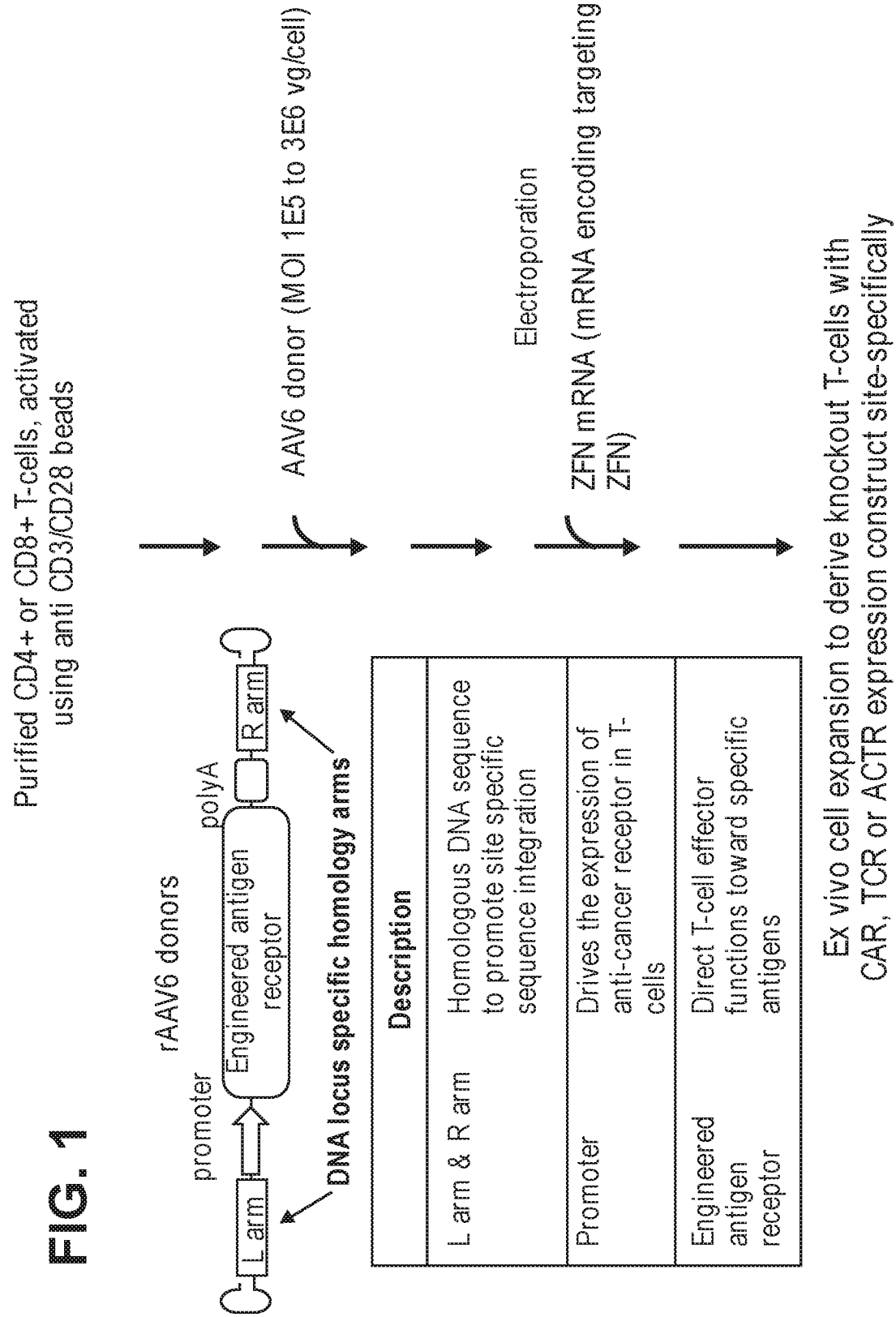
administering a donor molecule comprising the nucleic acid into the cell;

administering a nuclease to the cell, wherein the nuclease is administered 1 to 48 hours after or within 4 hours before the donor molecule and further wherein the donor molecule is integrated into the genome of the cell following cleavage by the nuclease.

15. A method of introducing one or more transgenes into a genome of a cell, the method comprising:

introducing at least one nuclease into the cell, wherein the at least one nuclease cleaves the genome of the cell such that the one or more transgenes are integrated into the genome of the cell, and further wherein (i) if the donor vector is introduced into the cell before the at least one nuclease, the at least one nuclease is introduced into the cell within 48 hours after donor vector is introduced and; (ii) if the at least one nuclease is introduced before the donor vector, the donor vector is introduced into the cell within 4 hours after the at least one nuclease is introduced.

FIG. 1



	Description
L arm & R arm	Homologous DNA sequence to promote site specific sequence integration
Promoter	Drives the expression of anti-cancer receptor in T-cells
Engineered antigen receptor	Direct T-cell effector functions toward specific antigens

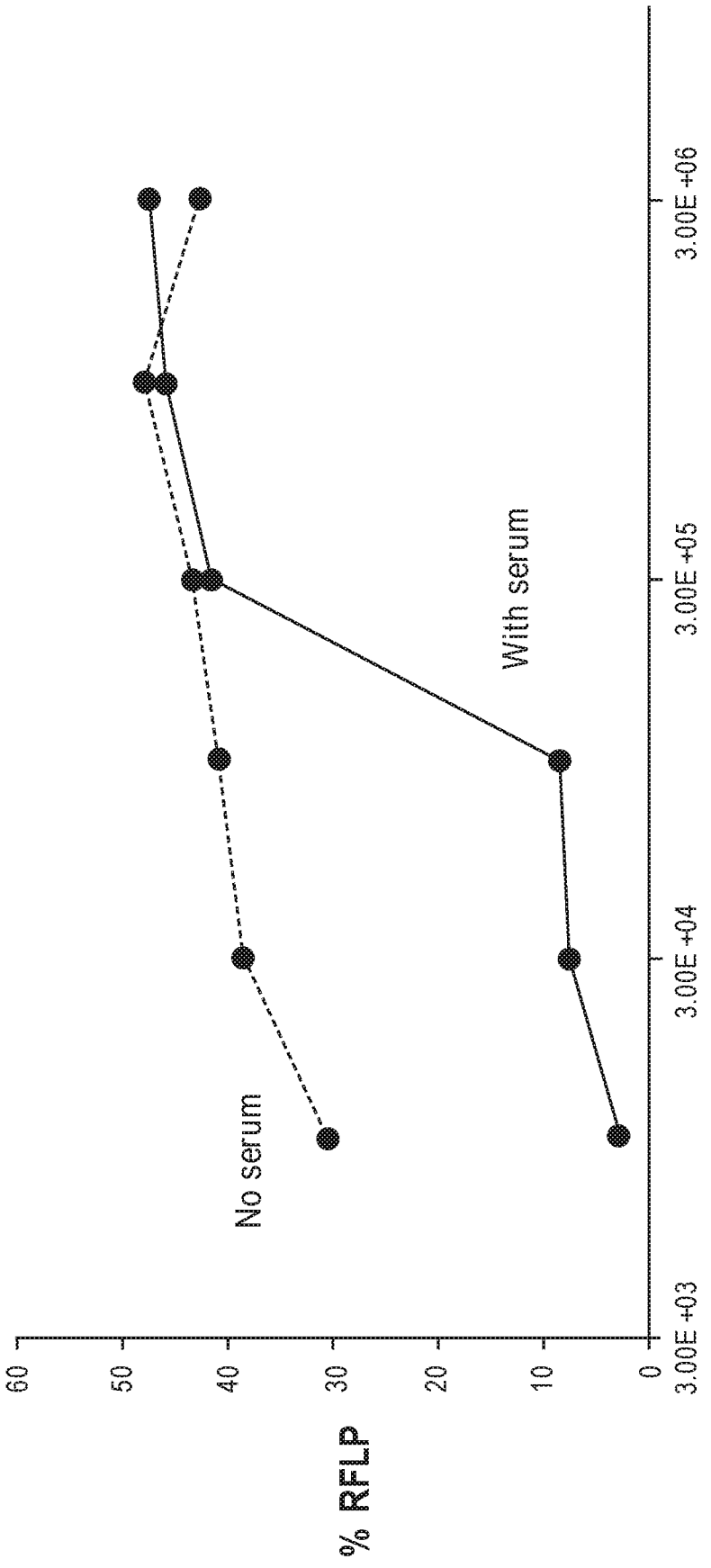

% RFLP
60
50
40
30
20
10
0
3.00E +03
3.00E +04
3.00E +05
3.00E +06
AAV6-CCR5-RFLP (vg/cell)
No serum
With serum

FIG. 2

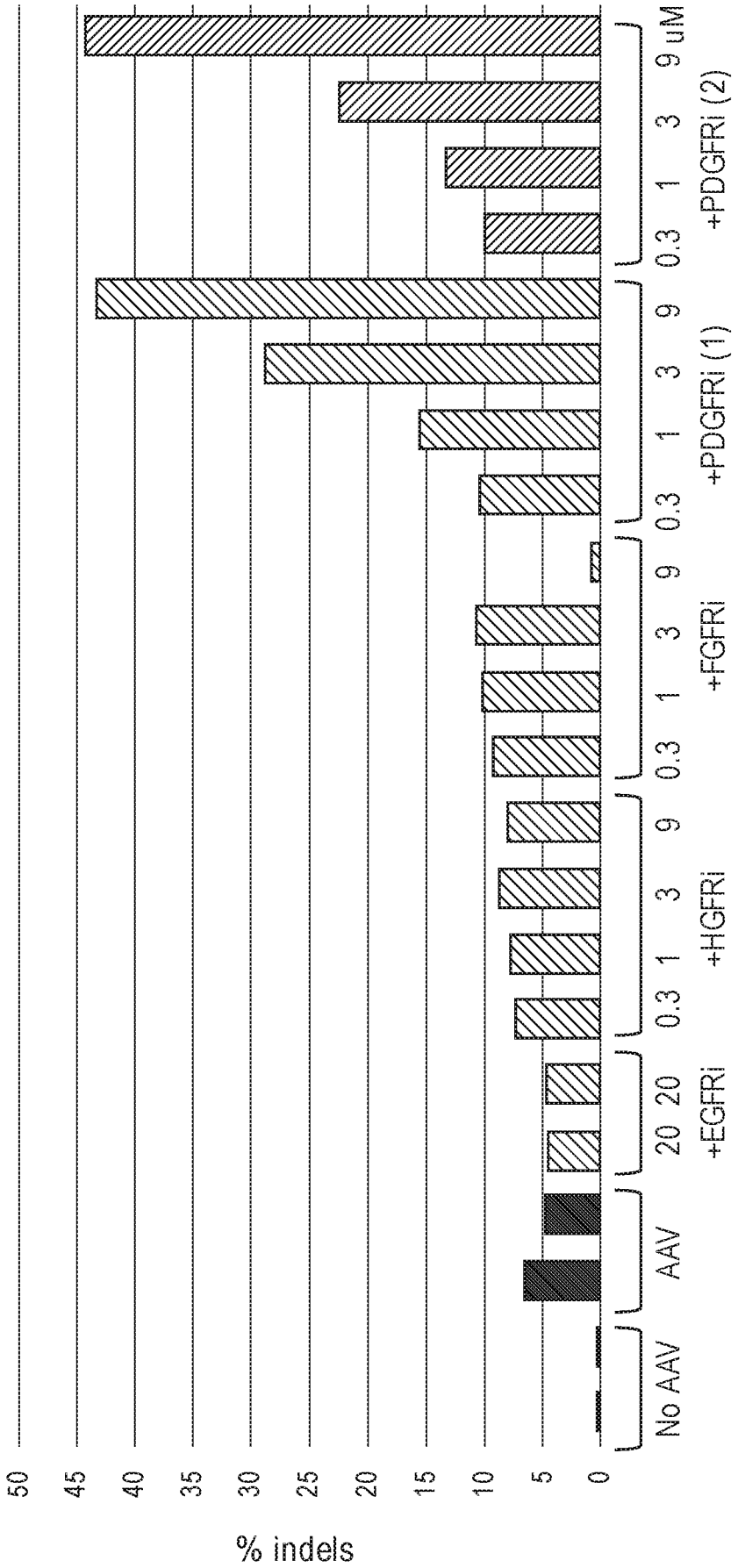

% indels
0
5
10
15
20
25
30
35
40
45
50
No AAV
AAV
20
20
+EGFRi
0.3
1
3
9
+HGFRi
0.3
1
3
9
+FGFRi
0.3
1
3
9
+PDGFRi (1)
0.3
1
3
9
+PDGFRi (2)
uM

FIG. 3

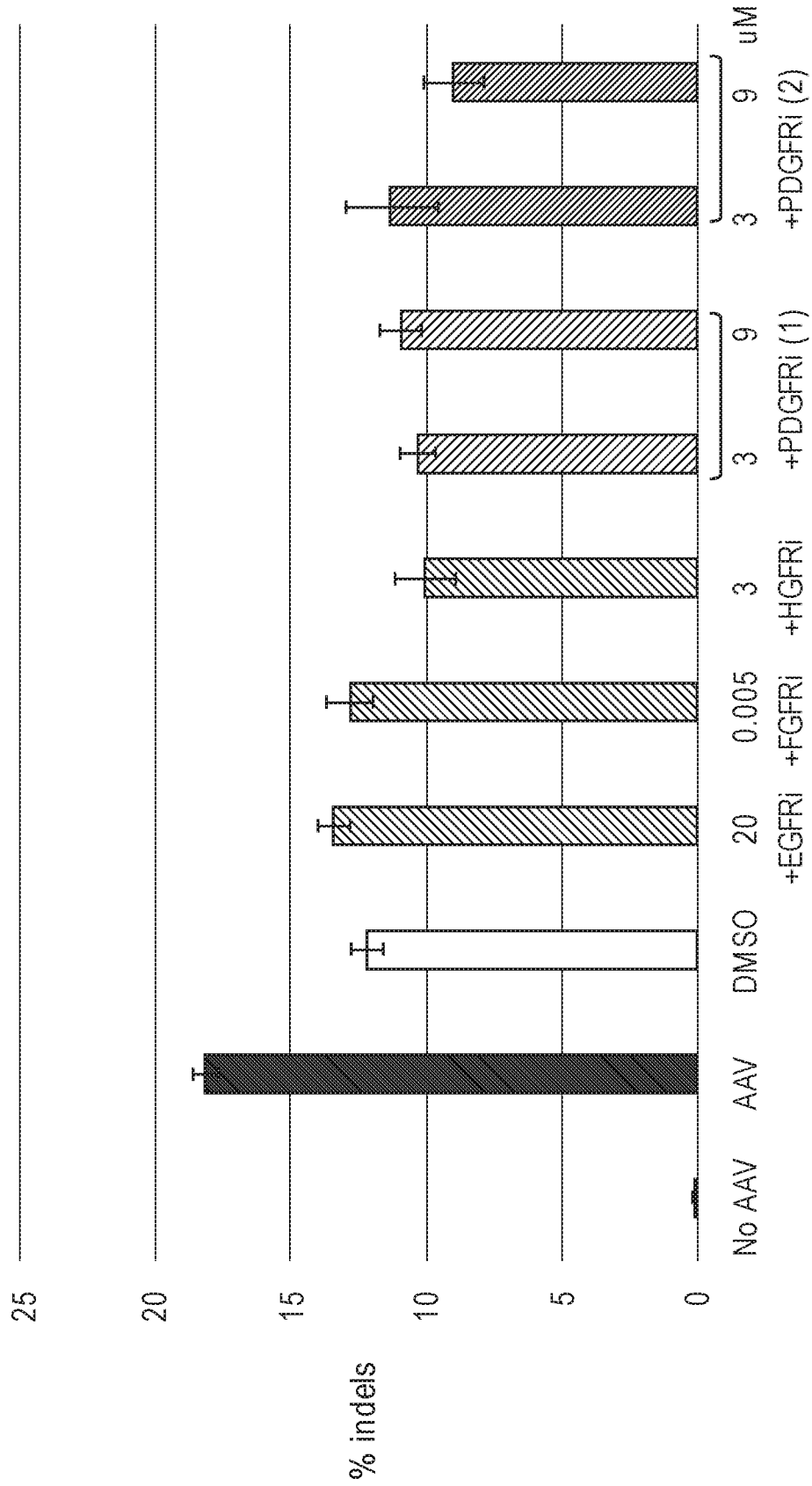

% indels
25
20
15
10
5
0
No AAV
AAV
DMSO
20
0.005
3
3
9
3
9
uM
+EGFRi
+FGFRi
+HGFRi
+PDGFRi (1)
+PDGFRi (2)

