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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
08 March 2018 (08.03.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/039783 A1

(51) International Patent Classification:
C12N 15/11 (2006.01) *C12N 15/90* (2006.01)
C07K 14/705 (2006.01) *C12N 9/14* (2006.01)
C12N 15/09 (2006.01) *C12N 9/22* (2006.01)
C12N 15/55 (2006.01) *C12Q 1/02* (2006.01)
C12N 15/63 (2006.01) *C12Q 1/68* (2006.01)
C12N 15/85 (2006.01)

(21) International Application Number:
PCT/CA2017/051011

(22) International Filing Date:
29 August 2017 (29.08.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/381,292 30 August 2016 (30.08.2016) US
62/434,650 15 December 2016 (15.12.2016) US

(71) Applicant: **UNIVERSITÉ LAVAL** [CA/CA]; Pavillon des Sciences de l'Éducation, 2320, rue des Bibliothèques, Suite 1434, Quebec, Québec G1V 0A6 (CA).

(72) Inventor: **DOYON, Yannick**; 1282 Roland-Desmeules, Quebec, Québec G1X 4Y3 (CA).

(74) Agent: **GOUDREAU GAGE DUBUC**; 2000 McGill College, Suite 2200, Montreal, Québec H3A 3H3 (CA).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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.

WO 2018/039783 A1

(54) Title: SELECTION SYSTEMS AND METHODS FOR GENOME EDITING

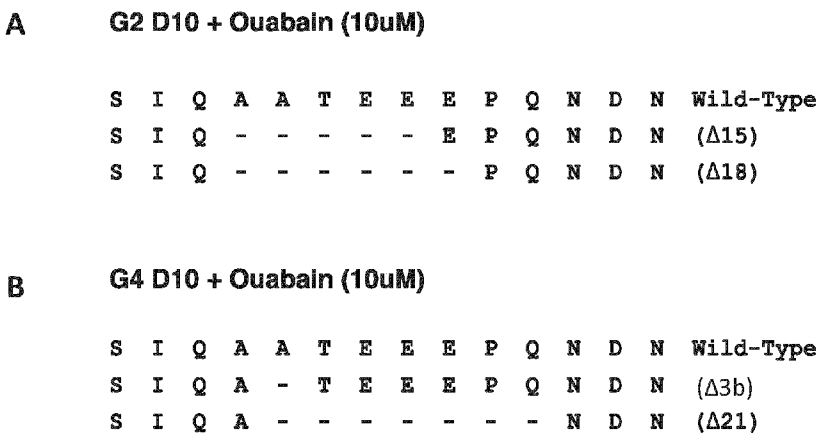


FIG. 10

(57) Abstract: Described herein are methods and products for modification (e.g., deletion, insertion or substitution) of a target gene in a cell. In an aspect, the method entails introduction of a modification into an endogenous gene (e.g., ATP1A) locus thereby to confer resistance to a drug (e.g., ouabain or a derivative thereof), and co-introduction of a modification in a target gene of interest. Drug-based selection may be used to enrich cells comprising said modified genes within a population of cells, thereby to select for such modified genes. Novel polypeptides, including novel ATP1A1 polypeptides, are also described, which confer drug (e.g., ouabain) resistance, as well as corresponding nucleic acids encoding same, and vectors and host cells comprising same.

[Continued on next page]

(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, 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))*
— *with sequence listing part of description (Rule 5.2(a))*

TITLE OF THE INVENTION

SELECTION SYSTEMS AND METHODS FOR GENOME EDITING

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the benefit of U.S. Provisional Application Serial No. 62/381,292, filed on August 30, 2016, and on U.S. Provisional Application Serial No. 62/434,650 filed on December 15, 2016. All documents above are incorporated herein by reference in their entirety.

REFERENCE TO SEQUENCE LISTING

[0002] This application contains a Sequence Listing in computer readable form entitled "11229_373_SeqList.txt", created August 29, 2017 and having a size of about 327 KB. The computer readable form is incorporated herein by reference.

FIELD OF THE INVENTION

[0003] The present invention relates to genome editing. More specifically, the present invention is concerned with reagents, products and methods for modifying the genome of a cell and for selecting and isolating cells bearing the targeted modification.

BACKGROUND OF THE INVENTION

[0004] Designer nucleases such as zinc-finger nucleases (ZFNs), TAL effector nucleases (TALENs), and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) systems greatly simplify the generation of cell lines with tailored modifications enabling the study of endogenous genetic elements in human cells (1-4). In particular, CRISPR-based technologies enable researchers to introduce double-strand breaks (DSBs) into virtually any DNA sequence using sequence-specific single guide RNAs (sgRNAs) that target Cas9 and Cpf1 nucleases to cleave the matching target. When targeting two or more genomic loci simultaneously, multiple DSBs can be created in the same cell via expression of distinct sgRNAs, a process referred to as multiplexing (5). The ensuing DSBs activate two competing repair pathways, namely non-homologous end joining (NHEJ) and homology-directed repair (HDR) to facilitate the precise modification of the targeted endogenous locus. NHEJ repairs the lesion by directly rejoining the two DSB ends without the need for a repair template resulting in either precise re-ligation or formation of small insertion or deletion mutations (indels) at the break site. In contrast, definite sequence changes can be introduced when the HDR machinery uses an exogenous DNA template with sequence homology to the DSB site to repair the lesion (reviewed in (6)).

[0005] At its most basic level, higher genome editing frequencies are associated with higher nuclease levels and activity. Consequently, several approaches have been implemented to capture and isolate these subpopulations of cells in order to expedite the generation of isogenic knockout and knock-in clones. First, direct coupling of expression of fluorescent protein and nucleases via 2A peptide sequences combined with fluorescence-activated cell sorting (FACS) allows for efficient isolation of cell populations with increasingly higher nuclease expression levels, which translates into increasingly higher genome editing rates (11-13). Fluorescence-based surrogate target gene reporters have also been successfully used to enrich for cells with high nuclease activity (14, 15). A limitation of both methods is that single-cell FACS enrichment is not possible for certain sensitive cell lines whereas construction of a surrogate reporter for each target site decreases the throughput of genome editing. Finally, enhancing HDR-mediated processes by altering cell cycle parameters, the timing of nuclease expression, chemical inhibition of NHEJ and use of HDR agonists show promising results but their general applicability and effects on specificity need to be thoroughly evaluated (16-22).

[0006] Other genetic approaches based on the creation of "classical" gain-of-function alleles have been developed in the worm *Caenorhabditis elegans*, referred to as "co-conversion"/"co-CRISPR" (23-26). A related approach has been described to isolate human cells harboring NHEJ-driven mutations by co-targeting the X-linked hypoxanthine phosphoribosyltransferase (*HPRT1*) gene (27). Further, "co-incidental insertion" (COIN) has been described in mouse embryonic stem cells (28).

[0007] Thus, there remains a need for improving genome editing methods and the selection of targeted genomic modification events.

[0008] The present description refers to a number of documents, the content of which is herein incorporated by reference in their entirety.

SUMMARY OF THE INVENTION

[0009] The present invention makes use of an endogenous gene and a drug (e.g., a highly potent small molecule (e.g., cardiotonic steroids such as ouabain)) to prepare and rapidly select and enrich for genome edited cells. In an embodiment, the process does not involve the introduction of superfluous "exogenous" DNA material. Compared to other systems (e.g., the HPRT/6-TG system), the method disclosed herein has the advantage of using a selection agent (ouabain or ouabain derivatives) that is non-mutagenic. Furthermore, the strategy employed was shown to provide the unexpected advantage of permitting the selection of NHEJ or HDR-based gene modification (i.e., favoring NHEJ or HDR-based gene modification), the latter being a highly sought-for characteristic in the field.

[0010] The genome editing approach described herein is based in part on Na+,K+-ATPase and inhibitors

thereof, notably cardiotonic steroids such as ouabain. Cardiotonic steroids (CTSs), such as ouabain (PubChem CID: 439501), constitute a broad class of specific Na+,K+-ATPase inhibitors prescribed for congestive heart failure for more than 2 centuries (29). Ouabain is a highly potent plant-derived inhibitor of the Na+,K+-ATPase (Enzyme Entry E.C. 3.6.1.96) (29-31). The Na+,K+-ATPase, encoded by the ATP1A1 gene (GeneCards ID: GC01P116372; HGNC: 799; Entrez Gene: 476; Ensembl: ENSG00000163399; OMIM: 182310; UniProtKB: P05023) is the ion pump responsible for maintenance of the electrochemical gradients of Na+ and K+ across the plasma membrane of animal cells. It is an essential and ubiquitously expressed enzyme functioning as a heteromeric complex consisting of a large α-subunit (ATP1A1), which is responsible for ATP hydrolysis and ion transport, and a β-subunit, acting as a chaperone (29, 30). Mutagenesis and modeling studies have been carried out on the structure-activity relationships of ouabain (29-34 and 47). There are four isoforms of the α subunit (α1, α2, α3, and α4), each encoded by separate genes (ATP1A1 (GCID GC01P116372, NM_000702.3 (SEQ ID NO: 1) and NP_0011533705.1 (SEQ ID NO: 3)), ATP1A2 (GCID GC01P160085, NM_000702.3 (SEQ ID NO: 459) and NP_000693.1 (SEQ ID NO: 460)), ATP1A3 (GCID GC19M041966, NM_152296.4 (SEQ ID NO: 461) and NP_689509.1 (SEQ ID NO: 462)) and ATP1A4 (GCID GC01P160151, NM_144699.3 (SEQ ID NO: 463) and NP_653300.2 (SEQ ID NO: 464)), respectively). The α1 isoform is ubiquitously expressed, whereas the α2 isoform is expressed mainly in skeletal, heart, and smooth muscle, brain, lung, and adipocytes. The α3 isoform occurs mainly in neurons and ovaries as well as in developing hearts of rat and in adult human heart. This isoform also occurs in white blood cells. The α4 isoform of the Na,K-ATPase is found in sperm. In humans, all ATP1A isoforms are sensitive to ouabain and specific mutations within the ATP1A1 isoform have been described to confer resistance to the drug. Based on the sequence conservation of ATP1A isoforms (FIG. 37), it is predicted they can be converted to ouabain-resistant proteins by introducing mutations in corresponding residues. As such, the ouabain sensitivity of ATP1A2 has been abolished by introducing two amino-acid mutations, Q116R and N127D, in the first extracellular loop, which is part of the ouabain-binding site of Na+/K+ ATPase (see (58) and FIGs 1 and 37).

[0011] As shown herein, using for example CRISPR, gain-of-function alleles in an endogenous gene (e.g., ATP1A1) were generated via either the NHEJ or the HDR DNA repair pathways and used to co-select for mechanistically related editing events at a second locus of interest with selected drugs (e.g., ouabain). This strategy was portable to many sgRNAs, independent of cell type and co-targets, and provides a general solution for facilitating the isolation of genome-edited human cells.

[0012] Accordingly, the present invention is useful for increasing the efficiency of selection, enrichment and identification of cells in which genome editing has likely occurred. It can be used in basic research contexts as well as for therapeutic applications.

[0013] Disclosed herein are methods, reagents and products for modification (e.g., a deletion, an insertion (e.g.,

of a transgene), or a substitution) of a genetic locus within a cell, and/or increasing the frequency of gene modification in a target polynucleotide locus. More specifically, the present invention relates to nucleases and methods derived therefrom for the targeted modification of an endogenous gene locus (e.g., an ATP1A gene locus, e.g., ATP1A1, ATP1A2, ATP1A3 or ATP1A4) and for increasing the frequency of gene modification at a second, target polynucleotide locus.

[0014] In an aspect, in accordance with the present invention, there is provided a method for nuclease-based modification of a target polynucleotide in one or more target cells within a population of cells and for enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide within the population, the method comprising:

(i) introducing, in the one or more target cells, by homology-directed repair (HDR) or non-homologous end joining (NHEJ), a mutation in (a) an endogenous gene locus, wherein the mutation in the endogenous gene locus confers resistance to a drug, and (b) the target polynucleotide, in the one or more target cells;

(ii) contacting the population of cells with the drug; and

(iii) selecting the one or more target cells by virtue of increased tolerance to the drug, thereby enriching the number of target cells comprising the nuclease-based modification in the target polynucleotide within the population of cells.

[0015] The mutations (e.g., a first mutation in the target polynucleotide and a second mutation in the endogenous gene locus) are introduced by nuclease(s) specific for the target polynucleotide and endogenous gene locus (e.g., CRISPR nucleases, TALENs, Meganucleases, zinc finger nucleases, etc.).

[0016] In embodiments, the nuclease-based modification in the target polynucleotide (the first mutation) in methods of the present invention is introduced by HDR and step (i) further comprises introducing into the one or more target cells a donor nucleic acid comprising the modification. In embodiments, the nuclease-based modification in the endogenous gene locus (the second mutation) in methods of the present invention is introduced by HDR and step (i) further comprises introducing into the one or more target cells a second donor nucleic acid comprising the second modification.

[0017] In embodiments, introducing a mutation in the endogenous gene locus involves introducing one or more double-stranded breaks (DSBs) in an exon of the gene locus using a nuclease specific for the gene locus (e.g., a zinc finger nuclease, a Meganuclease, a TALEN or a CRISPR nuclease). The introduction of a DSB in an exon of the gene locus was found to advantageously favor nuclease-based modifications by NHEJ.

[0018] In other embodiments, introducing a mutation in the endogenous gene locus involves introducing a double-stranded break (DSB) in an intron of the gene locus using a nuclease specific for the gene locus (e.g., a zinc

finger nuclease, a meganuclease, a TALEN or a CRISPR nuclease). The introduction of a DSB in an intron of the gene locus (together with the provision of a donor nucleic acid (ssODN) or vector comprising the donor nucleic acid) was found to advantageously favor nuclease-based modifications by HDR.

[0019] The present invention further provides a method for nuclease-based modification of a target polynucleotide in one or more target cells within a population of cells and for enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide within the population, the method comprising:

(i) providing the cell population with a CRISPR nuclease and one or more sgRNAs comprising:
  (a) a sgRNA guide sequence having a target sequence in an endogenous gene, such that the sgRNA guide sequence directs the cleavage of the endogenous gene for introduction of a mutation (i.e., a second nuclease-based modification) in the endogenous gene conferring resistance to a drug, and a CRISPR nuclease recognition sequence;
  (b) a sgRNA guide sequence having a target sequence in the target polynucleotide and a CRISPR nuclease recognition sequence;

wherein the target sequences in (a) and (b) are contiguous to a protospacer adjacent motif (PAM) recognized by the CRISPR nuclease;

(ii) contacting the population of cells with the drug; and

(iii) selecting the one or more target cells by virtue of increased tolerance to the drug, thereby enriching the number of target cells comprising the nuclease-based modification in the target polynucleotide in (b) within the population of cells.

[0020] The present invention further provides a method of inducing drug resistance in one or more target cells within a population of cells, the method comprising:

(i) introducing a double stranded break with a nuclease in the one or more target cells;

(ii) contacting the cell population with the drug; and

(iii) selecting the one or more target cells by virtue of increased tolerance to the drug, thereby inducing drug resistance in the one or more target cells.

[0021] The present invention further provides a method for enriching for target cells modified at an endogenous gene locus, wherein the modification confers resistance to a drug, the method comprising:

(i) cleaving with a nuclease the endogenous gene in one or more target cells within a population of cells;

(ii) contacting the population of cells with the drug; and

(iii) selecting the one or more target cells by virtue of increased tolerance to the drug, thereby enriching for target cells in which the endogenous gene has been modified.

[0022] In an embodiment, the endogenous gene and drug are selected from the endogenous genes and corresponding drugs set forth in Table 1 herein.

[0023] In an embodiment, the endogenous gene and associated drug is selected from the genes and drugs listed in Table 1. In an embodiment, the endogenous gene is (encodes) a Na/K-ATPAse and the drug is an Na/K-ATPase inhibitor. In an embodiment, the endogenous gene is (encodes) an ATP1A gene, and the drug is ouabain or a derivative thereof. In an embodiment, the endogenous gene is (encodes) MGMT and the drug is O6-benzylguanine.

[0024] In embodiments, the ouabain or a derivative thereof is used for selection (i.e., is comprises in the selection medium) at a concentration of less than about 5μM, in a further embodiment at a concentration of about 100μM or less, in a further embodiment at a concentration of about 1000μM (1 mM) or less, in a further embodiment from about 5μM to about 100μM, in a further embodiment from about 5μM to about 1000μM, in a further embodiment from about 100μM to about 1000μM, in a further embodiment from about 5μM to about 500μM, in a further embodiment from about 100μM to about 500μM, in a further embodiment from about 5μM to about 800μM, in a further embodiment from about 100μM to about 800μM in a further embodiment from about 75μM to about 125μM, in a further embodiment from about 75μM to about 500μM, in a further embodiment from about 75μM to about 800μM, in a further embodiment from about 75μM to about 1000μM.

[0025] The present invention further provides a method for nuclease-based modification of a target polynucleotide in one or more target cells within a population of cells comprising: (i) introducing (e.g., (co-introducing), by HDR or NHEJ, a mutation (a nuclease-based modification) in (a) an ATP1A gene locus and (b) the target polynucleotide, in the one or more target cells; (ii) contacting the population of cells with ouabain or a derivative thereof; and (iii) selecting the one or more target cells by virtue of increased tolerance to ouabain or derivative thereof, thereby enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide within the population of cells.

[0026] The present invention further provides a method for nuclease-based modification of a target polynucleotide in one or more target cells within a population of cells comprising: (i) providing the cell population with a CRISPR nuclease and one or more sgRNAs comprising: (a) a sgRNA guide sequence having a target sequence in the ATP1A gene and a CRISPR nuclease recognition sequence; (b) a sgRNA guide sequence having a target sequence in the target polynucleotide and a CRISPR nuclease recognition sequence; wherein the target sequences in (a) and (b) are contiguous to a protospacer adjacent motif recognized by the CRISPR nuclease; (ii) contacting the population of cells with ouabain or a derivative thereof; and (iii) selecting the one or more target cells by virtue of increased tolerance to ouabain or derivative thereof, thereby enriching the number of target cells comprising the mutation (CRISPR nuclease-based modification) within the population of cells.

[0027] The present invention further provides a method for nuclease-based modification of a target polynucleotide in one or more target cells within a population of cells and for enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide within the population comprising: (i) introducing, by HDR or NHEJ, a mutation (nuclease-based modification) in (a) an ATP1A gene locus and (b) the target polynucleotide, in the one or more target cells; (ii) contacting the population of cells with ouabain or a derivative thereof; and (iii) selecting the one or more target cells by virtue of increased tolerance to ouabain or derivative thereof, thereby enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide within the population of cells.

[0028] The present invention further provides a method for nuclease-based modification of a target polynucleotide in one or more target cells within a population of cells and for enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide within the population comprising: (i) providing the cell population with a CRISPR nuclease and one or more sgRNAs comprising: (a) a sgRNA guide sequence having a target sequence in the ATP1A gene and a CRISPR nuclease recognition sequence; (b) a sgRNA guide sequence having a target sequence in the target polynucleotide and a CRISPR nuclease recognition sequence; wherein the target sequences in (a) and (b) are contiguous to a protospacer adjacent motif recognized by the CRISPR nuclease; (ii) contacting the population of cells with ouabain or a derivative thereof; and (iii) selecting the one or more target cells by virtue of increased tolerance to ouabain or derivative thereof, thereby enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide within the population of cells.

[0029] The present invention further provides a method of inducing ouabain resistance in one or more target cells within a population of cells, the method comprising: (i) introducing a double stranded break with a nuclease in the one or more target cells; (ii) contacting the cell population with ouabain or a derivative thereof; and (iii) selecting the one or more target cells by virtue of increased tolerance to ouabain or derivative thereof, thereby inducing ouabain resistance in the one or more target cells.

[0030] The present invention further provides a method for enriching for target cells modified at an ATP1A gene locus, the method comprising: (i) cleaving with a nuclease the ATP1A gene in one or more target cells within a population of cells; (ii) contacting the population of cells with ouabain or a derivative thereof; and (iii) selecting the one or more target cells by virtue of increased tolerance to ouabain or derivative thereof, thereby enriching for target cells in which the ATP1A gene has been modified.

[0031] In embodiments, the nuclease is a CRISPR nuclease, and the method comprises providing the one or more target cells with a sgRNA comprising a sgRNA guide sequence having a target sequence in the ATP1A gene and a CRISPR nuclease recognition sequence.

[0032] In embodiments, the CRISPR nuclease is Cas9 or Cpf1 nuclease. In embodiments, the Cas9 is SpCas9 (or a functional derivative thereof) from *Streptococcus pyogenes*. In embodiments, the Cpf1 is AsCpf1 from *Acidaminococcus sp*.

[0033] In embodiments, the ATP1A gene is ATP1A2 and the sgRNA target sequence is selected within SEQ ID NO: 459. In embodiments, the ATP1A gene is ATP1A3 and the sgRNA target sequence is selected within SEQ ID NO: 461. In embodiments, the ATP1A gene is ATP1A4 and the sgRNA target sequence is selected within SEQ ID NO: 463. In embodiments, the sgRNA target sequence is selected within a region of the ATP1A2, ATP1A 3 or ATP1A4 gene corresponding to exon 4 of ATP1A (see FIG. 37).

[0034] In embodiments, the ATP1A gene is ATP1A1 and the sgRNA target sequence comprises or consists of a sequence as set forth in Table 3 and FIG. 28. In embodiments, the sgRNA comprises or consists of the sequence set forth in any one of SEQ ID NOs: 24-26 (sgRNA G2, G3, G4) and 33-36 (sgRNA G6, G7, G8, G9).

[0035] In embodiments, the sgRNA target sequence is within an exon of the ATP1A gene. In embodiments, the gRNA target sequence is within an intron of the ATP1A gene.

[0036] In embodiments, the ATP1A gene is ATP1A1 and the exon is exon 4. In embodiments, the target sequence comprises or consists of the following target sequence in the ATP1A1 gene: (i) ATCCAAGCTGCTACAGAAG (G2); (ii) GTTCCTCTTCTGTAGCAGCT (G4); (iii) TTGGCTTATAGCATCCAAGC (G6); (iv) AGGTTCCTCTTCTGTAGCAG (G7); or (v) GAGGTTCCTCTTCTGTAGCA (G8).

[0037] In embodiments, the ATP1A gene is ATP1A1 and the intron is intron 3 or intron 4. In embodiments, the target sequence comprises or consists of the following target sequence on the ATP1A1 gene: (i) GAGTTCTGTAATTCAGCATA (G3) or (ii) TAGTACACATCAGATATCTT (G9).

[0038] In embodiments, methods of the present invention further comprise providing the cell with an ATP1A donor nucleic acid conferring ouabain resistance. In embodiments, the ATP1A donor nucleic acid conferring ouabain resistance is an ATP1A1 donor nucleic acid.

[0039] In embodiments, the ATP1A1 donor nucleic acid comprises or consists of the following nucleic acid sequence: (i) CAATGTTACTGTGGATTGGAGCGATTCTTTGTTTCTTGGCTTATAGCATCGATGCTGCTACAGAAGAGGAACCT CAAAACGATCGTGTGAGTTCTGTAATTCAGCATATCGATTTGTAGTACACATCAGATATCTT (DR); or (ii) CAATGTTACTGTGGATTGGAGCGATTCTTTGTTTCTTGGCTTATAGCATCAGAGCTGCTACAGAAGAGGAACCT CAAAACGATGACGTGAGTTCTGTAATTCAGCATATCGATTTGTAGTACACATCAGATATCTT (RD).

[0040] In embodiments, the ATP1A1 donor nucleic acid confers ouabain resistance by replacing at least one of amino acids Q118 and N129 in the ATP1A1 gene by a charged amino acid. In embodiments, the ATP1A1 donor nucleic acid confers ouabain resistance by replacing amino acid Q118 by an aspartic acid and amino acid N129 by an arginine. In embodiments, the ATP1A1 donor nucleic acid confers ouabain resistance by replacing amino acid Q118 by an arginine and amino acid N129 by an aspartic acid.

[0041] In embodiments, the above-mentioned one or more target cells are selected from the group consisting of T-cells, B-cells or stem cells. In embodiments, the stem cells comprise embryonic stem cells (ESCs), induced pluripotent stem cell (iPSCs), CD34+ hematopoietic stem cells or hepatic stem cells. In an embodiment, the one or more target cells are primary cells.

[0042] In a further aspect, the present invention provides a ouabain resistant ATP1A polypeptide. In embodiments, the ouabain resistant ATP1A polypeptide is a ouabain resistant ATP1A1 polypeptide. In embodiments, the ouabain resistant ATP1A polypeptide is a ouabain resistant ATP1A2 polypeptide. In embodiments, the ouabain resistant ATP1A polypeptide is a ouabain resistant ATP1A3 polypeptide. In embodiments, the ouabain resistant ATP1A polypeptide is a ouabain resistant ATP1A4 polypeptide.

[0043] In embodiments, the ouabain resistant ATP1A1 polypeptide comprises or consists of an amino acid sequence as set forth in SEQ ID NOs: 5, 8, 10, 12, 14, 16, 18, 20 or 22. In embodiments, the ouabain resistant ATP1A1 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 5, wherein: (i) the amino acid at position 119 is any amino acid except an alanine or is absent; (ii) the amino acid at position 120 is any amino acid except an alanine or is absent; (iii) the amino acid at position 121 is any amino acid except a threonine or is absent; (iv) the amino acid at position 122 is any amino acid except a glutamic acid or is absent; (v) the amino acid at position 123 is any amino acid except a glutamic acid or is absent; (vi) the amino acid at position 124 is any amino acid except a glutamic acid or is absent; (vii) the amino acid at position 125 is any amino acid except a proline or a lysine or is absent; (viii) the amino acid at position 126 is any amino acid except a glutamine or is absent; (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii). In embodiments, the ouabain resistant ATP1A1 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 6, wherein: (i) the amino acid at position 119 is any amino acid except an alanine or is absent; (ii) the amino acid at position 120 is any amino acid except an alanine or is absent; (iii) the amino acid at position 121 is any amino acid except a threonine or is absent; (iv) the amino acid at position 122 is any amino acid except a glutamic acid or is absent; (v) the amino acid at position 123 is any amino acid except a glutamic acid or is absent; (vi) the amino acid at position 124 is any amino acid except a glutamic acid or is absent; (vii) the amino acid at position 125 is any amino acid except a proline or a lysine or is absent; (viii) the amino acid at position 126 is any amino acid except a glutamine or is absent; (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0044] In embodiments, the ouabain resistant ATP1A1 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 5 or 6, wherein: (i) the amino acid at position 119 is any amino acid except an alanine or an isoleucine, or is absent; (ii) the amino acid at position 120 is any amino acid except an alanine, a glycine or a tyrosine, or is absent; (iii) the amino acid at position 121 is any amino acid except a threonine, a methionine or a phenylalanine, or is absent; (iv) the amino acid at position 122 is any amino acid except a glutamic acid or an asparagine, or is absent; (v) the amino acid at position 123 is any amino acid except a glutamic acid or an aspartic acid, or is absent; (vi) the amino acid at position 124 is any amino acid except a glutamic acid or an aspartic acid, or is absent; (vii) the amino acid at position 125 is any amino acid except a proline or a lysine, or is absent; (viii) the amino acid at position 126 is any amino acid except a glutamine, a serine or a threonine, or is absent; or (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0045] In embodiments, the ouabain resistant ATP1A1 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 5 or 6, wherein (i) the amino acid at position 119 is absent; (ii) the amino acid at position 120 is absent; (iii) the amino acid at position 121 is a lysine or is absent; (iv) the amino acid at position 122 is absent; (v) the amino acid at position 123 is absent; (vi) the amino acid at position 124 is absent; (vii) the amino acid at position 125 is an alanine or is absent; (viii) the amino acid at position 126 is absent; or (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0046] In embodiments, the present invention provides an ATP1A1 ouabain resistant polypeptide, wherein the polypeptide comprises at least one deletion or non-conservative substitution at amino acid: (i) A119, A120, T121, E122, E123, E124, P125 and/or Q126 of SEQ ID NO: 3 or 4. In embodiments the ouabain resistant ATP1A1 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 6, wherein the polypeptide comprises at least one deletion or non-conservative substitution at amino acid: (i) A119, A120, T121, E122, E123, E124, P125 and/or Q126. In embodiments, the ouabain resistant ATP1A1 polypeptide further comprises at least one further amino acid mutation set forth in Table 4.

[0047] In embodiments, the present invention provides an ATP1A2 ouabain resistant polypeptide, wherein the polypeptide comprises at least one deletion or non-conservative substitution at amino acid A117, A118, M119, E120, D121, E122, P123 and/or S124 of SEQ ID NO: 460. In embodiments, the ouabain resistant ATP1A2 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 485, wherein: (i) the amino acid at position 117 is any amino acid except an alanine or an isoleucine, or is absent; (ii) the amino acid at position 118 is any amino acid except an alanine, a glycine or a tyrosine, or is absent; (iii) the amino acid at position 119 is any amino acid except a threonine, a methionine or a phenylalanine, or is absent; (iv) the amino acid at position 120 is any amino acid except a glutamic acid or an asparagine, or is absent; (v) the amino acid at position 121 is any amino acid except a glutamic acid or an aspartic acid, or is absent; (vi) the amino acid at position 122 is any amino

acid except a glutamic acid or an aspartic acid, or is absent; (vii) the amino acid at position 123 is any amino acid except a proline or a lysine, or is absent; (viii) the amino acid at position 126 is any amino acid except a glutamine, a serine or a threonine, or is absent; or (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0048] In embodiments, the ouabain resistant ATP1A2 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 485, wherein (i) the amino acid at position 117 is absent; (ii) the amino acid at position 118 is absent; (iii) the amino acid at position 119 is a lysine or is absent; (iv) the amino acid at position 120 is absent; (v) the amino acid at position 121 is absent; (vi) the amino acid at position 122 is absent; (vii) the amino acid at position 123 is an alanine or is absent; (viii) the amino acid at position 124 is absent; or (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0049] In embodiments, the present invention provides an ATP1A3 ouabain resistant polypeptide, wherein the polypeptide comprises at least one deletion or non-conservative substitution at amino acid: A109, G110, T111, E112, D113, D114, P115 and/or S116 of SEQ ID NO: 462. In embodiments, the ouabain resistant ATP1A1 polypeptide comprises or consists of a polypeptide as set forth in 486, wherein: (i) the amino acid at position 109 is any amino acid except an alanine or an isoleucine, or is absent; (ii) the amino acid at position 110 is any amino acid except an alanine, a glycine or a tyrosine, or is absent; (iii) the amino acid at position 111 is any amino acid except a threonine, a methionine or a phenylalanine, or is absent; (iv) the amino acid at position 112 is any amino acid except a glutamic acid or an asparagine, or is absent; (v) the amino acid at position 113 is any amino acid except a glutamic acid or an aspartic acid, or is absent; (vi) the amino acid at position 114 is any amino acid except a glutamic acid or an aspartic acid, or is absent; (vii) the amino acid at position 115 is any amino acid except a proline or a lysine, or is absent; (viii) the amino acid at position 116 is any amino acid except a glutamine, a serine or a threonine, or is absent; or (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0050] In embodiments, the ouabain resistant ATP1A3 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 486, wherein (i) the amino acid at position 109 is absent; (ii) the amino acid at position 110 is absent; (iii) the amino acid at position 111 is a lysine or is absent; (iv) the amino acid at position 112 is absent; (v) the amino acid at position 113 is absent; (vi) the amino acid at position 114 is absent; (vii) the amino acid at position 115 is an alanine or is absent; (viii) the amino acid at position 116 is absent; or (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0051] In embodiments, the present invention provides an ATP1A4 ouabain resistant polypeptide, wherein the polypeptide comprises at least one deletion or non-conservative substitution at amino acid: I127, Y128, F129, N130, E131, E132, P133 and/or T134 of SEQ ID NO: 464. In embodiments, the ouabain resistant ATP1A1 polypeptide comprises or consists of a polypeptide as set forth in 486, wherein: (i) the amino acid at position 127 is any amino acid except an alanine or an isoleucine, or is absent; (ii) the amino acid at position 128 is any amino acid except an

alanine, a glycine or a tyrosine, or is absent; (iii) the amino acid at position 129 is any amino acid except a threonine, a methionine or a phenylalanine, or is absent; (iv) the amino acid at position 130 is any amino acid except a glutamic acid or an asparagine, or is absent; (v) the amino acid at position 131 is any amino acid except a glutamic acid or an aspartic acid, or is absent; (vi) the amino acid at position 132 is any amino acid except a glutamic acid or an aspartic acid, or is absent; (vii) the amino acid at position 133 is any amino acid except a proline or a lysine, or is absent; (viii) the amino acid at position 134 is any amino acid except a glutamine, a serine or a threonine, or is absent; or (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0052] In embodiments, the ouabain resistant ATP1A4 polypeptide comprises or consists of a polypeptide as set forth in SEQ ID NO: 487, wherein (i) the amino acid at position 127 is absent; (ii) the amino acid at position 128 is absent; (iii) the amino acid at position 129 is a lysine or is absent; (iv) the amino acid at position 130 is absent; (v) the amino acid at position 131 is absent; (vi) the amino acid at position 132 is absent; (vii) the amino acid at position 133 is an alanine or is absent; (viii) the amino acid at position 134 is absent; or (ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

[0053] The present invention also provides a sgRNA for inducing or conferring ouabain resistance by NHEJ or HDR comprising (i) a sgRNA guide sequence having a target sequence in an ATP1A gene; and (ii) a CRISPR nuclease recognition sequence.

[0054] In embodiments, the sgRNA guide sequence has a target sequence in the ATP1A1 gene. In embodiments, the sgRNA guide sequence has a target sequence comprising or consisting of a target sequence (without the PAM) as set forth in Table 3 and FIG. 28. In embodiments, the sgRNA comprises a sgRNA guide sequence having a target sequence comprising or consisting of the following target sequence in the ATP1A1 gene: (i) GATCCAAGCTGCTACAGAAG (G2, SEQ ID NO: 28); (ii) GTTCCTCTTCTGTAGCAGCT (G4, SEQ ID NO: 30); (iii) TTGGCTTATAGCATCCAAGC (G6, SEQ ID NO: 38); (iv) AGGTTCCTCTTCTGTAGCAG (G7, SEQ ID NO: 39); or (v) GAGGTTCCTCTTCTGTAGCA (G8, SEQ ID NO: 40). In embodiments, the sgRNA guide sequence has a target sequence comprising or consisting of the following target sequence in the ATP1A1 gene: (i) GAGTTCTGTAATTCAGCATATGG (G3, SEQ ID NO: 29) or (ii) TAGTACACATCAGATATCTT (G9, SEQ ID NO: 41).

[0055] The present invention also provides nucleic acids encoding drug resistant polypeptides (e.g., Na/K ATPase inhibitor resistant, ouabain resistant ATP1A polypeptides) and sgRNAs described herein, vectors for expressing nucleic acids of the present invention and cells comprising such vectors and/or nucleic acids. In embodiments, the vector is a viral vector.

[0056] The present invention also provides a kit comprising one or more: sgRNAs; drug resistant polypeptides

(e.g., Na/K ATPase inhibitor resistant, ouabain resistant ATP1A polypeptides, etc.); nucleic acids, vectors; cells and/or donor nucleic acids (e.g., donor nucleic acids) described herein. In embodiments, the kit further comprises instructions to use the kit according to methods disclosed herein and/or drug(s) for selecting (enriching) cells expressing drug resistant polypeptides generated using methods disclosed herein.