FIG. 4

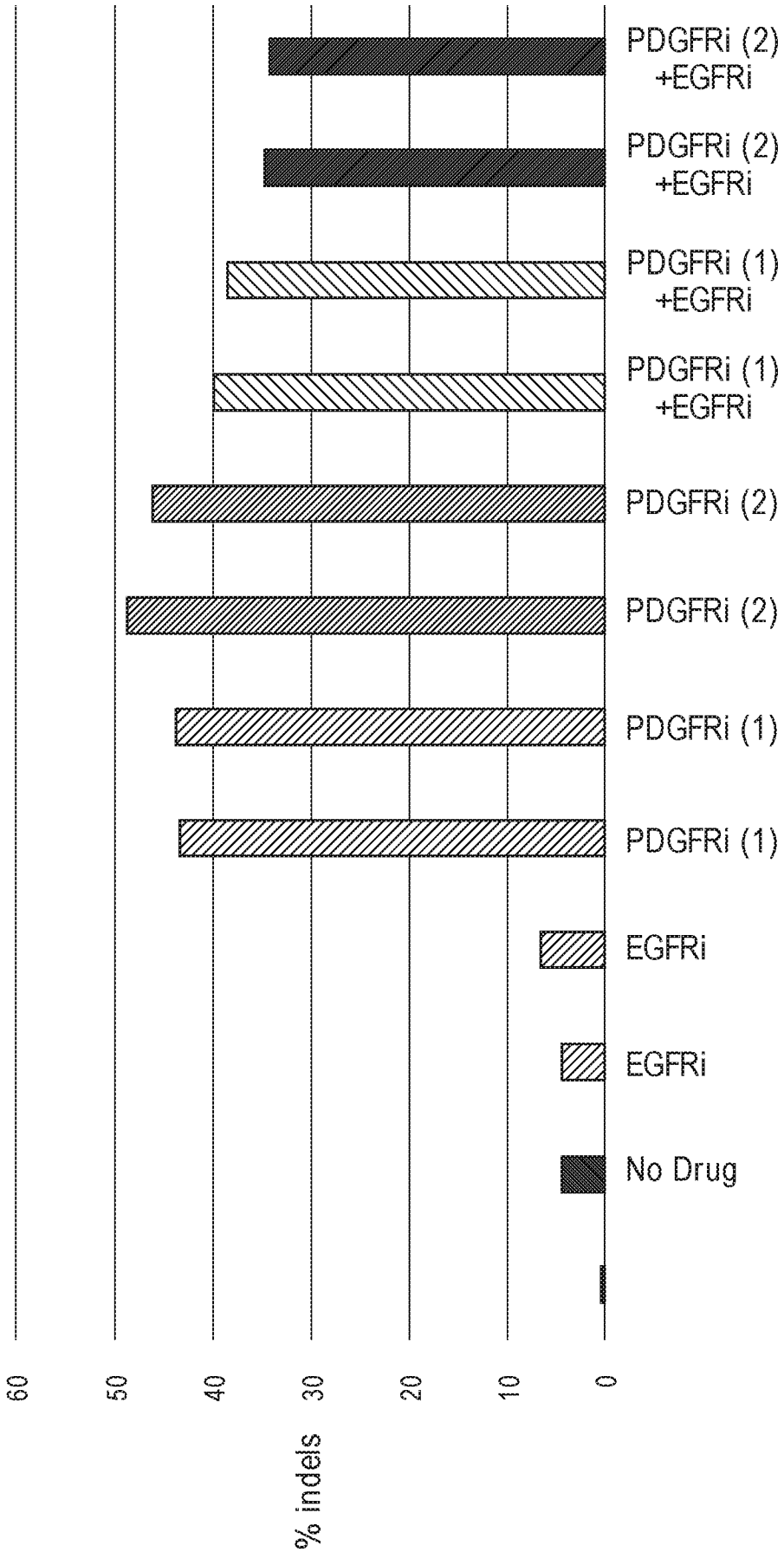

PDGFRi (2) +EGFRi
PDGFRi (2) +EGFRi
PDGFRi (1) +EGFRi
PDGFRi (1) +EGFRi
PDGFRi (2)
PDGFRi (2)
PDGFRi (1)
PDGFRi (1)
EGFRi
EGFRi
No Drug
60
50
40
30
20
10
0
% indels

FIG. 5

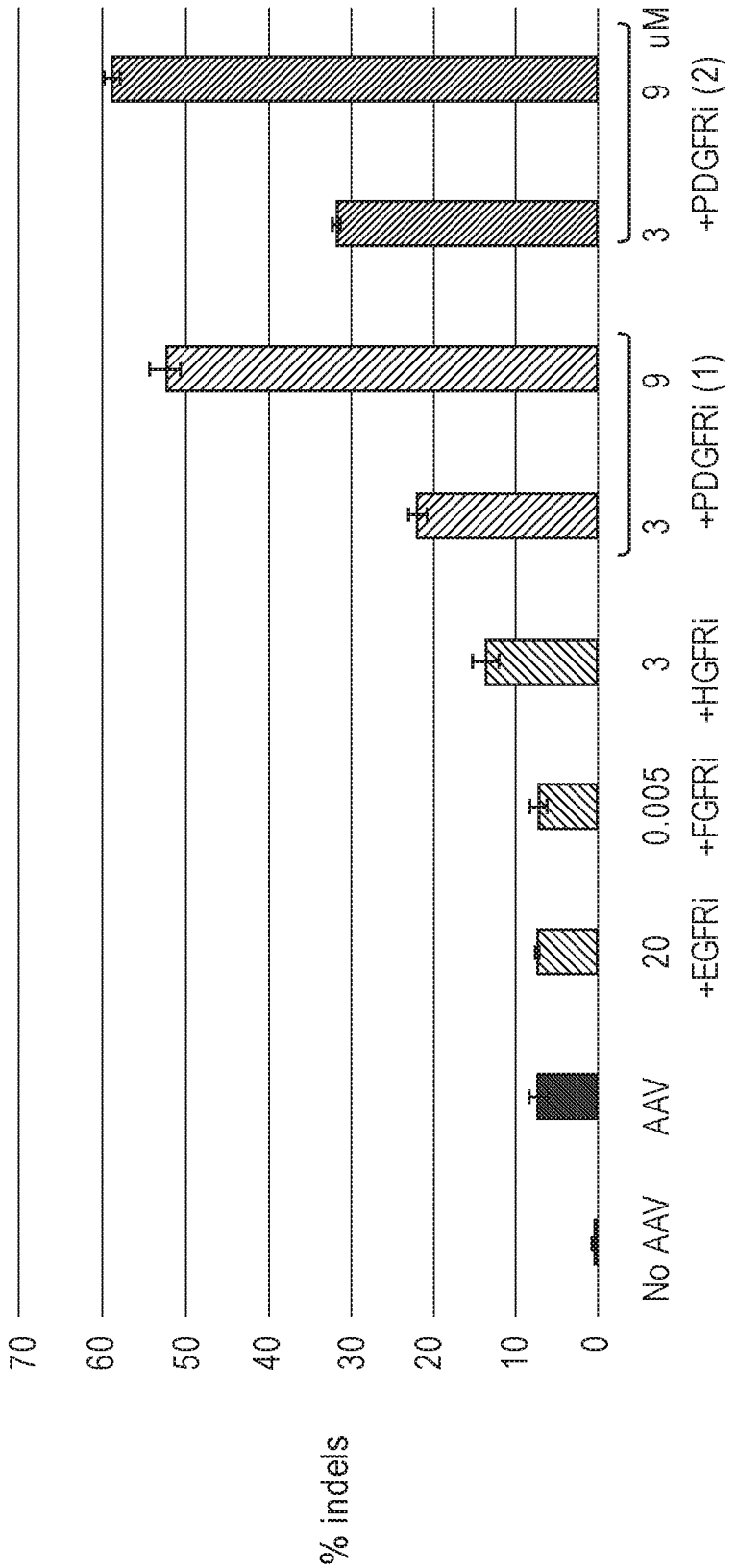

% indels
70
60
50
40
30
20
10
0
No AAV
AAV
+EGFRi
20
+FGFRi
0.005
+HGFRi
3
+PDGFRi (1)
3
9
+PDGFRi (2)
3
9
uM

FIG. 6

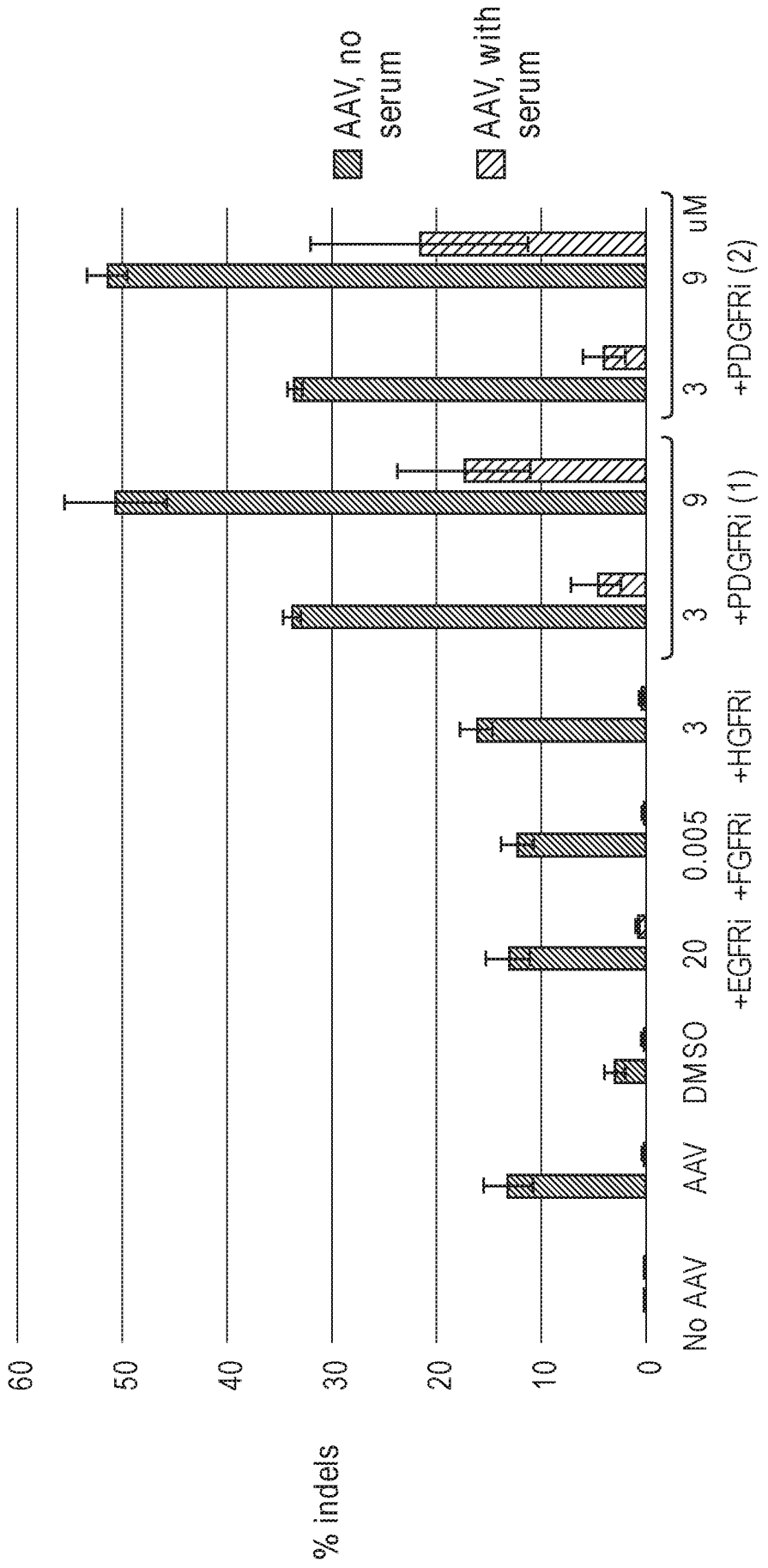

AAV, no serum
AAV, with serum
% indels
0
10
20
30
40
50
60
No AAV
AAV
DMSO
20
0.005
3
3
9
3
9
uM
+EGFRi
+FGFRi
+HGFRi
+PDGFRi (1)
+PDGFRi (2)

FIG. 7

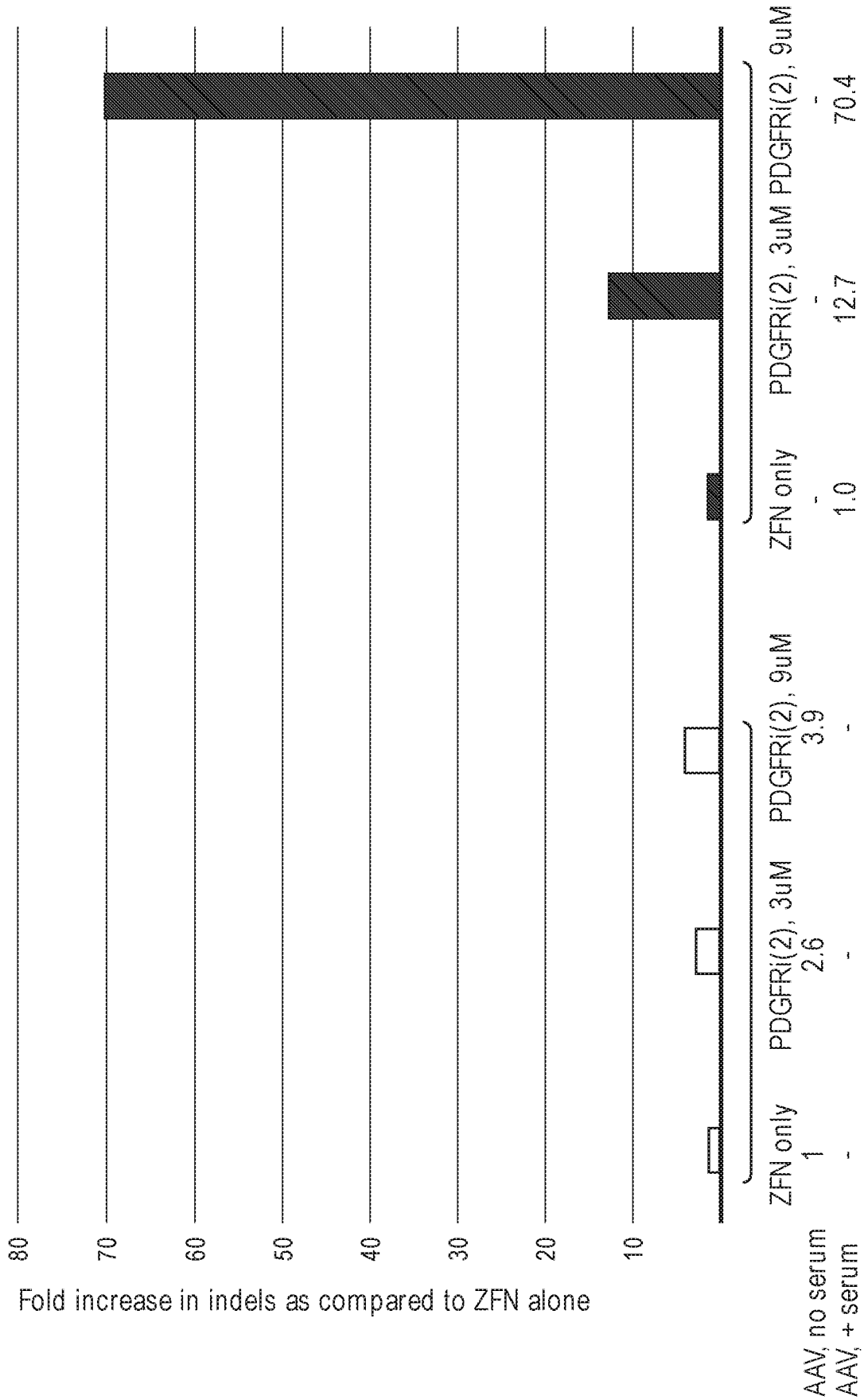

Fold increase in indels as compared to ZFN alone
80
70
60
50
40
30
20
10
AAV, no serum
AAV, + serum
ZFN only
1
-
PDGFRi(2), 3uM
2.6
-
PDGFRi(2), 9uM
3.9
-
ZFN only
-
1.0
PDGFRi(2), 3uM
-
12.7
PDGFRi(2), 9uM
-
70.4

FIG. 8

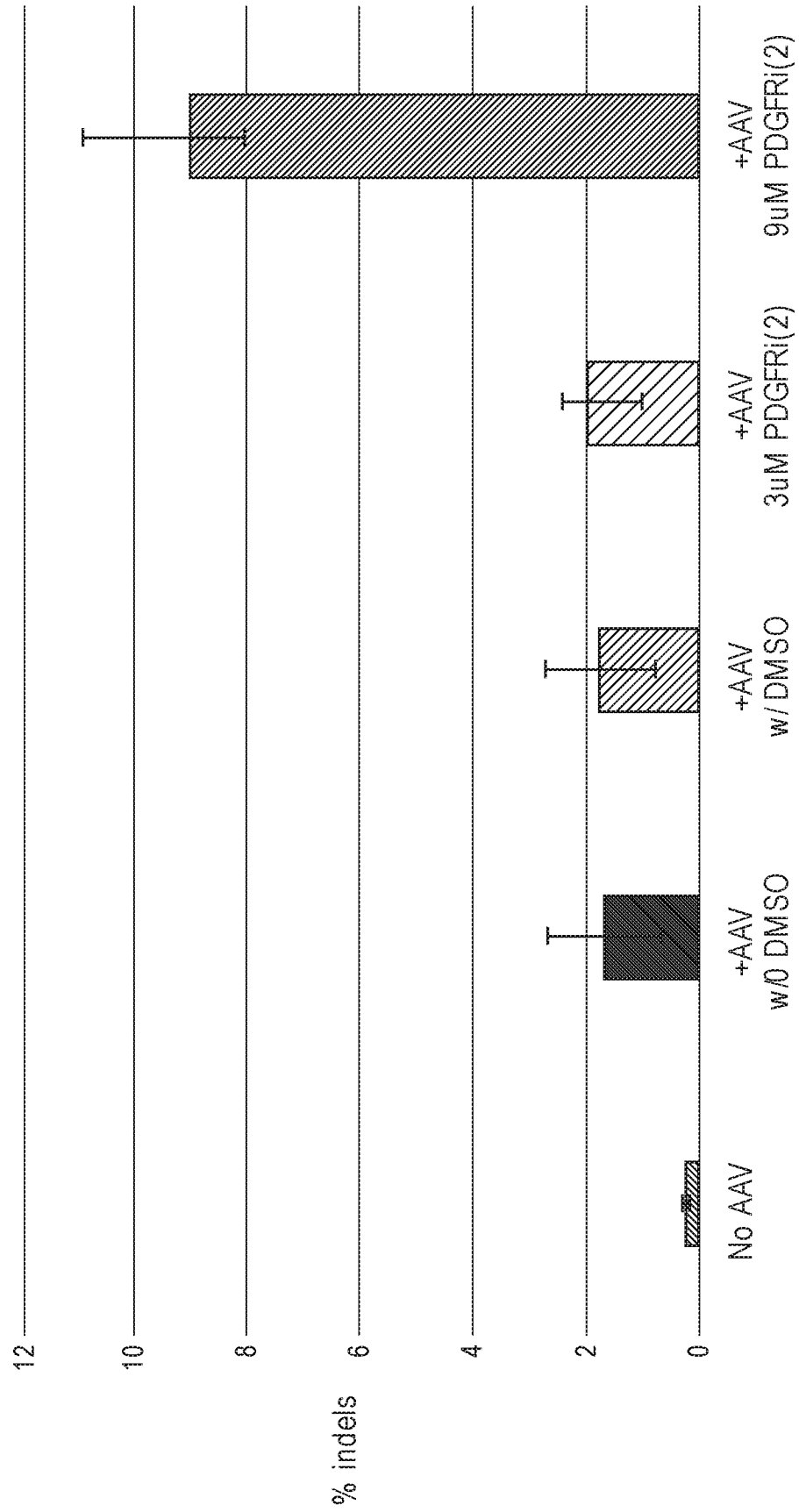

% indels
12
10
8
6
4
2
0
No AAV
+AAV w/0 DMSO
+AAV w/ DMSO
+AAV 3uM PDGFRi(2)
+AAV 9uM PDGFRi(2)