[0057] The present invention also relates to the use of one or more sgRNAs; drug resistant polypeptides (e.g., Na/K ATPase inhibitor resistant, ouabain resistant ATP1A polypeptides, etc.); nucleic acids, vectors; cells, donor nucleic acids and/or kits described herein for (a) inducing or conferring drug resistance (e.g., Na/K ATPase inhibitor resistance, ouabain resistance, etc.) in one or more target cells; (b) nuclease-based modification of a target polynucleotide in one or more target cells and for enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide; and/or or (c) for enriching for target cells modified by a nuclease at a drug resistance gene locus (e.g., an Na/K ATPase gene locus, ATP1A gene locus, MGMT gene locus, etc.).

[0058] The present invention also provides one or more sgRNAs; drug resistant polypeptides (e.g., Na/K ATPase inhibitor resistant, ouabain resistant ATP1A polypeptides, etc.); nucleic acids, vectors; cells and/or donor nucleic acids described herein for use in (a) inducing or conferring drug resistance (e.g., Na/K ATPase inhibitor resistant, ouabain resistance etc.) in one or more target cells; (b) for nuclease-based modification of a target polynucleotide in one or more target cells for enriching the number of target cells comprising the nuclease-based modification of the target polynucleotide; and/or or (c) for enriching for target cells modified by a nuclease at a drug resistance gene locus (e.g., an ATP1A gene locus).

[0059] The present invention also provides a composition comprising one or more sgRNAs; drug resistant polypeptides (e.g., Na/K ATPase inhibitor resistant, ouabain resistant ATP1A polypeptides, O6-benzylguanine resistant, etc.); nucleic acids, vectors; cells and/or donor nucleic acids described herein. In embodiments, the composition may further comprise a suitable carrier (e.g., pharmaceutically acceptable).

[0060] The present invention further concerns the above-described composition or kit for (a) inducing or conferring drug resistance (e.g., Na/K ATPase inhibitor resistance, ouabain resistance, O6-benzylguanine resistance, etc.) in one or more target cells; (b) for modifying with a nuclease a target polynucleotide in one or more target cells and for enriching the number of target cells comprising the modified target polynucleotide; and/or or (c) for enriching for target cells modified by a nuclease at an ATP1A gene locus.

[0061] In embodiments, the above-mentioned ATP1A gene is an ATP1A1, ATP1A2, ATP1A3, or ATP1A4 gene. In particular embodiments, the ATP1A gene is ATP1A1.

BRIEF DESCRIPTION OF THE DRAWINGS

[0062] FIG. 1 shows a schematic representation of SpCas9 target sites surrounding DNA encoding the first extracellular loop of human *ATP1A1 (*SEQ ID NOs: 1 and 3). Annotated are the positions of residues Q118 and N129, exon/intron boundary, protospacer adjacent motifs (PAM) and four potential SpCas9 target sequences (Targets 1-4, SEQ ID NOs: 59 (Target 1), 70 (Target 2), 29 (Target 3), and 30 (Target 4));

[0063] FIG. 2 shows the identification of highly active sgRNA/nuclease combinations targeting the *ATP1A1* locus. The indicated single-guide RNA (sgRNA G1-G4: SEQ ID NOs: 23-26) expression vectors (500 ng) were transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1 (*adeno-associated virus integration site 1; 19q13-qter; HGNC:22; OMIM:102699; GenBank: S51329.1) safe harbor locus. Genomic DNA was harvested 3 and 10 days post-transfection and the Surveyor assay was used to determine the frequency of SpCas9-induced insertions and deletions indicated as the % Indels at the base of each lane. An expression vector encoding EGFP (500ng) was used as a negative control.

[0064] FIG. 3 shows active sgRNAs/nucleases targeting the coding sequence surrounding Q118 and N129 of *ATP1A1* induce cellular resistance to ouabain. The indicated single-guide RNA (sgRNA, G1-G4: SEQ ID NOs: 23-26) expression vectors (500 ng) were transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus. Cells were treated with 0.5μM ouabain for 7 days starting 3 days post transfection. Only cells transfected with sgRNA G2 (SEQ ID NO: 24) and G4 (SEQ ID NO: 26) survived treatment and grew robustly. Genomic DNA was harvested 3 and 10 days post-transfection and the Surveyor assay was used to determine the frequency of SpCas9-induced insertions and deletions indicated as the % Indels at the base of each lane. An expression vector encoding EGFP (500ng) was used as a negative control;

[0065] FIG. 4 shows selection for short in-frame deletions in the coding sequence surrounding Q118 and N129 of *ATP1A1* following DNA cleavage driven by nuclease G2 (sgRNA G2 + SpCas9) expression and ouabain selection. K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus were transfected with the G2 sgRNA (SEQ ID NO: 24) expression vector and treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection. Genomic DNA was harvested from these cells and the TIDE (Tracking of Indels by Decomposition) assay was performed to determine the spectrum and frequency of targeted mutations generated in the pool of cells;

[0066] FIG. 5 shows DNA cleavage driven by nuclease G3 (sgRNA G3 + spCas9) within the intron downstream of Q118 and N129 of *ATP1A1* does not confer ouabain resistance nor selects for specific insertions or deletions. K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus were transfected with the G3 sgRNA (SEQ ID NO: 25) expression vector and treated or not with 0.5μM ouabain for 7 days starting 3 days post-

transfection. Genomic DNA was harvested from untreated cells and the TIDE (Tracking of Indels by Decomposition) assay was performed to determine the spectrum and frequency of targeted mutations generated in the pool of cells;

[0067] FIG. 6 shows selection for short in-frame deletions in the coding sequence surrounding Q118 and N129 of *ATP1A1* following DNA cleavage driven by nuclease G4 expression and ouabain selection. K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus were transfected with the G4 sgRNA (SEQ ID NO: 26) expression vector and treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection. Genomic DNA was harvested from these cells and the TIDE (Tracking of Indels by Decomposition) assay was performed to determine the spectrum and frequency of targeted mutations generated in the pool of cells;

[0068] FIG. 7 shows partial amino acid sequences of mutant *ATP1A1* polypeptides (SEQ ID NO: 3 (wild-type), SEQ ID NO: 8 (Δ3), SEQ ID NO: 10 (Δ6), SEQ ID NO: 12 (Δ9), SEQ ID NO: 14 (Δ12) and SEQ ID NO: 16 (Δ3b)) generated following nuclease treatment and ouabain selection. The region from amino acids S116 to N129 is shown. The indicated single-guide RNA (sgRNA G2 and G4, SEQ ID NOs: 24 and 26) expression vectors (500 ng) were transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus. Cells were treated with 0.5μM ouabain for 7 days starting 3 days post-transfection. Genomic DNA was harvested 10 days post-transfection and the region surrounding the cleavage site was amplified by PCR, TOPO cloned and sequenced;

[0069] FIG. 8 shows selection for larger in-frame deletions in the coding sequence surrounding Q118 and N129 of *ATP1A1* following DNA cleavage driven by nuclease G2 (sgRNA G2 +spCas9) expression and high-dose ouabain selection. K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus were transfected with the G2 sgRNA (SEQ ID NO: 24) expression vector and treated or not with increasing doses of ouabain for 7 days starting 3 days post-transfection. Genomic DNA was harvested from these cells and the TIDE (Tracking of Indels by Decomposition) assay was performed to determine the spectrum and frequency of targeted mutations generated in the pool of cells;

[0070] FIG. 9 shows selection for larger in-frame deletions in the coding sequence surrounding Q118 and N129 of *ATP1A1* following DNA cleavage driven by nuclease G4 expression and high-dose ouabain selection. K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus were transfected with the G4 sgRNA (SEQ ID NO: 26) expression vector and treated or not with increasing doses of ouabain for 7 days starting 3 days post-transfection. Genomic DNA was harvested from these cells and the TIDE (Tracking of Indels by Decomposition) assay was performed to determine the spectrum and frequency of targeted mutations generated in the pool of cells;

[0071] FIG. 10 shows amino acid sequences of mutant *ATP1A1* polypeptides (SEQ ID NO: 3 (wild-type), SEQ ID NO: 18 (Δ15), SEQ ID NO: 20 (Δ18), SEQ ID NO: 16 (Δ3b), and SEQ ID NO: 22 (Δ21)) generated following nuclease treatment and high-dose ouabain selection. The region from amino acids S116 to N129 of the ATP1A1

polypeptide (SEQ ID NO: 3 is shown. The indicated single-guide RNA (sgRNA G2 (A) and G4 (B), SEQ ID NOs: 24 and 26) expression vectors (500 ng) were transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus. Cells were treated with 10μM ouabain for 7 days starting 3 days post-transfection. Genomic DNA was harvested 10 days post-transfection and the region surrounding the cleavage site was amplified by PCR, TOPO cloned and sequenced;

[0072] FIGs. 11A-B show that active nucleases targeting the coding sequence surrounding Q118 and N129 of *ATP1A1* induce cellular resistance to ouabain. **(A)** U2OS cells were transfected with a wild-type SpCas9 expression vector (500ng) and the indicated sgRNAs (G2-G4 (SEQ ID NOs: 24-26); 500ng). Only cells transfected with sgRNA G2 survived treatment and grew robustly. Genomic DNA was harvested 10 days post-transfection and the Surveyor assay was used to determine the frequency of SpCas9-induced insertions and deletions indicated as the % Indels at the base of each lane. Where indicated cells were treated with 0.5μM ouabain for 7 days starting 3 days post transfection. **(B)** hTERT-RPE-1 cells were transfected with eSpCas9(1.1)_No_FLAG vectors expressing ATP1A1 sgRNAs (G2 (i) or G3 (ii): 500ng) along with ssODN RD (1ul of a 10μM stock, (SEQ ID NO: 44)) where indicated. Cells were treated as in (A) and the frequency of mutated alleles was determined using the TIDE assay. In addition, a ClaI restriction fragment length polymorphism (RFLP) assay (B (ii)) was used to determine the frequency of SpCas9-induced HDR at the cleavage site indicated as the % HDR at the base of panel (B)(ii). An expression vector encoding EGFP (-) was used as a negative control in (A) and (B);

[0073] FIG. 12 shows enrichment of cells with targeted mutagenesis at the *AAVS1* locus by co-editing the *ATP1A1* gene. K562 cells were transfected with two sgRNAs that target respectively *ATP1A1* (G2 (SEQ ID NO: 24): 5ng, 50ng, 500ng, panel (A)) and *AAVS1* (T2, SEQ ID NO: 468), 5ng and 50ng, panel (B)), in combination with a wild-type SpCas9 expression vector, and treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection. Genomic DNA was harvested 10 days post-transfection and the Surveyor assay was used to determine the frequency of SpCas9-induced insertions and deletions indicated as the % Indels at the base of each lane. An expression vector encoding EGFP (500ng) was used as a negative control. *ATP1A1* targeted cells selected by resistance to ouabain show an increased level of indels at the *AAVS1* site;

[0074] FIG. 13 shows enrichment of cells with targeted mutagenesis at the *HPRT1* locus by co-editing the *ATP1A1* gene. K562 cells were transfected with two sgRNAs that target respectively *ATP1A1* (G2 (SEQ ID NO: 24): 5ng, 50ng, 500ng, panel (A)) and *HPRT1* -S2L (SEQ IS NO: 456) 5ng and 50ng, panel (B)), in combination with a wild-type SpCas9 expression vector, and treated or not with 0.5μM ouabain for 7 days starting 3 days post transfection. Genomic DNA was harvested 10 days post-transfection and the Surveyor assay was used to determine the frequency of SpCas9-induced insertions and deletions indicated as the % Indels at the base of each lane. An expression vector encoding EGFP (500ng) was used as a negative control. *ATP1A1* targeted cells

selected by resistance to ouabain show an increased level of indels at the *HPRT1* site;

[0075] FIG. 14 shows a schematic representation of the intronic SpCas9 target site G3 (Target 3, SEQ ID NO: 29) and partial sequence of single-stranded oligodeoxynucleotides (ssODNs, RD (SEQ ID NO: 44) and DR (SEQ ID NO: 43)) donors used to introduce the Q118D/R and N129D/R mutations conferring ouabain resistance. Annotated are the positions of residues Q118 and N129 in the ATP1A1 polypeptide (SEQ ID NO: 3), exon/intron boundary, protospacer adjacent motifs (PAM) and novel restriction sites introduced to monitor the insertion of ssODN-specified mutations. (A) Wild-type ATP1A1 nucleic acid and amino acid sequences (see SEQ ID NOs: 2 and 3; (B) Mutation introduced by ssODNs, RD (SEQ ID NO: 44) and (C) mutation introduced by ssODN DR (SEQ ID NO: 43);

[0076] FIG. 15 shows the introduction of the Q118R/D and N129D/R mutations via nuclease G3-driven homology-directed repair (HDR) confer cellular resistance to ouabain. The G3 sgRNA (SEQ ID NO: 25) expression vector (500 ng) was co-transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus along with the indicated ssODNs (RD (SEQ ID NO: 44) and DR (SEQ ID NO: 43)). Cells were treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection. Only cells co-transfected with sgRNA G3 and ssODNs survived treatment and grew robustly. Genomic DNA was harvested 3 and 10 days post-transfection and the Surveyor assay was used to determine the frequency of SpCas9-induced insertions and deletions indicated as the % Indels at the base of each lane. An expression vector encoding EGFP (500ng) was used as a negative control;

[0077] FIG. 16 shows that ouabain selection increases the levels of nuclease-driven HDR at the *ATP1A1* locus. The G3 sgRNA (SEQ ID NO: 25) expression vector (500 ng) was co-transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus along with the indicated ssODNs (RD (SEQ ID NO: 44) and DR (SEQ ID NO: 43)). Cells were treated or not with 0.5μM ouabain for 7 days starting 3 days post transfection. Only cells co-transfected with sgRNA G3 and ssODNs survived treatment and grew robustly. Genomic DNA was harvested 3 and 10 days post-transfection and a ClaI restriction fragment length polymorphism (RFLP) assay was used to determine the frequency of SpCas9-induced HDR at the cleavage site indicated as the % HDR at the base of each lane. An expression vector encoding EGFP (500ng) was used as a negative control;

[0078] FIG. 17 shows that ouabain selection increases the levels of nuclease-driven HDR at the *ATP1A1* locus. The G3 sgRNA expression vector (500 ng) was co-transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus along with the indicated ssODNs (RD (SEQ ID NO: 44) and DR (SEQ ID NO: 43)). Cells were treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection. Only cells co-transfected with sgRNA G3 (SEQ ID NO: 25) and ssODNs survived treatment and grew robustly. Genomic DNA was harvested 3 and 10 days post-transfection and BmgBI (A) and PvuI (B) restriction fragment length polymorphism (RFLP) assays were used to determine the frequency of SpCas9-induced HDR of the N129D/R

mutations indicated as the % HDR at the base of each lane. An expression vector encoding EGFP (500ng) was used as a negative control;

[0079] FIG. 18 shows that ouabain selection increases the levels of nuclease-driven HDR at the *ATP1A1* locus. The G3 sgRNA expression vector (500 ng) was co-transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus along with the indicated ssODNs RD (SEQ ID NO: 44) and DR (SEQ ID NO: 43)). (A) Cells were treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection. Only cells co-transfected with sgRNA G3 (SEQ ID NO: 25) and ssODNs survived treatment and grew robustly. Genomic DNA was harvested 3 and 10 days post-transfection and a Hpy188I restriction fragment length polymorphism (RFLP) assay was used to detect SpCas9-induced HDR of the Q118R mutation. As shown in (B), increasing the dose of ouabain results in a higher frequency of insertion of the Q118R mutation. An expression vector encoding EGFP (500ng) was used as a negative control.

[0080] FIG. 19 shows enrichment of cells with targeted insertion of a point mutation at the *HPRT1* locus by co-editing the *ATP1A1* gene. K562 cells were transfected with a wild-type SpCas9 expression vector, two sgRNAs that target respectively *ATP1A1* (G3 (SEQ ID NO: 25): 50ng and 500ng) and *HPRT1-S1R* (SEQ ID NO: 465, 50ng) along with ssODNs RD (SEQ ID NO: 44) and a plasmid donor containing homology arms to the *HPRT1* locus and a mutation introducing a BamHI restriction site. Cells were treated or not with 0.5μM ouabain for 7 days starting 3 days post transfection. Genomic DNA was harvested 10 days post-transfection and a BamHI restriction fragment length polymorphism (RFLP) assay was used to determine the frequency of SpCas9-induced HDR of the point mutation. The % HDR at the *HPRT1* locus with and without ouabain treatment is shown on the bar graph;

[0081] FIG. 20 shows enrichment of cells with targeted integration of a PGK-EGFP expression cassette at the *HPRT1* locus by co-editing the *ATP1A1* gene. K562 cells were transfected with a wild-type SpCas9 expression vector, two sgRNAs that target respectively *ATP1A1* (G3 (SEQ ID NO: 25): 50ng and 500ng) and *HPRT1-S1R* (SEQ ID NO: 465, 50ng) along with ssODNs RD (SEQ ID NO: 44) and a plasmid donor containing a PGK-EGFP cassette and homology arms to the *HPRT1* locus. Cells were treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection. Flow cytometry was used to determine the % of EGFP positive cells in each population;

[0082] FIG. 21 shows enrichment of cells with in-frame insertion of mAG at the *HIST1H2BK* c-terminus by co-editing the *ATP1A1* gene. (A) K562 cells were transfected with a wild-type SpCas9 expression vector, two sgRNAs that target respectively *ATP1A1* (G3 (SEQ ID NOs: 25 (sgRNA) and 29 (target)): (500ng) and *HIST1H2BK* (SEQ ID NOs: 467 (sgRNA) and 457 (target)) (500ng) along with ssODNs RD (SEQ ID NO: 44) and a plasmid donor containing a mAG cassette and homology arms to the *HIST1H2BK* locus (donor sequence SEQ ID NO: 474). Cells were treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection. Flow cytometry was used to determine the % of mAG positive cells in each population. (B) Exemplary fluorescence imaging of ouabain treated

cells expressing the H2BK-mAG fusion;

[0083] FIG. 22 shows the enrichment of cells with targeted mutagenesis at the AAVS1 and HPRT1 loci by co-editing the ATP1A1 gene via NHEJ (A) K562 cells stably expressing wild-type SpCas9 were transfected with two sgRNAs that target respectively ATP1A1 (G2 (SEQ ID NOs: 24 (sgRNA) and 28 (target)): 5ng, 50ng, 500ng) and AAVS1 (T2 (SEQ ID NOs: 468 (target) and 469 (sgRNA)): 5ng) and treated or not with 0.5µM ouabain for 7 days starting 3 days post-transfection. Genomic DNA was harvested 10 days post-transfection and the Surveyor assay was used to determine the frequency of SpCas9-induced insertions and deletions. The data are expressed as fold change of % Indels in ouabain-treated versus non-treated cells. (B) Same as in (A) but using the HPRT1-*S2L* sgRNA (SEQ ID NO: 466). See also FIGs. 12 and 13 above;

[0084] FIG. 23 shows a robust enrichment of cells with targeted integrations at endogenous loci by co-editing the *ATP1A1* gene via HDR. (A) K562 cells were transfected with combinations of wild-type SpCas9 expression vector (500ng), sgRNAs that target *ATP1A1* (G3 (SEQ ID NO: 25): 50ng and 500ng) and *HPRT1* (S1R (SEQ ID NO: 465): 50ng), ssODN RD ((SEQ ID NO: 44), 1ul of a 10µM stock) and a plasmid donor containing a PGK-EGFP expression cassette (1ug) and homology arms to the *HPRT1* locus (SEQ ID NO: 480) (cells were transfected with (i) PGK-EGFP donor; (ii) PGK-EGFP donor, *HPRT1* sgRNA and ssODN RD; (iii) PGK-EGFP donor, *HPRT1* sgRNA, ssODN RD and *ATP1A1* sgRNA G3 (50ng); (iv) PGK-EGFP donor, *HPRT1* sgRNA, ssODN RD and *ATP1A1* sgRNA G3 (500ng); (v) PGK-EGFP donor, ssODN RD and *ATP1A1* sgRNA G3 (500ng); (vi) PGK-EGFP donor, *HPRT1* sgRNA, ssODN RD and ATP1A1 sgRNA G3 (50ng); and (vii) PGK-EGFP donor, *HPRT1* sgRNA, ssODN RD, and *ATP1A1* sgRNA G3 (500ng)). Cells were treated ((v) to (vii)) or not ((i) to (iv)) with 0.5µM ouabain for 7 days starting 3 days post-transfection. Flow cytometry diagram displaying relative fluorescence intensity on the x-axis and the number of events on the y-axis. (B) K562 cells were transfected with eSpCas9(1.1)_No_FLAG expressing *AAVS1*-T2 sgRNA (SEQ ID NO: 469,100ng), a pUC19-based *ATP1A1* sgRNA vector (G3 (SEQ ID NO: 25): 100ng) along with ssODN RD ((SEQ ID NO: 44) 0.5ul of a 10µM stock) and a splicing acceptor-SA-2A-EGFP-pA donor (SEQ ID NO: 472) containing homology arms to *AAVS1* (250ng) (cells transfected with (i) eSpCas9(1.1); (ii) eSpCas9(1.1), *AAVS1*-T2 sgRNA, ssODN RD, SA-2A-EGFP and *ATP1A1* sgRNA G3; (iii) same as (ii) but cells were treated 0.5µM ouabain for 7 days starting 3 days post-transfection). Cells were treated and analyzed as in (A). (C) K562 cells were transfected with a pX330-based vector expressing the *LMNA* sgRNA (G2 (SEQ ID NO: 471 and 470 (target sequence)): 100 ng), a pUC19-based *ATP1A1* sgRNA vector (G3 (SEQ ID NO: 25): 100 ng), along with ssODN RD (0.5ul of a 10µM stock) and a plasmid donor containing a Clover cassette (250ng) and homology arms to the LMNA locus (SEQ ID NO: 473). Cells were treated (ii) or not (i) with ouabain (0.5 µM) and analyzed as in (A);

[0085] FIG. 24 shows enrichment of cells with targeted integrations at endogenous loci by co-editing the ATP1A1 gene via HDR. (A) K562 cells were transfected with a pX330-based vector expressing the LMNA sgRNA

(G2: 100 ng, SEQ ID NOs: 470 (target) and 471 (sgRNA)), a pUC19-based ATP1A1 sgRNA vector (G3: 100 ng, (SEQ ID NO: 25)), along with ssODNs RD (0.5ul of a 10μM stock) and a plasmid donor containing a Clover cassette (250ng, SEQ ID NO: 473) and homology arms to the LMNA locus. Cells were treated as in FIG. 20 (treated or not with 0.5μM ouabain for 7 days starting 3 days post-transfection) and flow cytometry was used to determine the % of Clover positive cells in each population. (B) Exemplary fluorescence imaging of ouabain treated cells expressing the Clover-LMNA fusion. See also FIG. 23C;

[0086] FIG. 25 shows that high specificity CRISPR nucleases AsCpf1 and eSpCas9 (1.1) are compatible with the ouabain selection strategy. (A) Schematic representation of AsCpf1 target sites surrounding DNA encoding the first extracellular loop of human *ATP1A1*. Annotated are the positions of residues Q118 and N129, exon/intron boundary, protospacer adjacent motifs (PAM) and five potential AsCpf1 target sequences (Targets 1-5, SEQ ID NOs: 37-41). (B) The indicated crRNA expression vectors (G5-G9 SEQ ID NOs: 32-36 (RNA sequence) (500 ng) were transfected into K562 cells stably expressing AsCpf1 from the *AAVS1* safe harbor locus. Genomic DNA was harvested 10 days post-transfection and the Surveyor assay was used to determine the frequency of AsCpf1-induced insertions and deletions indicated as the % Indels at the base of each lane. Where indicated cells were treated with 0.5μM ouabain for 7 days starting 3 days post transfection. An expression vector encoding EGFP (-) was used as a negative control. (C) K562 cells were transfected with eSpCas9(1.1)_No_FLAG vectors expressing sgRNA G2 or G3 along with the indicated ssODN (1ul of a 10μM stock) where indicated. Cells were treated and the % Indels was determined as in (B). (D) Genomic DNA was extracted from cells prepared in (B) and the region surrounding the cleavage site was amplified by PCR and analyzed by TIDE assay. (E) Same as in (D) but for samples presented in (C);

[0087] FIG. 26 shows the enrichment of NHEJ- and HDR-driven events in primary human cord blood (CB) CD34+ cells upon selection with ouabain. (A) Cultured CD34+ cells (100 000-200 000 per sample) were electroporated with Cas9 RNP G2 and grown for the indicated period of time in presence of 0.5μM ouabain starting 6 days post-transfection. Genomic DNA was harvested and the Surveyor assay was used to determine the frequency of SpCas9-induced insertions and deletions indicated as the % Indels at the base of each lane. Recombinant Cas9 (-) was used as a negative control. (B) The region surrounding the cleavage site was amplified by PCR and analyzed by TIDE assay. (C) CD34+ cells were treated as in (A) but ssODN RD (1ul of a 50μM stock) was added to the transfection mixes. A BmgBI restriction fragment length polymorphism (RFLP) assay was used to determine the frequency of SpCas9-induced HDR at the cleavage site indicated as the % HDR at the base of each lane;

[0088] FIG. 27 shows the strategy for CRISPR/Cas9-driven HDR at the *MGMT* locus to protect hematopoietic cells from the combination of O6-benzylguanine and BCNU. Schematic representation of putative intronic SpCas9 target

sites and partial sequences of single-stranded oligodeoxynucleotides (ssODNs) donors (SEQ ID NO: 484) to introduce the P140K and G156A mutations. Annotated are the positions of residues P140 and G156, exon/intron boundary and protospacer adjacent motifs (PAM). The effects of the above mentioned mutations are independent in the sense that combination of the mutations leads to an additional increase in resistance. SEQ ID NO: 483 corresponds to the partial genomic sequence of MGMT intron 4 and exon 5 shown in the figure;

[0089] FIGs. 28A-I show non-limiting examples of SpCas9 target sequences and their associated PAM for *ATP1A.* (A) Exon 5 (SEQ ID NOs: 90-123); (B) Exon 8 (SEQ ID NOs: 124-182); (C) Exon 9 (SEQ ID NOs: 183-218); (D) Exon 17 (SEQ ID NOs: 219-252); Exon 18 (SEQ ID NOs:253-293); (F-G) Exons 19-20 (SEQ ID NOs: 294-381; (H) Exon 21 (SEQ ID NOs: 382-414); and (I) Exon 22 (SEQ ID NOs: 415-448);

[0090] FIG. 29 shows enrichment of cells with in-frame insertion of mAG at the *HIST1H2BK* c-terminus by co-editing the *ATP1A1* gene under increasing doses of ouabain. K562 cells were transfected with eSpCas9(1.1)_No_Flag_ATP1A1_G3_dual_H2BK (100ng), a vector that drives the expression of eSpCas9(1.1) along with two sgRNAs that target respectively *ATP1A1* and *HIST1H2BK* (SEQ ID NOs: 29/25 and 457/467). The transfection mix was supplemented with ssODNs RD (5pmol) and a plasmid donor containing a mAG cassette and homology arms to the *HIST1H2BK* locus (250ng, SEQ ID NO: 474). Cells were treated with 0.5μM ouabain for 7 days starting 3 days post-transfection, and then the dose was increased to 100μM for another seven days. Finally, cells selected with 100μM ouabain were exposed to 1000μM for another 7 days. Flow cytometry was used to determine the % of mAG positive cells in each population;

[0091] FIG. 30 shows a schematic of the strategy for targeted parallel integration of an exogenous sequence into the *ATP1A1* locus. Targeted integration of the exogenous sequence occurs within intron 4 (the polynucleotide sequence of intron 4 corresponds to nts 305-404 of SEQ ID NO: 2), in proximity to exon 4 (the polynucleotide sequence of exon 4 corresponds to nts 101-304 of SEQ ID NO: 2) that encodes Q118 and N129 (vertical black bars). The positions of the DNA double-strand break (DSB) generated within intron 4 by either SpCas9 (G3- SEQ ID NO: 29 (target sequence)) or AsCpf1 (G9-SEQ ID NO: 41 (target sequence)) are predicted to be 19bp and 49bp downstream of exon 4, respectively. In this example, the donor DNA molecule directs the integration of an expression cassette encoding a fluorescent protein (FP) in the same direction as the *ATP1A1* transcript (parallel orientation). The left homology arm (L-HA) of the donor DNA molecule contains the Q118R and N129D (RD) mutations (*asterisks). h*PGK1*: human phosphoglycerate kinase 1 gene promoter, bGHpA: polyadenylation signal from the bovine growth hormone gene;

[0092] FIG. 31 shows a schematic of the strategy for targeted antiparallel integration of an exogenous sequence into the *ATP1A1* locus. Targeted integration of the exogenous sequence occurs within intron 4 (the polynucleotide sequence of intron 4 corresponds to nts 305-404 of SEQ ID NO: 2), in proximity to exon 4 (the polynucleotide

sequence of exon 4 corresponds to nts 101-304 of SEQ ID NO: 2) that encodes Q118 and N129 (vertical black bars). The positions of the DNA double-strand break (DSB) generated within intron 4 by either SpCas9 (G3- SEQ ID NO: 29 (target sequence)) or AsCpf1 (G9-SEQ ID NO: 41 (target sequence)) are predicted to be 19bp and 49bp downstream of exon 4, respectively. In this example, the donor DNA molecule directs the integration of an expression cassette encoding a fluorescent protein (FP) in opposite direction to the *ATP1A1* transcript (antiparallel orientation). The left homology arm (L-HA) of the donor DNA molecule contains the Q118R and N129D (RD) mutations (*asterisks). h*PGK1*: human phosphoglycerate kinase 1 gene promoter, bGHpA: polyadenylation signal from the bovine growth hormone gene;

[0093] FIG. 32 shows the expression level of the TurboGFP protein (GFP-A) upon targeted integration of the expression cassette in the parallel (A) versus antiparallel (B) orientation as detailed in FIGs. 30-31. K562 cells were transfected with eSpCas9(1.1)_No_Flag_ATP1A1_G3 (500ng), a vector that drives the expression of eSpCas9(1.1) along with the sgRNA that targets intron 4 of *ATP1A1* (G3 -SEQ ID NOs:25 (sgRNA) and 29 (target)). The transfection mix was supplemented with the plasmid donors (500ng, SEQ ID NOs: 475 (parallel) and 476 (antiparallel), as indicated. Cells were left untreated or were treated with 0.5µM ouabain for 10 days starting 3 days post-transfection. Flow cytometry was used to determine the % of TurboGFP positive cells in each population;

[0094] FIG. 33 shows the expression level of the mNeonGreen protein upon targeted integration of the expression cassette in the antiparallel orientation as directed by two donor molecules (v1 (A, SEQ ID NO: 477) and v2 (B, SEQ ID NO: 478)) differing by the extent of homology present at the 3' end of the left homology arm. The left homology arm of donor molecule v1 spans and extends 27bp 3' of the predicted DSB. The left homology arm of donor molecule v2 spans and extends 6bp 3' of the predicted DSB. Otherwise, the structure of the donors corresponds to FIGs. 30-31. K562 cells were transfected with eSpCas9(1.1)_No_Flag_ATP1A1_G3 (500ng), a vector that drives the expression of eSpCas9(1.1) along with the sgRNA that targets intron 4 of *ATP1A1*. The transfection mix was supplemented with the plasmid donors (1000ng), as indicated. Cells were left untreated or were treated with 0.5µM ouabain for 10 days starting 3 days post-transfection. Flow cytometry was used to determine the % of mNeonGreen positive cells in each population;

[0095] FIG. 34 shows the expression level of the mNeonGreen protein upon targeted integration of the expression cassette in the antiparallel orientation as directed by donor molecule v2 (see description of FIG. 33 above) and either the eSpCas9 (G3, SEQ ID NOs: 25 (sgRNA) and 29 (target sequence)) or AsCpf1 (G9, SEQ ID NO: 36 (sgRNA) and 41 (target)) nucleases. K562 cells were transfected with either eSpCas9(1.1)_No_Flag_ATP1A1_G3 (500ng) or pY036_ATP1A1_G9 (500ng), vectors that drive the expression of eSpCas9(1.1) or AsCpf1 along with the sgRNA/crRNA that targets intron 4 of *ATP1A1*. The transfection mix was supplemented with the plasmid donor (1000ng), as indicated. Cells were left untreated or were treated with 0.5µM

ouabain for 10 days starting 3 days post-transfection. Flow cytometry was used to determine the % of mNeonGreen positive cells in each population;

[0096] FIG. 35 shows the expression level of the mNeonGreen protein upon targeted integration of the expression cassette in the antiparallel orientation as directed by donor molecule v2 and the AsCpf1 (G9) nucleases. K562 cells were transfected with pY036_ATP1A1_G9 (SEQ ID NOs: 36/41, 500ng), a vector that drives the expression of AsCpf1 along with the crRNA that targets intron 4 of *ATP1A1 (see SEQ ID NO: 2)*. The transfection mix was supplemented with the plasmid donor at two different doses (500ng (A) and 1000ng (B)), as indicated. Cells were left untreated or were treated with 0.5µM ouabain for 10 days starting 3 days post-transfection. Flow cytometry was used to determine the % of mNeonGreen positive cells in each population;

[0097] FIG. 36 shows enrichment of cells with targeted integrations at the endogenous *LMNA* locus by co-editing the *ATP1A1* gene via homology-driven targeted integration of an exogenous expression cassette encoding a fluorescent protein (FP). K562 cells were transfected with eSpCas9(1.1)_No_Flag_ATP1A1_G3_dual_LMNA (100ng), a vector that drives the expression of eSpCas9(1.1) along with two sgRNAs that target respectively *ATP1A1* and *LMNA* (G3 (SEQ ID NOs: 23 (target) and 29 (sgRNA) and G2 (SEQ ID NOs:470 (target) and 471 (sgRNA)). The transfection mix was supplemented with plasmid donors containing an mScarlet (red fluorescent protein) cassette and homology arms to the *LMNA* locus (250ng, SEQ ID NO: 479) in addition to the *ATP1A1* mNeonGreen antiparallel v1 donor (see FIG. 33 and Examples 8 and 9, SEQ ID NO: 477) 250ng) (See FIGs 31-32). Cells were left untreated or were treated with 0.5µM ouabain for 10 days starting 3 days post-transfection. Flow cytometry was used to determine the % of mScarlet and mNeonGreen positive cells in each population; and

[0098] FIGs. 37A-B show a protein sequence alignment between ATP1A1 (SEQ ID NO: 3), ATP1A2 SEQ ID NO: 460), ATP1A3 SEQ ID NO: 462) and ATP1A4 (SEQ ID NO: 464). (A) Alignment of amino acids 1-598 of ATP1A1. Exon 4 region (aa 124-138-) is shown between lines; (B) Alignment of amino acids 599-1023 of ATP1A1.

DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0099] Targeted genome editing using engineered nucleases facilitates the creation of bona fide cellular models for biological research and may be applied to human cell-based therapies. Broadly applicable and versatile methods for increasing editing efficiencies in a wide variety of biomedically important cell types would facilitate the analysis of gene function. Harnessing the multiplexing capabilities of the clustered, regularly interspaced, short palindromic repeat (CRISPR) system, described herein is a robust co-selection strategy enabling enrichment of cells harboring nuclease-driven non-homologous end joining (NHEJ) and homology-directed repair (HDR) events. Creation and selection for dominant alleles of the endogenous genes such as the sodium-potassium pump (Na+,K+-ATPase) resistant to inhibition by ouabain is used to enrich for custom modifications at another unlinked locus of interest,

effectively increasing the recovery of engineered cells.

[00100] An aspect of the system disclosed herein is that the co-selection process can be initiated through NHEJ or HDR processes independently. One only has to switch between sgRNAs targeting juxtaposed regions of gene (e.g., in an intron of an endogenous gene locus conferring resistance to a drug such as *ATP1A1*) and include a donor or patch nucleic acid (e.g., ssODN) to co-select for HDR- instead of NHEJ-driven events. Methods of the present invention are expected to be compatible with any engineered nuclease platforms and any cell type when ubiquitously expressed genes (e.g., *ATP1A1*) are used. The potent and rapid action of the drug combined with the ability to increase the frequency of HDR or NHEJ-driven modification(s) facilitates the incorporation of (e.g., marker-free) genetic changes in human cells while avoiding extensive screening and optimization. Furthermore, the absence of use of exogenous markers is advantageous in the case of therapeutic applications.

[0100] Described herein is a method that can be broadly applied to increase the frequency of genome editing events. The method comprises creating and selecting for dominant gain-of-function alleles of endogenous genes using a selected protein/drug combination. The modified protein/drug combination can then effectively be used to enrich for custom modifications at unlinked loci of interest.

[0101] In embodiments, described herein are methods comprising creating a mutated form of an endogenous gene which in turn encodes a mutated form of the protein that is more resistant to a corresponding drug. The mutated form of the protein is more resistant to a corresponding protein lacking the mutation, which is less resistant/more sensitive to the drug. Thus the mutation(s) or modification(s) in the endogenous gene and in turn the encoded protein confer resistance or greater resistance to the drug, relative to a corresponding form of the protein lacking the mutation(s) or modification(s).

[0102] Non-limiting examples of useful protein/drug combinations are listed in Table 1. Note that several combinations of mutations can confer resistance to a given drug, and thus other possible mutations or mutation combinations can be identified. Specific combinations may be selected for use in a tissue/cell type specific manner.

Table 1: Example of mutants of endogenous human enzymes resistant to cytotoxic drugs that could be used for *in vitro* or *in vivo* co-selection.

Endogenous Gene	GenBank Acc. # of WT protein (Gene ID)	Other sequence/database references	Mutation (s)	Drug resistance
ATP1A1	See Table 9 476	ATP1A1 : HGNC: 799; Ensembl: ENSG00000163399; OMIM: 182310; UniProtKB:	See Tables 5 and 9 and Figs 7 and 10	Ouabain or a derivative thereof

Endogenous Gene	GenBank Acc. # of WT protein (Gene ID)	Other sequence/database references	Mutation (s)	Drug resistance
		P05023; See also Table 9		
ATP1A2	477	ATP1A2 : HGNC: 800; Ensembl: ENSG00000018625; OMIM: 182340; UniProtKB: P50993; See also Table 9	See Table 9	
ATP1A3	478	ATP1A3: HGNC: 801; Ensembl: ENSG00000105409; OMIM: 182350; UniProtKB: P13637; See also Table 9		
ATP1A4	480	ATP1A4 : HGNC: 14073; Ensembl: ENSG00000132681; OMIM: 607321; UniProtKB: Q13733; See also Table 9		
MGMT O-6-methylguanine-DNA methyltransferase	4255	HGNC:7059; Entrez Gene: 4255; Ensembl: ENSG00000170430 C; GenBank: M29971.1 (cDNA); UniProt: P16455	P140K	O6-benzylguanine (BG)
MGMT	4255		G156A	BG
DHFR Dihydrofolate reductase	1719	HGNC:2861; Ensembl:ENSG00000228716; UniProt:P00374	L22Y	Antifolates
IMPDH2 Inosine monophosphate dehydrogenase 2	3615	HGNC:6053; Ensembl:ENSG00000178035; UniProt: P12268	T333I and S351Y	Mycophenolic acid (MPA)
EIF4A1 Eukaryotic translation initiation factor 4A1	1973	HGNC:3282; Ensembl:ENSG00000161960 UniProt: P60842	F163L	Rocaglates

[0103] In addition to the ATP1A1/ouabain combination used herein as a first proof of concept, co-enrichment of edited cells either *in vitro* or *in vivo* may be achieved for example after chemo-selection of cells modified to express the P140K and/or G156A variants of human O(6)-methylguanine-DNA-methyltransferase (MGMT wild type sequence: NM_002412.4 (SEQ ID NO: 481), NP_002403.2 (SEQ ID NO: 482); see FIG. 27). MGMT repairs DNA damage by removing adducts from the O6 position of guanine, decreasing the cytotoxic effects of alkylating agents such as dacarbazine, temozolomide (TMZ), procarbazine, and nitrosoureas such as 1,3-bis-(2-chloroethyl)-1-nitrosourea (BCNU). Inhibition of endogenous MGMT using O6-benzylguanine (BG), a pseudosubstrate that irreversibly inactivates the enzyme, is used to increase the efficacy of alkylating agents and results in better *in vivo* selection of cells overexpressing mutant forms of MGMT that are resistant to inactivation. Technically, it is

interesting to note that these residues are encoded next to an intron-exon boundary and could be targeted in a similar fashion as demonstrated herein for ATP1A1 to favor enrichment of cells edited via HDR (see FIG. 27).

[0104] Another example of protein/drug combination that may be used in accordance with the present invention is DHFR/methotrexate. Protection of cells during administration of antifolate chemotherapy, such as methotrexate (MTX) and trimetrexate (TMTX), can be achieved by overexpression of a mutant form of the DHFR enzyme (DHFR L22Y) that is resistant to MTX inhibition. Co-selection could be achieved through introduction of the L22Y mutation at the endogenous *DHFR* gene via HDR.

[0105] Modification of inosine monophosphate dehydrogenase 2 (IMPDH2) to create mycophenolic acid (MPA) resistance in T and B cells represents another alternative for *ex vivo* and *in vivo* selection of these cells. Mutants of IMPDH2, such as the double mutant T333I and S351Y, have diminished binding affinity for MPA and confer resistance to MPA upon overexpression.

[0106] Mutations within the translation machinery, specifically the eukaryotic initiation factor EIF4A1, can confer resistance to small molecule inhibitors of the rocaglate family.

[0107] Introduction of the eIF4A1 F163L mutation via overexpression or CRISPR/Cas9-mediated genome editing confers resistance to rocaglates in mouse cells.

[0108] In the context of clinical applications, co-selection achieved via the modification of an endogenous gene bypasses the need for introducing foreign DNA and overexpressing transgenes into host cells. Negative aspects of random gene integration and overexpression include (i) insertion mutagenesis, (ii) perturbation of the expression of nearby genes (classic example = activation of oncogene), (iii) immunogenicity of xenogenous enzymes, (iv) silencing of transgene leading to loss of therapeutic effect.

DEFINITIONS

[0109] In order to provide clear and consistent understanding of the terms in the instant application, the following definitions are provided.

[0110] The articles "a," "an" and "the" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article.

[0111] As used in this specification and claim(s), the words “comprising” (and any form of comprising, such as “comprise” and “comprises”), “having” (and any form of having, such as “have” and “has”), “including” (and any form of including, such as “includes” and “include”) or “containing” (and any form of containing, such as “contains” and “contain”) are inclusive or open-ended and do not exclude additional, un-recited elements or method steps and are

used interchangeably with the phrase "including but not limited to".

[0112] For the recitation of numeric ranges herein, each intervening number there between with the same degree of precision is explicitly contemplated. For example, for the range of 18-20, the numbers 18, 19 and 20 are explicitly contemplated, and for the range 6.0-7.0, the number 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, and 7.0 are explicitly contemplated. The terms "such as" are used herein to mean, and is used interchangeably with, the phrase "such as but not limited to".

[0113] Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. For example, any nomenclatures used in connection with, and techniques of, cell and tissue culture, molecular biology, immunology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those that are well known and commonly used in the art. The meaning and scope of the terms should be clear; in the event however of any latent ambiguity, definitions provided herein take precedent over any dictionary or extrinsic definition. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular.

[0114] Practice of the methods, as well as preparation and use of the products and 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.

[0115] 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.

[0116] 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 corresponding naturally-occurring amino acids.

[0117] As used herein, the term “non-conservative mutation” or “non-conservative substitution” in the context of polypeptides refers to a mutation in a polypeptide that changes an amino acid to a different amino acid with different biochemical properties (i.e., charge, hydrophobicity and/or size). Although there are many ways to classify amino acids, they are often sorted into six main groups on the basis of their structure and the general chemical characteristics of their R groups. (i) Aliphatic (Glycine, Alanine, Valine, Leucine, Isoleucine); (ii) Hydroxyl or Sulfur/Selenium-containing (also known as polar amino acids) (Serine, Cysteine, Selenocysteine, Threonine, Methionine); (iii) Cyclic (Proline); (iv) Aromatic (Phenylalanine, Tyrosine, Tryptophan); (v) Basic (Histidine, Lysine, Arginine) and (vi) Acidic and their Amide (Aspartate, Glutamate, Asparagine, Glutamine). Thus, a non-conservative substitution includes one that changes an amino acid of one group with another amino acid of another group (e.g., an aliphatic amino acid for a basic, a cyclic, an aromatic or a polar amino acid; a basic amino acid for an acidic amino acid, a negatively charged amino acid (aspartic acid or glutamic acid) for a positively charged amino acid (lysine, arginine or histidine) etc.

[0118] Conversely, a “conservative substitution” or “conservative mutations” in the context of polypeptides are mutations that change an amino acid to a different amino acid with similar biochemical properties (e.g. charge, hydrophobicity and size). For example, a leucine and isoleucine are both aliphatic, branched hydrophobes. Similarly, aspartic acid and glutamic acid are both small, negatively charged residues. Therefore, changing a leucine for an isoleucine (or vice versa) or changing an aspartic acid for a glutamic acid (or vice versa) are examples of conservative substitutions.

[0119] "Coding sequence" or "encoding nucleic acid" as used herein means the nucleic acids (RNA or DNA molecule) that comprise a nucleotide sequence which encodes a protein or sgRNA. The coding sequence can further include initiation and termination signals operably linked to regulatory elements including a promoter and polyadenylation signal capable of directing expression in the cells of an individual or mammal to which the nucleic acid is administered. The coding sequence may be codon optimized.

[0120] "Complement" or "complementary" as used herein refers to Watson-Crick (e.g., A-T/U and C-G) or Hoogsteen base pairing between nucleotides or nucleotide analogs of nucleic acid molecules. "Complementarity" refers to a property shared between two nucleic acid sequences, such that when they are aligned antiparallel to each other, the nucleotide bases at each position will be complementary.

Sequence similarity

[0121] "Homology" and “homologous” refers to sequence similarity between two peptides or two nucleic acid molecules. Homology can be determined by comparing each position in the aligned sequences. A degree of homology between nucleic acid or between amino acid sequences is a function of the number of identical or matching nucleotides or amino acids at positions shared by the sequences. As the term is used herein, a nucleic acid sequence is "substantially homologous" to another sequence if the two sequences are substantially identical and the functional activity of the sequences is conserved (as used herein, the term “homologous” does not infer evolutionary relatedness, but rather refers to substantial sequence identity, and thus is interchangeable with the terms “identity”/“identical”). Two nucleic acid sequences are considered substantially identical if, when optimally aligned (with gaps permitted), they share at least about 50% sequence similarity or identity, or if the sequences share defined functional motifs. In alternative embodiments, sequence similarity in optimally aligned substantially identical sequences may be at least 60%, 70%, 75%, 80%, 85%, 90% or 95%. For the sake of brevity, the units (e.g., 66, 67…81, 82,…91, 92%….) have not systematically been recited but are considered, nevertheless, within the scope of the present invention.

[0122] Substantially complementary nucleic acids are nucleic acids in which the complement of one molecule is substantially identical to the other molecule. Two nucleic acid or protein sequences are considered substantially identical if, when optimally aligned, they share at least about 70% sequence identity. In alternative embodiments, sequence identity may for example be at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 98% or at least 99%. Optimal alignment of sequences for comparisons of identity may be conducted using a variety of algorithms, such as the local homology algorithm of Smith and Waterman, 1981, Adv. Appl. Math 2: 482, the homology alignment algorithm of Needleman and Wunsch, 1970, J. Mol. Biol. 48:443, the search for similarity method of Pearson and Lipman (Pearson and Lipman 1988), and the computerized implementations of these algorithms (such as GAP, BESTFIT, FASTA and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, Madison, WI, U.S.A.). Sequence identity may also be determined using the BLAST algorithm, described in Altschul et al. (Altschul et al. 1990) 1990 (using the published default settings). Software for performing BLAST analysis may be available through the National Center for Biotechnology Information (through the internet at http://www.ncbi.nlm.nih.gov/). The BLAST algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold. Initial neighborhood word hits act as seeds for initiating searches to find longer HSPs. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension of the word hits in each direction is halted when the following parameters are met: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the

sensitivity and speed of the alignment. One measure of the statistical similarity between two sequences using the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. In alternative embodiments of the invention, nucleotide or amino acid sequences are considered substantially identical if the smallest sum probability in a comparison of the test sequences is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

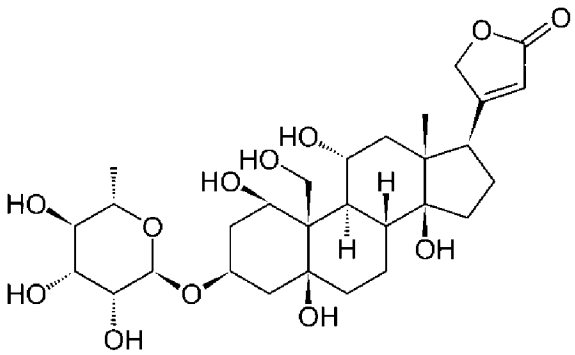
[0123] An alternative indication that two nucleic acid sequences are substantially complementary is that the two sequences hybridize to each other under moderately stringent, or preferably stringent, conditions. Hybridization to filter-bound sequences under moderately stringent conditions may, for example, be performed in 0.5 M NaHPO4, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.2 x SSC/0.1% SDS at 42°C (Ausubel 2010). Alternatively, hybridization to filter-bound sequences under stringent conditions may, for example, be performed in 0.5 M NaHPO4, 7% SDS, 1 mM EDTA at 65°C, and washing in 0.1 x SSC/0.1% SDS at 68°C (Ausubel 2010). Hybridization conditions may be modified in accordance with known methods depending on the sequence of interest (Tijssen 1993). Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point for the specific sequence at a defined ionic strength and pH.

[0124] "Binding" refers to a sequence-specific, non-covalent interaction between macromolecules (e.g., between a protein and a nucleic acid or between a sgRNA and a target polynucleotide or between a sgRNA and a CRISPR nuclease (e.g., Cas9, Cpf1). 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. “Affinity” refers to the strength of binding: increased binding affinity being correlated with a lower Kd.

[0125] A "binding protein" is a protein that is able to bind non-covalently 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.

[0126] “Ouabain” (Cas No: 630-60-4), also known as g-strophanthin, is a cardiac glycoside that inhibits Na+/K+-ATPase sodium-potassium ion pump. It is a plant derived toxic substance that was traditionally used as an arrow poison in eastern Africa for both hunting and warfare. Ouabain is a cardiac glycoside and in lower doses, can be used medically to treat hypotension and some arrhythmias. It acts by inhibiting the Na+/K+-ATPase sodium-potassium ion pump. Once ouabain binds to this enzyme, the enzyme ceases to function, leading to an increase of intracellular sodium. This increase in intracellular sodium reduces the activity of the sodium-calcium exchanger

(NCX), which pumps one calcium ion out of the cell and three sodium ions into the cell down their concentration gradient. Therefore, the decrease in the concentration gradient of sodium into the cell which occurs when the Na/K-ATPase is inhibited reduces the ability of the NCX to function. This in turn elevates intracellular calcium. It is a highly toxic substance which kills cells rapidly. Ouabain has the following chemical structure:



[0127] As used herein, the term "ouabain derivative" or "ouabain-like agent" refers to an agent which, like ouabain, binds to the ATP1A1 polypeptide, inhibits the Na+/K+-ATPase sodium-potassium ion pump and is toxic to cells (may kill cells at a certain concentration). It includes certain cardiotonic steroids (CTS) which may be sufficiently toxic to cells to enable proper selection of target cells in accordance with methods of the present invention.

[0128] Accordingly, as used herein, the term "ouabain resistance" refers to cellular resistance (as compared to normal cells, e.g., cells comprising an ATP1A1 gene lacking the modification conferring the ouabain resistance) to ouabain and/or ouabain derivatives. Such resistant cells are also referred to herein as being tolerant to a toxin such as a cardiotonic steroid (CTS), e.g., ouabain and/or a ouabain derivative. "Tolerance" or "tolerant" as used herein refers to the capacity of a first cell to be less affected by a toxin than a second cell that does not have a modified ATP1A1 gene that confers resistant to the toxin. Tolerant cells grow and develop better in the presence of a toxin when compared to intolerant cells, i.e. have increased growth and viability relative to intolerant cells. Such tolerance can be used as a basis of selection of tolerant cells over intolerant cells, e.g. based on increased growth of tolerant cells relative to intolerant cells when treated with the toxin, in an embodiment the viability of tolerant cells relative to the viability of intolerant cells when treated with the toxin.

[0129] As used herein, "a nuclease-based modification" refers to a modification in a polynucleotide e.g., an endogenous gene locus or genomic sequence) which involves the introduction of a cut (e.g., a double-stranded break in the polynucleotide) which ultimately will trigger a repair mechanism by the cell involving (Non-homologous-end-joining) NHEJ or homologous recombination (HDR). The nuclease-based modification is made by site specific nucleases targeting the polynucleotide of interest (i.e., an endogenous gene locus or genomic sequence). Site-specific nucleases (engineered) are well known and include (but are not limited to) Zinc finger nucleases, meganucleases, Mega-Tals, CRISPR nucleases, TALENs, etc.

[0130] 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.

[0131] 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 55 length and exhibits at least some sequence homology with other TALE repeat sequences within a naturally occurring TALE protein. See, also, U.S. Patent Publication No. 20110301073.

[0132] Zinc finger 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 protein. Similarly, TALEs can be "engineered" to bind to a predetermined nucleotide sequence, for example by engineering of the amino acids involved in DNA binding (the "Repeat Variable Diresidue" or "RVD" region). Therefore, engineered zinc finger proteins or TALE proteins are proteins that are non naturally occurring. Non-limiting examples of methods for engineering zinc finger proteins and TALEs are design and selection. A designed 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 or TALE designs and binding data. See, for example, U.S. Pat. Nos. 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 and U.S. application Ser. No. 13/068,735.

[0133] "Recombination" refers to a process of exchange of genetic information between two polynucleotides. For the purposes of this disclosure, "homologous recombination (HR)" refers to the specialized form of such exchange that takes place, for example, during repair of double-strand breaks in cells via homology-directed repair (HDR) mechanisms. This process requires nucleotide sequence homology, uses a "donor" molecule as a template for 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 re-synthesize 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.

[0134] In the methods described herein, one or more targeted (site-specific) nucleases (e.g., sgRNA/CRISPR nuclease) create a double-stranded break in the target sequence (e.g., cellular chromatin) at a predetermined site. A "donor" polynucleotide, having homology to the nucleotide sequence in the region of the break, may be introduced into the cell if desired (e.g., to introduce a specific modification in the ATP1A1 gene to generate a ouabain resistant ATP1A1 polypeptide and into one or more further target genes to introduce one or more further desired modifications). 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. In any of the methods described herein, additional sgRNA/CRISPR nucleases, pair zinc-finger, Meganucleases, Mega-Tals, and/or additional TALEN proteins can be used for additional double-stranded cleavage of additional target sites within the cell.

[0135] As used herein, the terms “donor” or “patch” nucleic acid are used interchangeably and refers to a nucleic acid that includes a fragment of the endogenous targeted gene of a cell (in some embodiments the entire targeted gene), but which includes desired modification(s) at specific nucleotides (e.g., a fragment of the ATP1A1 gene comprising targeted mutations conferring resistance to cardiotonic steroids (CTS) such as ouabain). The donor (patch) nucleic acid must be of sufficient size and similarity (e.g., in the right and left homology arms) to permit homologous recombination with the targeted gene. Preferably, the donor/patch nucleic acid is (or is flanked at the 5’ end and at the 3’ end by sequences) at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% identical to the endogenous targeted polynucleotide gene sequence. The patch nucleic acid may be provided for example as a ssODN, as a PCR product (amplicon) or within a vector. Preferably, the patch/donor nucleic acid will include modifications with respect to the endogenous gene which i) precludes it from being cut by a sgRNA once integrated in the genome of a cell and/or which facilitate the detection of the introduction of the patch nucleic acid by homologous recombination.

[0136] As used herein, a “targeted gene” or “targeted polynucleotide” corresponds to the polynucleotide within a cell that will be modified by the introduction of the patch nucleic acid. It corresponds to an endogenous gene naturally present within a cell. The targeted gene may comprise one or more mutations associated with a risk of developing a disease or disorder which may be corrected by the introduction of the patch/donor nucleic acid (e.g., will be modified to correspond to the WT gene or to a form which is no longer associated with increased risk of developing a disease or condition). One or both alleles of a targeted gene may be corrected or modified within a cell

in accordance with the present invention.

[0137] A “target polynucleotide” as used herein refers to any endogenous polynucleotide or nucleic acid present in the genome of a cell and encoding or not a known gene product. "Target gene" as used herein refers to any endogenous polynucleotide or nucleic acid present in the genome of a cell and encoding a known or putative gene product (siRNA, protein, etc.). The target gene or target polynucleotide further corresponds to the polynucleotide within a cell that will be modified by a nuclease of the present invention, alone or in combination with the introduction of one or more donor nucleic acid or patch nucleic acids. The target gene or target polynucleotide may be a mutated gene involved in a genetic disease.

[0138] "Promoter" as used herein means a synthetic or naturally-derived nucleic acid molecule which is capable of conferring, modulating or controlling (e.g., activating, enhancing and/or repressing) expression of a nucleic acid in a cell. A promoter may comprise one or more specific transcriptional regulatory sequences to further enhance or repress expression and/or to alter the spatial expression and/or temporal expression of same. A promoter may also comprise distal enhancer or repressor elements, which may be located as much as several thousand base pairs from the start site of transcription. A promoter may be derived from sources including viral, bacterial, fungal, plants, insects, and animals. A promoter may regulate the expression of a gene component constitutively or differentially with respect to cell, the tissue or organ in which expression occurs or, with respect to the developmental stage at which expression occurs, or in response to external stimuli such as physiological stresses, pathogens, metal ions, or inducing agents. Representative examples of promoters include the bacteriophage T7 promoter, bacteriophage T3 promoter, SP6 promoter, lac operator-promoter, tac promoter, SV40 late promoter, SV40 early promoter, RSV-LTR promoter, CMV IE promoter, SV40 early promoter or SV40 late promoter and the CMV IE promoter. In embodiments, the U6 promotor is used to express one or more sgRNAs in a cell.

[0139] “Vector" as used herein means a nucleic acid sequence containing an origin of replication. A vector may be a viral vector, bacteriophage, bacterial artificial chromosome or yeast artificial chromosome. A vector may be a DNA or RNA vector. A vector may be a self- replicating extrachromosomal vector, and preferably, is a DNA plasmid. For example, the vector may comprise nucleic acid sequence(s) that/which encode(s) a sgRNA, a donor (or patch) nucleic acid, and/or a CRISPR nuclease (e.g., Cas9 or Cpf1) of the present invention. A vector for expressing one or more sgRNA will comprise a “DNA” sequence of the sgRNA.

[0140] "Adeno-associated virus" or "AAV" as used interchangeably herein refers to a small virus belonging to the genus Dependovirus of the Parvoviridae family that infects humans and some other primate species. AAV is not currently known to cause disease and consequently the virus causes a very mild immune response.

Gene editing and selection methods

[0141] Gene editing and selection methods of the present invention allow to rapidly select and enrich for genome edited cells. In embodiments, methods of the present invention further allow to enrich for certain types of modifications (i.e., increase the efficiency of a cell to generate a specific type of modification(s) e.g., knock out/deletion/inactivation vs targeted insertion/mutation(s)) by selectively favoring Homologous directed repair (HDR) or Non-homologous end joining (NHEJ). Methods of the present invention involve generating a ATP1A1 polypeptide resistant to cardiotonic steroids (CTS), such as ouabain by cleaving with a nuclease within the endogenous ATP1A1 gene. Ouabain resistance was shown to allow for rapid selection and efficient enrichment of genome-edited cells. The sole introduction of targeted double stranded breaks (DSB) within the APT1A1 gene sequence was also surprisingly found to generate spontaneous deletion mutants/variants resistant to ouabain. Furthermore, introduction of targeted DSB within an intron of the ATP1A1 gene together with a specific ATP1A1 donor polynucleotide comprising targeted modifications conferring ouabain resistance, was shown to enable to rapidly and efficiently select for cells modified by HDR at a second targeted locus. Thus, Applicant provides herein methods to increase the frequency of specific genome editing events and to rapidly and efficiently select target cells comprising the desired modification(s) (editing events).

[0142] Various types of endonucleases or nickases may be used to induce a DSB at selected site(s) in one or more targeted polynucleotides. Non-limiting examples of useful endonucleases and nickases include meganucleases, Zinc Finger Nucleases (ZFNs), Transcription Activator-Like Effector nucleases (TALENs), Mega-Tals, or the CRISPR nucleases or nickases used in combination with at least one (e.g., one or two) guide RNA(s) (sgRNA) in the Clustered regularly interspaced short palindrome repeat (CRISPR) system. Each of these technologies can be used generate ouabain resistance and modify one or more target polynucleotides in accordance with the present invention.

[0143] Preferably, the present invention uses the CRISPR system (i.e., combination of sgRNA and CRISPR nuclease or nickase such as Cas9 and Cpf1), optionally together with a donor (patch) polynucleotide sequence to generate ouabain resistance and introduce one or more genetic modifications in a target endogenous polynucleotide/gene. Applicants demonstrate herein that sgRNAs can be used with a CRISPR nuclease to generate ouabain resistance and to increase frequency of specific genome editing events in one or more targeted polynucleotide human cells.

CRISPR system

[0144] CRISPR technology is a system for genome editing, e.g., for modification of the expression of a specific gene.

[0145] This system stems from findings in bacterial and archaea which have developed adaptive immune

defenses termed clustered regularly interspaced short palindromic repeats (CRISPR) systems, which use crRNAs and Cas proteins to degrade complementary sequences present in invading viral and plasmid DNA. Jinek et al. (46) and Mali et al. (45) have engineered a type II bacterial CRISPR system using custom guide RNA (sgRNA) to induce double strand break(s) in DNA. In one system, the Cas9 protein was directed to genomic target sites by a synthetically reconstituted "guide RNA" ("sgRNA"), corresponding to a crRNA-tracrRNA fusion that obviates the need for RNase III and crRNA processing in general. It comprises a “sgRNA guide sequence” or “sgRNA target sequence” and an RNA sequence (Cas recognition sequence)”, which is necessary for Cas (e.g., Cas9) binding to the targeted gene. The sgRNA guide sequence is the sequence that confers specificity. It hybridizes with (i.e., it is complementary to) the opposite strand of a target sequence (i.e., it corresponds to the RNA sequence of a DNA target sequence). Other CRISPR systems using different CRISPR nucleases have been developed and are known in the art (e.g., using the Cpf1 nuclease instead of a Cas9 nuclease).

[0146] Because the original Cas9 nuclease combined with a sgRNA may produce off-target mutagenesis, one may alternatively use in accordance with the present invention a pair of specifically designed sgRNAs in combination with a Cas9 nickase or in combination with a dCas9-FokI nuclease to cut both strands of DNA.

[0147] In embodiments, provided herein are CRISPR/nuclease-based engineered systems for use in modifying the *ATP1A1* gene and inducing ouabain resistance in cells. The CRISPR-based engineered systems of the present invention are designed to (i) target and cleave the *ATP1A1* gene (e.g., in *ATP1A1* intron 3, exon 4 or intron 4) to generate ATP1A1 variants conferring ouabain resistance (deletion or insertion variants (i.e., NHEJ-induced variants) or specifically selected and targeted variants (i.e., HDR-induced variants using a donor *ATP1A1* polynucleotide comprising the desired mutation(s)/modification(s)), (ii) increase the frequency of specific genome editing events (i.e., favoring NHEJ or HDR-directed modifications) in cells; and/or (iii) simplify and/or increase the speed of identification and selection of genome-edited target cells having the desired modification in one or more target polynucleotides.

[0148] The CRISPR/nuclease-based systems of the present invention include at least one CRISPR nuclease (e.g. a Cas9 or Cpf1 nuclease) and at least one sgRNA targeting the endogenous *ATP1A1* gene in target cells.

[0149] Accordingly, in an aspect, the present invention involves the design and preparation of one or more sgRNAs for inducing a DSB (or two single stranded breaks (SSB) in the case of a nickase) in *ATP1A1*. In embodiments, the present invention also involves the design and preparation of one or more sgRNAs for inducing a DSB (or two SSBs in the case of a nickase) in a target polynucleotide located at a different locus within the genome of target cells. The sgRNAs (targeting the *ATP1A1* gene or targeting the *ATP1A1* gene and the target polynucleotide(s)) and the nuclease are then used together to introduce the desired modification(s) (i.e., gene-editing events) by NHEJ or HDR within the genome of one or more target cells. When the desired modification(s)

include specific point mutation(s) or insertions/deletion(s) (e.g., in the *ATP1A1* gene or in the *ATP1A1* gene and target polynucleotide(s)), one or more donor or patch nucleic acids comprising the desired modification(s) are provided to introduce the modification(s) by HDR.

sgRNAs

[0150] In order to cut DNA at a specific site, CRISPR nucleases require the presence of a sgRNA and a of a protospacer adjacent motif (PAM) on the targeted gene. The PAM immediately follows (i.e., is adjacent to) the sgRNA target sequence in the targeted polynucleotide gene sequence. The PAM is located at the 3' end or 5' end of the sgRNA target sequence (depending on the CRISPR nuclease used) but is not included in the sgRNA guide sequence. For example, the PAM for Cas9 CRSIPR nucleases is located at the 3' end of the sgRNA target sequence on the target gene while the PAM for Cpf1 nucleases is located at the 5' end of the sgRNA target sequence on the target gene. Different CRISPR nucleases also require a different PAM. Accordingly, selection of a specific polynucleotide sgRNA target sequence (e.g., on the *ATP1A1* gene nucleic acid sequence) is generally based on the CRISPR nuclease used. The PAM for the *Streptococcus pyogenes* Cas9 CRISPR system is 5'-NRG-3', where R is either A or G, and characterizes the specificity of this system in human cells. The PAM of *S. aureus* Cas9 is NNGRR. The S. *pyogenes* Type II system naturally prefers to use an "NGG" sequence, where "N" can be any nucleotide, but also accepts other PAM sequences, such as "NAG" in engineered systems. Similarly, the Cas9 derived from *Neisseria meningitidis* (NmCas9) normally has a native PAM of NNNNGATT, but has activity across a variety of PAMs, including a highly degenerate NNNNGNNN PAM. The PAM for AsCpf1 or LbCpf1 CRISPR nuclease is TTTN. In an embodiment, the PAM for a Cas9 protein used in accordance with the present invention is a NGG trinucleotide-sequence (Cas9). In another embodiment, the PAM for a Cpf1 CRISPR nuclease used in accordance with the present invention is a TTTN nucleotide sequence. Table 2 below provides a list of non-limiting examples of CRISPR/nuclease systems with their respective PAM sequences.

Table 2: Non-exhaustive list of CRISPR-nuclease systems from different species (see. Mohanraju, P. et al., PMID 27493190; Shmakov, S et al., PMID: 26593719; and Zetsche, B. et al., PMID: 26422227). Also included are engineered variants recognizing alternative PAM sequences (see Kleinstiver, BP. et al., (Nature biotech 2015) PMID: 26524662 and Kleinstiver, BP. et al., (Nature 2015)).

CRISPR nuclease	PAM Sequence
Streptococcus pyogenes (SP); SpCas9	NGG + NAG
SpCas9 D1135E variant	NGG (reduced NAG binding)
SpCas9 VRER variant	NGCG
SpCas9 EQR variant	NGAG
SpCas9 VQR variant	NGAN or NGNG
Staphylococcus aureus (SA); SaCas9	NNGRRT or NNGRR(N)

CRISPR nuclease	PAM Sequence
SaCas9 KKH variant	NNNRRT
Neisseria meningitidis (NM)	NNNNGATT
Streptococcus thermophilus (ST)	NNAGAAW
Treponema denticola (TD)	NAAAAC
AsCpf1(*Acidominococcus*)	TTTN
AsCpf1 S542R/K607R	TYCV
AsCpf1 S542R/K548V/N552R	TATV
LbCpf1 (*Lachnospiraceae*)	TTTN
LbCpf1 G532R/K595R	TYCV

[0151] As used herein, the expression "sgRNA" refers to a guide RNA which works in combination with a CRISPR nuclease to introduce a cut into DNA. The sgRNA comprises a sgRNA guide sequence and a "CRISPR nuclease recognition sequence".

[0152] As used herein, the expression "sgRNA guide sequence" refers to the corresponding RNA sequence of the "sgRNA target sequence". Therefore, it is the RNA sequence equivalent of the protospacer on the target polynucleotide gene sequence. It does not include the corresponding PAM sequence in the genomic DNA. It is the sequence that confers target specificity. The sgRNA guide sequence is linked to a CRISPR nuclease recognition sequence which binds to the nuclease (e.g., Cas9/Cpf1). The sgRNA guide sequence recognizes and binds to the targeted gene of interest. It hybridizes with (i.e., is complementary to) the opposite strand of a target gene sequence, which comprises the PAM (i.e., it hybridizes with the DNA strand opposite to the PAM). As noted above, the "PAM" is the nucleic acid sequence, that immediately follows (is contiguous to) the target sequence on the *ATP1A1* gene or target polynucleotide but is not in the sgRNA. Exemplary PAM and target sequences (corresponding to the DNA sequence of the sgRNA guide sequence) targeting the *ATP1A1* gene are shown in Table 3, below.

[0153] A "CRISPR nuclease recognition sequence" as used herein refers broadly to one or more RNA sequences (or RNA motifs) required for the binding and/or activity (including activation) of the CRISPR nuclease on the target gene. Some CRISPR nucleases require longer RNA sequences than other to function. Also, some CRISPR nucleases require multiple RNA sequences (motifs) to function while others only require a single short RNA sequence/motif. For example, Cas9 proteins require a tracrRNA sequence in addition to a crRNA sequence to function while Cpf1 only requires a crRNA sequence (e.g., SEQ ID NO: 41 for AsCpf1). Thus, unlike Cas9, which requires both crRNA sequence and a tracrRNA sequence (or a fusion or both crRNA and tracrRNA) to mediate interference, Cpf1 processes crRNA arrays independent of tracrRNA, and Cpf1-crRNA complexes alone cleave target DNA molecules, without the requirement for any additional RNA species (see Zetsche et al., PMID:

26422227).