FIG. 9

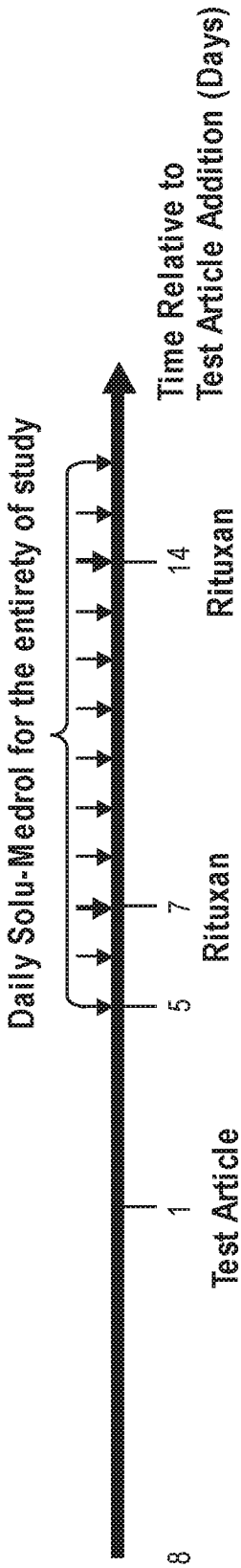


FIG. 10A

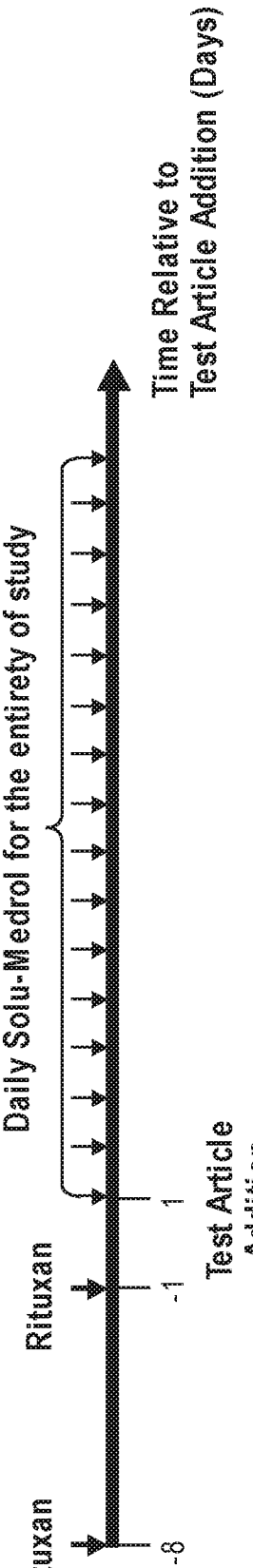


FIG. 10B

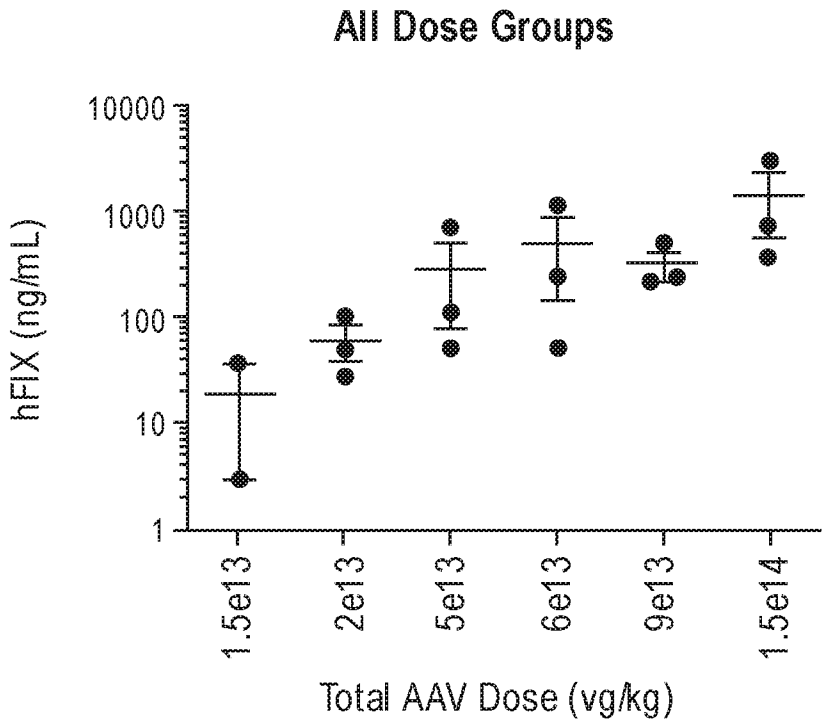


FIG. 11A

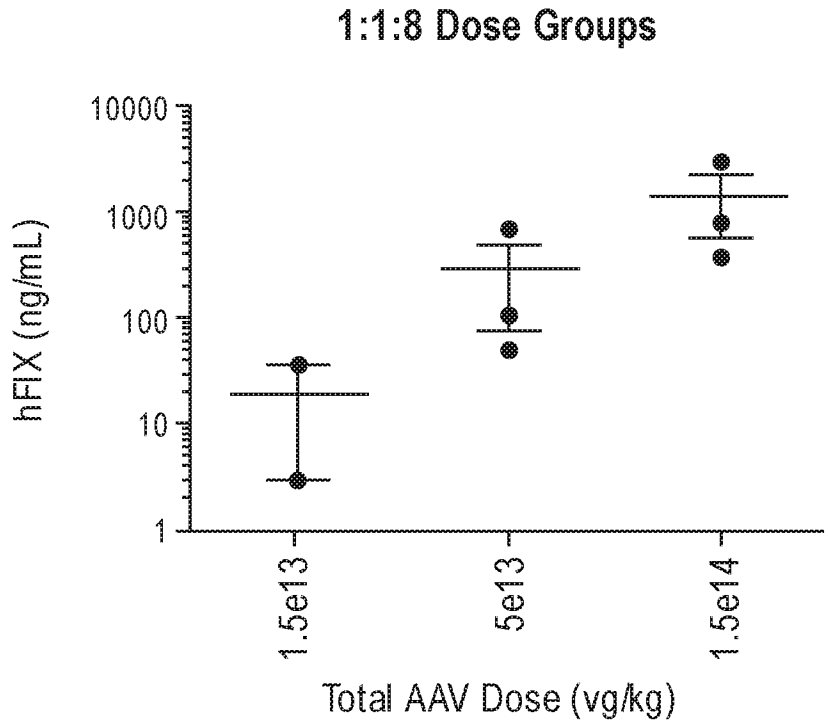


FIG. 11B

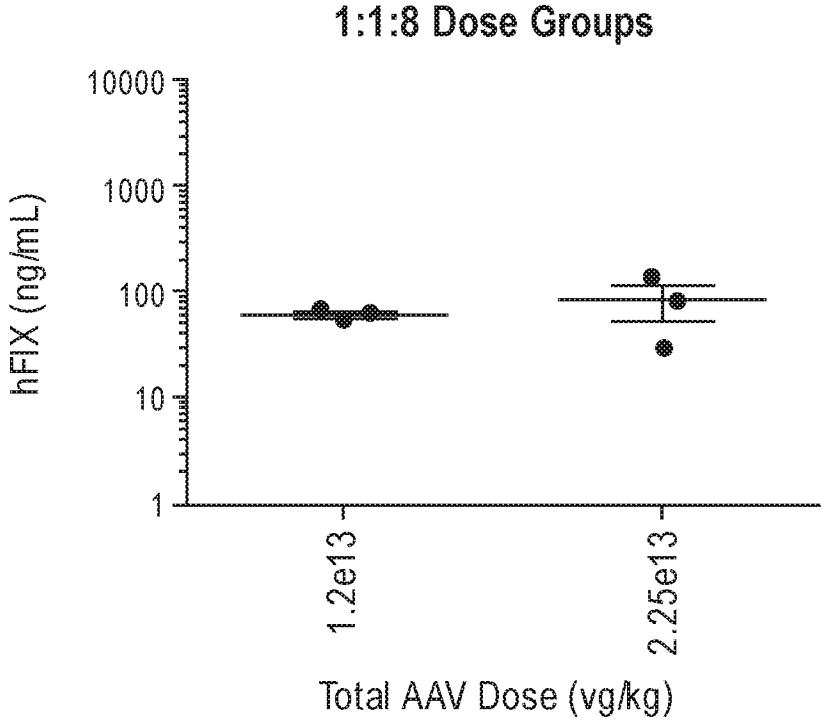


FIG. 11C

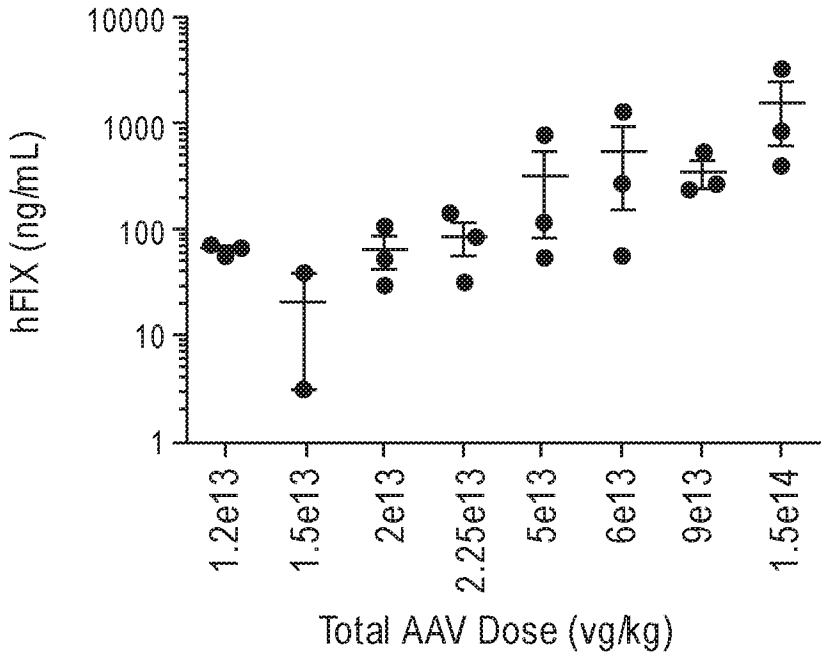


FIG. 11D

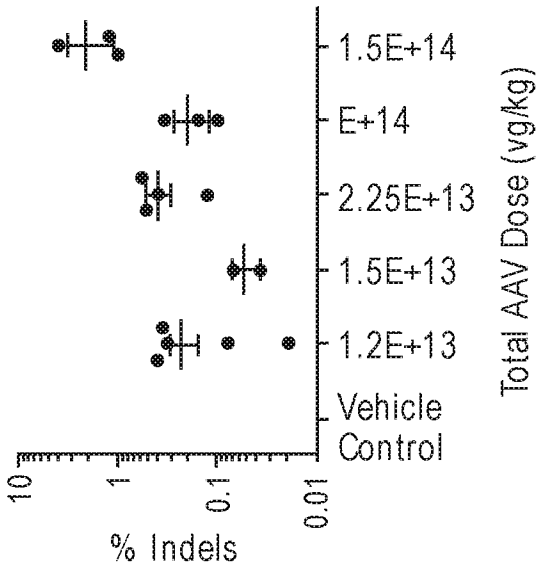


FIG. 12C

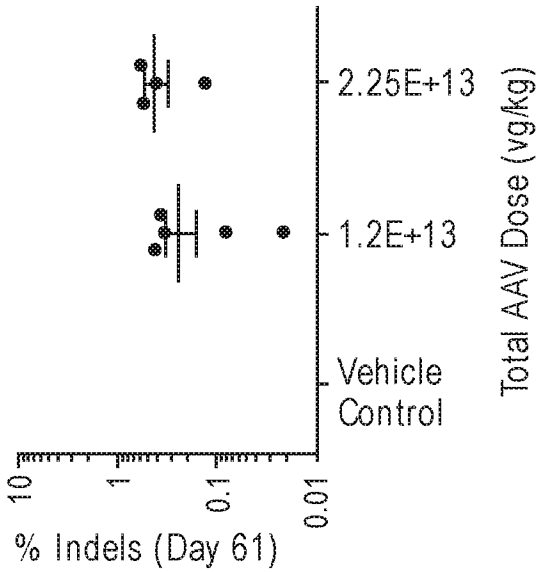


FIG. 12B