[0154] The "CRISPR nuclease recognition sequence" included in the sgRNA described herein is thus selected based on the specific CRISPR nuclease used. It includes direct repeat sequences and any other RNA sequence known to be necessary for the selected CRISPR nuclease binding and/or activity. Various RNA sequences which can be fused to an RNA guide sequence to enable proper functioning of CRISPR nucleases (referred to herein as CRISPR nuclease recognition sequence) are well known in the art and can be used in accordance with the present invention. The "CRISPR nuclease recognition sequence" may thus include a crRNA sequence only (e.g., for AsCpf1 activity, such as the CRISPR nuclease recognition sequence UAAUUUCUAC UCUUGUAGAU set forth in SEQ ID NO: 41) or may include additional sequences (e.g., tracrRNA sequence necessary for Cas9 activity, such as the CRISPR nuclease recognition sequence set forth in SEQ ID NO: 31 which includes both crRNA and tracrRNA sequences). Furthermore, in accordance with the present invention and as well known in the art, RNA motifs necessary for CRISPR nuclease binding and activity may be provided separately (e.g., (i) RNA guide sequence-crRNA CRISPR recognition sequence" (also known as crRNA) in one RNA molecule and (ii) a tracrRNA CRISPR recognition sequence on another, separate RNA molecule. Alternatively, all necessary RNA sequences (motifs) may be fused together in a single RNA guide. The CRISPR recognition sequence is preferably fused directly to the sgRNA guide sequence (in 3' (e.g., Cas9) or 5' (Cpf1) depending on the CRISPR nuclease used) but may include a spacer sequence separating two RNA motifs. In embodiments, the CRISPR nuclease recognition sequence is a Cas9 recognition sequence having at least 65 nucleotides. In embodiments, the CRISPR nuclease recognition sequence is a Cas9 CRISPR nuclease recognition sequence having at least 85 nucleotides. In embodiments, the CRISPR nuclease recognition sequence is a Cpf1 recognition sequence (5' direct repeat) having about 19 nucleotides. In a particular embodiment, the Cas9 recognition sequence comprises (or consists of) the sequence as set forth in SEQ ID NO: 31. In a particular embodiment, the AsCpf1 recognition sequence comprises (or consists of) the sequence UAAUUUCUAC UCUUGUAGAU (SEQ ID NO: 42). The sgRNA of the present invention may comprise any variant of the above noted sequences, provided that it allows for the proper functioning of the selected CRISPR nuclease (e.g., binding of the CRISPR nuclease protein to the *ATP1A1* gene and/or target polynucleotide sequence(s)).

[0155] Together, the RNA guide sequence and CRSIPR nuclease recognition sequence(s) provide both targeting specificity and scaffolding/binding ability for the CRISPR nuclease of the present invention. sgRNAs of the present invention do not exist in nature, i.e., is a non-naturally occurring nucleic acid(s).

[0156] A "target region", "target sequence" or "protospacer" in the context of sgRNAs and CRISPR system of the present invention are used herein interchangeably and refers to the region of the target gene, which is targeted by the CRISPR/nuclease-based system, without the PAM. It refers to the sequence corresponding to the nucleotides

that precede the PAM (i.e., in 5' or 3' of the PAM, depending of the CRISPR nuclease) in the genomic DNA. It is the sequence that is included into a sgRNA expression construct (e.g., vector/plasmid/AVV). The CRISPR/nuclease-based system may include at least one (i.e., one or more) sgRNAs, wherein each sgRNA target different DNA sequences on the target gene. The target DNA sequences may be overlapping. The target sequence or protospacer is followed or preceded by a PAM sequence at an (3' or 5' depending on the CRISPR nuclease used) end of the protospacer. Generally, the target sequence is immediately adjacent (i.e., is contiguous) to the PAM sequence (it is located on the 5' end of the PAM for SpCas9-like nuclease and at the 3' end for Cpf1-like nuclease).

Table 3: Exemplary sgRNA target sites and their adjacent PAM in *ATP1A1* exon 4 and surrounding introns. The position of PAM is provided with reference to SEQ ID NO: 2. Identification of SpCas9 PAM sites was performed using CCTop™ (Stemmer, M., Thumberger, T., del Sol Keyer, M., Wittbrodt, J. and Mateo, J.L. *CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool.* PLOS ONE (2015). doi: 10.1371/journal.pone.0124633) Cpf1 PAM were searched manually. Further non-limiting examples of SpCas9 PAM for additional exons are provided in FIGs 28A-I.

sgRNA	sgRNAs target sequences + PAM for Cas9	SEQ ID NO.	PAM	PAM Position (cs: coding sequence, in: intron)	Strand (+/-) ATP1A1 sequence
T1	CGTTGGGACCATCTCGCGCC	SEQ ID NO: 47	AGG	cs (130-132)	(-)
T2	GTAGCTGGACGGTTGCAATA	SEQ ID NO: 48	AGG	in (12-14)	(-)
T3	ATTGCAACCGTCCAGCTACC	SEQ ID NO: 49	AGG	in (36-38)	(+)
T4	GCTGAGATCCTGGCGCGAGA	SEQ ID NO: 50	TGG	cs (142-144)	(+)
T5	CAACCGTCCAGCTACCAGGT	SEQ ID NO: 51	AGG	in (40-42)	(+)
T6	GTATATTGCCTTGTAAGTGC	SEQ ID NO: 52	TGG	in (62-64)	(+)
T7	CAATATACCTACCTGGTAGC	SEQ ID NO: 53	TGG	in (27-29)	(-)
T8	TACAAGGCAATATACCTACC	SEQ ID NO: 54	TGG	in (34-36)	(-)
T9	CGACAAAACTTGATCCATTC	SEQ ID NO: 55	AGG	cs (176-178)	(-)
T10	TTGATCCATTCAGGAGTAGT	SEQ ID NO: 56	GGG	cs (167-169)	(-)
T11	AAGTTTTGTCGGCAGCTCTT	SEQ ID NO: 57	TGG	cs (208-210)	(+)
T12	AGTTTTGTCGGCAGCTCTTT	SEQ ID NO: 58	GGG	cs (209-210)	(+)
T13/G1	GAACTCACATTATCGTTTTG	SEQ ID NO: 59	AGG	cs (290-292)	(-)
T14	ATACCTACCTGGTAGCTGGA	SEQ ID NO: 60	CGG	in (23-25)	(-)
T15	CTTGATCCATTCAGGAGTAG	SEQ ID NO: 61	TGG	cs (168-170)	(-)
T16/G3	GAGTTCTGTAATTCAGCATA	SEQ ID NO: 29	TGG	in (327-329)	(+)
T17	CCCCTCCCACTACTCCTGAA	SEQ ID NO: 62	TGG	cs (182-184)	(+)
T18	ACGGTTGCAATAAGGATTTA	SEQ ID NO: 63	AGG	in (4-6)	(-)
T19	TCCATTCAGGAGTAGTGGGA	SEQ ID NO: 64	GGG	cs (163-165)	(-)
T20	TTTTGTCGGCAGCTCTTTGG	SEQ ID NO: 65	GGG	cs (211-213)	(+)
T21	CTGAATGGATCAAGTTTTGT	SEQ ID NO: 66	CGG	cs (197-199)	(+)
T22	AAACTGTACCAGCACTTACA	SEQ ID NO: 67	AGG	in (50-52)	(-)
T23	ATCCATTCAGGAGTAGTGGG	SEQ ID NO: 68	AGG	cs (164-166)	(-)
T24	AATTACCACTCATTACTTAA	SEQ ID NO: 69	TGG	in (402-404)	(+)
T25	CATCCAAGCTGCTACAGAAG	SEQ ID NO: 70	AGG	cs (285-287)	(+)

sgRNA	sgRNAs target sequences + PAM for Cas9	SEQ ID NO.	PAM	PAM Position (cs: coding sequence, in: intron)	Strand (+/-) ATP1A1 sequence
T26	CCATTCAGGAGTAGTGGGAG	SEQ ID NO: 71	GGG	cs (162-164)	(-)
T27	GTTTTGTCGGCAGCTCTTTG	SEQ ID NO: 72	GGG	cs (210-212)	(+)
T28/G4	GTTCCTCTTCTGTAGCAGCT	SEQ ID NO: 30	TGG	cs (268-270)	(-)
T29	TGGAGCGATTCTTTGTTTCT	SEQ ID NO: 73	TGG	cs (255-257)	(+)
T30	TTCTCAATGTTACTGTGGAT	SEQ ID NO: 74	TGG	cs (235-237)	(+)
T31	GGGGGTTCTCAATGTTACTG	SEQ ID NO: 75	TGG	cs (230-232)	(+)
T32	TGCTCGTGCAGCTGAGATCC	SEQ ID NO: 76	TGG	cs (132-134)	(+)
T33	TGTTAATCCCTGAGAAGCAG	SEQ ID NO: 77	TGG	in (87-89)	(-)
T34	AGTGGTAATTGAGAAGAAGT	SEQ ID NO: 78	GGG	in (369-371)	(-)
T35	TTTGTCGGCAGCTCTTTGGG	SEQ ID NO: 79	GGG	cs (212-214)	(+)
T36	GAGTGGTAATTGAGAAGAAG	SEQ ID NO: 80	TGG	in (370-372)	(-)
T37	CATTCAGGAGTAGTGGGAGG	SEQ ID NO: 81	GGG	cs (161-163)	(-)
T38	GAAGCAGTGGAATATAAATA	SEQ ID NO: 82	AGG	in (74-76)	(-)
T39	TTTATATTCCACTGCTTCTC	SEQ ID NO: 83	AGG	in (99-101)	(+)
T40	TTATATTCCACTGCTTCTCA	SEQ ID NO: 84	GGG	in (100-102)	(+)
T41	GGAGGGGGAGTGAGGGCGTT	SEQ ID NO: 85	GGG	cs (146-148)	(-)
T42	GGGAGGGGGAGTGAGGGCGT	SEQ ID NO: 86	TGG	cs (147-149)	(-)
T43	AGTAGTGGGAGGGGGAGTGA	SEQ ID NO: 87	GGG	cs (153-155)	(-)
T44	GAAGAAGTGGGAGACAAAGA	SEQ ID NO: 88	CGG	in (357-359)	(-)
T45	GAGTAGTGGGAGGGGGAGTG	SEQ ID NO: 89	AGG	cs (154-156)	(-)
G2	[G]ATCCAAGCTGCTACAGAAG	SEQ ID NO: 28	AGG	cs (285-287)	(+)
	sgRNAs target sequences for Cpf1				
G5	TTTCTTGGCTTATAGCATCC	SEQ ID NO: 37		cs (246-249)	(+)
G6	TTGGCTTATAGCATCCAAGC	SEQ ID NO: 38		cs (250-253)	(+)
G7	AGGTTCCTCTTCTGTAGCAG	SEQ ID NO: 39		cs (293-296)	(-)
G8	GAGGTTCCTCTTCTGTAGCA	SEQ ID NO: 40		cs (294-297)	(-)
G9	TAGTACACATCAGATATCTT	SEQ ID NO: 41		in (321-333)	(+)

[0157] In embodiments, the sgRNA of the present invention comprises a “sgRNA guide sequence” or has a “sgRNA target sequence” which corresponds to the target sequence on the *ATP1A1* gene or target polynucleotide sequence that is followed or preceded by a PAM sequence (is adjacent to a PAM). The sgRNA may comprise a "G" at the 5' end of its polynucleotide sequence. The presence of a “G” in 5’ is preferred when the sgRNA is expressed under the control of the U6 promoter (Taeyoung KooJungjoon Lee and Jin-Soo Kim Mol Cells. 2015 Jun 30; 38(6): 475–481). The CRISPR/nuclease system of the present invention may use sgRNAs of varying lengths. The sgRNA may comprise a sgRNA guide sequence of at least at least a 10, at least 12 nts, at least a 13 nts, at least a 14 nts, at least a 15 nts, at least a 16 nts, at least a 17 nts, at least a 18 nts, at least a 19 nts, at least a 20 nts, at least a 21 nts, at least a 22 nts, at least a 23 nts, at least a 24 nts, at least a 25 nts, at least a 30 nts, or at least a 35 nts of a target sequence on the *ATP1A1* gene or target polynucleotide (such target sequence is followed or preceded by a PAM on the *ATP1A1* gene or target polynucleotide but is not part of the sgRNA). The length of the sgRNA is selected based on the specific CRISPR nuclease used. In embodiments, the “sgRNA guide sequence” or “sgRNA

target sequence" may be least 17 nucleotides (17, 18, 19, 20, 21, 22, 23) long, preferably between 17 and 30 nts long, more preferably between 17-22 nucleotides long. In embodiments, the sgRNA guide sequence is between 10-40, 10-30, 12-30, 15-30, 18-30, or 10-22 nucleotides long. In embodiments, the PAM sequence is "NGG", where "N" can be any nucleotide. In embodiments, the PAM sequence is "TTTN", where "N" can be any nucleotide. sgRNAs may target any region of a target gene (e.g., *ATP1A1*) which is immediately adjacent (contiguous, adjoining, in 5' or 3') to a PAM (e.g., NGG/TTTN or CCN/NAAA for a PAM that would be located on the opposite strand) sequence. In embodiments, the sgRNA of the present invention has a target sequence that is located in an exon (the sgRNA guide sequence consists of the RNA sequence of the target (DNA) sequence which is located in an exon). In embodiments, the target sequence of the sgRNA is in exon 4 of the *ATP1A1* gene. In embodiments, the sgRNA of the present invention has a target sequence that is located in an intron (the sgRNA guide sequence consists of the RNA sequence of the target (DNA) sequence which is located in an intron). In embodiments, the target sequence of the sgRNA is in intron 3 or intron 4 of the *ATP1A1* gene. In embodiments, the sgRNA may target any region (sequence) which is followed (or preceded, depending on the CRISPR nuclease used) by a PAM in the *ATP1A1* gene or target polynucleotide and which may be used to confer ouabain resistance.

[0158] Although a perfect match between the sgRNA guide sequence and the DNA sequence on the targeted gene is preferred, a mismatch between a sgRNA guide sequence and target sequence on the gene sequence of interest is also permitted as along as it still allows hybridization of the sgRNA with the complementary strand of the sgRNA target polynucleotide sequence on the targeted gene. A seed sequence of between 8-12 consecutive nucleotides in the sgRNA, which perfectly matches a corresponding portion of the sgRNA target sequence is preferred for proper recognition of the target sequence. The remainder of the guide sequence may comprise one or more mismatches. In general, sgRNA activity is inversely correlated with the number of mismatches. Preferably, the sgRNA of the present invention comprises 7 mismatches, 6 mismatches, 5 mismatches, 4 mismatches, 3 mismatches, more preferably 2 mismatches, or less, and even more preferably no mismatch, with the corresponding sgRNA target gene sequence (less the PAM). Preferably, the sgRNA nucleic acid sequence is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% and 99% identical to the sgRNA target polynucleotide sequence in the gene of interest (e.g., ATP1A1). Of course, the smaller the number of nucleotides in the sgRNA guide sequence the smaller the number of mismatches tolerated. The binding affinity is thought to depend on the sum of matching sgRNA-DNA combinations.

[0159] Non-limiting examples of sgRNA target sequences are presented in Table 3 above. Any sgRNA guide sequence can be selected in the target gene, as long as it allows introducing at the proper location, the desired modification(s) (e.g., spontaneous insertions/deletions or selected target modification(s) using one or more patch/donor sequence(s)). Accordingly, the sgRNA guide sequence or target sequence of the present invention may be in coding or non-coding regions of the *ATP1A1* gene (i.e., introns or exons). In embodiments, the sgRNA target

sequence is selected in intron 3, exon 4 and/or intron 4 of the *ATP1A1* gene. In embodiments, the sgRNA guide sequence, in combination with the donor/patch sequence allows to replace at least a portion of an in intron and an exon of the endogenous *ATP1A1* polynucleotide gene sequence within a cell so as to confer ouabain resistance. SEQ ID NO: 2 (from gene ID 476) presents a fragment of the *ATP1A1* polynucleotide gene sequence comprising part of intron 3, exon 4 and part of intron 3 from which sgRNA guide sequence may easily be selected in accordance with the present invention. The complementary strand of the sequence may alternatively and equally be used to identify proper PAM and sgRNA target/guide sequences. In embodiments, the sgRNA of the present invention comprise a guide sequence having a target sequence/guide sequence as set forth in Table 3, FIG. 28 or any one of SEQ ID NOs: 27-30, 37-41.

[0160] The number of sgRNAs administered to or expressed in a target cell in accordance with the methods of the present invention may be at least 1 sgRNA, at least 2 sgRNAs, at least 3 sgRNAs at least 4 sgRNAs, at least 5 sgRNAs, at least 6 sgRNAs, at least 7 sgRNAs, at least 8 sgRNAs, at least 9 sgRNAs, at least 10 sgRNAs, at least 11 sgRNAs, at least 12 sgRNAs, at least 13 sgRNAs, at least 14 sgRNAs, at least 15 sgRNAs, at least 16 sgRNAs, at least 17 sgRNAs, or at least 18 sgRNAs. The number of sgRNAs administered to or expressed in a cell may be between at least 1 sgRNA and 15 sgRNAs, 1 sgRNA and least 10 sgRNAs, 1 sgRNA and 8 sgRNAs, 1 sgRNA and 6 sgRNAs, 1 sgRNA and 4 sgRNAs, 1 sgRNA and sgRNAs, 2 sgRNA and 5 sgRNAs, or 2 sgRNAs and 3 sgRNAs.

[0161] In embodiments a sgRNA having a sgRNA target sequence/guide sequence as shown in Table 3 (SEQ ID NOs: 27-30 and 37-41) or FIG. 28 is used to modify the *ATP1A1* gene and confer ouabain resistance in one or more target cells.

CRISPR nucleases

[0162] Recently, Tsai et al. (44). have designed recombinant dCas9-FoKI dimeric nucleases (RFNs) that can recognize extended sequences and edit endogenous genes with high efficiency in human cells. These nucleases comprise a dimerization-dependent wild type FokI nuclease domain fused to a catalytically inactive Cas9 (dCas9) protein. Dimers of the fusion proteins mediate sequence specific DNA cleavage when bound to target sites composed of two half-sites (each bound to a dCas9 (i.e., a Cas9 nuclease devoid of nuclease activity) monomer domain) with a spacer sequence between them. The dCas9-FoKI dimeric nucleases require dimerization for efficient genome editing activity and thus, use two sgRNAs for introducing a cut into DNA.

[0163] The recombinant CRISPR nuclease that may be used in accordance with the present invention is i) derived from a naturally occurring Cas; and ii) has a nuclease (or nickase) activity to introduce a DSB (or two SSBs in the case of a nickase) in cellular DNA when in the presence of appropriate sgRNA(s). Thus, as used herein, the term "CRISPR nuclease" refers to a recombinant protein which is derived from a naturally occurring Cas nuclease which

has nuclease or nickase activity and which functions with the sgRNAs of the present invention to introduce DSBs (or one or two SSBs) in the targets of interest, e.g., the *ATP1A1* gene and/or target polynucleotide. In embodiments, the CRISPR nuclease is SpCas9. In embodiments, the CRISPR nuclease is Cpf1. In another embodiment, the CRISPR nuclease is a Cas9 protein having a nickase activity. As used herein, the term "Cas9 nickase" refers to a recombinant protein which is derived from a naturally occurring Cas9 and which has one of the two nuclease domains inactivated such that it introduces single stranded breaks (SSB) into the DNA. It can be either the RuvC or HNH domain. In a further embodiment, the Cas protein is a dCas9 protein fused with a dimerization-dependant FokI nuclease domain.

[0164] Exemplary CRISPR nucleases that may be used in accordance with the present invention are provided in Table 2 above.

[0165] CRISPR nucleases such as Cas9/nucleases cut 3-4bp upstream of the PAM sequence. CRISPR nucleases such as Cpf1 on the other hand, generate a 5' overhang. The cut occurs 19 bp after the PAM on the targeted (+) strand and 23 bp on the opposite strand (Zetsche et al., 2015, PMID 26422227). There can be some off-target DSBs using wildtype Cas9. The degree of off-target effects depends on a number of factors, including: how closely homologous the off-target sites are compared to the on-target site, the specific site sequence, and the concentration of nuclease and guide RNA (sgRNA). These considerations only matter if the PAM sequence is immediately adjacent to the nearly homologous target sites. The mere presence of additional PAM sequences should not be sufficient to generate off target DSBs; there needs to be extensive homology of the protospacer followed or preceded by PAM.

Optimization of codon degeneracy

[0166] Because CRISPR nuclease proteins are (or are derived from) proteins normally expressed in bacteria, it may be advantageous to modify their nucleic acid sequences for optimal expression in eukaryotic cells (e.g., mammalian cells) when designing and preparing CRISPR nuclease recombinant proteins. This has already been done for the embodiments of the present invention described in Examples 1-6. Similarly, donor or patch nucleic acids of the present invention used to introduce specific modifications in *ATP1A* or in the target polynucleotide may use codon degeneracy (e.g., to introduce new restriction sites for enabling easier detection of the targeted modification).

[0167] Accordingly, the following codon chart (Table 4) may be used, in a site-directed mutagenic scheme, to produce nucleic acids encoding the same or slightly different amino acid sequences of a given nucleic acid:

[0168] Table 4: Codons encoding the same amino acid

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

ATP1A1 mutations conferring ouabain resistance

[0169] Crystal structures of the sodium-potassium pump with bound ouabain and related CTSs have revealed that the inhibitors are wedged very deeply between transmembrane helices in addition to contacting the extracellular surface of the ATPase alpha-subunit (Laursen, M et al., PNAS (2015) PMID: 25624492; Laursen, M et al., PNAS (2013) PMID: 23776223; Ogawa, H. et al., PNAS (2009), PMID: 196666591). The high-affinity CTS-binding site is constituted by the transmembrane helices M1–6, forming a pocket exposed to the extracellular side. In addition, several residues located in the ten transmembrane domains and five extracellular loops impact binding ((Laursen, M et al., PNAS (2015) PMID: 25624492; Laursen, M et al., PNAS (2013) PMID: 23776223; Ogawa, H. et al., PNAS (2009), PMID: 196666591)). Accordingly, mutagenesis screens and structure-activity relationships of CTSs compounds have identified residues affecting ouabain binding in these regions of the enzyme (Ogawa, H. et al., PNAS (2009), PMID: 196666591; Croyle, ML. et al., Eur. J. Biochem. (1997) PMID: 9346307) (see Table 5 below).

Residues affecting ouabain sensitivity are localized throughout the alpha1 subunit with clusters of residues conferring the greatest resistance to ouabain occurring at the HI-H2 transmembrane and extracellular regions, as well as the H4-H8 transmembrane and extracellular regions (Croyle, ML. et al., Eur. J. Biochem. (1997) PMID: 9346307). Mutations in these residues may function through several mechanisms, including affecting physical binding directly or indirectly through conformational changes or might interfere with the ability of cardiac glycosides to inhibit catalytic activity. Previous studies have identified critical regions of susceptibilities of the enzyme. Replacement of the border residues (Q118 and N129) of the HI-H2 extracellular domain with charged amino acids generates enzymes resistant to ouabain when overexpressed in cells (Price, EM. et al., Biochem (1988) PMID: 2853965; and Price, EM. et al, JBC (1990) PMID: 2157705) (see also Table 5 below).

[0170] Table 5: Exemplary known amino acid modifications in ATP1A1 conferring ouabain resistance (Adapted from Croyle ML, Woo AL, Lingrel JB. Extensive random mutagenesis analysis of the Na+/K+-ATPase alpha subunit identifies known and previously unidentified amino acid residues that alter ouabain sensitivity--implications for ouabain binding. Eur J Biochem. 1997; 248(2):488-95. PubMed PMID: 9346307).

Mutation (ref. position: human ATP1A1 isoform C-SEQ ID NO: 3)	Protein region	Genomic context	Fold resistance
C111F	H1 transmembrane	Exon 4	23
C111A	H1 transmembrane	Exon 4	3.2
C111Y	H1 transmembrane	Exon 4	6.3
Y115A	H1 transmembrane	Exon 4	9.1
Q118R	H1/H2 extracellular	Exon 4	12.5
P125K	H1/H2 extracellular	Exon 4	3.9
D128E	H1/H2 extracellular	Exon 4	25-60
D128N	H1/H2 extracellular	Exon 4	65
D128A	H1/H2 extracellular	Exon 4	38
D128S	H1/H2 extracellular	Exon 4	60
N129D	H1/H2 extracellular	Exon 4	12.5
Q118R + N129D	H1/H2 extracellular	Exon 4	1250
Q118D + N129R	H1/H2 extracellular	Exon 4	≈ 50000
Y315C	H3/H4 extracellular	Exon 8	2.7
L337Q	H4 transmembrane	Exon 8	7.9
A338G	H4 transmembrane	Exon 8	3.1
T345N	H4 transmembrane	Exon 9	3.3
T345A	H4 transmembrane	Exon 9	3.2
F793N	H5 transmembrane	Exon 17	19
F793I	H5 transmembrane	Exon 17	11

L800P	H5/H6 extracellular	Exon 17	24
L800N	H5/H6 extracellular	Exon 17	2.8
L800K	H5/H6 extracellular	Exon 17	8.8
T804A	H5/H6 extracellular	Exon 17	66
T804V	H5/H6 extracellular	Exon 17	79
T804N	H5/H6 extracellular	Exon 17	80
F870L	H7 transmembrane	Exon 19	5.7
R887P	H7 /H8 extracellular	Exon 19	7.9
R887L	H7 /H8 extracellular	Exon 19	3.8
F989S	H10 transmembrane	Exon 22	6.3

[0171] In embodiments, methods of the present invention comprise introducing one or more modification(s) in the *ATP1A1* gene by NHEJ or HDR to confer or induce ouabain resistance. In some embodiments, ouabain resistance is conferred by introducing a DSB in ATP1A1 gene and by providing an ATP1A1 donor nucleic acid comprising (i.e., encoding) the desired amino acids modification(s). In such embodiments where ouabain resistance is introduced by HDR, any modification or combination of modifications known to confer ouabain resistance may be introduced in the ATP1A gene (i.e., any nucleic acid encoding an ATP1A1 polypeptide conferring ouabain resistance may be used to replace the corresponding endogenous ATP1A1 polypeptide).

[0172] In embodiments, the ATP1A1 donor nucleic acids confer ouabain resistance by replacing at least one of amino acid Q118 and N129 in the ATP1A1 gene by a charged amino acid. In embodiments, the charged amino acid is an arginine (R) or an aspartic acid (D). In embodiments, both of Q118 and N129 are replaced by an arginine (R) or an aspartic acid (D). In embodiments, Q118 is replaced by an aspartic acid and N129 is replaced by an arginine. In embodiments, Q118 is replaced by an arginine and N129 is replaced by an aspartic acid.

[0173] Exemplary sgRNA binding sites surrounding all transmembrane and extracellular regions of ATP1A1 are provided in Table 3 above and FIG. 28.

[0174] Thus, far, no deletion or insertion of amino acid sequences in the ATP1A1 polypeptide has been reported to confer ouabain resistance. All changes described in the literature are specific point mutations. By introducing a DSB in specific exon sequences of the ATP1A1 gene (e.g., exon 4 comprising the H1-H2 region of the APT1A1 polypeptide, several novel deletion and point mutations conferring ouabain resistance have been identified (e.g., variants set forth in SEQ ID NOs: 5, 8, 10, 12, 14, 16, 18, 20 and 22).

[0175] Accordingly, the present invention further provides a method of inducing ouabain resistance in one or more target cells by introducing a DSB in an exon of the ATP1A1 gene of one or more target cells. In embodiments, such

method comprises providing (e.g., transfecting or introducing) the one or more target cells with a sgRNA guide sequence having a target sequence in an exon of the ATP1A1 gene and a CRISPR-nuclease. In embodiments the sgRNA target sequence is located in an exon encoding the H1, H2; H3, H4, H5, H6, H7 H8 or H10 ATP1A1 polypeptide region. In embodiments, the target sequence is located in exon 4.

[0176] In embodiments, the ouabain resistant ATP1A1 polypeptide comprises a deletion of the amino acid corresponding to an alanine at position 119 (A_{119}) in the wild-type ATP1A1 polypeptide (SEQ ID NO: 3). In embodiments, the ouabain resistant ATP1A1 polypeptide comprises a deletion of the amino acid corresponding to an alanine at position 120 (A_{120}) in the wild-type ATP1A1 polypeptide. In embodiments, the ouabain resistant ATP1A1 polypeptide comprises a deletion of the amino acid corresponding to a threonine at position 121 (T_{121}) in the wild-type ATP1A1 polypeptide. In embodiments, the ouabain resistant ATP1A1 polypeptide comprises a deletion of the amino acids corresponding to a glutamic acid at positions 122 (E_{122}), 123 (E_{123}) and/or 124 (E_{124}) in the wild-type ATP1A1 polypeptide. In embodiments, the ouabain resistant ATP1A1 polypeptide comprises a deletion of the amino acid corresponding to a proline at position 125 (P_{125}) in the wild-type ATP1A1 polypeptide. In embodiments, the ouabain resistant ATP1A1 polypeptide comprises a deletion of the amino acid corresponding to a glutamine at position 126 (Q_{126}) in the wild-type ATP1A1 polypeptide. In embodiments, the ouabain resistant polypeptide comprises a lysine instead of a threonine at amino acid position 121 (T121K mutation). In embodiments, the ouabain resistant polypeptide comprises an alanine instead of a proline at amino acid position 125 (P125A mutation). In embodiments, the ouabain resistant ATP1A1 polypeptide comprises a deletion of at least one, at least two, at, least 3, at least 4, at east 5, at least 6, at least 7 or at least 8 amino acids in amino acids corresponding to positions 119 (A_{119}) and 126 (Q_{126}) in the wild-type ATP1A1 polypeptide. In embodiments, the ouabain resistant ATP1A1 polypeptide comprises a polypeptide sequence as set forth in SEQ ID NO: 5.

[0177] **Table 6: NHEJ-induced amino acid modifications present in ouabain resistant ATP1A1 polypeptides of the present invention.**

Amino acid residue in WT ATP1A1 polypeptide	Mutation	Ouabain resistant ATP1A1 polypeptides comprising the mutated residue
A119	deletion (del)	Δ12, Δ15, Δ18
A120	del	Δ12, Δ3b, Δ15, Δ18, Δ21
T121	del/K (lysine)	Δ12, Δ15, Δ18, Δ21/Δ3
E122	del	Δ6, Δ9, Δ15, Δ18, Δ21
E123	del	Δ6, Δ9, Δ15, Δ18, Δ21
E124	del	Δ6, Δ9, Δ18, Δ21
P125	del/A (alanine)	Δ21/Δ9
Q126	del	Δ21

[0178] The present invention further provides nucleic acids encoding sgRNAs, CRISPR nucleases and ouabain

resistant ATP1A1 polypeptides of the present invention.

[0179] The present invention is illustrated in further details by the following non-limiting examples.

EXAMPLE 1
NHEJ-MEDIATED DISRUPTION OF THE FIRST EXTRACELLULAR DOMAIN OF ATP1A1 CONFERS CELLULAR RESISTANCE TO OUABAIN

[0180] Crystal structures of the sodium-potassium pump with bound ouabain and related CTSs have revealed that the inhibitors are wedged very deeply between transmembrane helices in addition to contacting the extracellular surface of ATP1A1 (29-31). Accordingly, random mutagenesis and overexpression studies, residues affecting ouabain sensitivity have been localized throughout the protein with clusters of residues conferring the greatest resistance to ouabain occurring at the HI-H2 transmembrane and extracellular regions, as well as the H4-H8 transmembrane and extracellular regions (34). Most prominently, replacement of the border residues (Q118 and N129) of the HI-H2 extracellular domain with charged amino acids generates enzymes highly resistant to ouabain when overexpressed in cells (FIG. 1) (32, 33). However, it is unknown if deletions in these regions can disrupt ouabain binding while preserving the functionality of ATP1A1. Furthermore, it remains to be demonstrated if modification of the endogenous locus, as opposed to overexpression of the mutant protein, can result in cellular resistance to ouabain.

[0181] We tested whether targeting this region using the CRISPR/Cas9 system would induce cellular resistance to ouabain. We identified two highly active sgRNAs targeting the exon encoding the first extracellular loop (hereafter named G2 and G4) and one sgRNA targeting the adjacent intron (hereafter named G3) in K562 cells using the surveyor nuclease assay to determine the frequency of the small insertions and deletions (indels) characteristic of imprecise DSB repair by NHEJ (FIGs. 1, 2). We note that the seed region of the inactive G1 guide contains the GTTTT sequence that is identical to the sequence of the lower stem module formed by the CRISPR repeat:tracrRNA duplex which likely perturbs this sgRNA element. Cells were then treated with ouabain starting three days post-transfection and monitored for survival and growth. Both active nucleases targeting the coding sequence surrounding Q118 and N129 of ATP1A1 (G2 and G4) induced cellular resistance to ouabain with robust cell growth observed for G2 while cells cleaved in the intron with G3 all died within 48 hours (FIG. 3 and Table 7). We have not observed any spontaneous resistance to ouabain treatment. As ATP1A1 is essential for cell survival, these observations suggest that in frame insertions and deletions created in the first extracellular loop do not cripple enzymatic activity but prevent ouabain binding. Cloning and sequencing ATP1A1 alleles following ouabain treatment identified in frame deletion products resulting in disruption of the first extracellular loop of the pump (FIG. 7). We used the TIDE (Tracking of Indels by DEcomposition) method to assess the spectrum and frequency of targeted mutations generated in these pool of cells (35). The analysis revealed that in-frame deletions in the coding region are selected over time and predominantly upon ouabain treatment (FIGs. 4-6). Interestingly, sgRNA G2

generates a much more diversified set of mutations than G4, which correlates with a higher fraction of in frame deletions and more robust growth. There was no selection for a particular set of indels when sgRNA G3 that cleaves in the intron was transfected. Interestingly, titration of ouabain in the culture medium (up to 10μM) results in the selection of progressively larger deletions (FIGs. 8-10). This positive selection was also observed in U2OS cells and hTERT-RPE1 cells (FIGs. s 11A and B). Thus, we identified highly active nucleases capable of producing gain-of-function alleles at the ATP1A1 locus via NHEJ.

[0182] Table 7. Active sgRNAs/nucleases conferring cellular resistance to ouabain

sgRNA (G)	Ouabain resistance (0.5μM)
G1	-
G2	+
G3	-
G4	+

[0183] The RNA (sgRNA) expression vectors (G1-G4) (500 ng) indicated in Table 7 were transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus. Cells were treated with 0.5μM ouabain for 7 days starting 3 days post-transfection and monitored for survival and growth. An expression vector encoding EGFP (500ng) was used as a negative control.

EXAMPLE 2

CO-SELECTION STRATEGY FOR ENRICHING CELLS HARBORING CRISPR-DRIVEN MUTATIONS INDUCED BY NHEJ

[0184] To test whether selection for the above-mentioned gain-of-function alleles of *ATP1A1* can result in co-enrichment of NHEJ-driven mutations at a second locus, we co-transfected sgRNAs against *ATP1A1* and *AAVS1* in cells and treated them with ouabain, a process referred to as multiplexing. Cells were transfected with high and low doses of *AAVS1* sgRNA and increasing amounts of sgRNA targeting *ATP1A1* to simulate the impact of co-selection at a broad range of gene editing frequencies (FIG. 12). ATP1A1 targeted cells selected by resistance to ouabain showed an increased level of indels at the AAVS1 site (FIG. 22A). Low initial levels of NHEJ at *ATP1A1* resulted in the greatest increases of mutated alleles upon ouabain treatment while no apparent enrichment was observed at the highest dose, as judged by the surveyor assay. Under these conditions, the increase in NHEJ at the co-selected AAVS1 locus was moderate, reaching about 4 fold (from 6% indels to 22% upon ouabain treatment) (FIG. 22A and FIG. 12). Under the same conditions, co-selection increased gene disruption at HPRT1 from undetectable to 13% indels (FIG. 22B and FIG. 13). At higher doses of HPRT1 sgRNA, co-selection resulted in the rescue of on-target

signal with some conditions increasing from undetectable levels to as much as 22% indels (FIG. 22B and FIG. 13). We observed that transfection of high doses of the HPRT1 sgRNA in combination with increasing amounts of the ATP1A1 sgRNA vector negatively impacted on-target mutagenesis at both sites, implying that this HPRT1 sgRNA has low specificity and is toxic to the cells (FIG. 13). This observation suggests that the co-selection strategy may facilitate the isolation of cell lines even when the gene knockout negatively affects growth.

EXAMPLE 3

EDITING THE *ATP1A1* LOCUS VIA HDR PROVOKES ROBUST CELLULAR RESISTANCE TO OUABAIN

[0185] Having established that CRISPR-induced gain-of-function mutations at the endogenous ATP1A1 gene can be used efficiently for co-selection via NHEJ, we aimed to test whether it was possible to enrich for HDR-driven editing. Reaching a high threshold of HDR in human cells is a major challenge in the genome-editing field and could benefit both basic and therapeutic applications. At the population level, human cells favor DSB repair via NHEJ over HDR. Therefore, cleaving within the coding sequence of ATP1A1 would disfavor the recovery of cells edited through HDR at the expense of cells mutated via NHEJ since ouabain would select for both type of repair events. We took advantage of the highly active sgRNA G3 that targets the intron to achieve selection exclusively via HDR-driven events (FIGs. 1-3). We tested the highly active sgRNA (G3) that targets the intron just downstream of the crucial residues Q118 and N129 as it fails to induce resistance by itself (FIG. 1 and Table 7). We designed two single-stranded oligos (ssODNs) to create ATP1A1 alleles conferring ouabain resistance in the millimolar range when overexpressed in human cells (33). The objective is to replace border residues of the HI-H2 extracellular domain with the charged amino acids (33). The ssODN donors create the double replacements Q118R, N129D (RD) and Q118D, N129R (DR), destroy the PAM motif and include additional silent mutations to create restriction sites to facilitate genotyping (FIG. 14).

[0186] Cas9-expresing cells were co-transfected with sgRNA G3 along with ssODNs and growth was monitored following addition of ouabain. Cells survived and grew robustly only in the presence of the guide and either donors. Surveyor and restriction fragment length polymorphism (RFLP) assays confirmed the introduction of desired sequence changes and their enrichment upon ouabain treatment (FIGs. 15-18). In addition, increasing the dose of ouabain selected for the double mutants within the population (FIG. 18). Interestingly, the ClaI RFLP assay revealed that mutations that are proximal to the DSB are preferentially incorporated in the resistant cell population (FIG. 16). Single mutations at either position confer an intermediate level of resistance to ouabain as compared to the double mutant and that the dose used in these experiments is far below the IC50 for the single mutants. It is envisioned that it may be possible to select cells having experienced longer gene conversion tracks by increasing the dose of ouabain during selection.

[0187] Titration of ouabain in the culture medium indicates that cells modified through HDR are resistant to concentration of the drug of at least 1mM, which is more than 100 fold higher than for NHEJ-induced mutations (Table 8) and more than 2000 fold higher than what is needed to kill the cells (0.5µM). These values correspond to the level of resistance observed in cells overexpressing the mutant enzymes and highlight the wide range of doses that can be used for selection (33). In addition, increasing the dose of ouabain positively selects for the double mutants within the population (FIG. 18). Thus, optimization of the drug-based selection process for various cell lines is straightforward as it is not necessary to precisely titrate the amount of ouabain required for selection.

[0188] Table 8. Cellular resistance to ouabain conferred by the different ssODNs

ssODN	Ouabain resistance (1000uM)
Q118R N129D (RD)	++
Q118D N129R (DR)	+++

[0189] For results reported in Table 8 above, the G3 sgRNA expression vector (500 ng) was co-transfected into K562 cells stably expressing wild-type SpCas9 from the *AAVS1* safe harbor locus along with the indicated ssODNs. Cells were treated or not with increasing doses of ouabain for 7 days starting 3 days post-transfection and cell survival and growth was monitored.

EXAMPLE 4

CO-SELECTION STRATEGY FOR ENRICHING CELLS HARBORING CRISPR-DRIVEN MUTATIONS INDUCED BY HDR

[0190] We then tested if selection for cells having experienced a CRISPR-driven HDR event at ATP1A1 could provide a substantial enrichment for correctly targeted cells at a second locus. Cells were co-transfected with a Cas9 expression vector plus sgRNAs targeting ATP1A1 (G3) and HPRT1 (S1R) in combination with ssODN RD and a plasmid donor for targeted insertion of a PGK1-EGFP expression cassette at HPRT1. Upon co-selection with ouabain a ~6-fold increase in EGFP+ cells was observed by FACS analysis (FIG. 20). Importantly, co-transfection of the ATP1A1-targeting components did not negatively impact the targeting frequency at HPRT1 since the % of EGFP+ cells in absence of ouabain treatment were similar whether present or not in the transfection mixes (FIG. 23A). Likewise, selection for ouabain resistant cells did not result in the random integration of the PGK-EGFP donor indicating that the process specifically enriches for HDR-driven events (FIG. 23A). Next, we targeted the same locus to introduce a BamHI restriction site using a plasmid donor. RFLP assays detected a 3-fold increase in BamHI positive HPRT1 alleles following ouabain treatment (FIG. 19). Using the highly potent AAVS1 sgRNA T2 to target a gene trap cassette (Splicing acceptor-2A-EGFP) resulted in 17% EGFP+ cells, a ~2-fold increase after co-selection.

A caveat for the precise calculation of the fold increase in this experiment is due to the weak promoter activity of the PPP1R12C gene resulting in dim fluorescence of the cells and incomplete separation of the EGFP- and EGFP+ peaks during FACS analysis (FIG. 23B). Nevertheless, enrichment was clearly observed at this safe harbor locus. To demonstrate that co-selection could be generally applied, we targeted two additional endogenous genes to generate N- and C-terminal fusions with fluorescent proteins. To label chromatin, the HIST1H2BK locus was targeted to create a C-terminal fusion of H2B with monomeric Azami-Green (mAG). In absence of selection, only 0.5% of the cells expressed the fusion protein while ouabain treatment resulted in a ~26-fold increase of targeted cells (FIG. 21). Since the left homology arm of the donor contains a truncated and non-functional H2B ORF, specific targeting through HDR must have occurred to generate the labeling observed in mAG+ cells (FIG. 21B). Finally, we inserted the sequence for the green fluorescent protein Clover after the second codon of the LMNA gene, which encodes the lamin A and lamin C isoforms (22). In this case, 64% of cells displayed the distinct localization pattern of green fluorescence enriched at the nuclear periphery in ouabain-selected cells as opposed to 5% without co-selection (FIG. 24 and FIG. 23C). Taken together, these data demonstrate that co-selection for ouabain resistant cells markedly improved the outcome of HDR-driven gene editing experiments.

EXAMPLE 5

COMPATIBILITY WITH HIGH-FIDELITY CRISPR/CAS9 AND CRISPR/CPF1 NUCLEASES

[0191] A versatile and robust approach to increase specificity in a CRISPR nuclease-based system is to use nucleases with inherent or enhanced specificity of action. Currently, the Cpf1 and high-fidelity Cas9 variants provide an alternative to wild-type SpCas9 and have been shown to be highly specific for their target site based on genome-wide specificity assays in human cells (37-39 and 48-49).

[0192] We tested whether these engineered CRISPR nuclease systems could be used to create robust cellular resistance to ouabain as observed for wild-type SpCas9. First, we selected the type V CRISPR system from *Acidaminococcus sp*. Cpf1 (AsCpf1), a single-RNA-guided (crRNA) enzyme that recognizes a TTTN protospacer-adjacent motif (PAM) sequences and produces cohesive double-stranded breaks (DSBs). We identified two highly active crRNAs targeting the exon (hereafter named G6 and G7) and one crRNA targeting the adjacent intron (hereafter named G9) in K562 cells using the surveyor nuclease assay (FIGs. 25A, B). Interestingly, crRNAs with adjacent or overlapping PAM sequences displayed markedly different activities pointing to different requirements for efficient DNA cleavage for this nuclease (FIGs. 25A, B). Cells were then treated with ouabain starting three days post transfection and monitored for survival and growth. As observed for wild-type SpCas9, active nucleases targeting the coding sequence surrounding Q118 and N129 of ATP1A1 induced cellular resistance to ouabain with the most robust cell growth observed for G7. Selected cells also displayed a similar profile of in-frame deletions as determined by TIDE analysis (FIG. 25D). Cells cleaved within the intron with G9 all died within 48 hours, unless they

were co-transfected with the ssODN donor template to generate point mutations conferring resistance to ouabain.

[0193] Next, we evaluated the rationally engineered "enhanced specificity" SpCas9 (eSpCas9 1.1) variant (38). Again, this nuclease system was found to be fully compatible with our selection strategy (FIGs. 25C, E). These observations further increase the potential of the approach by reducing concerns over off-target mutagenesis generated by the combined use of user-specific and ATP1A1 CRISPR reagents.

EXAMPLE 6

EFFICIENT ENRICHMENT OF GENE-EDITED HUMAN HEMATOPOIETIC STEM AND PROGENITOR CELLS (HSPCS)

[0194] To explore the potential for clinical translation of our method we tested the ouabain selection strategy during *ex vivo* expansion of cord blood-derived human hematopoietic stem and progenitor cells (HSPCs). Genome editing in HSPCs by homologous recombination remains challenging but offers the possibility to treat several genetic diseases. Moreover, rapid selection for gene-edited cells using ouabain could alleviate the problems associated with differentiation and progressive loss of long-term repopulating capacity after culturing. Purified human CD34+ cells were electroporated with preformed SpCas9 ribonucleoprotein complexes (RNPs) containing synthetic crRNAs and tracrRNAs targeting ATP1A1 and expanded *ex vivo* in presence or absence of ouabain. Surveyor nuclease assays and TIDE analysis revealed the presence of in-frame deletions in up to 50% of alleles in presence of ouabain in comparison to ~10% in untreated cells (FIG. 26). Furthermore, cells edited via HDR at the ATP1A1 locus were efficiently enriched using ouabain (FIG. 26). These results support the notion that the process could be adapted and optimized for use in preclinical studies. Critically, these data demonstrate that the procedure is applicable to primary cells.

EXAMPLE 7

REAGENTS AND METHODS

[0195] **Cell culture and transfection**. K562 and U2OS cells were obtained from the ATCC and maintained at 37 °C under 5% CO_2 in RPMI medium supplemented with 10% FBS, penicillin-streptomycin and GlutaMAX. Cells were transfected using the Amaxa 4D-Nucleofector™ (Lonza) per manufacturer's recommendations. Ouabain (Sigma) was added directly to the culture medium.

[0196] **Surveyor nuclease, RFLP knock in assays, and TIDE analysis**. Genomic DNA from 2.5E5 cells was extracted with 250 µl of QuickExtract™ DNA extraction solution (Epicentre) per manufacturer's recommendations. The *ATP1A1* locus was amplified by 30 cycles of PCR using the following primers: 5'-CACTTGTAAGAGCATCTACAACG-3' (SEQ ID NO: 449) and 5'-GGATTAACATCTGCTCGTGCAGC-3' (SEQ ID

NO: 450). The *AAVS1* locus was amplified by 30 cycles of PCR using the following primers: 5'-CCCCTTACCTCTCTAGTCTGTGC-3' (SEQ ID NO: 451) and 5'- CTCAGGTTCTGGGAGAGGGTAG-3' (SEQ ID NO: 452). The *HPRT1* locus was amplified by 30 cycles of PCR using the following primers: 5'-GGTGTGGAAGTTTAATGACTAAGAGG-3' (SEQ ID NO: 453) and 5'-TCACTGTAACCAAGTGAA ATGAAAGC-3' (SEQ ID NO: 454). Assays were performed with the Surveyor™ mutation detection kit (Transgenomics) according to the manufacturer's protocol. Samples were separated on 10% PAGE gels in TBE buffer. For RFLP assays, the PCR products were purified and digested with the corresponding enzyme and resolved by 10% PAGE. Gels were imaged using a ChemiDoc™ MP (Bio-Rad) system and quantifications were performed using Image lab software (Bio-Rad) as described (13). TIDE analysis was performed as described (35).

[0197] CRISPR/Cas9 reagents. The CAG-driven human codon optimized Cas9 nuclease vectors pCas9_GFP (Addgene #44719) was used in all transient transfection experiments. The expression cassette was transferred to AAVS1_Puro_PGK1_3xFLAG_Twin_Strep (Addgene #68375) in order to establish the K562 cell line constitutively expressing Cas9. All sgRNA expression vectors were built in the MLM3636 (Addgene #43860) backbone. Target sequences for *ATP1A1-G1* (GAACTCACATTATCGTTTTG, SEQ ID NO: 27), *AT1A1-G2* (GATCCAAGCTGCTACAGAAG, SEQ ID NO: 28), *ATP1A1-G3* (GAGTTCTGTAATTCAGCATA, SEQ ID NO: 29), *ATP1A1-G4* (GTTCCTCTTCTGTAGCAGCT, SEQ ID NO: 30), *HPRT1-S1R* (FIGs. 19-20) (GATGTGATGAAGGAGATGGG, SEQ ID NO: 455), *HPRT1-S2L* (FIG. 13) (GACAGAGGGCTACAATGTGA SEQ ID NO: 456), and *HIST1H2BK* (GGGGCTTTAAGACGCTTACT, SEQ ID NO: 457) were chosen according to a web-based CRISPR design tool (39). The DNA sequence for the sgRNA G2 was modified at position 1 to encode a 'G' due to the transcription initiation requirement of the human U6 promoter. ssODN were synthesized as ultramer (IDT). To test the high specificity eSpCas9(1.1) variant, guide sequences were cloned into the eSpCas9(1.1)_No_FLAG vector (Addgene plasmid # 79877). The sgRNA sequences are provided as follows: ATP1A1-G1 (SEQ ID NO: 23); *AT1A1-G2* (SEQ ID NO: 24); *ATP1A1-G3* (SEQ ID NO: 25); *ATP1A1-G4* (SEQ ID NO: 26); *HPRT1-S1R* (SEQ ID NO: 465)*; HPRT1-S2L* (SEQ ID NO: 466) and *HIST1H2BK* (SEQ ID NO: 467).

[0198] DR: (CAATGTTACTGTGGATTGGAGCGATTCTTTGTTTCTTGGCTTATAGCATCGATGCTGCTACAGAAG AGGAACCTCAAAACGATCGTGTGAGTTCTGTAATTCAGCATATCGATTTGTAGTACACATCAGATATCTT SEQ ID NO: 43) and

[0199] RD: (CAATGTTACTGTGGATTGGAGCGATTCTTTGTTTCTTGGCTTATAGCATCAGAGCTGCTACAGAA GAGGAACCTCAAAACGATGACGTGAGTTCTGTAATTCAGCATATCGATTTGTAGTACACATCAGATATCTT, SEQ ID NO: 44)

[0200] CRISPR/Cpf1 reagents. The human codon optimized Cpf1 ORF from the nuclease vector pY010, a gift from Feng Zhang (Addgene plasmid # 69982) (41), was transferred to AAVS1_Puro_PGK1_3xFLAG_Twin_Strep

(Addgene plasmid # 68375) in order to establish the K562 cell line constitutively expressing Cpf1 from a CAG promoter. A crRNA expression vector was built in the pUC19 backbone by linking the AsCpf1 5'DR to a human U6 promoter (41). The guide sequences where cloned downstream of the 5'DR. Target sequences for ATP1A1-G5 (TTTCTTGGCTTATAGCATCC, SEQ ID NO: 37), ATP1A1-G6 (TTGGCTTATAGCATCCAAGC, SEQ ID NO: 38), ATP1A1-G7 (AGGTTCCTCTTCTGTAGCAG, SEQ ID NO: 39), ATP1A1-G8 (GAGGTTCCTCTTCTGTAGCA, SEQ ID NO: 40), and ATP1A1-G9 (TAGTACACATCAGATATCTT, SEQ ID NO: 41 were chosen manually by scanning the target regions for TTTN PAMs. Desalted ssODN were synthesized as ultramers (IDT) at 4 nmole scale

[0201] DR:(CAATGTTACTGTGGATTGGAGCGATTCTTTGTTTCTTGGCTTATAGCATCGATGCTGCTACAGAAG AGGAACCTCAAAACGATCGTGTGAGTTCTGTAATTCAGCATATCGATTTGTAGTACACATCAGATATCTT, SEQ ID NO: 43); and

[0202] RD:(CAATGTTACTGTGGATTGGAGCGATTCTTTGTTTCTTGGCTTATAGCATCAGAGCTGCTACAGAAG AGGAACCTCAAAACGATGACGTGAGTTCTGTAATTCAGCATATCGATTTGTAGTACACATCAGATATCTT, SEQ ID NO: 44).

[0203] All plasmid donor sequences contain short homology arms (<1 kb) and have been modified in order to prevent their cleavage by Cas9 as described (9).

[0204] **Flow cytometry** The frequency of EGFP or mAG expressing cells was assessed using a BD LSR II™ flow cytometer.

[0205] **Targeting in primary human cord blood (CB) CD34⁺ cells.** CB samples from volunteer donors were acquired after written informed consent according to the guidelines approved by Hema-Quebec and Laval University. Mononuclear cells were isolated using Ficoll-Paque Plus (GE Healthcare) density centrifugation. $CD34^+$ haematopoietic stem and progenitor cells (HSPCs) were enriched using anti-CD34 microbeads and EasySep™ magnet (StemCell Technologies) according to manufacturer's instructions. Purity of $CD34^+$ cells was assessed and cells were cryopreserved in Cryostor CS10 (StemCell Technologies). $CD34^+$ HSPCs were cultured in StemSpan ACF (StemCell Technologies) supplemented with 100ng/ml SCF, 100ng/ml FLT3-L, 50ng/ml TPO, 10µg/ml LDL, 35nM UM171. The medium was changed every 3-4 days. Ouabain (0.5µM) was added directly to the culture media change at day 6 post-transfection. For enrichment via HDR, ouabain was kept in the culture during the entire selection period. For enrichment via NHEJ, ouabain was added for the first week. Cultured $CD34^+$ cells (100 000-200 000 per sample) were spun at 300g for 10 minutes, washed in PBS and resuspended in P3 Nucleofector solution (Lonza). Cells were electroporated with Cas9 RNPs with or without ssODNs using the Amaxa 4D Nucleofector™ X unit (Lonza) and the E0-100 program as optimized by the manufacturer. Prior to nucleofection, 50pmol of Cas9 protein (Integrated DNA Technologies) was incubated with 150pmol sgRNA or crRNA:tracrRNA

complex at RT for 10 minutes to form the RNP complex. The Alt-R CRISPR system (Integrated DNA Technologies) was used to disrupt *ATP1A1* via NHEJ using the crRNA G2 (target sequence: ATCCAAGCTGCTACAGAAG, SEQ ID NO: 28). crRNA and tracrRNA were resuspended to 200μM stock solutions in Nuclease-Free IDTE Buffer. For crRNA:tracrRNA complex formation, the two RNA oligos were mixed in equimolar concentrations, heated at 95°C for 5 minutes and allowed to cool to RT on bench top. sgRNA G3 was transcribed *in vitro* using the EnGen™ sgRNA Synthesis kit (NEB) as per the manufacturer's instructions using the following oligo (5′-TTCTAATACGACTCACTATAGAGTTCTGTAATTCAGCATAGTTTTAGAGCTAGA-3′; SEQ ID NO: 458). *In vitro* transcription products were purified using RNA Clean & Concentrator-25 (Zymo Research) and eluted in nuclease-free water.

EXAMPLE 8

INCREASING THE DOSE OF OUABAIN DURING SELECTION RESULTS IN HIGHER LEVELS OF TARGETED INTEGRATION AT THE CO-SELECTED LOCUS

[0206] As shown in FIG. 18, increasing the dose of ouabain selected for the *ATP1A1* double mutants (Q118R, N129D) within the population indicated that a subpopulation of cells having undergone HDR events could be further selected this way. This result prompted us to test if this was also true at the co-selected locus. Thus, we performed a co-selection experiment to tag HIST1H2BK at its c-terminus with the mAG1 green fluorescent protein (see FIG. 29). The drug was applied at an initial concentration of 0.5μM and then increased in a stepwise manner. Quantification of the % mAG positive cells in each population by flow cytometry revealed a greater than 2 fold increase in the fraction of targeted cells. These data indicate that the frequency of HDR-driven events at both the selected and the co-selected loci can be amplified by simply increasing the amount of ouabain during selection.

EXAMPLE 9

THE *ATP1A1* LOCUS CAN BE USED AS A "GENOMIC SAFE HARBOR" FOR TARGETED INTEGRATION OF AN EXOGENOUS SEQUENCE, FOR EXAMPLE, FOR EXPRESSION OF A POLYPEPTIDE OF INTEREST

[0207] Broadly speaking, a genomic safe harbor (GSH) is a chromosomal site where transgenes can be stably and reliably expressed in the tissue of interest without adversely affecting endogenous gene structure or expression (55). An ideal site for transgene insertion should allow (i) robust and stable transgene expression across different cell types; (ii) no transcriptional perturbation owing to the transgene cassette and (iii) no disruption of essential regulatory or coding sequences due to the transgene cassette, and (iv) no physiological impact on the organism (56,57). The standard locus for transgenesis via gene targeting in mouse embryonic stem cells, rosa26, is a prototypical GSH (55). In human cells, the site known as *AAVS1*, found within the first intron of the *PPP1R12C* gene is considered a "safe harbor" locus (56, 54, see also DeKelver et al. US 8,110,379 B2). However, the efficacy

of targeted integration at this locus is subject to the same limitations imposed by the cell type and cell-cycle stage as these factors play a major role in determining the fate of genome editing. Thus, in order to obtain pure populations of cells targeted at *AAVS1*, one requires the use of classical selection markers. We tested if we could use the *ATP1A1* locus as a potential safe harbor and link the targeted integration of exogenous transgenes to the introduction of *ATP1A1* mutations conferring ouabain resistance. We designed DNA donor molecules containing expression cassettes for green fluorescent proteins flanked by homology arms for targeted integration within intron 4 of *ATP1A1* using the eSpCas9 (G3) and AsCpf1 (G9) nucleases (FIGs. 30-31). Critically, the left homology arm (L-HA) was engineered to contain the Q118R and N129D (RD) mutations. Thus, the outcome of HDR would be the co-introduction of a ouabain resistant allele and of the exogenous DNA sequences. In other terms, the two genetic modifications would be linked *in cis*. If such a strategy was practical, all ouabain-resistant cells should also express the exogenous transgene. We first established which orientation of the transgene, if any, was permissive to the expression of the transgene without perturbing *ATP1A1* expression (FIG. 32). Note that the cell will die if *ATP1A1* function is abrogated since it is an essential gene. We observed that the targeted integration of an expression cassette in the same orientation as the *ATP1A1* transcriptional unit (FIG. 30, parallel, SEQ ID NO: 475) precluded cell growth as the vast majority of the cells died upon addition of ouabain. The few cells that survived treatment were likely able to use the L-HA to introduce only the RD mutations, not the transgene, as indicated by the very small fraction of cells expressing the fluorescent protein (FP) (FIG. 32A). We speculate that this configuration of the donor resulted in a "gene trap" leading to the termination of the *ATP1A1* transcript enforced by the bGH polyadenylation signal. In contrast, targeted integration of the expression cassette in the opposite orientation to the *ATP1A1* transcriptional unit (FIG. 31, antiparallel) was well tolerated and robust growth was observed upon ouabain treatment. The ouabain-resistant cell population was highly enriched for FP-expressing cells (FIG. 32B panel). However, the dim fluorescence of the TurboGFP protein used in this assay made it difficult to clearly discriminate between the FP positive and FP negative cell populations and to obtain an accurate quantification. Substitution for the bright FP (mNeonGreen) in the context of the antiparallel donor (FIG. 31) led to the unambiguous discrimination between the two cell populations indicating that >85% of the ouabain-resistant cells were also FP positive (FIG. 33A). However, a small fraction of the cells appeared to be able to bypass the integration of the transgene and decouple ouabain resistance and FP expression. Modification of the left homology arm to decrease the amount of DNA homology found in the vicinity and spanning the predicted location of the DSB lead to an improvement in the fraction of ouabain resistant cells expressing the FP (FIG. 33B) (compare antiparallel donor v1 to v2). Using this improved donor molecule (v2), we compared the efficacy of eSpCas9 (G3) and AsCpf1 (G9) nucleases at generating FP-expressing cell populations. Strikingly, AsCpf1-driven targeted integration into intron 4 of *ATP1A1* led to >97% of the cells expressing the FP after ouabain selection (FIG. 34). Note that AsCpf1 cleaves farther downstream of exon 4, as compared to eSpCas9, and that there is no homology to sequences around the DSB site generated by AsCpf1 in donor v2. This likely prevents the uncoupling of the introduction of the two gene editing

events. We also observed that increasing the dose of the donor in the transfection mix increases the % of cells expressing the FP at higher levels (FIG. 35, see the cell population on the right side of the main peak, shifted away from the main distributionin FIG. 35B compared to FIG. 35A). We speculate that this population corresponds to multiple alleles of *ATP1A1* being modified in the same cell. Thus, this method allows one to obtain pure populations of cells expressing the transgene of interest without the need for exogenous selection marker.

EXAMPLE 10

THE USE OF *ATP1A1* AS A "GENOMIC SAFE HARBOR" CAN BE COUPLED TO THE CO-SELECTION PROCESS IN ORDER TO INCREASE THE LEVELS OF TARGETED INTEGRATION AT A DISTINCT LOCUS OF INTEREST

[0208] The use of *ATP1A1* as a "genomic safe harbor" (see Example 9) can be coupled to the co-selection process in order to increase the levels of targeted integration at a distinct locus of interest. We co-targeted the *LMNA* locus with a red fluorescent protein (mScarlet) and the *ATP1A1* gene with the mNeonGreen antiparallel v1 donor (see Example 9). Upon selection with ouabain, >67% of the cells were double positive for both FPs indicating that selection for targeted integration at *ATP1A1* led to the co-selection for targeted integration at *LMNA* (FIG. 36). Thus, it is possible to co-select for targeted integration of a transgene of interest at *ATP1A1* and a second genetic modification.

[0209] Here we show that the creation of gain-of-function alleles at the *ATP1A1* locus with CRISPR can be robustly selected for using a highly potent therapeutic small molecule. Ouabain treatment rapidly kills cells within 48 hours of exposure and targeted cells display no apparent growth delay resulting from the selection process. The turnover of ATP1A1 at the plasma membrane appears to be rapid since ouabain can be added to the culture medium 15 hours post transfection of the CRISPR components encoded on plasmids. The well-defined mechanism of action of ouabain acting on a nonsignaling ion pump independently of proliferation offers an additional advantage over chemoselection strategies.

[0210] An aspect of the system described herein is that the co-selection process can be initiated through NHEJ or HDR processes independently. One only has to switch between two sgRNAs targeting juxtaposed regions of *ATP1A1* and include an ssODN in the transfection reaction to co-select for HDR- instead of NHEJ-driven events. Perhaps more importantly, we show that the co-selection strategy markedly enriches for cells modified via homologous recombination, a clear limitation in the field. This was demonstrated at endogenous loci for four different types of HDR-driven genome editing events; (i) RFLP knock-in, (ii) gene trapping, (iii) targeted integration of an autonomous expression cassette, and (iv) protein tagging with fluorescent markers.

[0211] Methods of the present invention should be compatible with any engineered nuclease platforms and any cell type when using a ubiquitously expressed gene such as *ATP1A1*. Perhaps more importantly, robust selection is

achieved without the use of exogenous markers making the process compatible with therapeutic applications.

[0212] Based on the results presented herein, the concept of creating and selecting for dominant gain-of-function alleles of endogenous genes and their use to enrich for custom modifications at unlinked loci of interest has been demonstrated and may be broadened to many other protein/drug combinations. For example, one could achieve the co-enrichment of edited cells *in vivo* after chemo-selection of transplanted cells modified to express the P140K mutant of human O(6)-methylguanine-DNA-methyltransferase (MGMT, FIG. 27) (50). Technically, it is interesting to note that the P140 residue in human MGMT is encoded next to an intron-exon boundary and could be targeted in a similar fashion as ATP1A1. The use of endogenous promoters and regulatory elements while avoiding the expression of xenogenic enzymes would be a great advantage in the context of clinical applications.

[0213] Table 9: Sequences described herein

SEQ ID NO(s)	aa/nts	Description
SEQ ID NO: 1	nts	Human ATP1A1 wild-type ORF (cDNA) (NM_001160233.1). ATPase Na+/K+ transporting subunit alpha 1 isoform c. This variant represents the longest transcript and encodes the longest isoform.
SEQ ID NO: 2	nts	Part of genomic polynucleotide sequence of human ATP1A1, from Gene ID 476. 1-nts 1-100 : intron 3; 101-304: exon 4; and 305-404 intron 4
SEQ ID NO: 3	aa	Human ATP1A1 isoform C wild-type protein (NP_001153705.1) ATPase Na+/K+ transporting subunit alpha 1 isoform c. This variant represents the longest isoform
SEQ ID NO: 4	aa	Human ouabain resistant ATP1A1 from HDR experiments (consensus sequence). wherein each X is independently R or D
SEQ ID NO: 5	aa	Ouabain resistant human ATP1A1 consensus sequence for variants (mutants) in exon 4. wherein each X is independently any amino acid or is absent
SEQ ID NO: 6	aa	Ouabain resistant human ATP1A1 consensus sequence including all amino acids which can be mutated to confer ouabain resistance (includes WT aa and mutant aa in Table 4 which lists known amino acid mutations in ATP1A1 conferring ouabain resistance. Includes aa in novel variants identified herein).
SEQ ID NO: 7	nts	Ouabain resistant human ATP1A1 variant 1 (Δ3bp): ORF (cDNA)
SEQ ID NO: 8	aa	Ouabain resistant human ATP1A1 variant 1 (Δ3bp)
SEQ ID NO: 9	nts	Ouabain resistant human ATP1A1 variant 2 (Δ6bp): ORF (cDNA)
SEQ ID NO: 10	aa	Ouabain resistant human ATP1A1 variant 2 (Δ6bp)
SEQ ID NO: 11	nts	Ouabain resistant human ATP1A1 variant 3 (Δ9bp): ORF (cDNA)
SEQ ID NO: 12	aa	Ouabain resistant human ATP1A1 variant 3 (Δ9bp)
SEQ ID NO: 13	nts	Ouabain resistant human ATP1A1 variant 4 (Δ12bp): ORF (cDNA)
SEQ ID NO: 14	aa	Ouabain resistant human ATP1A1 variant 4 (Δ12bp)
SEQ ID NO: 15	nts	Ouabain resistant human ATP1A1 variant 5 (Δ3b): ORF (cDNA)
SEQ ID NO: 16	aa	Ouabain resistant human ATP1A1 variant 5 (Δ3b)
SEQ ID NO: 17	nts	Ouabain resistant human ATP1A1 variant 6 (Δ15bp): ORF (cDNA)
SEQ ID NO: 18	aa	Ouabain resistant human ATP1A1 variant 6 (Δ15bp)

SEQ ID NO: 19	nts	Ouabain resistant human ATP1A1 variant 7 (△18bp): ORF (cDNA)
SEQ ID NO: 20	aa	Ouabain resistant human ATP1A1 variant 7 (△18bp)
SEQ ID NO: 21	nts	Ouabain resistant human ATP1A1 variant 8 (△21bp): ORF (cDNA)
SEQ ID NO: 22	aa	Ouabain resistant human ATP1A1 variant 8 (△21bp): Protein
SEQ ID NO: 23	nts	Single guide RNA (sgRNA) 1 (ATP1A1-G1) sequence including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)
SEQ ID NO: 24	nts	Single guide RNA (sgRNA) 2 (ATP1A1-G2) including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)
SEQ ID NO: 25	nts	Single guide RNA (sgRNA) 3 (ATP1A1-G3) including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)
SEQ ID NO: 26	nts	Single guide RNA (sgRNA) 4 (ATP1A1-G4) including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)
SEQ ID NO: 27	nts	target sequence (sgRNA) 1 (ATP1A1-G1)
SEQ ID NO: 28	nts	target sequence (sgRNA) 2 (ATP1A1-G2) ATP1A1 (lacking added G in 5' for U6 promoter)
SEQ ID NO: 29	nts	target sequence (sgRNA) 3 (ATP1A1-G3)
SEQ ID NO: 30	nts	target sequence (sgRNA) 4 (ATP1A1-G4)
SEQ ID NO: 31	nts	Cas9 CRISPR nuclease recognition sequence (chimeric guide RNA scaffold/tracr RNA sequence)
SEQ ID NO: 32	nts	Single guide RNA (sgRNA) 5 (ATP1A1-G5) including CRISPR nuclease (AsCpf1) recognition sequence (Single CRISPR guide RNA -5' Direct repeat)
SEQ ID NO: 33	nts	Single guide RNA (sgRNA) 6 (ATP1A1-G6) including CRISPR nuclease (AsCpf1) recognition sequence (Single CRISPR guide RNA -5' Direct repeat)
SEQ ID NO: 34	nts	Single guide RNA (sgRNA) 7 (ATP1A1-G7) including CRISPR nuclease recognition (AsCpf1) sequence (Single CRISPR guide RNA -5' Direct repeat)
SEQ ID NO: 35	nts	Single guide RNA (sgRNA) 8 (ATP1A1-G8) including CRISPR nuclease recognition (AsCpf1) sequence (Single CRISPR guide RNA -5' Direct repeat)
SEQ ID NO: 36	nts	Single guide RNA (sgRNA) 9 (ATP1A1-G9) including CRISPR nuclease (AsCpf1) recognition sequence (Single CRISPR guide RNA -5' Direct repeat)
SEQ ID NO: 37	nts	target sequence (sgRNA) 5 (ATP1A1-G5) (DNA)- (AsCpf1)
SEQ ID NO: 38	nts	target sequence (sgRNA) 6 (ATP1A1-G6) (DNA)- (AsCpf1)
SEQ ID NO: 39	nts	target sequence (sgRNA) 7 (ATP1A1-G7) (DNA)- (AsCpf1)
SEQ ID NO: 40	nts	target sequence (sgRNA) 8 (ATP1A1-G8) (DNA)- (AsCpf1)
SEQ ID NO: 41	nts	target sequence (sgRNA) 9 (ATP1A1-G9) (DNA) - (AsCpf1)
SEQ ID NO: 42	nts	CRISPR recognition sequence for AsCpf1 (5' direct-repeat RNA sequence)
SEQ ID NO: 43	nts	ssODN DR- (Q118R/N129D)
SEQ ID NO: 44	nts	ssODN RD (Q118D/N129R)
SEQ ID NO: 45	aa	SpCas9 protein sequence
SEQ ID NO: 46	aa	AsCpf1 protein sequence (derived from *Acidaminococcus* species-basic sequence, no histag or NLS)
SEQ ID NOs: 47-89	nts	Target sequences for SpCas9 or Cpf1 (DNA) T1 to T337 (except T16/G3 and T28/G4) listed in Table 3
SEQ ID NOs: 90-448	nts	Target sequences for SpCas9 on ATP1A1 shown on FIG. 28
SEQ ID NOs:	nts	Primers for amplification of (i) ATP1A1 locus (SEQ ID NOs:449-450); (ii) The *AAVS1*

449-454		locus (SEQ ID NOs: 451-452); (iii) The *HPRT1* locus (SEQ ID NOs: 453-454) (Example 7)
SEQ ID NO: 455	nts	Target sequence for *HPRT1-S1R* ((FIG. 19-20), Example 7
SEQ ID NO: 456	nts	Target sequence for *HPRT1-S2L* ((FIG. 13), Example 7
SEQ ID NO: 457	nts	Target sequence for *HIST1H2BK,* Example 7 – FIG. 21
SEQ ID NO: 458	nts	Oligo for *in vitro* transcription of sgRNA G3-Example 7
SEQ ID NO: 459	nts	Human ATP1A2 wild-type ORF (cDNA): NM_000702.3. ATPase Na+/K+ transporting subunit alpha 2
SEQ ID NO: 460	aa	Human ATP1A2 wild-type protein NP_000693.1. ATPase Na+/K+ transporting subunit alpha 2 precursor
SEQ ID NO: 461	nts	Human ATP1A3 wild-type ORF (cDNA): NM_152296.4. ATPase Na+/K+ transporting subunit alpha 3 isoform 1
SEQ ID NO: 462	aa	Human ATP1A3 wild-type protein: NP_689509.1. ATPase Na+/K+ transporting subunit alpha 3 isoform 1
SEQ ID NO: 463	nts	Human ATP1A4 wild-type ORF (cDNA): NM_144699.3. ATPase Na+/K+ transporting subunit alpha 4 isoform 1
SEQ ID NO: 464	aa	Human ATP1A4 isoform 1 wild-type protein NP_653300.2. ATPase Na+/K+ transporting subunit alpha 4 isoform 1
SEQ ID NO: 465	nts	Single guide RNA (sgRNA) *HPRT1-S1R* sequence including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)- FIGs. 19 and 20
SEQ ID NO: 466	nts	Single guide RNA (sgRNA) *HPRT1-S2L* sequence including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)- FIG. 13
SEQ ID NO: 467	nts	Single guide RNA (sgRNA) *HIST1H2BK* sequence including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)- FIGs. 21 and 29
SEQ ID NO: 468	nts	AAVS1 (T2) target sequence- FIGs. 12 and 22
SEQ ID NO: 469	nts	Single guide RNA (sgRNA) *AAVS1-T2* sequence including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)- FIGs. 12 and 22
SEQ ID NO: 470	nts	LMNA-G2 target sequence- FIG. 24
SEQ ID NO: 471	nts	Single guide RNA (sgRNA) LMNA-G2 sequence including CRISPR nuclease recognition sequence (chimeric guide RNA scaffold)- FIG. 24
SEQ ID NO: 472	nts	Splicing acceptor-SA-2A-EGFP-pA donor sequence containing homology arms to AAVS1- FIG. 23
SEQ ID NO: 473	nts	Donor sequence containing a Clover cassette (250ng) and homology arms to the LMNA locus - FIG. 23
SEQ ID NO: 474	nts	Donor sequence for experiments shown in FIGs. 21 and 29 for *HIST1H2BK* locus
SEQ ID NO: 475	nts	ATP1A1 donor sequence (TurboGFP, parallel orientation, FIGs. 30 and 32 and Example 9)
SEQ ID NO: 476	nts	ATP1A1 donor sequence (TurboGFP, antiparallel orientation, Figures 31 and 32 and Example 9)
SEQ ID NO: 477	nts	DNA donor sequence (V1-NeonGreen) containing expression cassettes for green fluorescent protein flanked by homology arms for targeted integration within intron 4 of *ATP1A1*- FIG. 33
SEQ ID NO: 478	nts	DNA donor sequence (V2-NeonGreen) containing expression cassettes for green fluorescent protein flanked by homology arms for targeted integration within intron 4 of *ATP1A1*-Exemple 9 and FIG. 33
SEQ ID NO: 479	nts	DNA donor sequence (LMNA-mScarlet)-Example 10 and FIG. 36
SEQ ID NO: 480	nts	DNA donor sequence for HPRT (PGK-EGFP)- FIG. 23
SEQ ID NO: 481	nts	Polynucleotide sequence of wild type human MGMT (NM_002412.4)

SEQ ID NO: 482	aa	Amino acid sequence of wild type human MGMT (GenBank acc. No. 4255, NP_002403.2)
SEQ ID NO: 483	nts	Partial genomic sequence of wild-type human MGMT at the junction of intron 4 and exon 5 and shown on FIG. 27
SEQ ID NO: 484	nts	Sequence of ssODN of MGMT shown in FIG. 27
SEQ ID NO: 485	aa	Human ATP1A2 mutant. Mutations within amino acids 117-124
SEQ ID NO: 486	aa	Human ATP1A3 mutant. Mutations within amino acids 109-116
SEQ ID NO: 487	aa	Human ATP1A4 mutant. Mutations within amino acids 127-134

[0214] The scope of the claims should not be limited by the preferred embodiments set forth in the examples, but should be given the broadest interpretation consistent with the description as a whole.