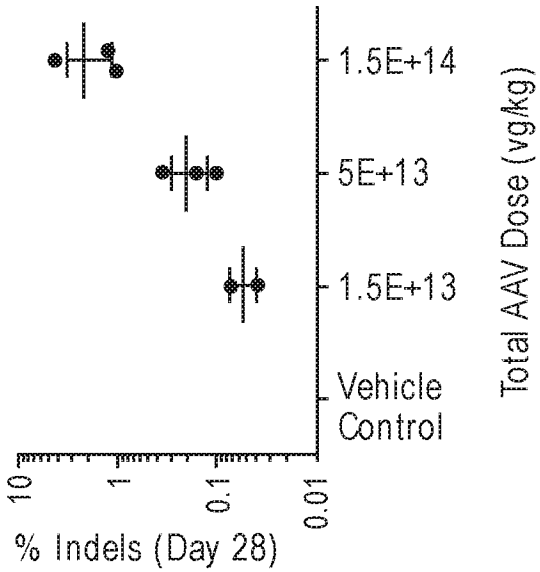


FIG. 12A

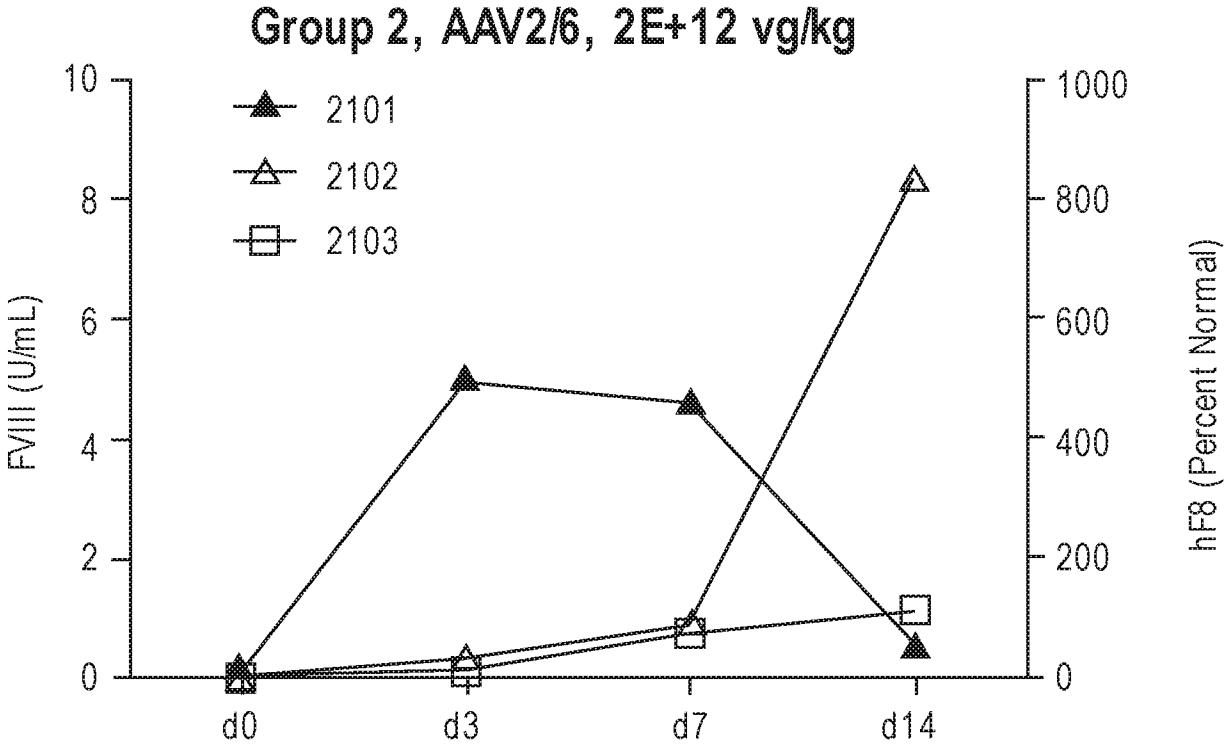


FIG. 13A

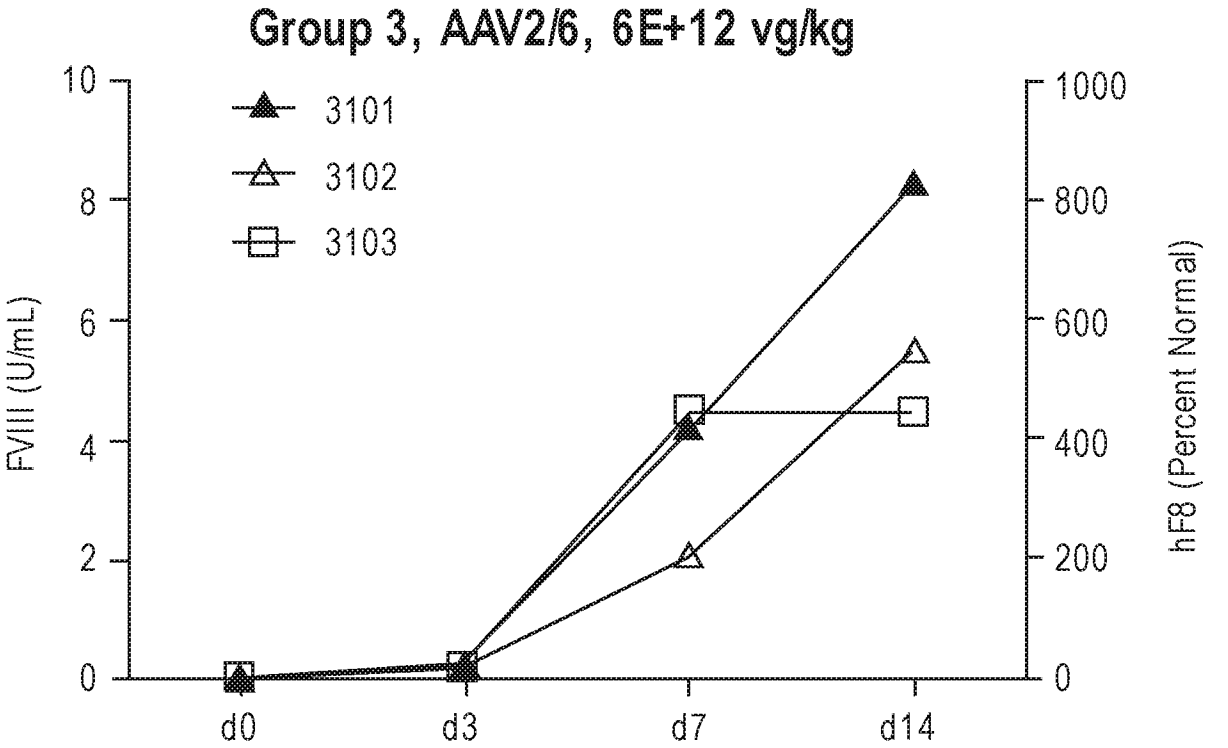


FIG. 13B

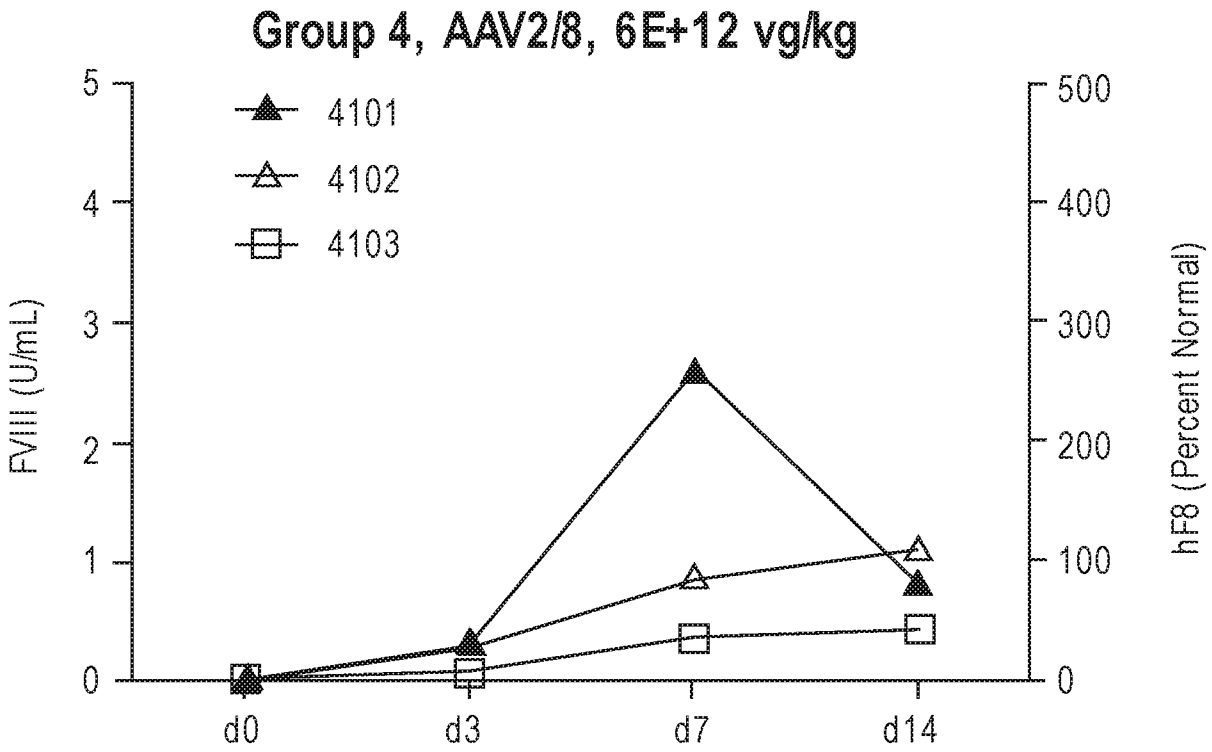


FIG. 13C

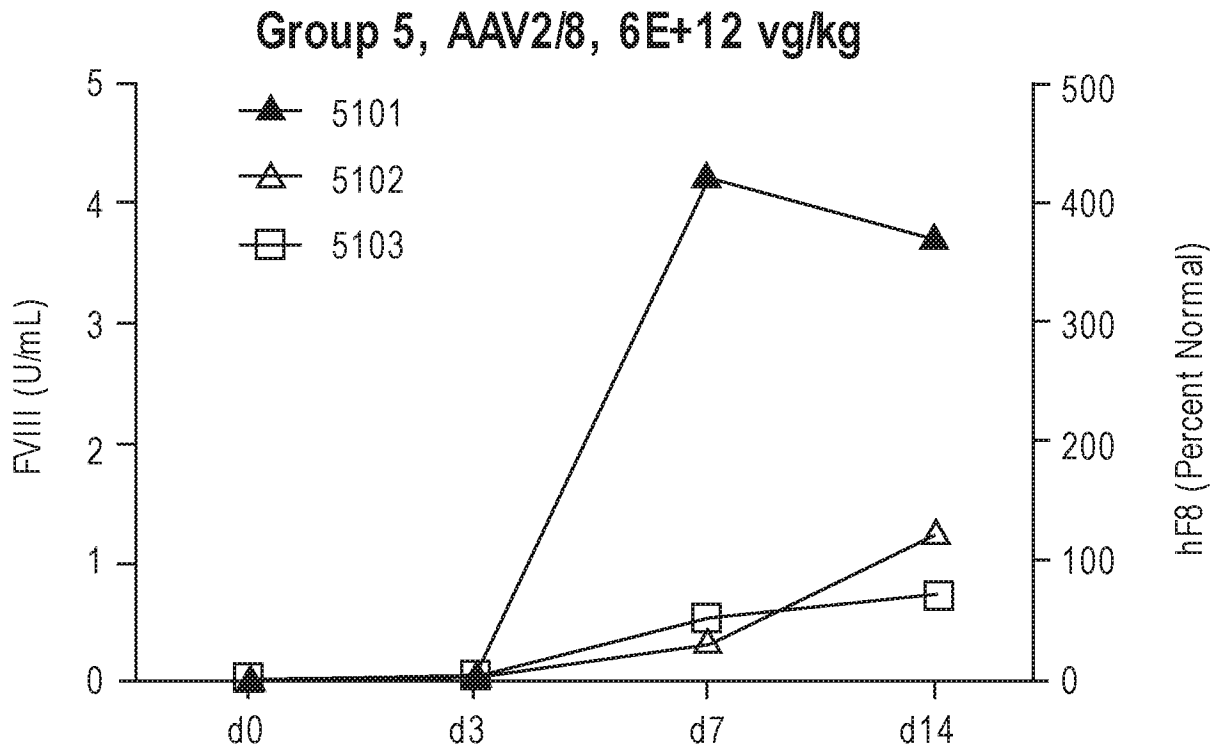


FIG. 13D

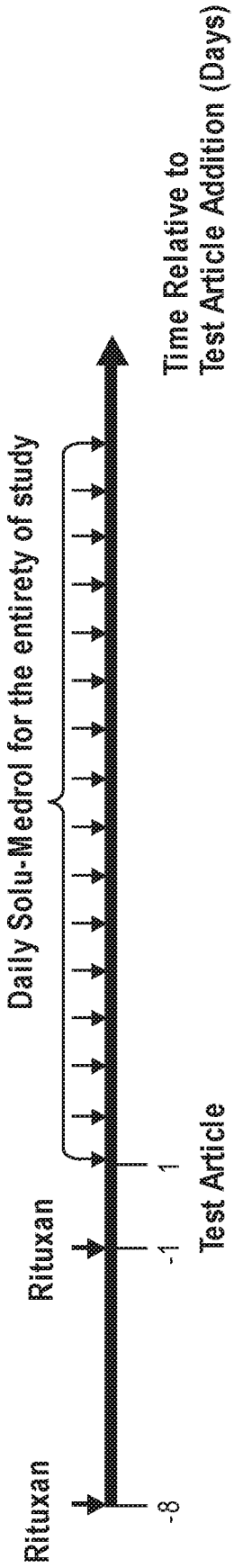

Time Relative to
Test Article Addition (Days)
Daily Solu-Medrol for the entirety of study
Rituxan
Rituxan
Test Article
Addition
-8
-1
1