REFERENCES

1. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014;157(6):1262-78. doi: 10.1016/j.cell.2014.05.010. PubMed PMID: 24906146; PMCID: PMC4343198.

2. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol. 2013;14(1):49-55. Epub 2012/11/22. doi: 10.1038/nrm3486. PubMed PMID: 23169466; PMCID: 3547402.

3. Sternberg SH, Doudna JA. Expanding the Biologist's Toolkit with CRISPR-Cas9. Mol Cell. 2015;58(4):568-74. doi: 10.1016/j.molcel.2015.02.032. PubMed PMID: 26000842.

4. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 2010;11(9):636-46. doi: 10.1038/nrg2842. PubMed PMID: 20717154.

5. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819-23. doi: 10.1126/science.1231143. PubMed PMID: 23287718; PMCID: PMC3795411.

6. Jasin M, Haber JE. The democratization of gene editing: Insights from site-specific cleavage and double-strand break repair. DNA Repair (Amst). 2016;44:6-16. doi: 10.1016/j.dnarep.2016.05.001. PubMed PMID: 27261202.

7. Graham DB, Root DE. Resources for the design of CRISPR gene editing experiments. Genome Biol. 2015;16:260. doi: 10.1186/s13059-015-0823-x. PubMed PMID: 26612492; PMCID: PMC4661947.

8. Tycko J, Myer VE, Hsu PD. Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. Mol Cell. 2016;63(3):355-70. doi: 10.1016/j.molcel.2016.07.004. PubMed PMID: 27494557; PMCID: PMC4976696.

9. Dalvai M, Loehr J, Jacquet K, Huard CC, Roques C, Herst P, Cote J, Doyon Y. A Scalable Genome-Editing-Based Approach for Mapping Multiprotein Complexes in Human Cells. Cell Rep. 2015;13(3):621-33. doi: 10.1016/j.celrep.2015.09.009. PubMed PMID: 26456817.

10. Lopes R, Korkmaz G, Agami R. Applying CRISPR-Cas9 tools to identify and characterize transcriptional enhancers. Nat Rev Mol Cell Biol. 2016;17(9):597-604. doi: 10.1038/nrm.2016.79. PubMed PMID: 27381243.

11. Chiang TW, le Sage C, Larrieu D, Demir M, Jackson SP. CRISPR-Cas9(D10A) nickase-based genotypic and phenotypic screening to enhance genome editing. Sci Rep. 2016;6:24356. doi: 10.1038/srep24356. PubMed PMID: 27079678; PMCID: PMC4832145.

12. Duda K, Lonowski LA, Kofoed-Nielsen M, Ibarra A, Delay CM, Kang Q, Yang Z, Pruett-Miller SM, Bennett EP, Wandall HH, Davis GD, Hansen SH, Frodin M. High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases or CRISPR/Cas9 nickase pairs. Nucleic Acids Res. 2014;42(10):e84. doi: 10.1093/nar/gku251. PubMed PMID: 24753413; PMCID: PMC4041425.

13. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-308. doi: 10.1038/nprot.2013.143. PubMed PMID: 24157548; PMCID: PMC3969860.

14. Certo MT, Ryu BY, Annis JE, Garibov M, Jarjour J, Rawlings DJ, Scharenberg AM. Tracking genome engineering outcome at individual DNA breakpoints. Nat Methods. 2011;8(8):671-6. doi: 10.1038/nmeth.1648. PubMed PMID: 21743461; PMCID: PMC3415300.

15. Kim H, Um E, Cho SR, Jung C, Kim H, Kim JS. Surrogate reporters for enrichment of cells with nuclease-induced mutations. Nat Methods. 2011;8(11):941-3. doi: 10.1038/nmeth.1733. PubMed PMID: 21983922.

16. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, Kuhn R. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol. 2015;33(5):543-8. doi: 10.1038/nbt.3198. PubMed PMID: 25803306.

17. Doyon Y, Choi VM, Xia DF, Vo TD, Gregory PD, Holmes MC. Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. Nat Methods. 2010;7(6):459-60. Epub 2010/05/04. doi: 10.1038/nmeth.1456 nmeth.1456 [pii]. PubMed PMID: 20436476.

18. Gutschner T, Haemmerle M, Genovese G, Draetta GF, Chin L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. Cell Rep. 2016;14(6):1555-66. doi: 10.1016/j.celrep.2016.01.019. PubMed PMID: 26854237.

19. Howden SE, McColl B, Glaser A, Vadolas J, Petrou S, Little MH, Elefanty AG, Stanley EG. A Cas9 Variant for Efficient Generation of Indel-Free Knockin or Gene-Corrected Human Pluripotent Stem Cells. Stem Cell Reports. 2016. doi: 10.1016/j.stemcr.2016.07.001. PubMed PMID: 27499201.

20. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife. 2014;3:e04766. doi: 10.7554/eLife.04766. PubMed PMID: 25497837; PMCID: PMC4383097.

21. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol. 2015;33(5):538-42. doi: 10.1038/nbt.3190. PubMed PMID: 25798939; PMCID: PMC4618510.

22. Pinder J, Salsman J, Dellaire G. Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. Nucleic Acids Res. 2015;43(19):9379-92. doi: 10.1093/nar/gkv993. PubMed PMID: 26429972; PMCID: PMC4627099.

23. Arribere JA, Bell RT, Fu BX, Artiles KL, Hartman PS, Fire AZ. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. Genetics. 2014;198(3):837-46. doi: 10.1534/genetics.114.169730. PubMed PMID: 25161212; PMCID: PMC4224173.

24. Farboud B, Meyer BJ. Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. Genetics. 2015;199(4):959-71. doi: 10.1534/genetics.115.175166. PubMed PMID: 25695951; PMCID: PMC4391549.

25. Kim H, Ishidate T, Ghanta KS, Seth M, Conte D, Jr., Shirayama M, Mello CC. A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. Genetics. 2014;197(4):1069-80. doi:

10.1534/genetics.114.166389. PubMed PMID: 24879462; PMCID: PMC4125384.

26. Ward JD. Rapid and precise engineering of the Caenorhabditis elegans genome with lethal mutation co-conversion and inactivation of NHEJ repair. Genetics. 2015;199(2):363-77. doi: 10.1534/genetics.114.172361. PubMed PMID: 25491644; PMCID: PMC4317648.

27. Liao S, Tammaro M, Yan H. Enriching CRISPR-Cas9 targeted cells by co-targeting the HPRT gene. Nucleic Acids Res. 2015;43(20):e134. doi: 10.1093/nar/gkv675. PubMed PMID: 26130722; PMCID: PMC4787791.

28. Shy BR, MacDougall MS, Clarke R, Merrill BJ. Co-incident insertion enables high efficiency genome engineering in mouse embryonic stem cells. Nucleic Acids Res. 2016. doi: 10.1093/nar/gkw685. PubMed PMID: 27484482.

29. Ogawa H, Shinoda T, Cornelius F, Toyoshima C. Crystal structure of the sodium-potassium pump (Na+,K+-ATPase) with bound potassium and ouabain. Proc Natl Acad Sci U S A. 2009;106(33):13742-7. doi: 10.1073/pnas.0907054106. PubMed PMID: 19666591; PMCID: PMC2728964.

30. Laursen M, Yatime L, Nissen P, Fedosova NU. Crystal structure of the high-affinity Na+K+-ATPase-ouabain complex with Mg2+ bound in the cation binding site. Proc Natl Acad Sci U S A. 2013;110(27):10958-63. doi: 10.1073/pnas.1222308110. PubMed PMID: 23776223; PMCID: PMC3704003.

31. Laursen M, Gregersen JL, Yatime L, Nissen P, Fedosova NU. Structures and characterization of digoxin- and bufalin-bound Na+,K+-ATPase compared with the ouabain-bound complex. Proc Natl Acad Sci U S A. 2015;112(6):1755-60. doi: 10.1073/pnas.1422997112. PubMed PMID: 25624492; PMCID: PMC4330780.

32. Price EM, Lingrel JB. Structure-function relationships in the Na,K-ATPase alpha subunit: site-directed mutagenesis of glutamine-111 to arginine and asparagine-122 to aspartic acid generates a ouabain-resistant enzyme. Biochemistry. 1988;27(22):8400-8. PubMed PMID: 2853965.

33. Price EM, Rice DA, Lingrel JB. Structure-function studies of Na,K-ATPase. Site-directed mutagenesis of the border residues from the H1-H2 extracellular domain of the alpha subunit. J Biol Chem. 1990;265(12):6638-41. PubMed PMID: 2157705.

34. Croyle ML, Woo AL, Lingrel JB. Extensive random mutagenesis analysis of the Na+/K+-ATPase alpha subunit identifies known and previously unidentified amino acid residues that alter ouabain sensitivity--implications for ouabain binding. Eur J Biochem. 1997;248(2):488-95. PubMed PMID: 9346307.

35. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res. 2014;42(22):e168. doi: 10.1093/nar/gku936. PubMed PMID: 25300484; PMCID: PMC4267669.

36. Elliott B, Richardson C, Winderbaum J, Nickoloff JA, Jasin M. Gene conversion tracts from double-strand break repair in mammalian cells. Mol Cell Biol. 1998;18(1):93-101. PubMed PMID: 9418857; PMCID: PMC121458.

37. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, Joung JK. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature. 2016;529(7587):490-5. doi:

10.1038/nature16526. PubMed PMID: 26735016; PMCID: PMC4851738.

38. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. Science. 2016;351(6268):84-8. doi: 10.1126/science.aad5227. PubMed PMID: 26628643; PMCID: PMC4714946.

39. Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin EV, van der Oost J. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. Science. 2016;353(6299):aad5147. doi: 10.1126/science.aad5147. PubMed PMID: 27493190.

40. Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, Semenova E, Minakhin L, Joung J, Konermann S, Severinov K, Zhang F, Koonin EV. Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. Mol Cell. 2015;60(3):385-97. doi: 10.1016/j.molcel.2015.10.008. PubMed PMID: 26593719; PMCID: PMC4660269.

41. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015;163(3):759-71. doi: 10.1016/j.cell.2015.09.038. PubMed PMID: 26422227; PMCID: PMC4638220.

42. Kleinstiver BP, Prew MS, Tsai SQ, Nguyen NT, Topkar VV, Zheng Z, Joung JK. Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. Nat Biotechnol. 2015;33(12):1293-8. doi: 10.1038/nbt.3404. PubMed PMID: 26524662; PMCID: PMC4689141.

43. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales AP, Li Z, Peterson RT, Yeh JR, Aryee MJ, Joung JK. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. 2015;523(7561):481-5. doi: 10.1038/nature14592. PubMed PMID: 26098369; PMCID: PMC4540238.

44. Tsai Q. et al., Nature Biotechnology, Volume: 32, pages: 569–576 (2014) DOI: doi:10.1038/nbt.2908.

45. Mali P*, et al.* (2013) RNA-guided human genome engineering via Cas9. *Science* 339(6121):823-826.

46. Jinek M*, et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816-821.

47. Lingrel JB. The physiological significance of the cardiotonic steroid/ouabain-binding site of the Na,K-ATPase. Annu Rev Physiol. 2010;72:395-412. doi: 10.1146/annurev-physiol-021909-135725. PubMed PMID: 20148682; PMCID: PMC3079441.

48. Kim D, Kim J, Hur JK, Been KW, Yoon SH, Kim JS. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. Nat Biotechnol. 2016;34(8):863-8. doi: 10.1038/nbt.3609. PubMed PMID: 27272384.

49. Kleinstiver BP, Tsai SQ, Prew MS, Nguyen NT, Welch MM, Lopez JM, McCaw ZR, Aryee MJ, Joung JK. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. Nat Biotechnol. 2016;34(8):869-74. doi: 10.1038/nbt.3620. PubMed PMID: 27347757; PMCID: PMC4980201.

50. Nagree MS, Lopez-Vasquez L, Medin JA. Towards *in vivo* amplification: Overcoming hurdles in the use of hematopoietic stem cells in transplantation and gene therapy. World J Stem Cells. 2015;7(11):1233-50. doi: 10.4252/wjsc.v7.i11.1233. PubMed PMID: 26730268; PMCID: PMC4691692.

51. Dever DP, Bak RO, Reinisch A, Camarena J, Washington G, Nicolas CE, Pavel-Dinu M, Saxena N, Wilkens AB, Mantri S, Uchida N, Hendel A, Narla A, Majeti R, Weinberg KI, Porteus MH. CRISPR/Cas9 beta-globin gene targeting in human haematopoietic stem cells. Nature. 2016. doi: 10.1038/nature20134. PubMed PMID: 27820943.

52. Genovese P, Schiroli G, Escobar G, Di Tomaso T, Firrito C, Calabria A, Moi D, Mazzieri R, Bonini C, Holmes MC, Gregory PD, van der Burg M, Gentner B, Montini E, Lombardo A, Naldini L. Targeted genome editing in human repopulating haematopoietic stem cells. Nature. 2014;510(7504):235-40. doi: 10.1038/nature13420. PubMed PMID: 24870228; PMCID: PMC4082311.

53. Guschin DY, Waite AJ, Katibah GE, Miller JC, Holmes MC, Rebar EJ. A rapid and general assay for monitoring endogenous gene modification. Methods Mol Biol. 2010;649:247-56. doi: 10.1007/978-1-60761-753-2_15. PubMed PMID: 20680839.

54. DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing SH, Sancak Y, Cui X, Steine EJ, Miller JC, Tam P, Bartsevich VV, Meng X, Rupniewski I, Gopalan SM, Sun HC, Pitz KJ, Rock JM, Zhang L, Davis GD, Rebar EJ, Cheeseman IM, Yamamoto KR, Sabatini DM, Jaenisch R, Gregory PD, Urnov FD. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome Res. 2010;20(8):1133-42. doi: 10.1101/gr.106773.110. PubMed PMID: 20508142; PMCID: PMC2909576.

55. Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. Nat Rev Cancer. 2012;12(1):51-8.

56. Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF, Neri M, Magnani Z, Cantore A, Lo Riso P, Damo M, Pello OM, Holmes MC, Gregory PD, Gritti A, Broccoli V, Bonini C, Naldini L. Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nature methods. 2011;8(10):861-9.

57. Papapetrou EP, Lee G, Malani N, Setty M, Riviere I, Tirunagari LM, Kadota K, Roth SL, Giardina P, Viale A, Leslie C, Bushman FD, Studer L, Sadelain M. Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nature biotechnology. 2011;29(1):73-8.

58. De Fusco M. et al., Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. Nat. Genet., Feb 2003, 33(2):192-196 (PMID: 12539047).

CLAIMS:

1. A method for nuclease-based modification of a target polynucleotide in one or more target cells within a population of cells and for enriching the number of said target cells comprising said nuclease-based modification of said target polynucleotide within the population, the method comprising:

   (i) introducing in the one or more target cells, by homology-directed repair (HDR) or non-homologous end joining (NHEJ), (a) said nuclease-based modification in the target polynucleotide and (b) a second nuclease-based modification in an endogenous gene locus, wherein the second nuclease-based modification in the endogenous gene locus confers resistance to a drug;

   (ii) contacting said population of cells with the drug; and

   (iii) selecting said one or more target cells by virtue of increased tolerance to the drug, thereby enriching the number of said target cells comprising said nuclease-based modification in said target polynucleotide within the population of cells.

2. The method of claim 1, wherein said nuclease-based modification in the target polynucleotide is introduced by HDR and step (i) further comprises introducing into the one or more target cells a donor nucleic acid comprising said modification.

3. The method of claim 1 or 2, wherein said second nuclease-based modification in the endogenous gene is introduced by HDR and step (i) further comprises introducing into the one or more target cells a second donor nucleic acid comprising said second modification.

4. A method for nuclease-based modification of a target polynucleotide in one or more target cells within a population of cells and for enriching the number of said target cells comprising said nuclease-based modification of said target polynucleotide within the population, the method comprising:

   (i) providing said cell population with a CRISPR nuclease and one or more sgRNAs comprising:

      (a) a sgRNA guide sequence having a target sequence in an endogenous gene, such that the sgRNA guide sequence directs the cleavage of the endogenous gene for introduction of a second nuclease-based modification in the endogenous gene conferring resistance to a drug, and a CRISPR nuclease recognition sequence;

      (b) a sgRNA guide sequence having a target sequence in the target polynucleotide and a CRISPR nuclease recognition sequence;

wherein the target sequences in (a) and (b) are contiguous to a protospacer adjacent motif recognized by the CRISPR nuclease;

(ii) contacting said population of cells with the drug; and

(iii) selecting said one or more target cells by virtue of increased tolerance to the drug, thereby enriching the number of said target cells within the population of cells.

5. The method of claim 4, wherein said nuclease-based modification in the target polynucleotide is introduced by homology-directed repair (HDR) and step (i) further comprises providing said cell population with a donor nucleic acid comprising said modification.

6. The method of claim 4 or 5, wherein said nuclease-based modification in the endogenous gene is introduced by homology-directed repair (HDR) and step (i) further comprises providing said cell population with a second donor nucleic acid comprising said second modification.

7. A method for enriching for target cells modified at an endogenous gene locus, wherein the modification confers resistance to a drug, the method comprising:

(i) cleaving with a nuclease said endogenous gene in one or more target cells within a population of cells;

(ii) contacting said population of cells with the drug; and

(iii) selecting said one or more target cells by virtue of increased tolerance to the drug,

thereby enriching for target cells in which the endogenous gene has been modified.

8. The method of any one of claims 1-3 and 7, wherein the nuclease is a CRISPR nuclease, and wherein said method comprises providing said one or more target cells with a sgRNA comprising a sgRNA guide sequence having a target sequence in the endogenous gene and a CRISPR nuclease recognition sequence.

9. The method of any one of claims 4-6 or 8, wherein said CRISPR nuclease is Cas9 or Cpf1.

10. The method of claim 9, wherein said Cas9 is spCas9 from *Streptococcus pyogenes*.

11. The method of claim 9, wherein said Cpf1 is AsCpf1 from *Acidaminococcus sp*.

12. The method of any one of claims 1 to 11, wherein said endogenous gene and drug are selected from the endogenous genes and corresponding drugs set forth in Table 1.

13. The method of any one of claims 1 to 12, wherein said endogenous gene encodes a Na/K-ATPase and said drug is an Na/K-ATPase inhibitor.

14. The method of claim 13, wherein the Na/K-ATPase inhibitor is ouabain or a derivative thereof.

15. The method of any one of claims 1 to 14, wherein said endogenous gene is an ATP1A gene and said drug is ouabain or a derivative thereof.

16. The method of claim 15, wherein said ATP1A gene is ATP1A1 and said sgRNA target sequence comprises or consists of a sequence as set forth in Table 3 and FIG. 28.

17. The method of claim 15, wherein said sgRNA comprises or consists of the sequence set forth in any one of SEQ ID NOs: 24-26 and 33-36.

18. The method of claim 15, wherein said sgRNA target sequence is within an exon of said ATP1A gene.

19. The method of claim 18, wherein said ATP1A gene is ATP1A1 and said exon is exon 4.

20. The method of claim 19, wherein said target sequence comprises or consists of the following target sequence (i) ATCCAAGCTGCTACAGAAG (G2); (ii) GTTCCTCTTCTGTAGCAGCT (G4); (iii) TTGGCTTATAGCATCCAAGC (G6); (iv) AGGTTCCTCTTCTGTAGCAG (G7); or (v) GAGGTTCCTCTTCTGTAGCA (G8).

21. The method of claim 15, wherein said gRNA target sequence is within an intron of said ATP1A gene.

22. The method of claim 21, wherein said ATP1A is ATP1A1 and said intron is intron 3 or intron 4.

23. The method of claim 22, wherein said target sequence comprises or consists of the following target sequence (i) GAGTTCTGTAATTCAGCATA (G3) or (ii) TAGTACACATCAGATATCTT (G9).

24. The method of claim 21 or 22, wherein said method further comprises providing said cell with an ATP1A donor nucleic acid conferring ouabain resistance.

25. The method of claim 24, wherein said ATP1A donor nucleic acid conferring ouabain resistance is an ATP1A1 donor nucleic acid which comprises or consists of the following nucleic acid sequence:

(i) CAATGTTACTGTGGATTGGAGCGATTCTTTGTTTCTTGGCTTATAGCATCGATGCTGCTACAGA AGAGGAACCTCAAAACGATCGTGTGAGTTCTGTAATTCAGCATATCGATTTGTAGTACACATC AGATATCTT (DR); or

(ii) CAATGTTACTGTGGATTGGAGCGATTCTTTGTTTCTTGGCTTATAGCATCAGAGCTGCTACAG AAGAGGAACCTCAAAACGATGACGTGAGTTCTGTAATTCAGCATATCGATTTGTAGTACACAT CAGATATCTT (RD).

26. The method of claim 24, wherein said ATP1A donor nucleic acid is an ATP1A1 donor nucleic acid which confers ouabain resistance by replacing at least one of amino acids Q118 and N129 in the ATP1A gene by a charged amino acid.

27. The method of claim 24, wherein said ATP1A donor nucleic acid is an ATP1A1 donor nucleic acid which confers ouabain resistance by replacing amino acid Q118 by an aspartic acid and amino acid N129 by an arginine.

28. The method of claim 24, wherein said ATP1A donor nucleic acid is an ATP1A1 donor nucleic acid which confers ouabain resistance by replacing amino acid Q118 by an arginine and amino acid N129 by an aspartic acid.

29. The method of any one of claims 1 to 28, wherein the one or more target cells are selected from the group consisting of T-cells, a B-cells or stem cells.

30. The method of claim 29, wherein the stem cell is an embryonic stem cell (ESC), an induced pluripotent stem cell (iPSC), a CD34+ hematopoietic stem cell or a hepatic stem cell.

31. A ouabain resistant ATP1A1 polypeptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NOs: 5, 8, 10, 12, 14, 16, 18, 20 or 22.

32. A ouabain resistant polypeptide comprising or consisting of a polypeptide as set forth in:
    (a) SEQ ID NO: 5 (ATP1A1), wherein:
        (i) the amino acid at position 119 is any amino acid except an alanine or an isoleucine or is absent;
        (ii) the amino acid at position 120 is any amino acid except an alanine, , a glycine or a tyrosine, or is absent;
        (iii) the amino acid at position 121 is any amino acid except a threonine, a methionine or a phenylalanine, or is absent;
        (iv) the amino acid at position 122 is any amino acid except a glutamic acid or an asparagine, or is absent;
        (v) the amino acid at position 123 is any amino acid except a glutamic acid or an aspartic acid, or is absent;

(vi) the amino acid at position 124 is any amino acid except a glutamic acid or an aspartic acid, or is absent;

(vii) the amino acid at position 125 is any amino acid except a proline or a lysine or is absent;

(viii) the amino acid at position 126 is any amino acid except a glutamine, a serine or a threonine or is absent; or

(ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii);

(b) SEQ ID NO: 6 (ATP1A1), wherein:

(i) the amino acid at position 119 is any amino acid except an alanine or an isoleucine, or is absent;

(ii) the amino acid at position 120 is any amino acid except an alanine, a glycine or a tyrosine, or is absent;

(iii) the amino acid at position 121 is any amino acid except a threonine, a methionine or a phenylalanine, or is absent;

(iv) the amino acid at position 122 is any amino acid except a glutamic acid or an asparagine, or is absent;

(v) the amino acid at position 123 is any amino acid except a glutamic acid or an aspartic acid, or is absent;

(vi) the amino acid at position 124 is any amino acid except a glutamic acid or an aspartic acid, or is absent;

(vii) the amino acid at position 125 is any amino acid except a proline or a lysine or is absent;

(viii) the amino acid at position 126 is any amino acid except a glutamine, a serine or a threonine, or is absent; or

(ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii);

(c) SEQ ID NO: 485 (ATP1A2), wherein:

(i) the amino acid at position 117 is any amino acid except an alanine or an isoleucine, or is absent;

(ii) the amino acid at position 118 is any amino acid except an alanine, a glycine or a tyrosine, or is absent;

(iii) the amino acid at position 119 is any amino acid except a threonine, a methionine or a phenylalanine, or is absent;

(iv) the amino acid at position 120 is any amino acid except a glutamic acid or an asparagine, or is absent;

(v) the amino acid at position 121 is any amino acid except a glutamic acid or an aspartic acid, or is absent;

(vi) the amino acid at position 122 is any amino acid except a glutamic acid or an aspartic acid, or is absent;

(vii) the amino acid at position 123 is any amino acid except a proline or a lysine or is absent;

(viii) the amino acid at position 124 is any amino acid except a glutamine, a serine or a threonine, or is absent; or

(x) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii);

(d) SEQ ID NO: 486 (ATP1A3), wherein:

(i) the amino acid at position 109 is any amino acid except an alanine or an isoleucine, or is absent;

(ii) the amino acid at position 110 is any amino acid except an alanine, a glycine or a tyrosine, or is absent;

(iii) the amino acid at position 111 is any amino acid except a threonine a methionine or a phenylalanine, or is absent;

(iv) the amino acid at position 112 is any amino acid except a glutamic acid or an asparagine, or is absent;

(v) the amino acid at position 113 is any amino acid except a glutamic acid or an aspartic acid, or is absent;

(vi) the amino acid at position 114 is any amino acid except a glutamic acid or an aspartic acid or is absent;

(vii) the amino acid at position 115 is any amino acid except a proline or a lysine or is absent;

(viii) the amino acid at position 116 is any amino acid except a glutamine, a serine or a threonine or is absent; or

(ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii); or

(e) SEQ ID NO: 487 (ATP1A4), wherein:

(i) the amino acid at position 127 is any amino acid except an alanine or an isoleucine, or is absent;

(ii) the amino acid at position 128 is any amino acid except an alanine, a glycine or a tyrosine, or is absent;

(iii) the amino acid at position 129 is any amino acid except a threonine a methionine or a phenylalanine, or is absent;

(iv) the amino acid at position 130 is any amino acid except a glutamic acid or an asparagine, or is absent;

(v) the amino acid at position 131 is any amino acid except a glutamic acid or an aspartic acid, or is absent;

(vi) the amino acid at position 132 is any amino acid except a glutamic acid or an aspartic acid or is absent;

(vii) the amino acid at position 133 is any amino acid except a proline or a lysine or is absent;

(viii) the amino acid at position 134 is any amino acid except a glutamine, a serine or a threonine or is absent; or

(ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

33. A ouabain resistant polypeptide, wherein said polypeptide comprises a deletion or a non-conservative substitution at amino acid:

(i) A119, A120, T121, E122, E123, E124, P125 and/or Q126 of SEQ ID NO: 3 or 4 (ATP1A1),

(ii) A117, A118, M119, E120, D121, E122, P123 and/or S124 of SEQ ID NO: 460 (ATP1A2);

(iii) A109, G110, T111, E112, D113, D114, P115 and/or S116 of SEQ ID NO: 462 (ATP1A3); or

(iv) I127, Y128, F129, N130, E131, E132, P133 and/or T134 of SEQ ID NO: 464 (ATP1A4).

34. The ouabain resistant polypeptide of claim 32 (a) or (b), wherein

(i) the amino acid at position 119 is absent;

(ii) the amino acid at position 120 is absent;

(iii) the amino acid at position 121 is a lysine or is absent;

(iv) the amino acid at position 122 is absent;

(v) the amino acid at position 123 is absent;

(vi) the amino acid at position 124 is absent;

(vii) the amino acid at position 125 is an alanine or is absent;

(viii) the amino acid at position 126 is absent; or

(ix) any combination of (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii).

35. A sgRNA for inducing or conferring ouabain resistance by NHEJ or HDR comprising (i) a sgRNA guide sequence having a target sequence in the ATP1A1 gene comprising or consisting of a target sequence as set forth in Table 3 and FIG. 22; and (ii) a CRISPR nuclease recognition sequence.

36. A sgRNA for inducing or conferring ouabain resistance by NHEJ or HDR comprising (i) a sgRNA guide sequence having a target sequence comprising or consisting of the following target sequence in the ATP1A gene: (i) GATCCAAGCTGCTACAGAAG (G2); (ii) GTTCCTCTTCTGTAGCAGCT (G4); (iii)

TTGGCTTATAGCATCCAAGC (G6); (iv) AGGTTCCTCTTCTGTAGCAG (G7); or (v) GAGGTTCCTCTTCTGTAGCA (G8).

37. A sgRNA for inducing or conferring ouabain resistance by NHEJ or HDR comprising (i) a sgRNA guide sequence having a target sequence comprising or consisting of the following target sequence in the ATP1A1 gene: (i) GAGTTCTGTAATTCAGCATATGG (G3) or (ii) TAGTACACATCAGATATCTT (G9).

38. A nucleic acid encoding the ouabain resistant ATP1A polypeptide of any one of claims 31 to 34 or the sgRNA of any one of claims 35 to 37.

39. A vector comprising the nucleic acid of claim 38.

40. The vector of claim 39, wherein the vector is a viral vector.

41. A cell comprising the nucleic acid of claim 38 or the vector of claim 39 or 40.

42. A kit comprising:
    (i) the sgRNA of any one of claims 35 to 37,
    (ii) the ouabain resistant ATP1A polypeptide of any one of claims 31 to 34;
    (iii) the vector of claim 39 or 40;
    (iv) the cell of claim 41; and/ or
    (v) the ATP1A donor nucleic acid defined in any one of claims 24 to 28.

43. Use of (i) the sgRNA of any one of claims 35 to 37, (ii) the ouabain resistant ATP1A1 polypeptide of any one of claims 31 to 34; (iii) the vector of claim 39 or 40; (iv) the cell of claim 41; and/ or (iv) the ATP1A donor nucleic acid defined in any one of claims 24 to 28 for (a) inducing or conferring ouabain resistance in one or more target cells; (b) nuclease-based modification a target polynucleotide in one or more target cells and for enriching the number of said target cells; and/or or (c) for enriching for target cells modified by a nuclease at an ATP1A gene locus.

44. The sgRNA of any one of claims 35 to 37, the ouabain resistant ATP1A1 polypeptide of any one of claims 31 to 34; the vector of claim 39 or 40; the cell of claim 41; and/ or the ATP1A donor nucleic acid defined in any one of claims 24 to 28 for use in (a) inducing or conferring ouabain resistance in one or more target cells; (b) for nuclease-based modification of a target polynucleotide in one or more target cells for enriching the number of said target cells; and/or or (c) for enriching for target cells modified by a nuclease at an ATP1A gene locus.

45. A composition comprising (i) the sgRNA of any one of claims 35 to 37, (ii) the ouabain resistant ATP1A1 polypeptide of any one of claims 31 to 34; (iii) the vector of claim 39 or 40; (iv) the cell of claim 41; and/ or (iv) the ATP1A donor nucleic acid defined in any one of claims 24 to 28.

46. The composition of claim 45, further comprising a suitable carrier.

47. The composition of claim 45 or 46 or the kit of claim 42 for (a) inducing or conferring ouabain resistance in one or more target cells; (b) for modifying with a nuclease a target polynucleotide in one or more target cells and for enriching the number of said target cells; and/or or (c) for enriching for target cells modified by a nuclease at an ATP1A gene locus.

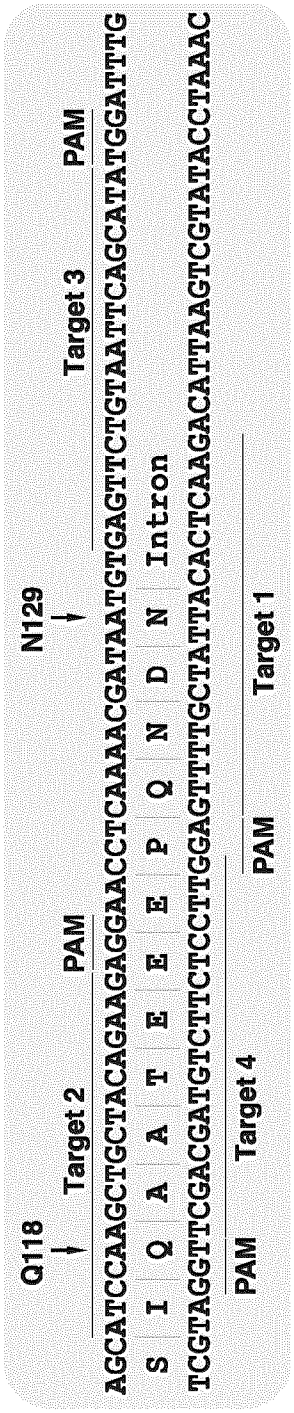

Q118
N129
Target 2
PAM
Target 3
PAM
AGCATCCAAGCTGCTACAGAAGAGGAACCTCAAAACGATAATGTGAGTTCTGTAATTCAGCATATGGATTTG
S I Q A A T E E E P Q N D N Intron
TCGTAGGTTCGACGATGTCTTCTCCTTGGAGTTTTGCTATTACACTCAAGACATTAAGTCGTATACCTAAAC
PAM
Target 4
PAM
Target 1

FIG. 1

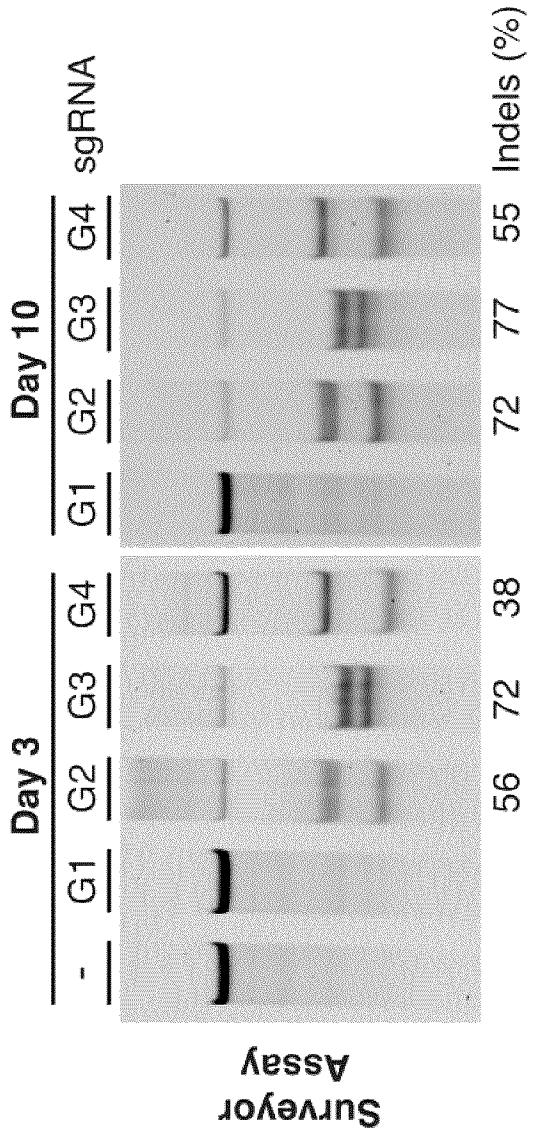

Surveyor Assay
Day 3
Day 10
sgRNA
-
G1
G2
G3
G4
G1
G2
G3
G4
Indels (%)
56
72
38
72
77
55

FIG. 2

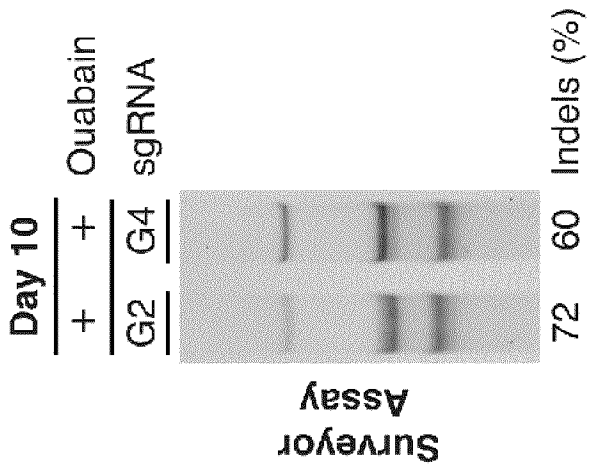


FIG. 3

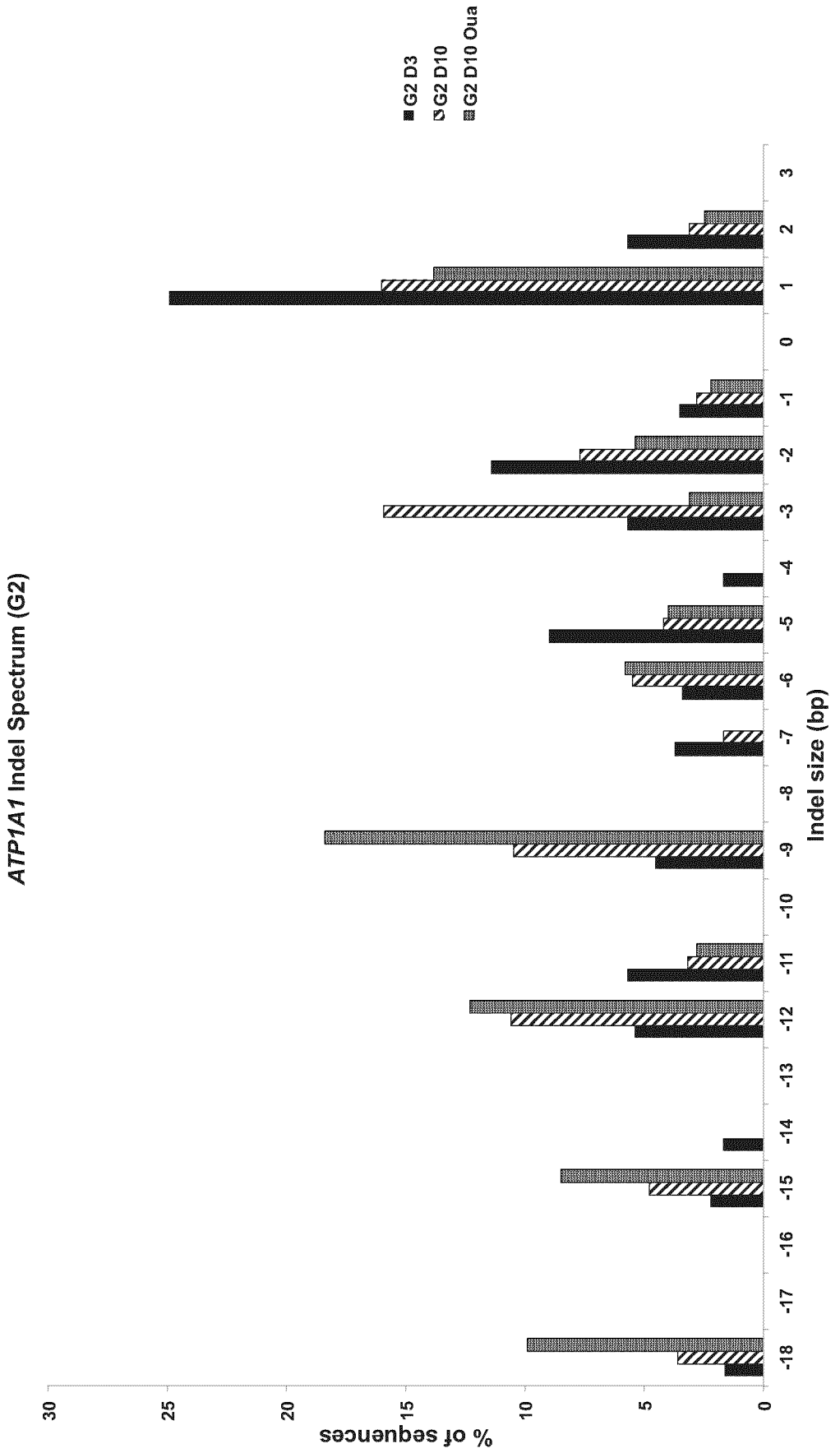

ATP1A1 Indel Spectrum (G2)
G2 D3
G2 D10
G2 D10 Oua
% of sequences
0
5
10
15
20
25
30
Indel size (bp)
3
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
-9
-10
-11
-12
-13
-14
-15
-16
-17
-18

FIG. 4

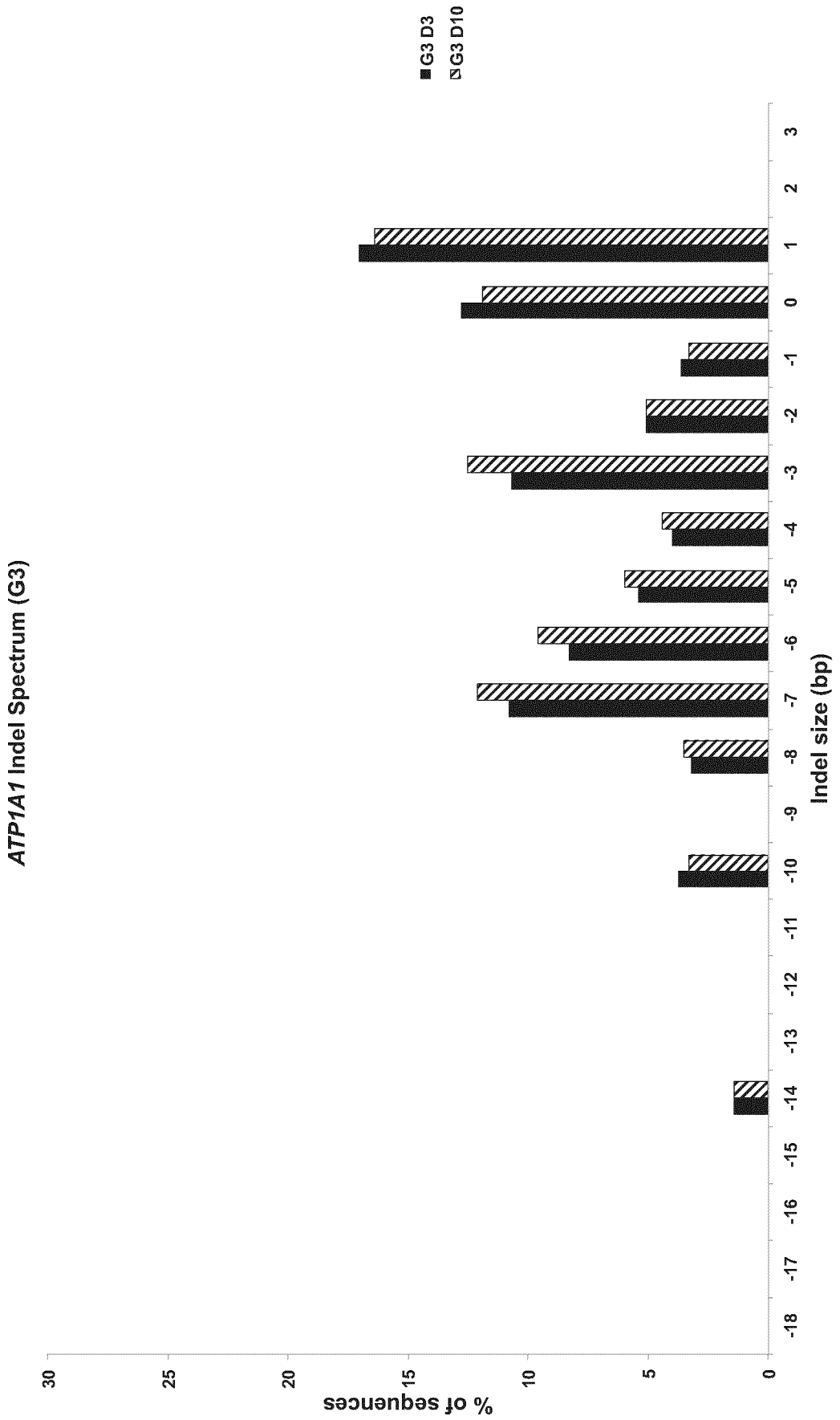

ATP1A1 Indel Spectrum (G3)
G3 D3
G3 D10
% of sequences
0
5
10
15
20
25
30
Indel size (bp)
3
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
-9
-10
-11
-12
-13
-14
-15
-16
-17
-18

FIG. 5

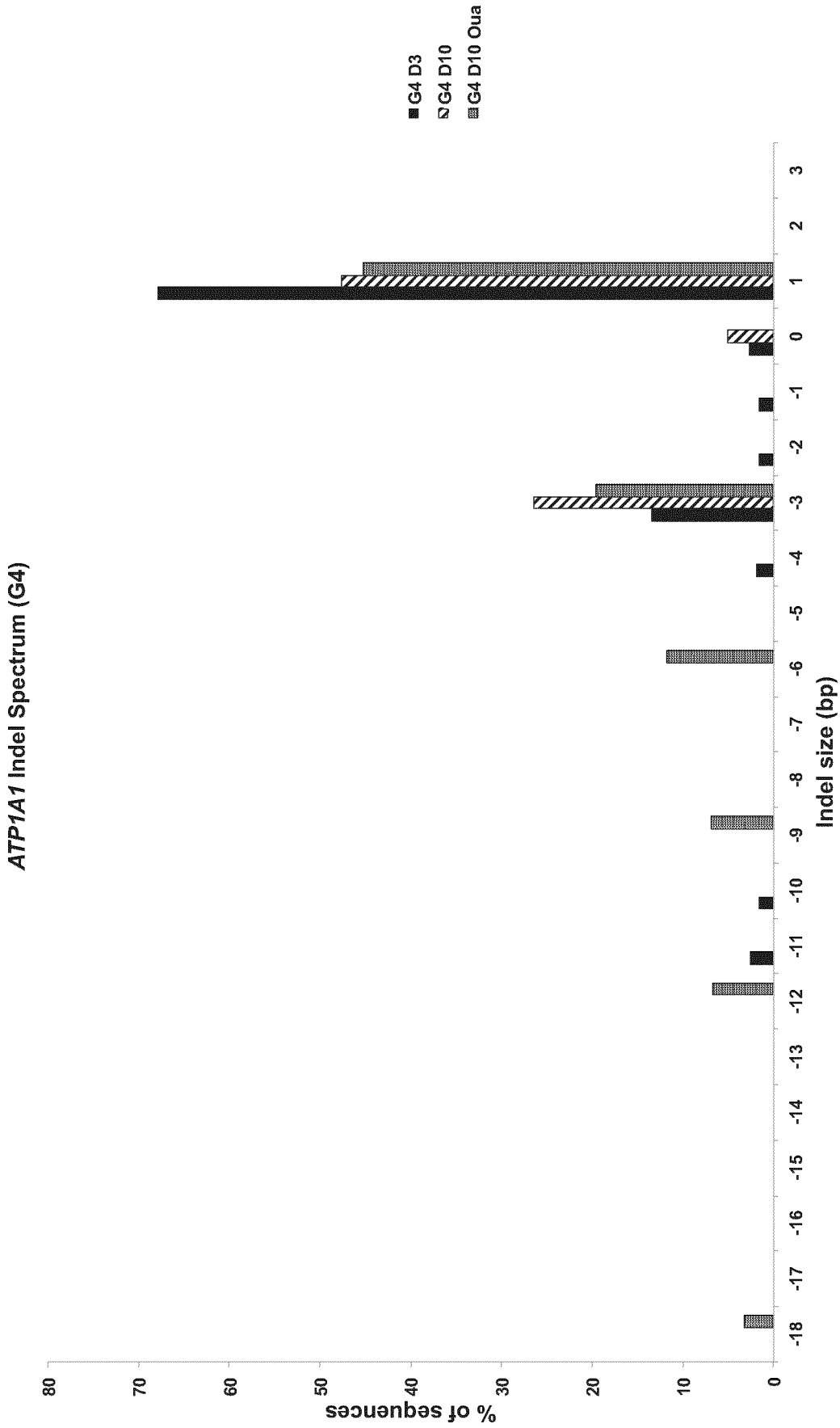

ATP1A1 Indel Spectrum (G4)
G4 D3
G4 D10
G4 D10 Oua
% of sequences
0
10
20
30
40
50
60
70
80
Indel size (bp)
-18
-17
-16
-15
-14
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3

FIG. 6

G2 D10 + Ouabain (0.5uM)
S I Q A A T E E E P Q N D N Wild-Type
S I Q A A K - E E P Q N D N (Δ3)
S I Q A A T E - - P Q N D N (Δ6)
S I Q A A T - - - A Q N D N (Δ9)
S I Q - - - - E E P Q N D N (Δ12)
G4 D10 + Ouabain (0.5uM)
S I Q A A T E E E P Q N D N Wild-Type
S I Q A - T E E E P Q N D N (Δ3b)

FIG. 7

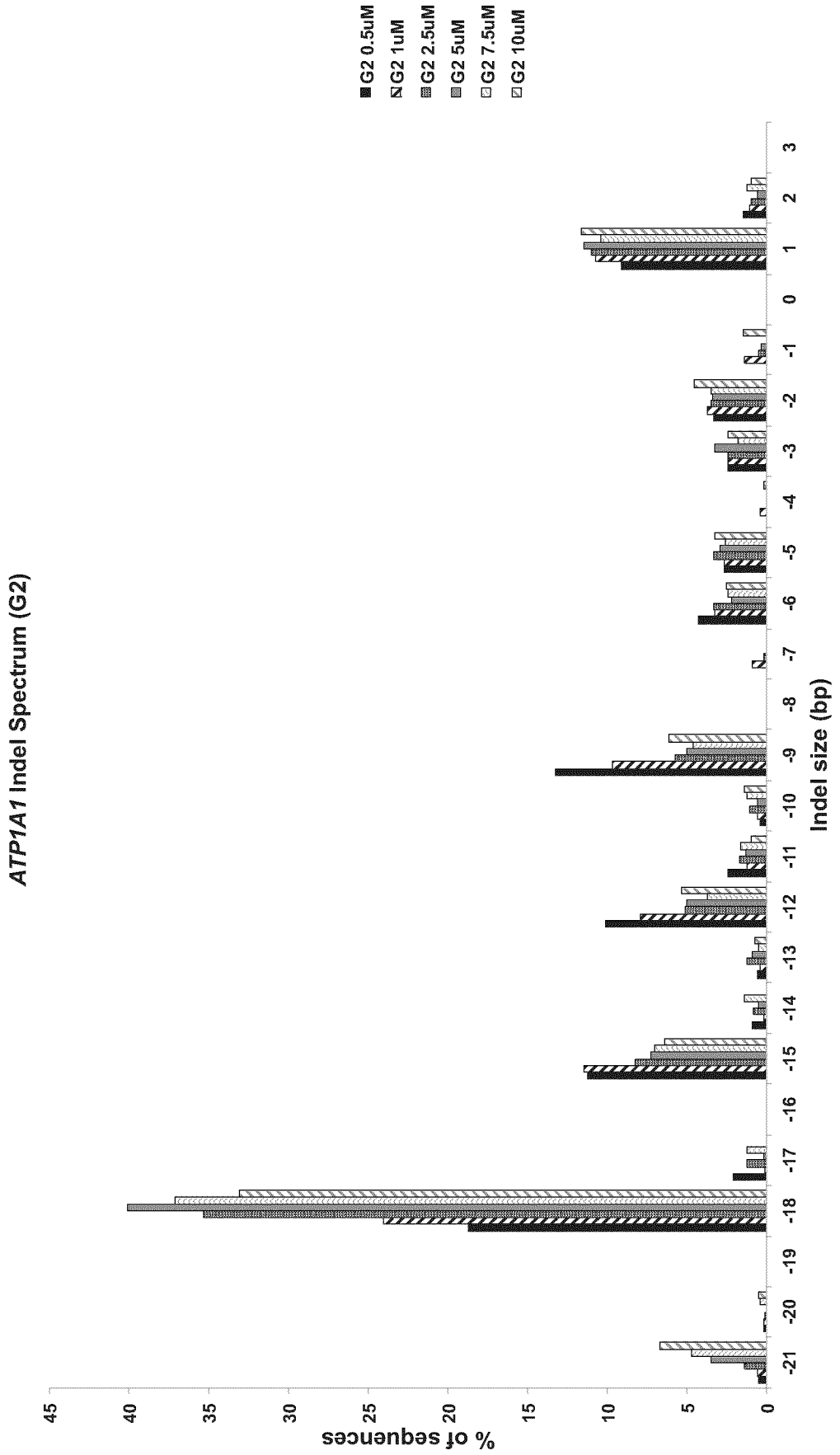

ATP1A1 Indel Spectrum (G2)
G2 0.5uM
G2 1uM
G2 2.5uM
G2 5uM
G2 7.5uM
G2 10uM
% of sequences
Indel size (bp)
45
40
35
30
25
20
15
10
5
0
-21
-20
-19
-18
-17
-16
-15
-14
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3

FIG. 8

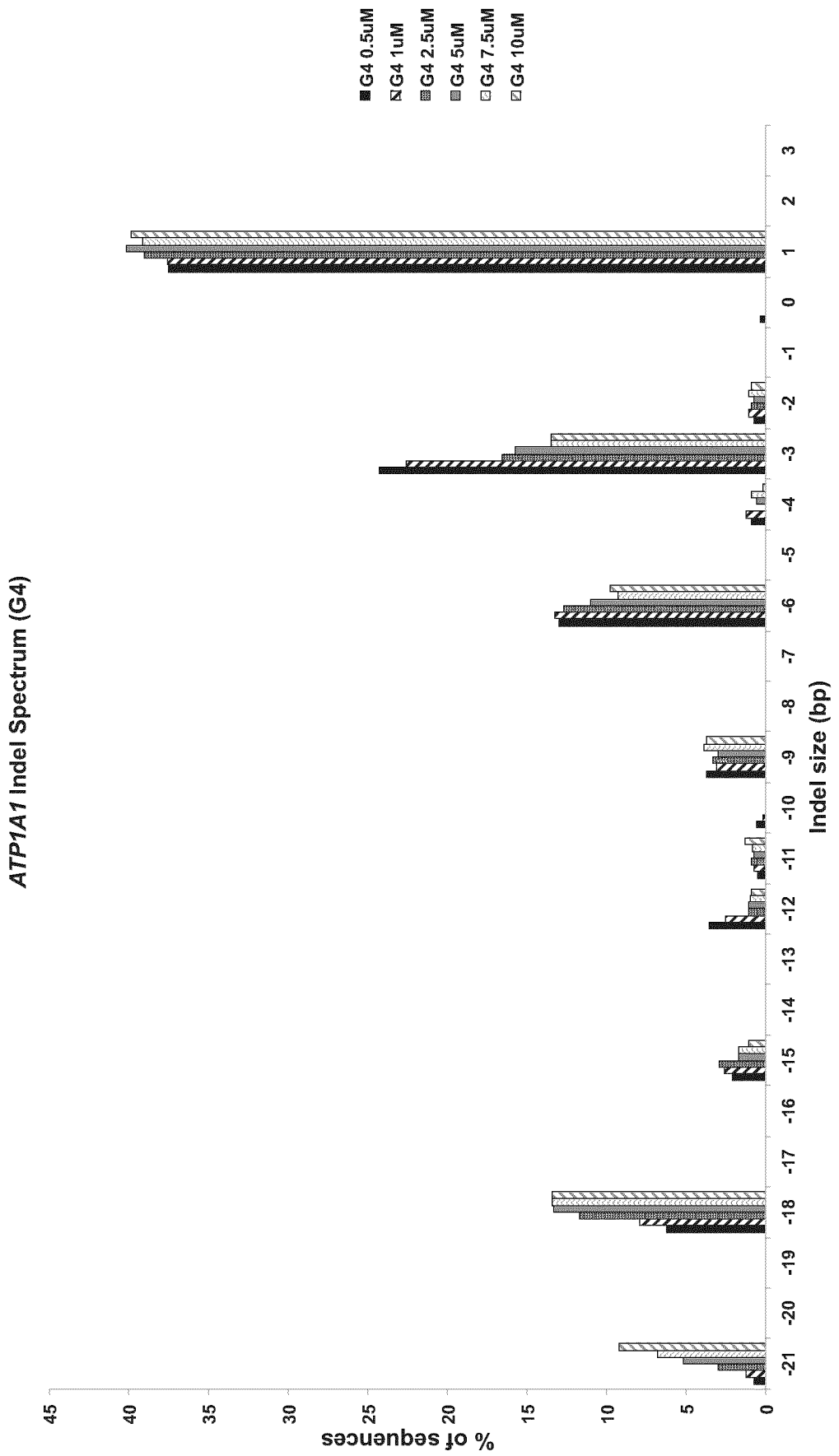

ATP1A1 Indel Spectrum (G4)
G4 0.5uM
G4 1uM
G4 2.5uM
G4 5uM
G4 7.5uM
G4 10uM
% of sequences
Indel size (bp)
45
40
35
30
25
20
15
10
5
0
-21
-20
-19
-18
-17
-16
-15
-14
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3

FIG. 9

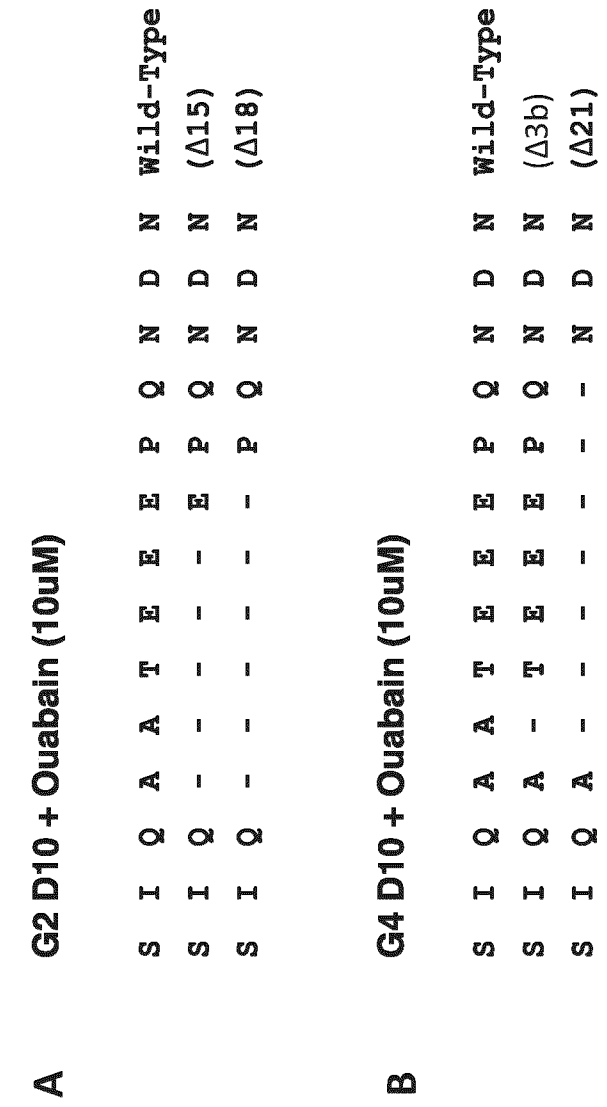

A
G2 D10 + Ouabain (10uM)
S I Q A A T E E E P Q N D N Wild-Type
S I Q - - - - - E P Q N D N (Δ15)
S I Q - - - - - - P Q N D N (Δ18)
B
G4 D10 + Ouabain (10uM)
S I Q A A T E E E P Q N D N Wild-Type
S I Q A - T E E E P Q N D N (Δ3b)
S I Q A - - - - - - - - N D N (Δ21)

FIG. 10

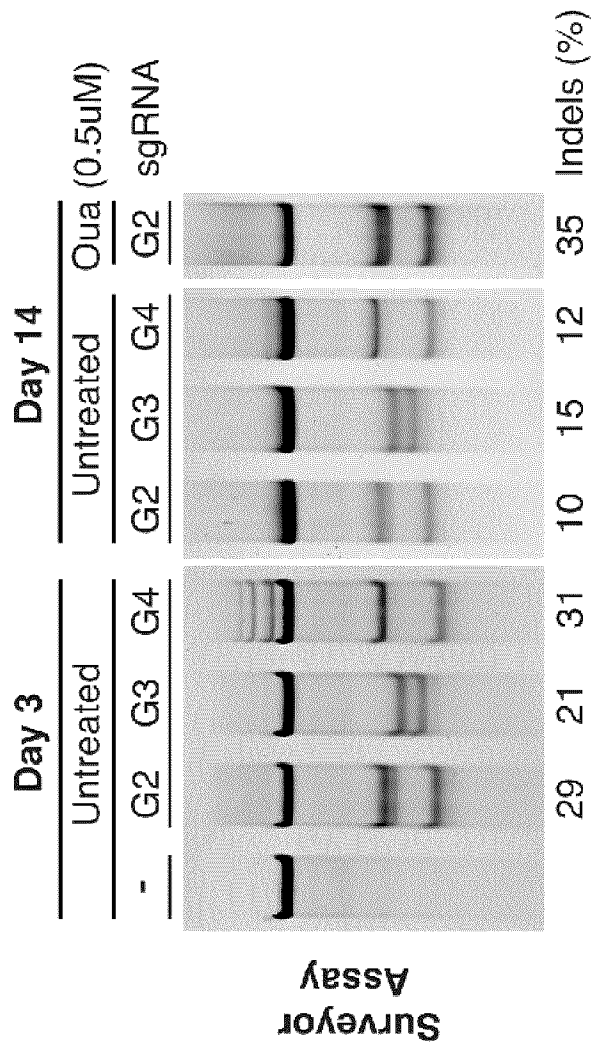

Day 3
Day 14
Untreated
Untreated
Oua (0.5uM)
-
G2
G3
G4
G2
G3
G4
G2
sgRNA
Surveyor Assay
29
21
31
10
15
12
35
Indels (%)

FIG. 11A

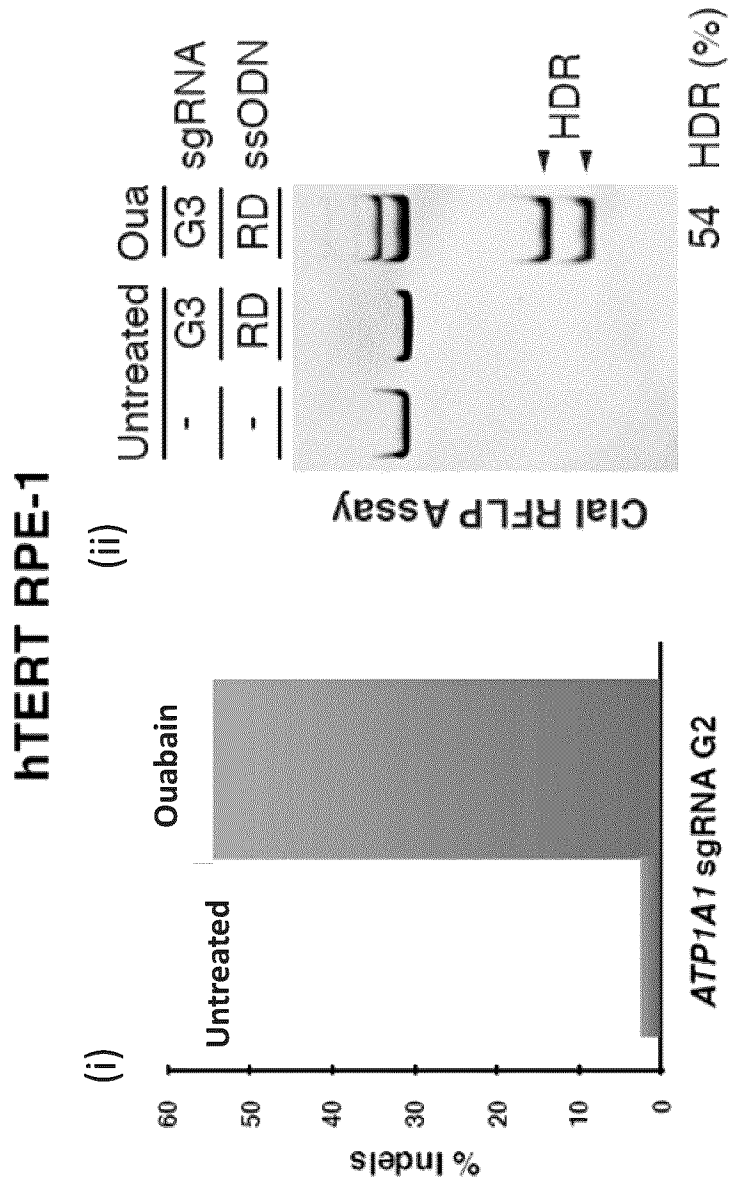

hTERT RPE-1
(i)
Ouabain
Untreated
% Indels
60
50
40
30
20
10
0
ATP1A1 sgRNA G2
(ii)
Untreated
-
G3
-
RD
Oua
G3
RD
sgRNA
ssODN
ClaI RFLP Assay
HDR
54 HDR (%)

FIG. 11B

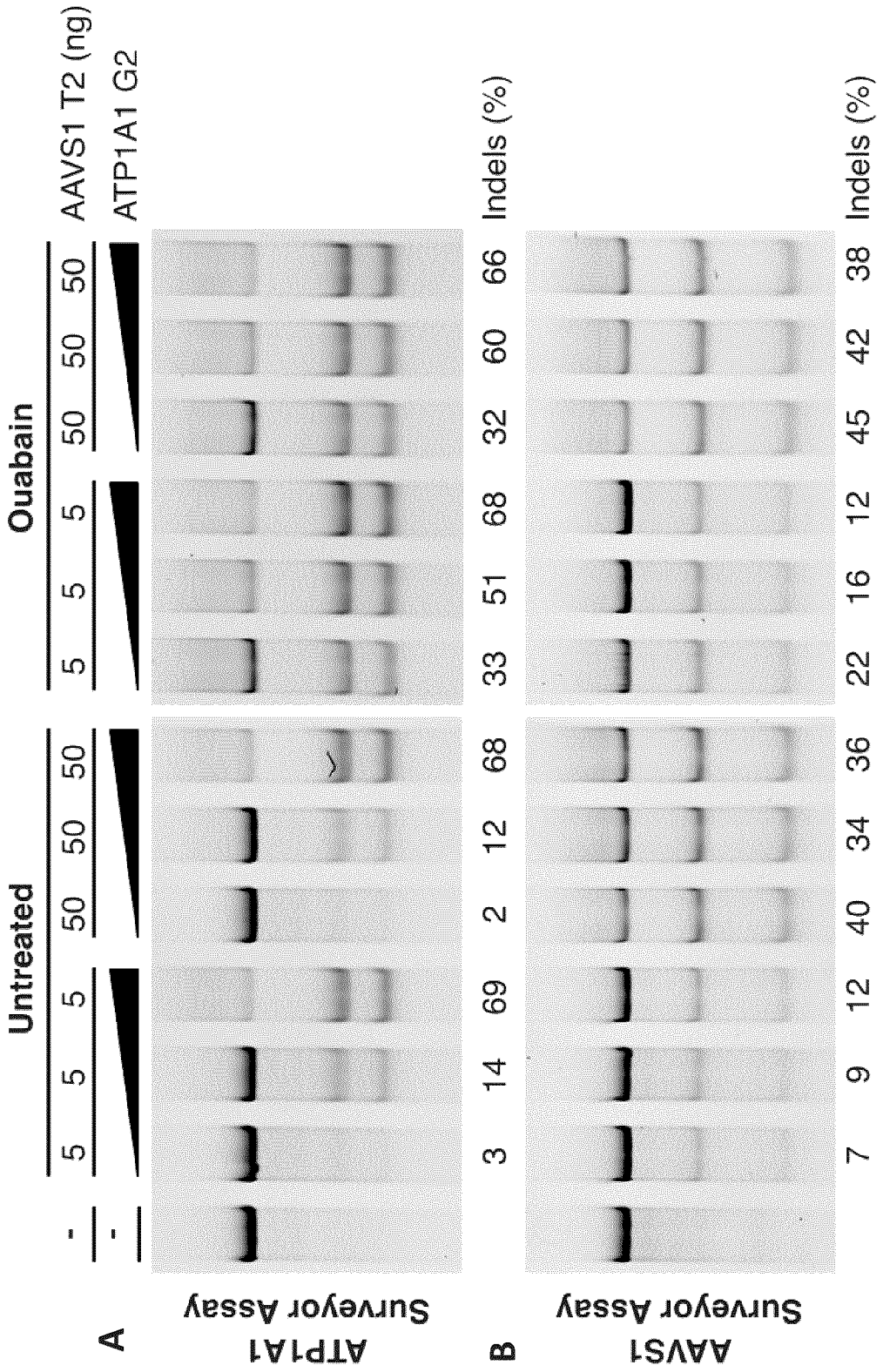

A
ATP1A1 Surveyor Assay
B
AAVS1 Surveyor Assay
Untreated
Ouabain
AAVS1 T2 (ng)
ATP1A1 G2
- - 5 5 5 50 50 50
5 5 5 50 50 50
Indels (%)
3 14 69 2 12 68
33 51 68 32 60 66
7 9 12 40 34 36
22 16 12 45 42 38