FIG. 14

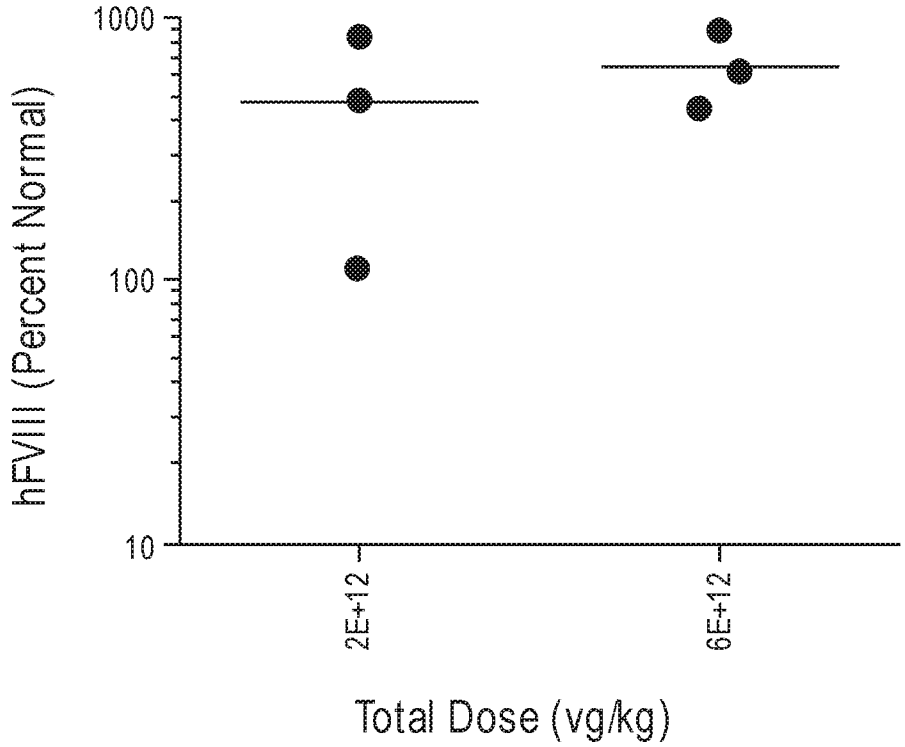


FIG. 15

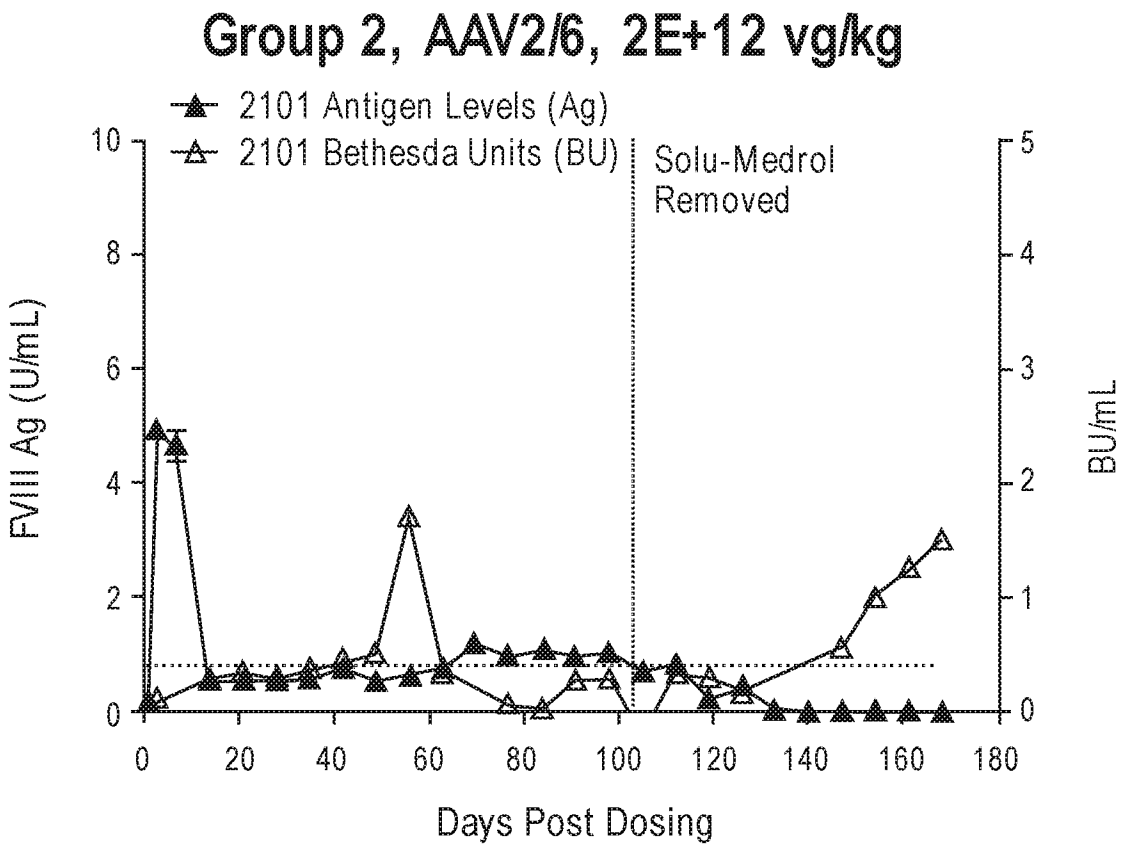


FIG. 16A

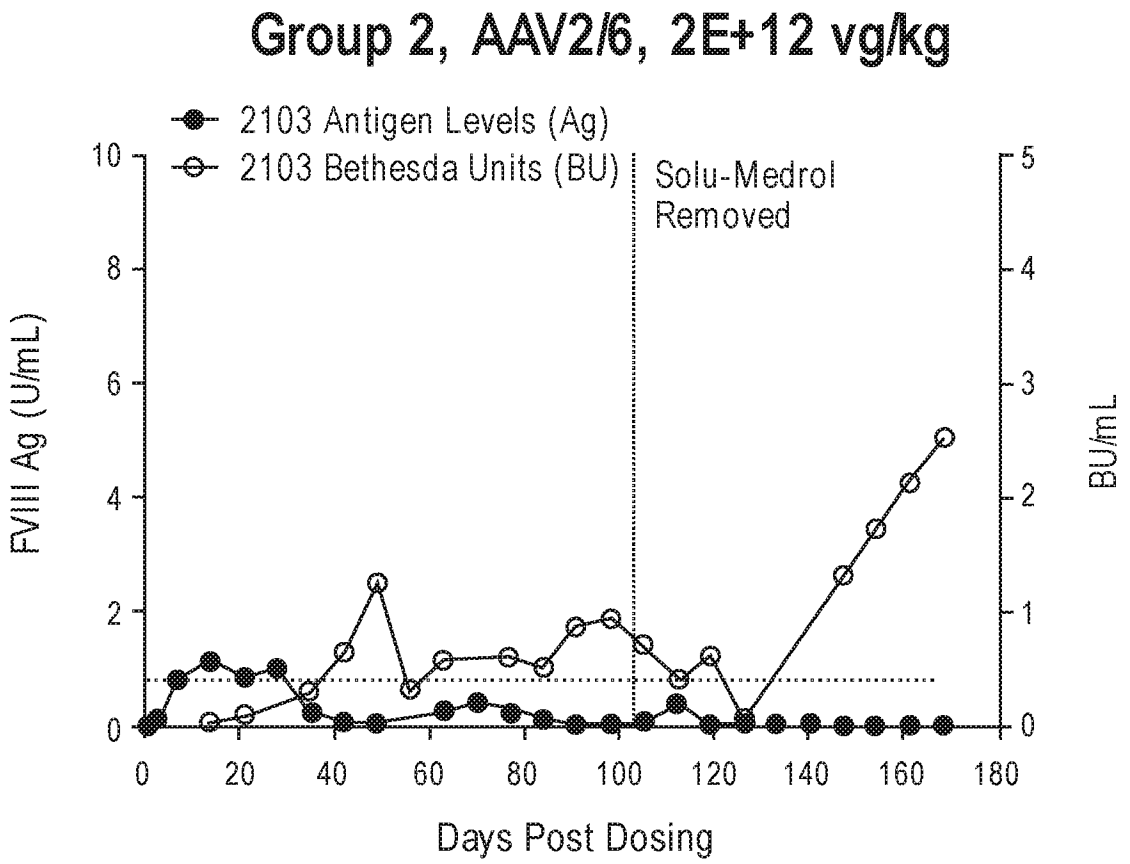


FIG. 16C

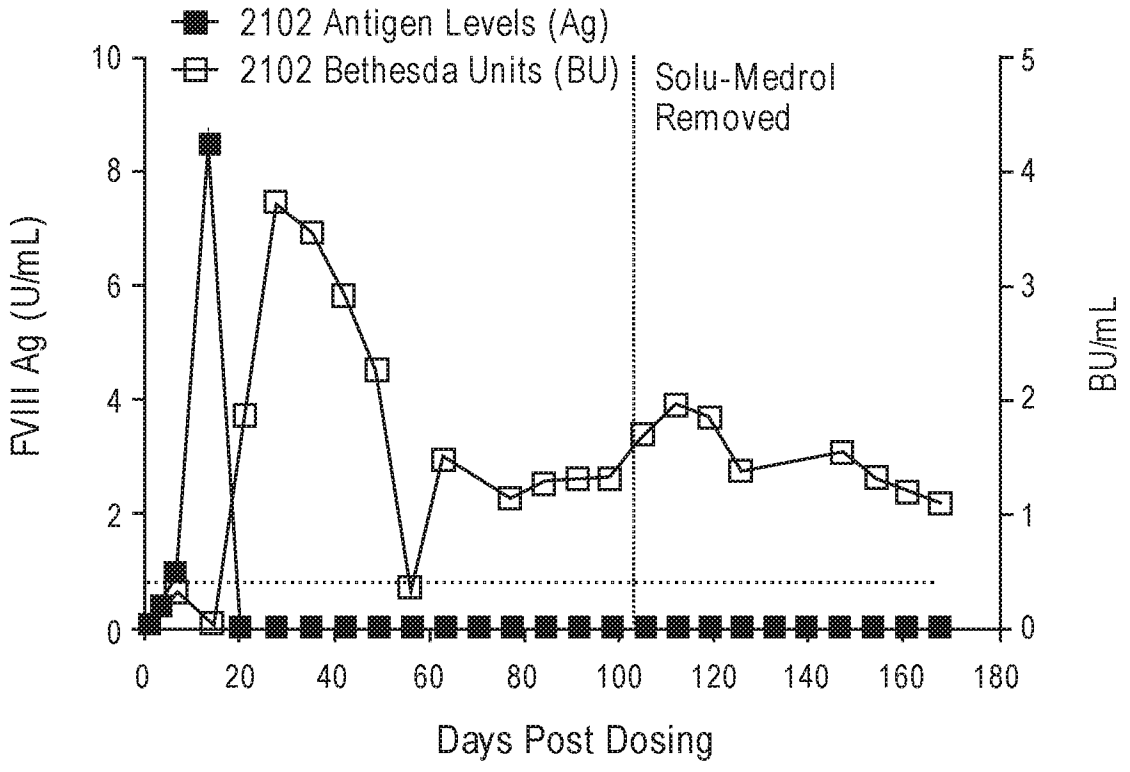


FIG. 16B

Group 3, AAV2/6, 6E+12 vg/kg

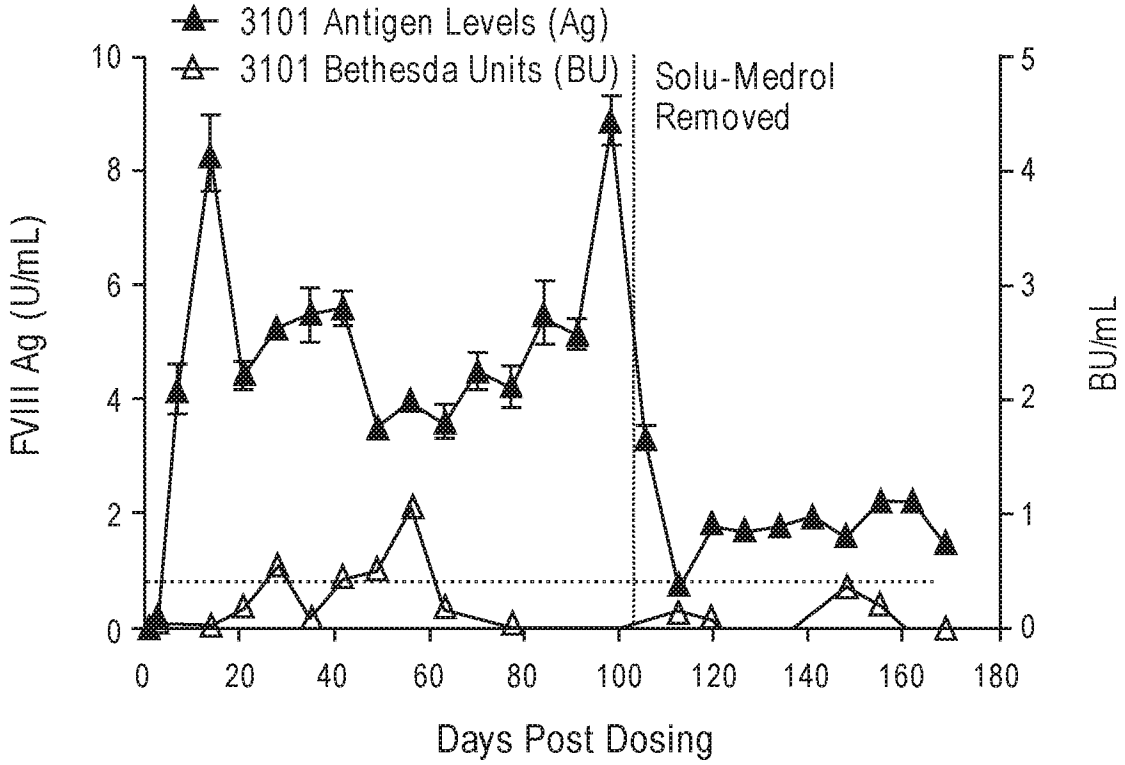


FIG. 17A

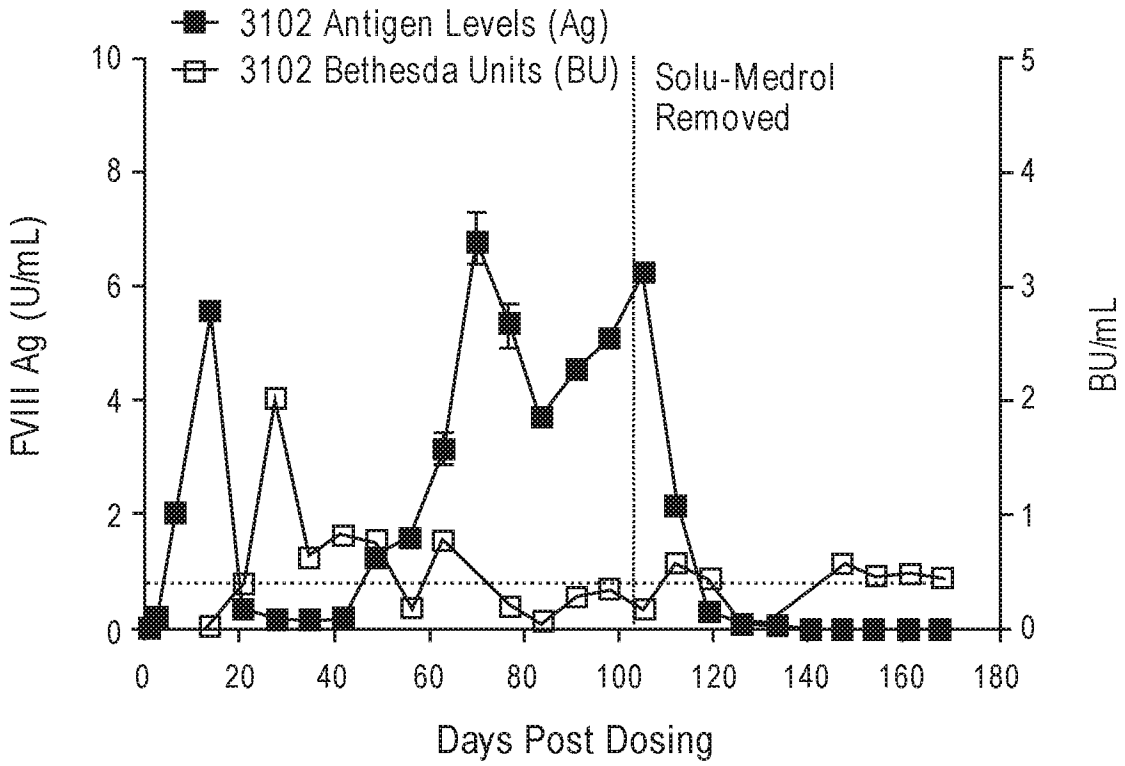


FIG. 17B

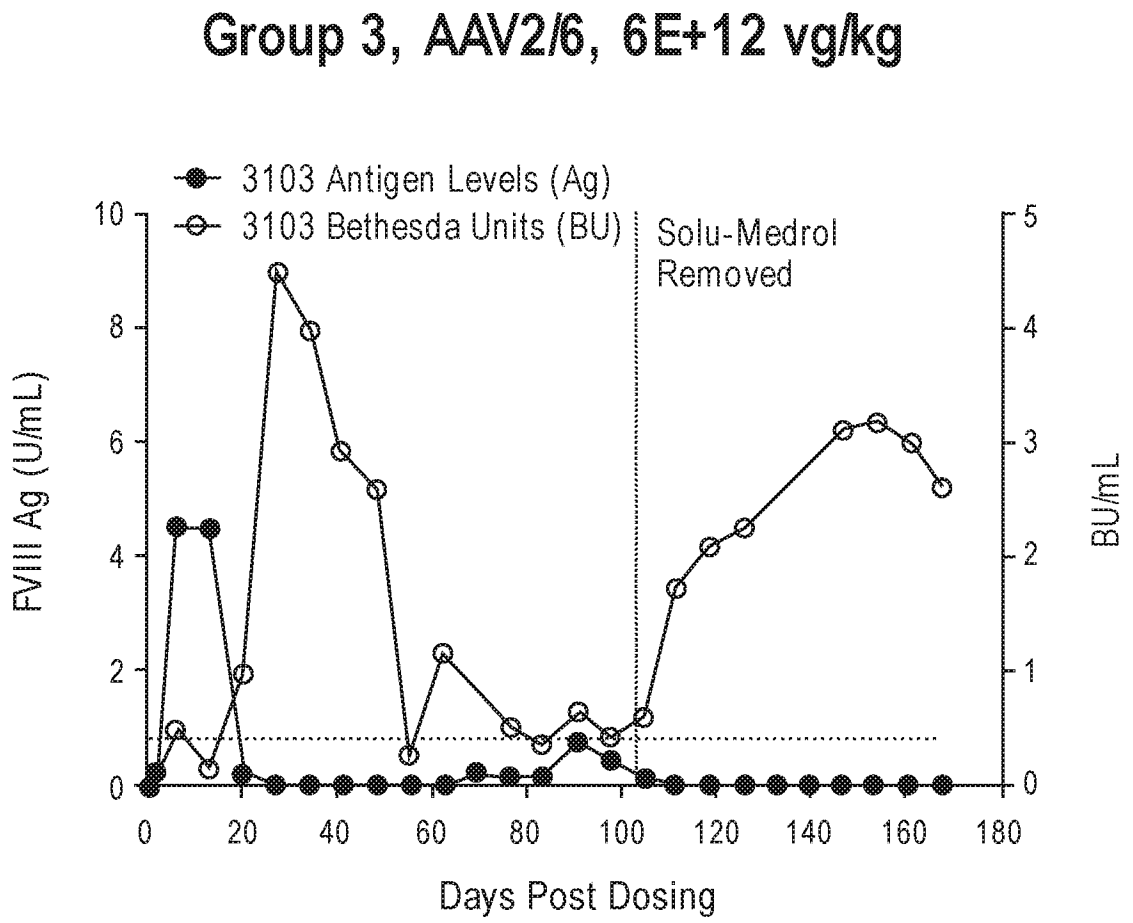

Group 3, AAV2/6, 6E+12 vg/kg
3103 Antigen Levels (Ag)
3103 Bethesda Units (BU)
Solu-Medrol Removed
FVIII Ag (U/mL)
BU/mL
Days Post Dosing
0
2
4
6
8
10
0
1
2
3
4
5
0
20
40
60
80
100
120
140
160
180

FIG. 17C

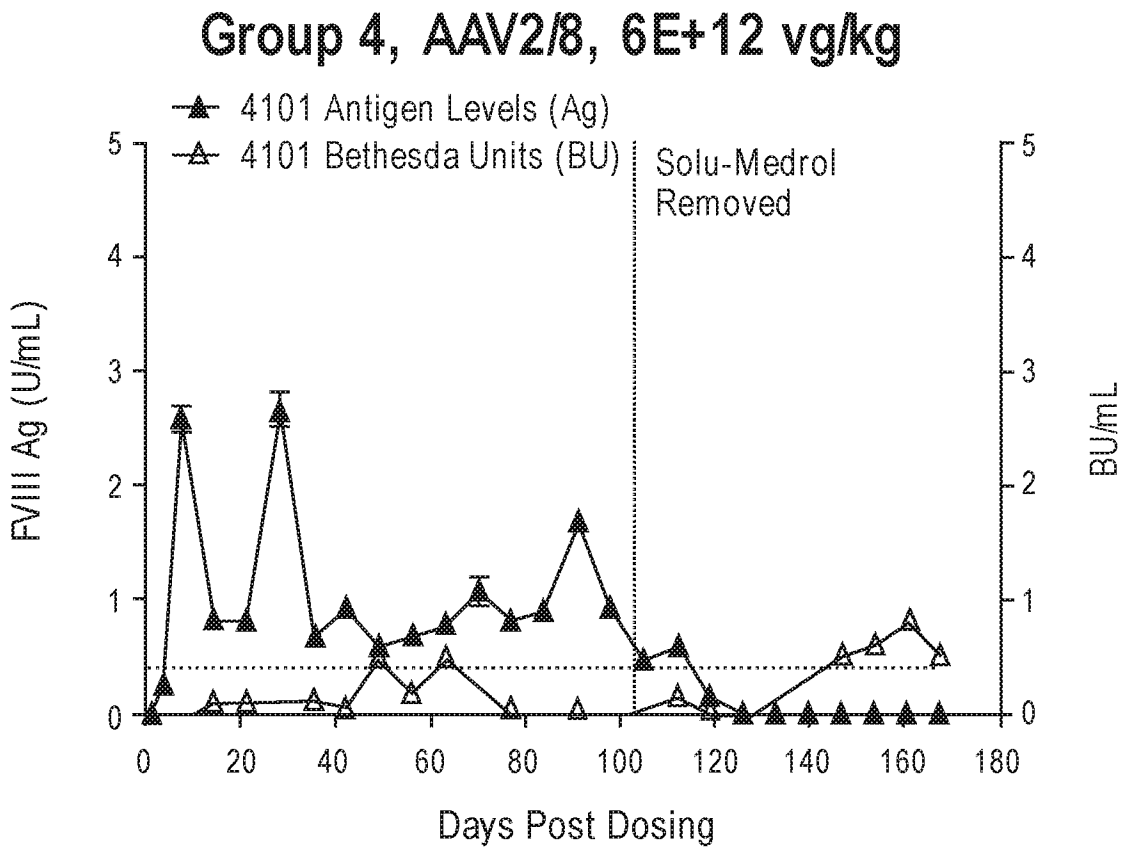


FIG. 18A

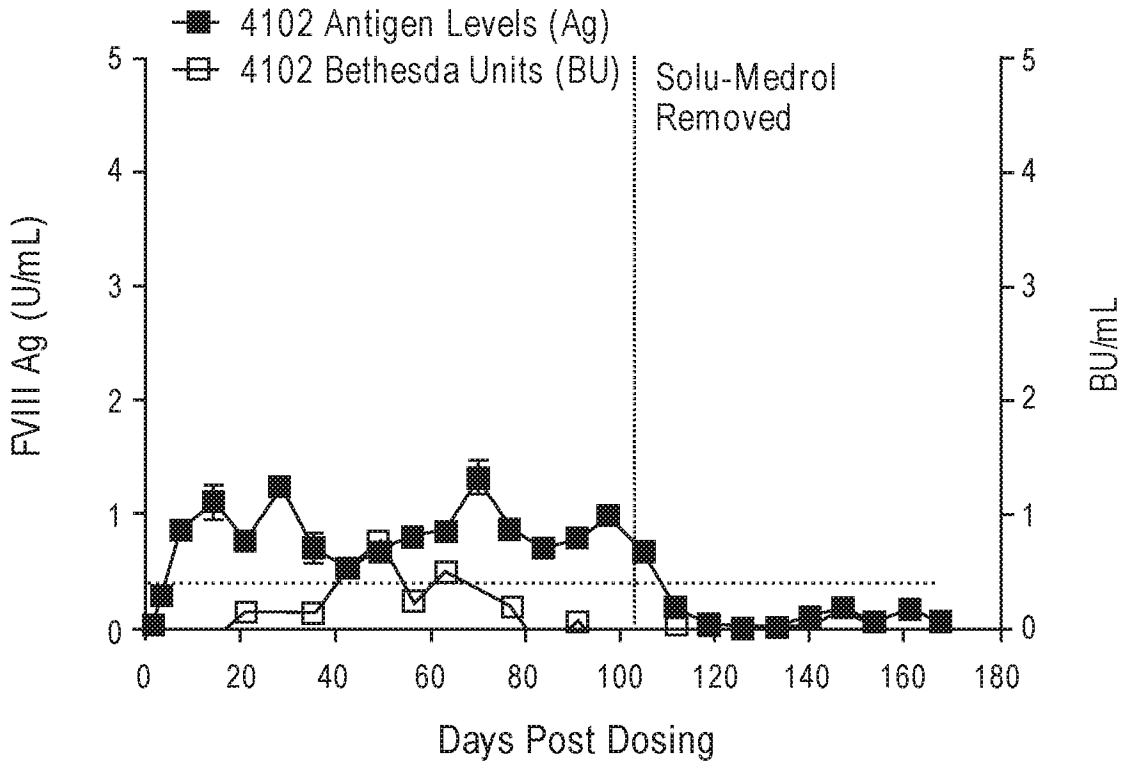


FIG. 18B

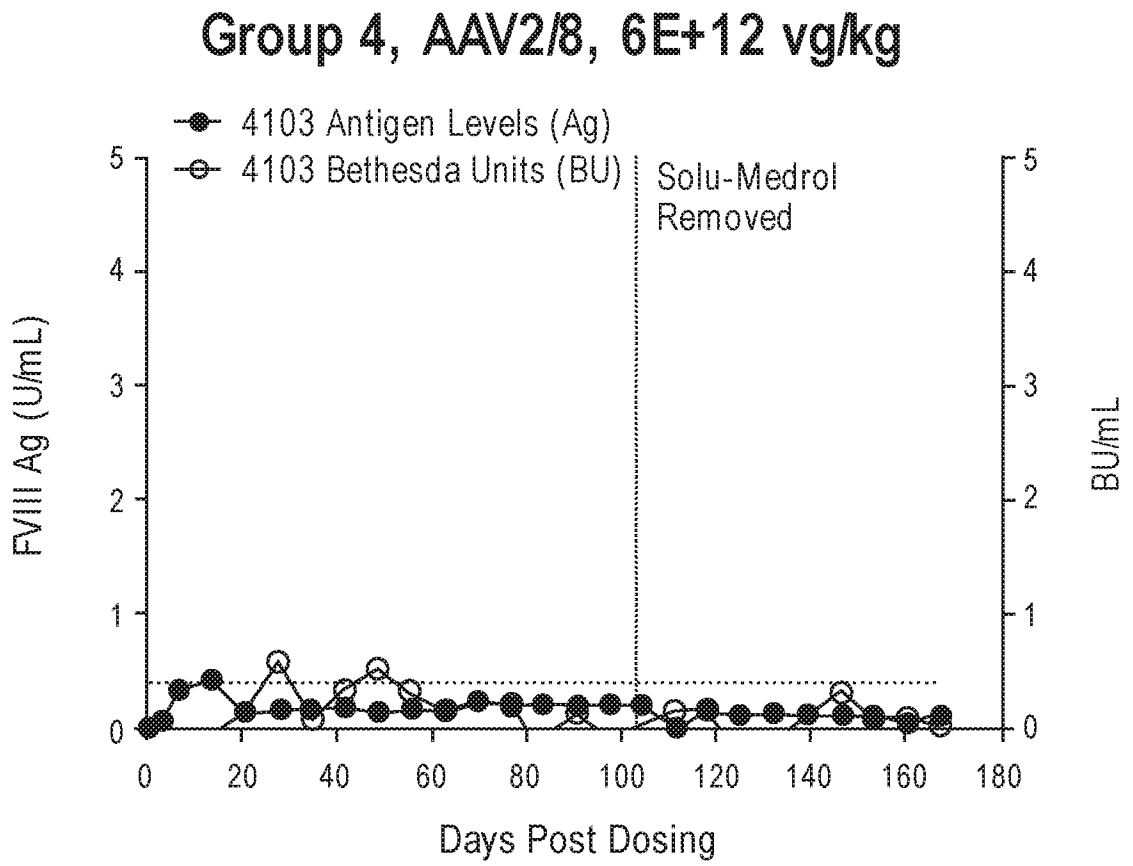


FIG. 18C

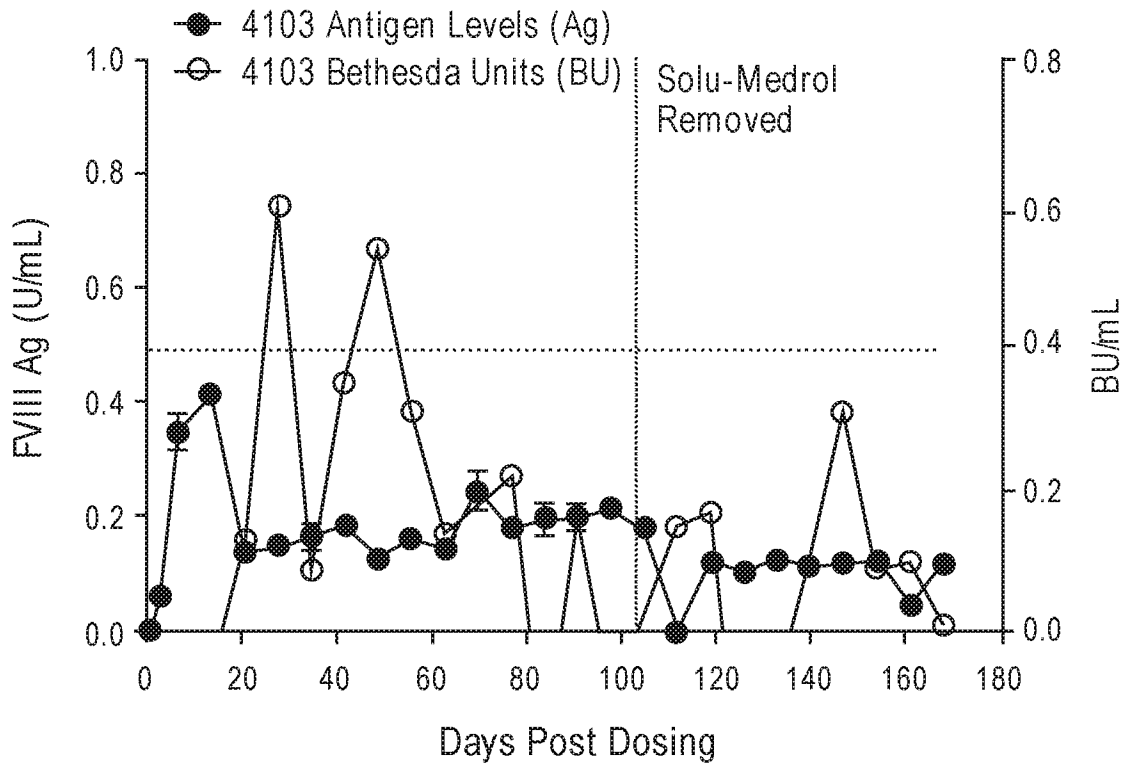


FIG. 18D

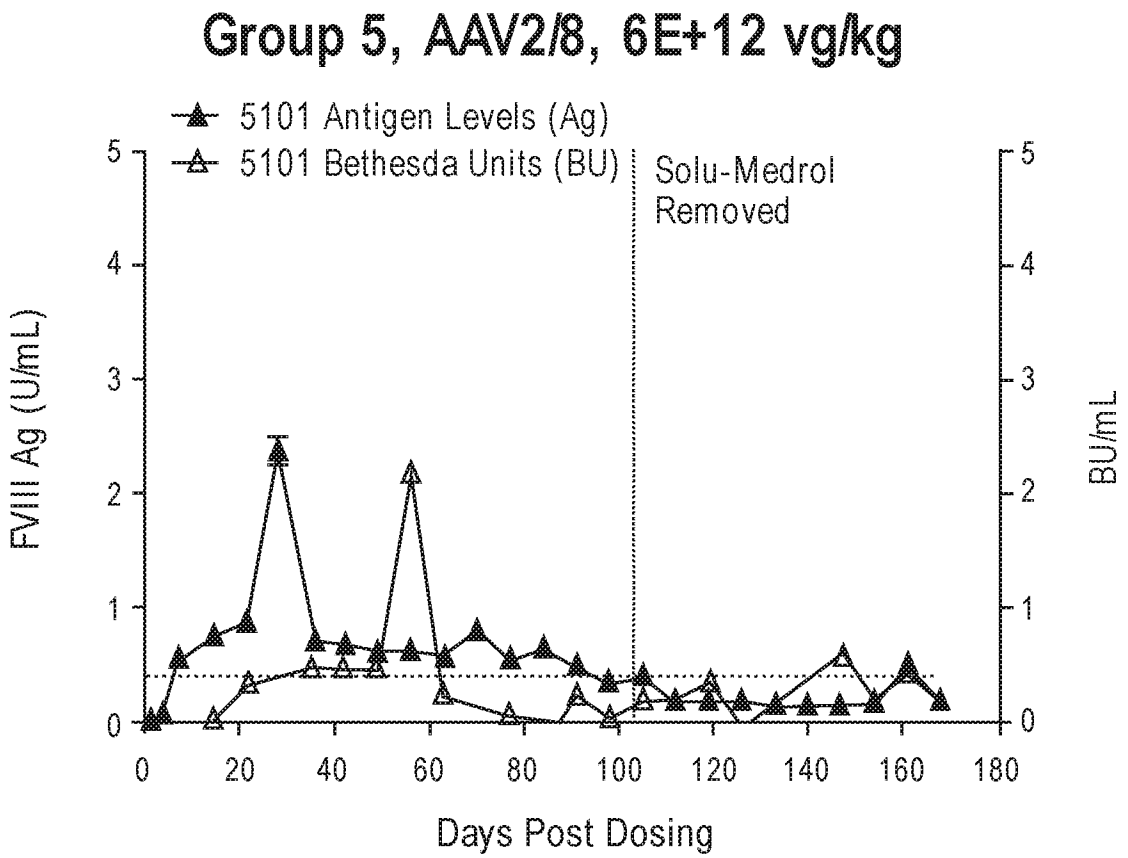


FIG. 19A

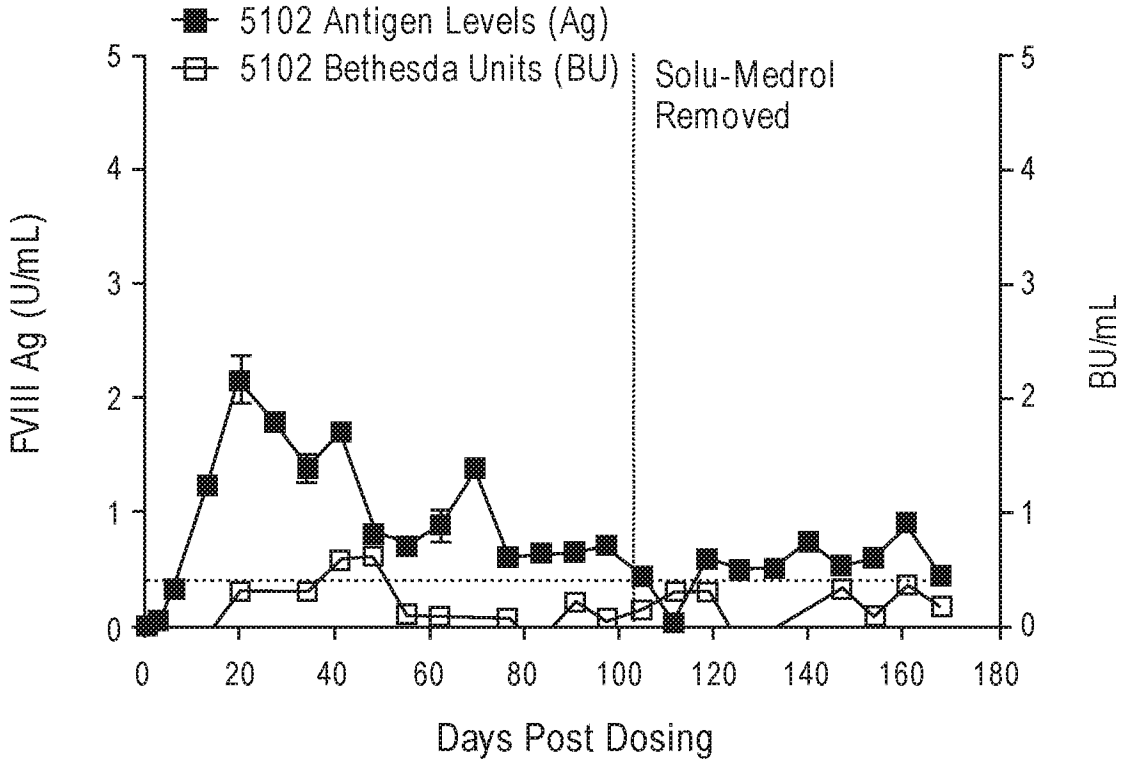


FIG. 19B

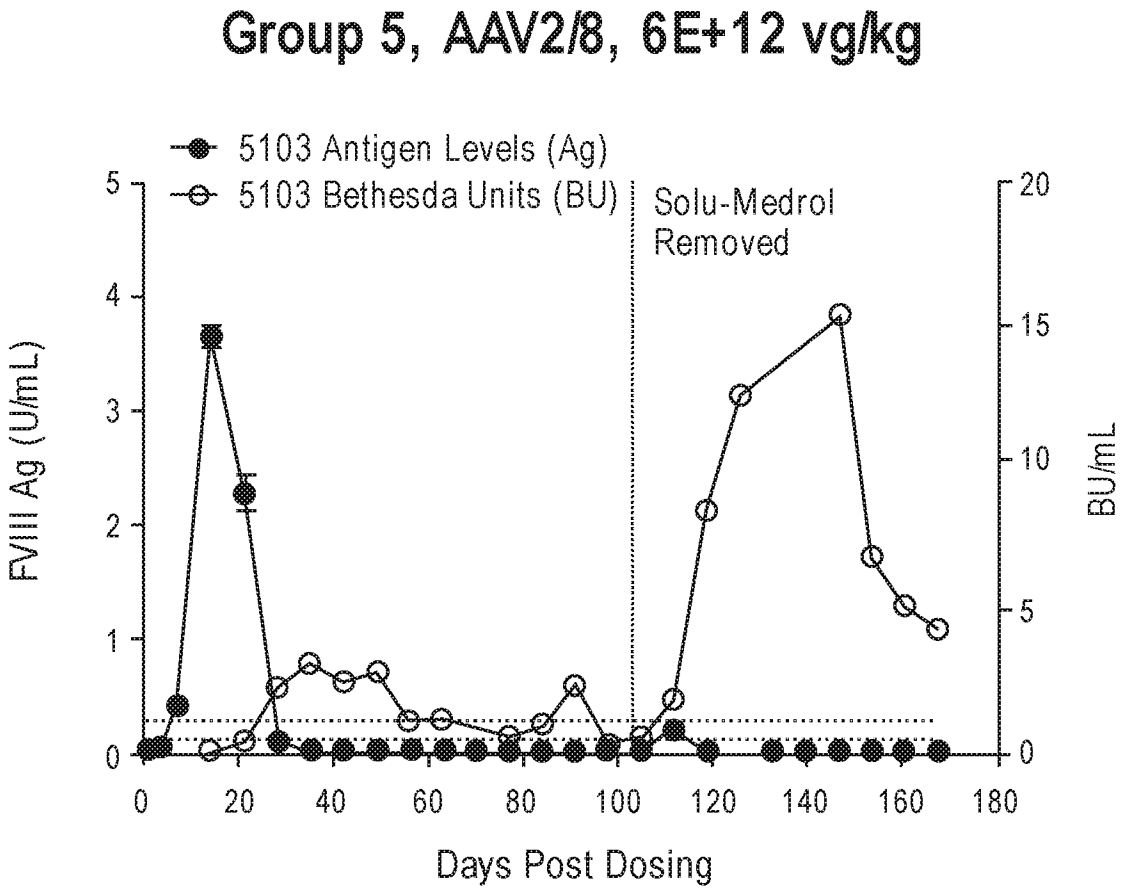

Group 5, AAV2/8, 6E+12 vg/kg
5103 Antigen Levels (Ag)
5103 Bethesda Units (BU)
Solu-Medrol Removed
FVIII Ag (U/mL)
BU/mL
Days Post Dosing
0
1
2
3
4
5
0
5
10
15
20
0
20
40
60
80
100
120
140
160
180

FIG. 19C

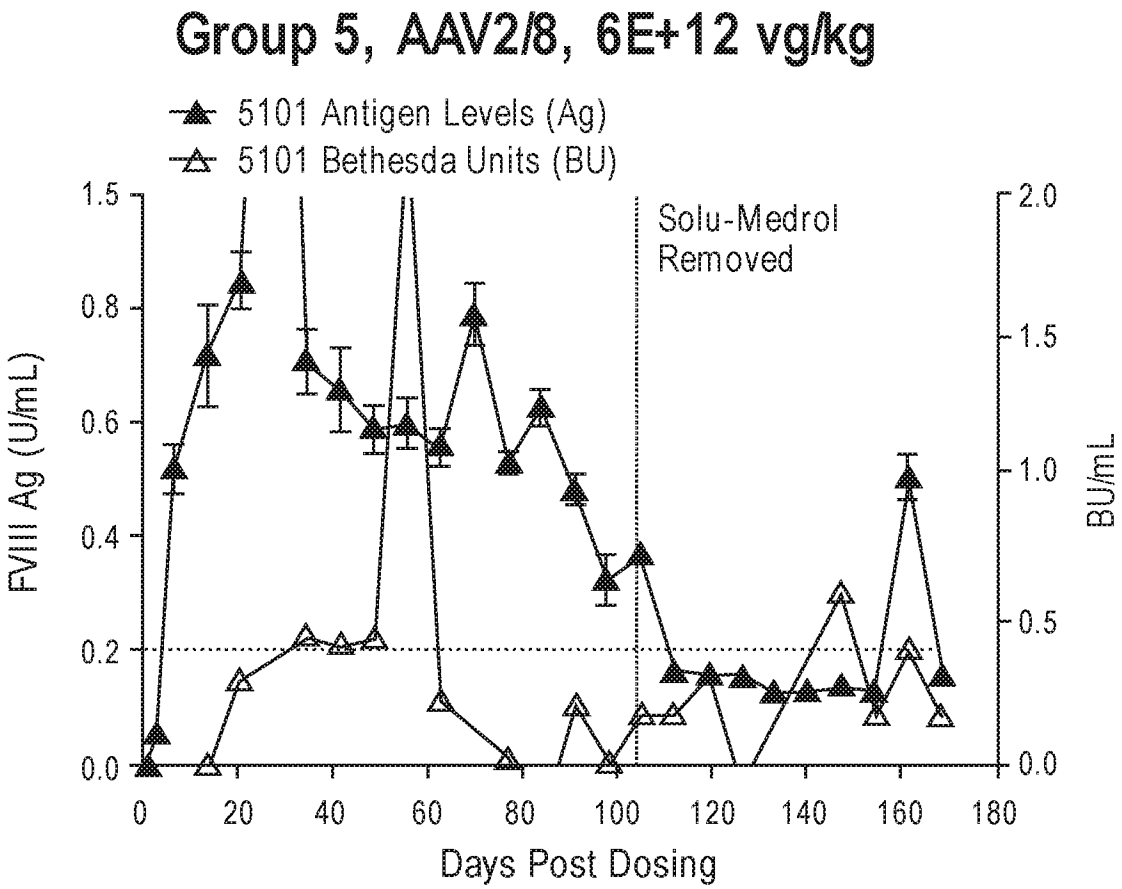


FIG. 19D

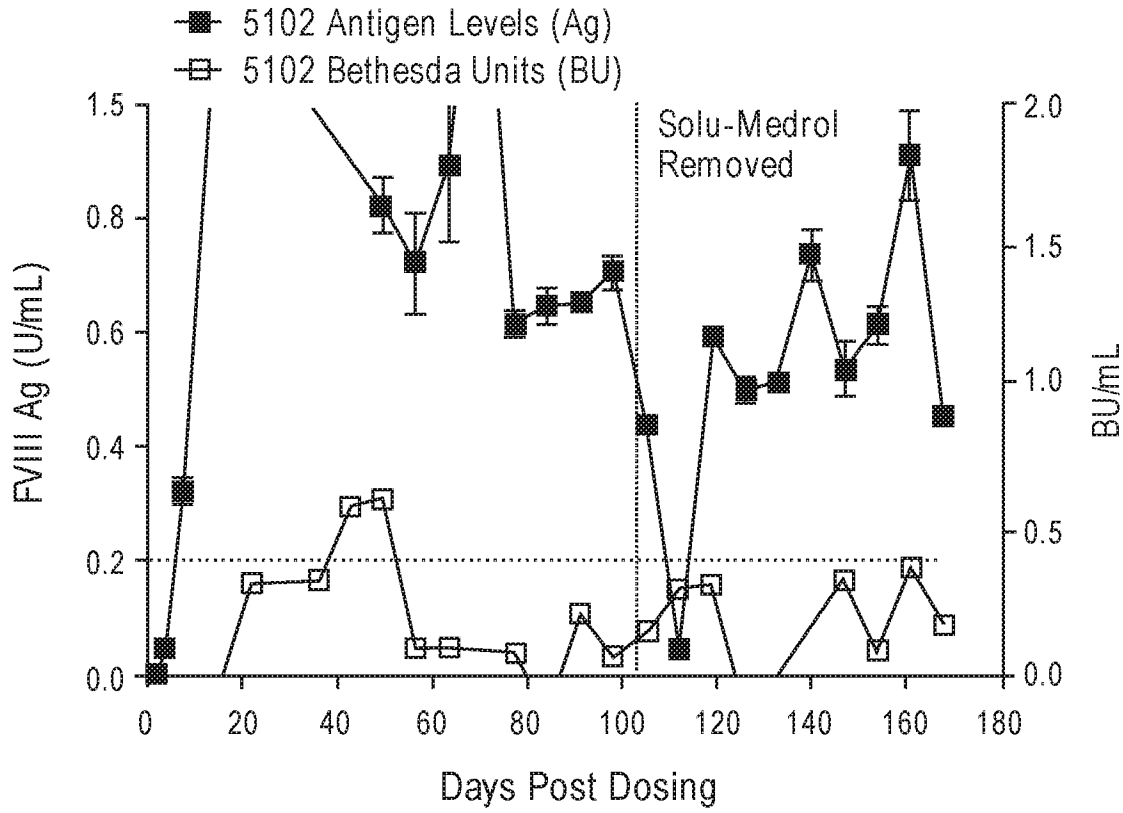


FIG. 19E

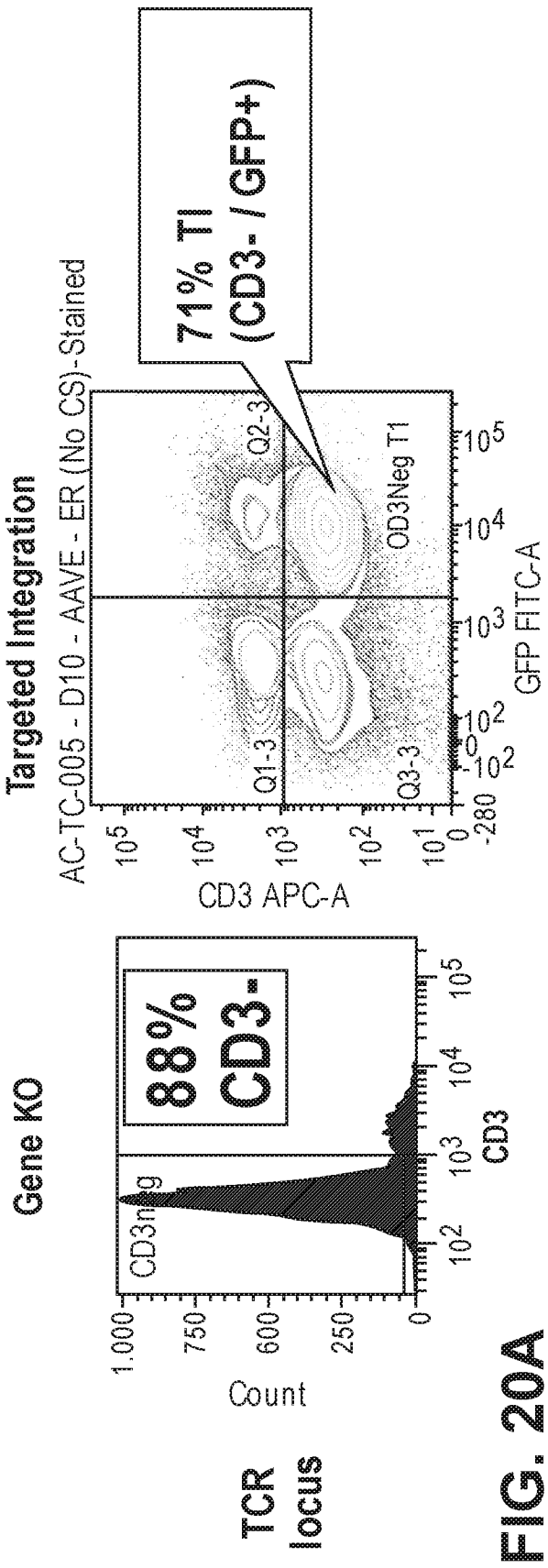


FIG. 20A

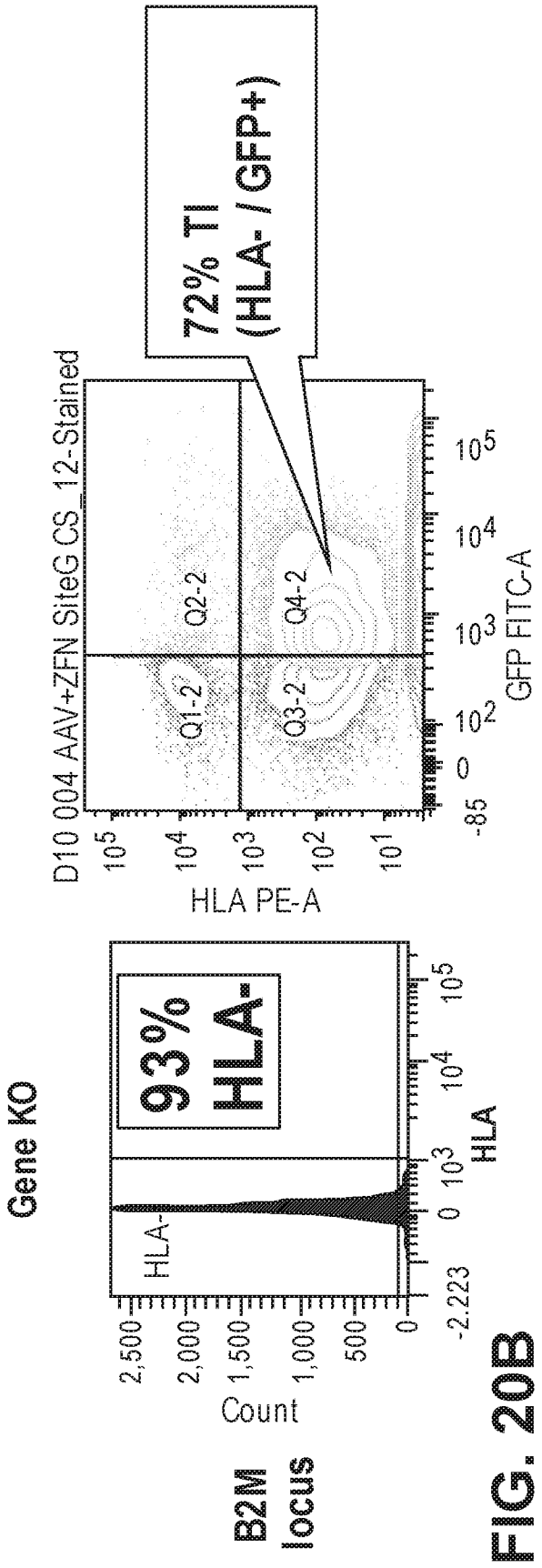


FIG. 20B

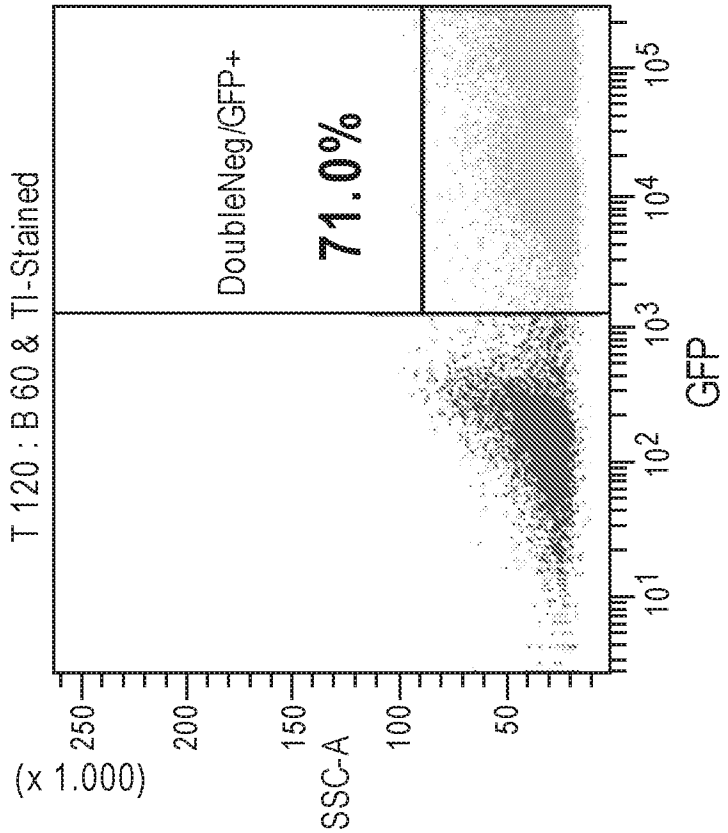


FIG. 20D

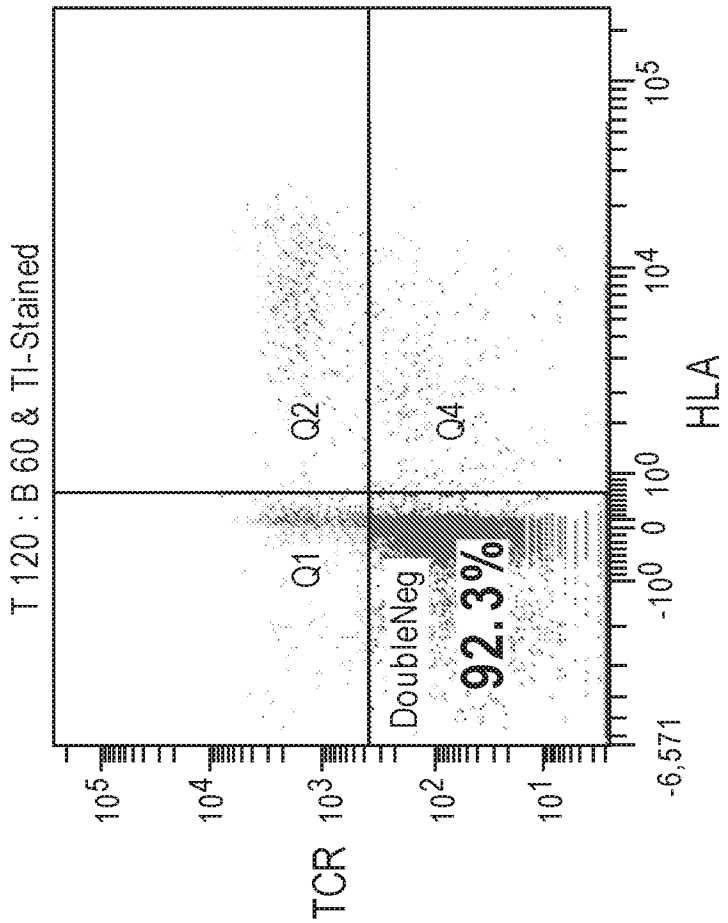


FIG. 20C

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/042024

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A01N 63/00; A61K 38/46; A61K 48/00 (2016.01)
CPC - A61K 48/0066; C12N 15/90; C12N 2830/42 (2016.08)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC - A01N 63/00; A61K 38/46; A61K 48/00
CPC - A61K 48/0066; C12N 15/90; C12N 2830/42

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 424/278.1; 435/325; 435/375; 514/44R (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Patbase, Google Patents, PubMed, Google

Search terms used: rAAV adeno associated virus vector transgene FVIII B-cell EGFR inhibitor AAV6 transfect

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2014/0296326 A1 (THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERV) 02 October 2014 (02.10.2014) entire document	1-3
X	WO 2014/089541 A2 (HAPLOMICS, INC. et al) 12 June 2014 (12.06.2014) entire document	10-15
A	BEANE et al. "Clinical Scale Zinc Finger Nuclease-mediated Gene Editing of PD-1 in Tumor Infiltrating Lymphocytes for the Treatment of Metastatic Melanoma," The American Society of Gene & Cell Therapy, 02 June 2015 (02.06.2015), Vol. 23, No. 8, Pgs. 1380-1390. entire document	1-3, 10-15
A	US 2003/0003583 A1 (HIRSCH et al) 02 January 2003 (02.01.2003) entire document	1-3, 10-15
P, X	WO 2015/164758 A1 (UNIVERSITY OF FLORIDA RESEARCH FOUNDATION, INC.) 29 October 2015 (29.10.2015) entire document	1-3, 10-15

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 October 2016	24 OCT 2016
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/042024

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☐ forming part of the international application as filed:

      ☐ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☒ furnished subsequent to the international filing date for the purposes of international search only:

      ☒ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☒ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/042024

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 4-9
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.