FIG. 12

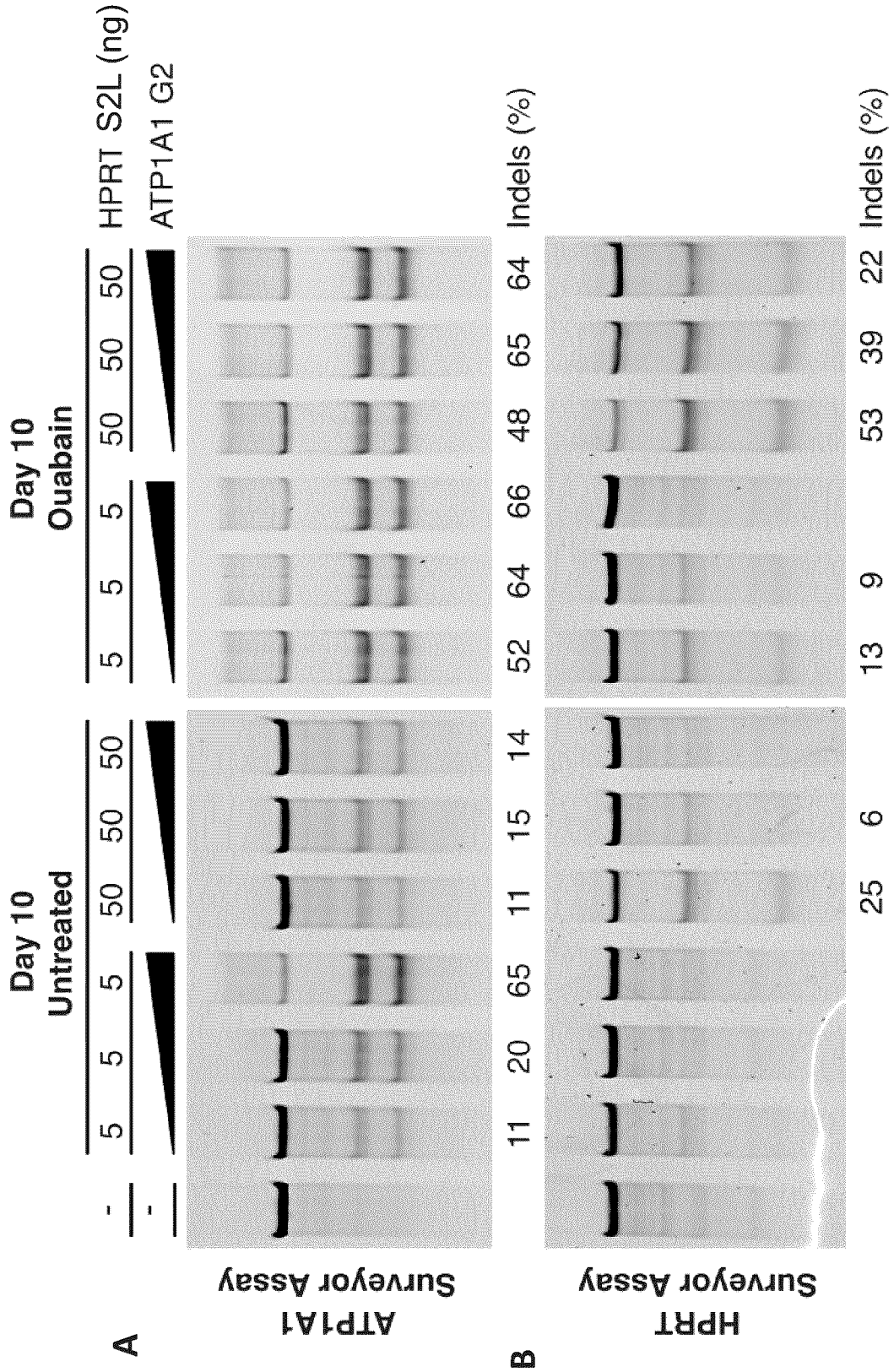

A
ATP1A1 Surveyor Assay
Day 10 Untreated
Day 10 Ouabain
HPRT S2L (ng)
ATP1A1 G2
- - 5 5 5 50 50 50 5 5 5 50 50 50
11 20 65 11 15 14 52 64 66 48 65 64 Indels (%)
B
HPRT Surveyor Assay
25 6 13 9 53 39 22 Indels (%)

FIG. 13

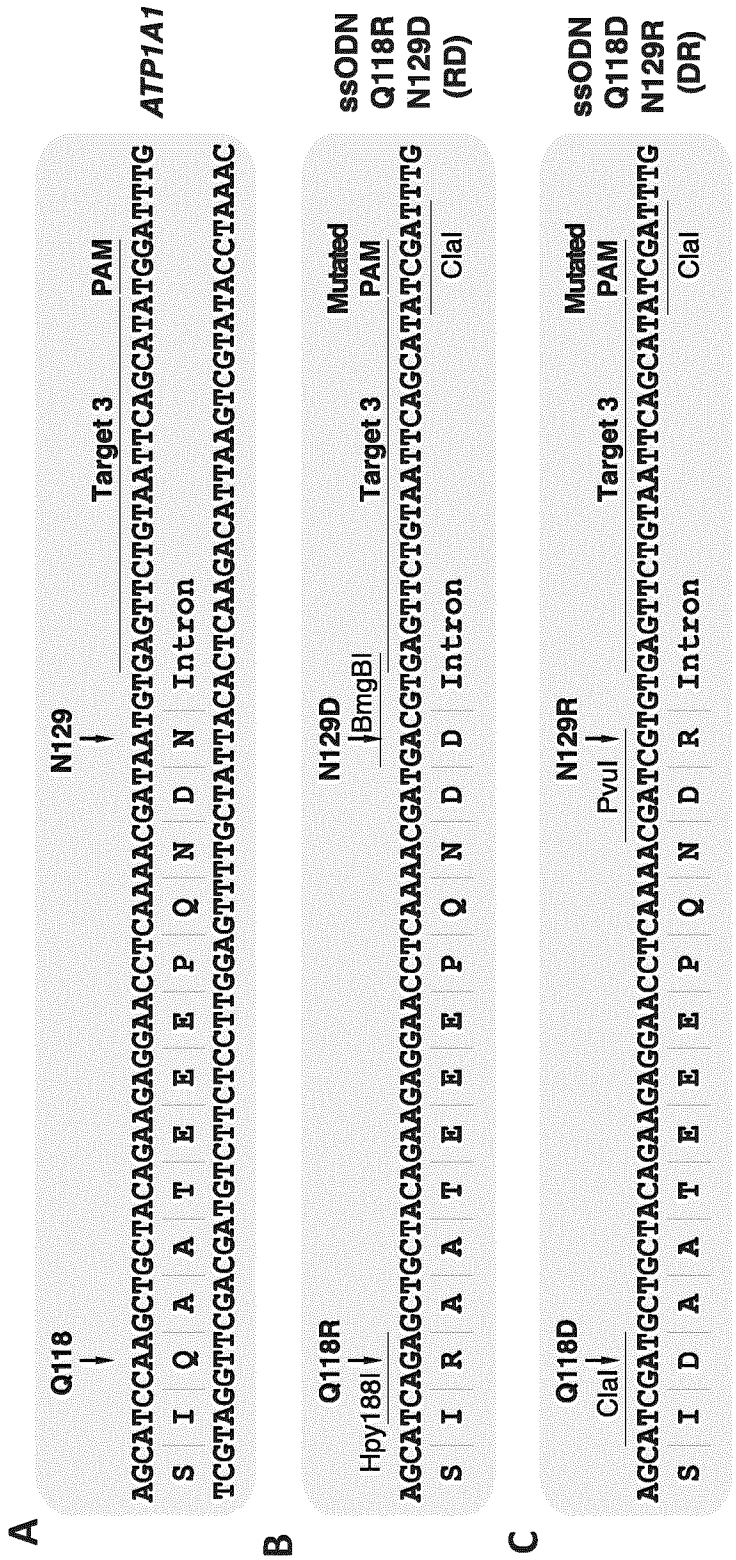

A
Q118
N129
Target 3
PAM
ATP1A1
S I Q A A T E E E P Q N D N Intron
B
Q118R
N129D
Hpy188I
BmgBI
Target 3
Mutated PAM
ssODN Q118R N129D (RD)
S I R A A T E E E P Q N D D Intron
ClaI
C
Q118D
ClaI
N129R
PvuI
Target 3
Mutated PAM
ssODN Q118D N129R (DR)
S I D A A T E E E P Q N D R Intron
ClaI

FIG. 14

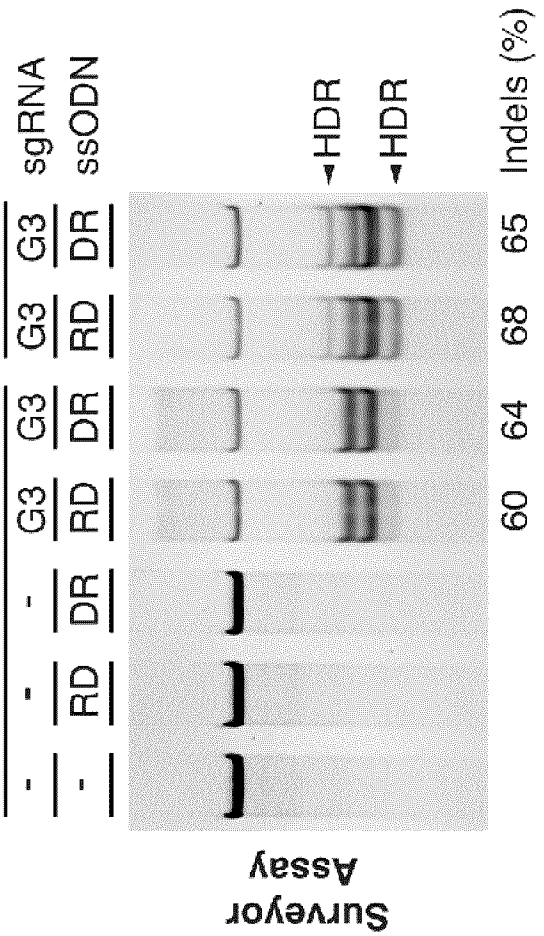

Surveyor Assay
Untreated
Ouabain
sgRNA
ssODN
- - - G3
- RD DR RD
- - G3 DR
- RD RD
- G3 G3
- DR DR
HDR
HDR
Indels (%)
60
64
68
65

FIG. 15

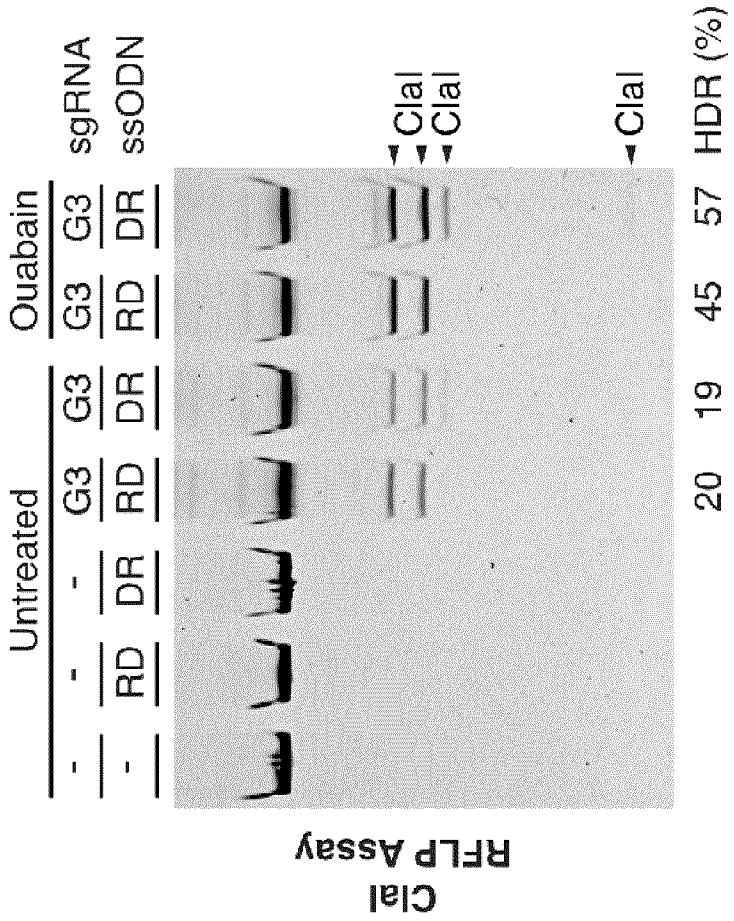

ClaI RFLP Assay
Untreated
Ouabain
sgRNA
ssODN
- - - G3 G3 G3 G3
- RD DR RD DR RD DR
ClaI
ClaI
ClaI
20 19 45 57
HDR (%)

FIG. 16

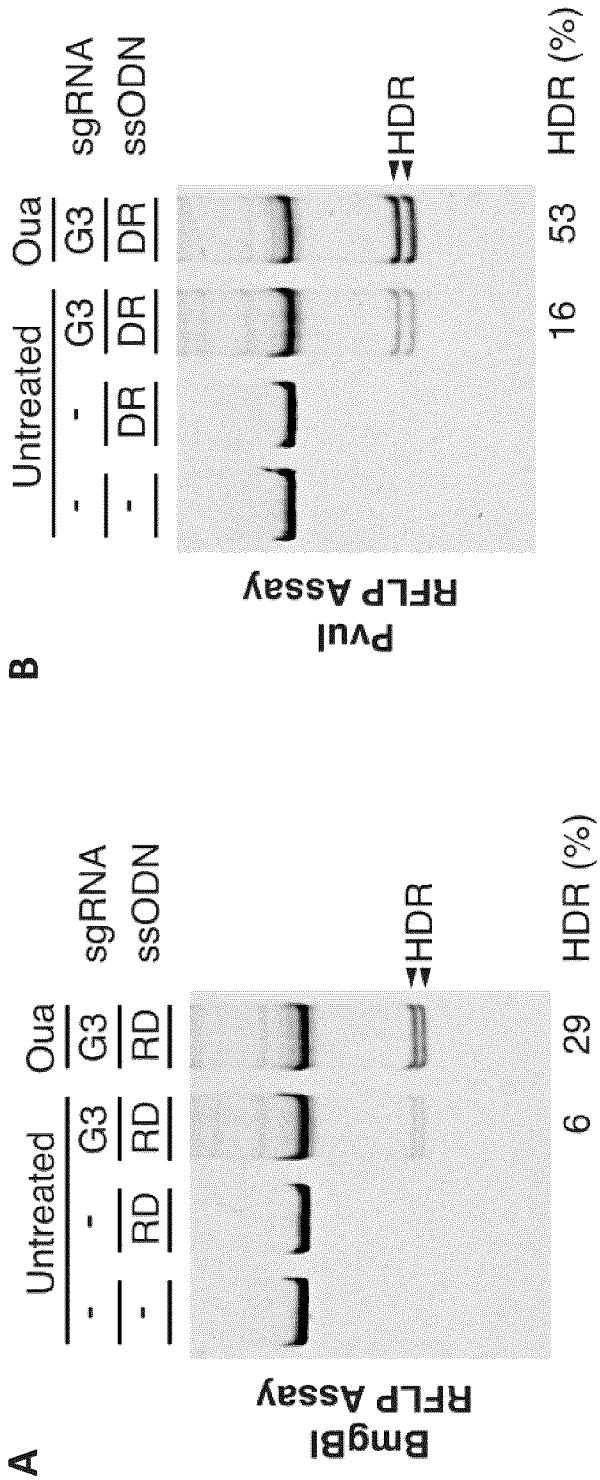

A
BmgBI
RFLP Assay
Untreated
Oua
sgRNA
ssODN
-
-
RD
G3
RD
G3
RD
HDR
6
29
HDR (%)
B
PvuI
RFLP Assay
Untreated
Oua
sgRNA
ssODN
-
-
DR
G3
DR
G3
DR
HDR
16
53
HDR (%)

FIG. 17

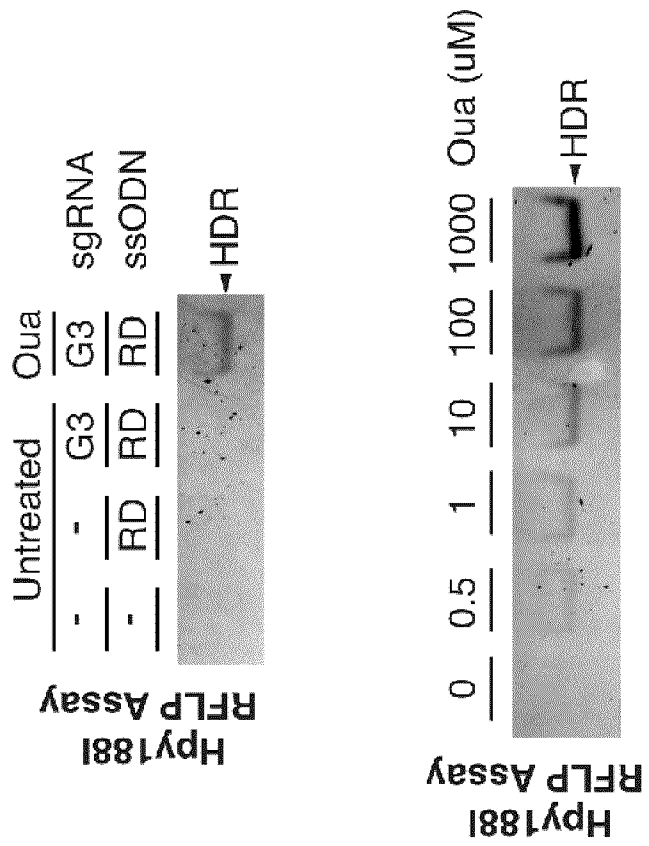

A
Hpy188I
RFLP Assay
Untreated
Oua
sgRNA
ssODN
-
-
G3
-
RD
RD
-
RD
RD
HDR
B
Hpy188I
RFLP Assay
0
0.5
1
10
100
1000
Oua (uM)
HDR

FIG. 18

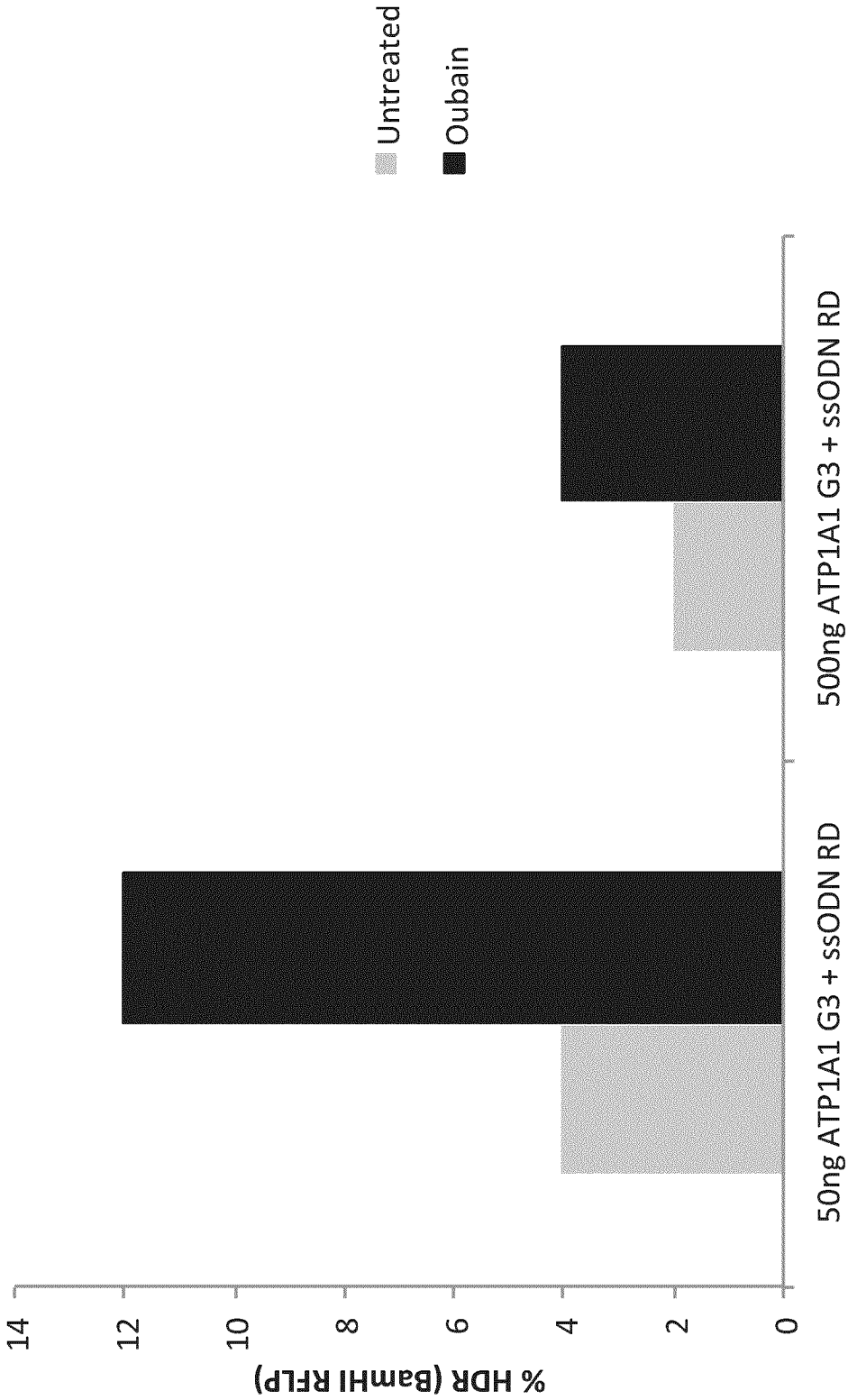

Untreated
Oubain
% HDR (BamHI RFLP)
0
2
4
6
8
10
12
14
50ng ATP1A1 G3 + ssODN RD
500ng ATP1A1 G3 + ssODN RD

FIG. 19

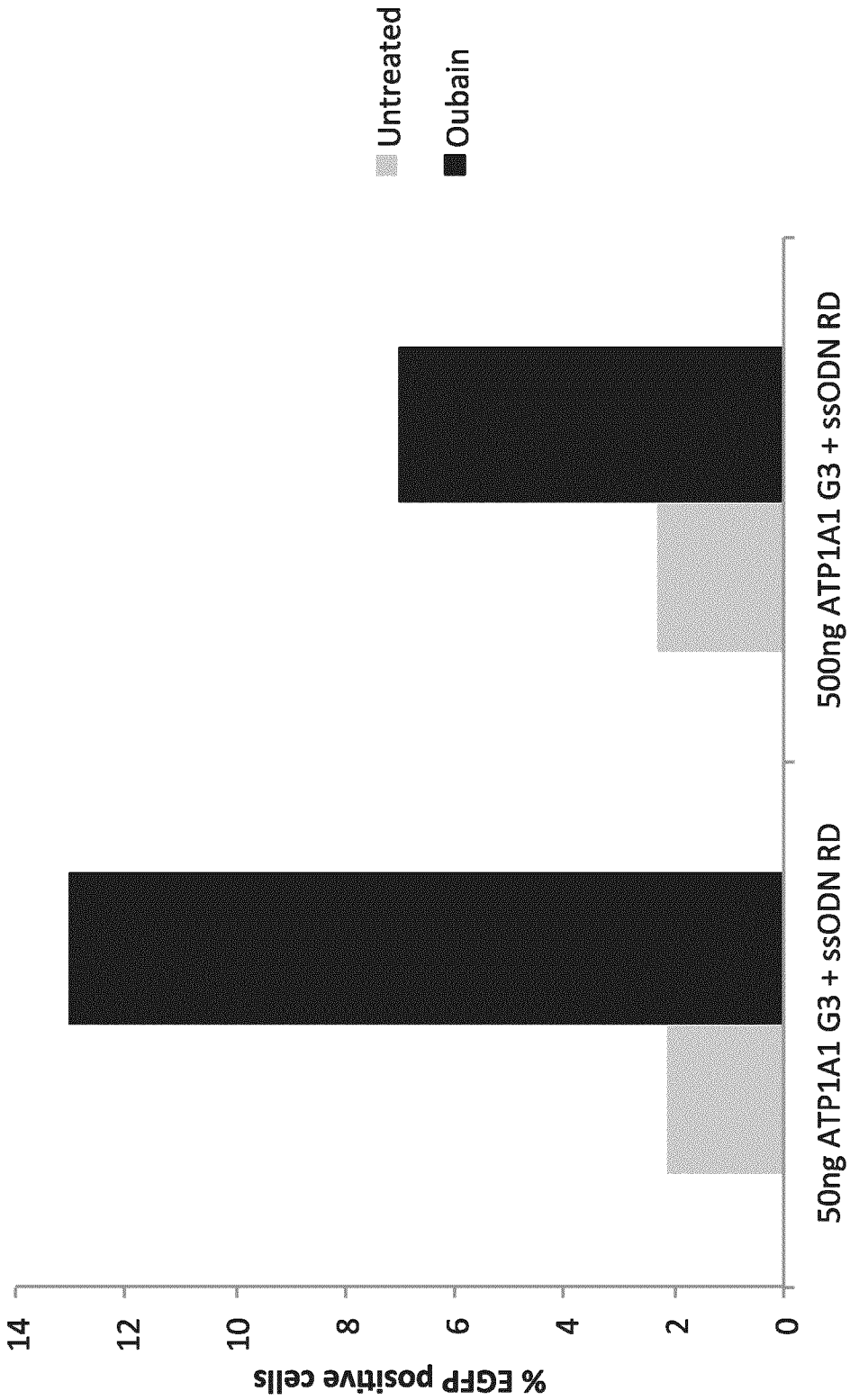

Untreated
Oubain
% EGFP positive cells
0
2
4
6
8
10
12
14
50ng ATP1A1 G3 + ssODN RD
500ng ATP1A1 G3 + ssODN RD

FIG. 20

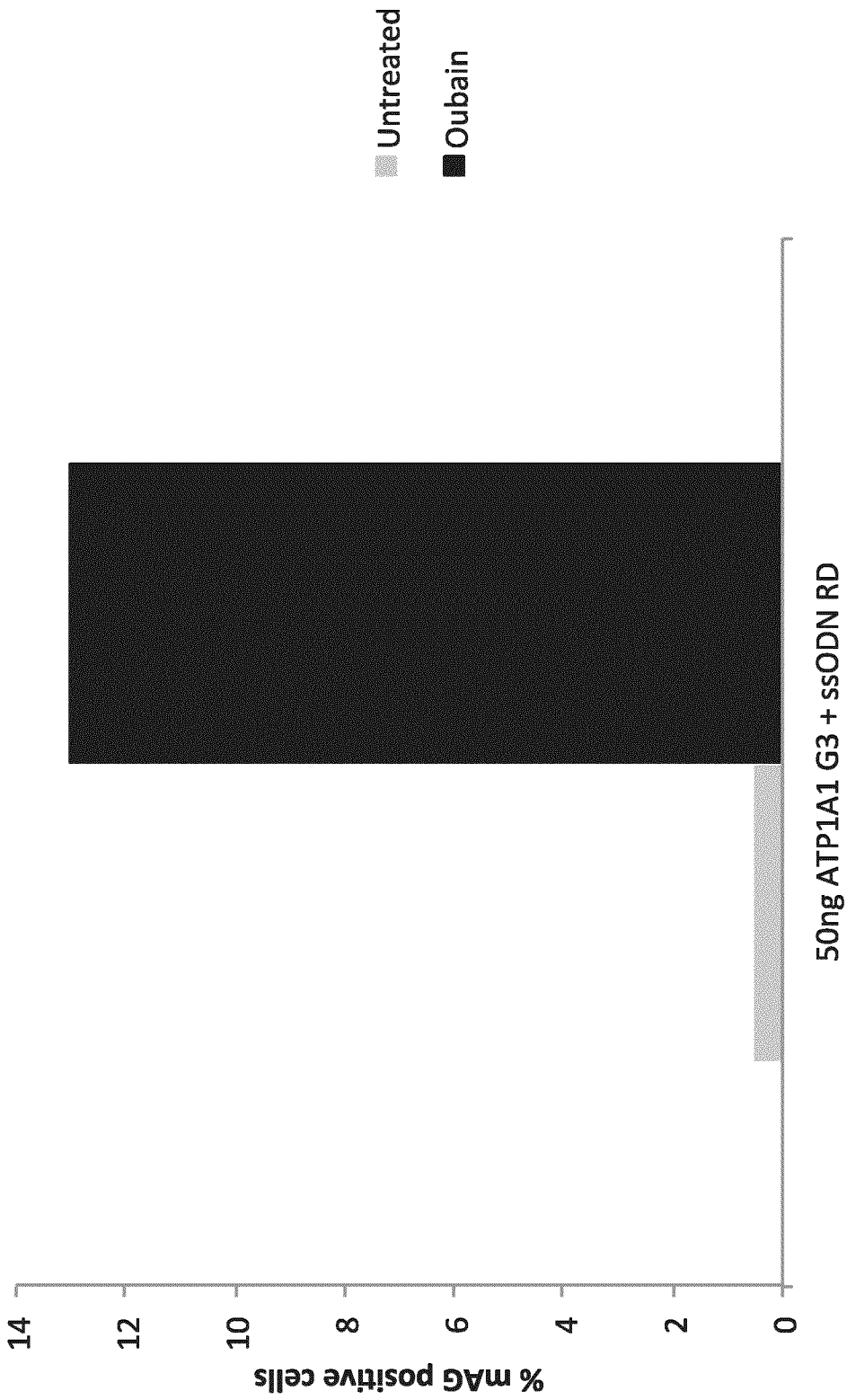

Untreated
Ouabain
% mAG positive cells
0
2
4
6
8
10
12
14
50ng ATP1A1 G3 + ssODN RD

FIG. 21A

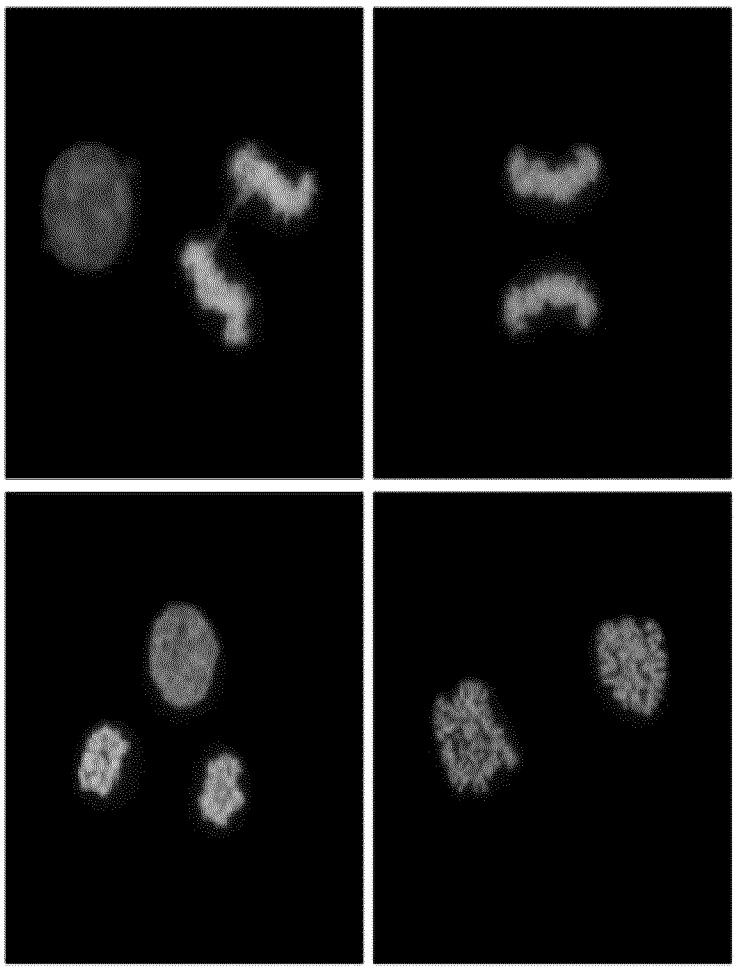


FIG. 21B

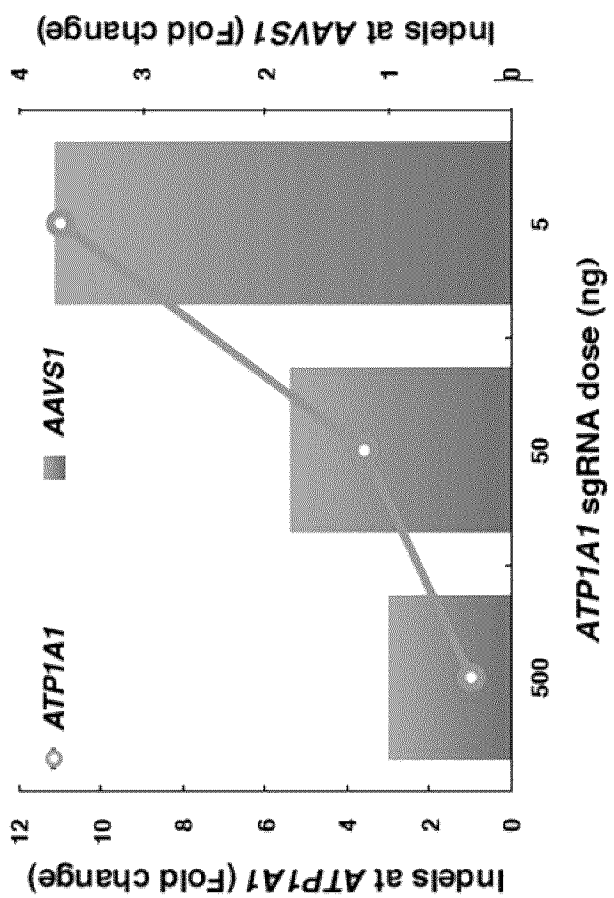

Indels at ATP1A1 (Fold change)
0
2
4
6
8
10
12
ATP1A1
AAVS1
500
50
5
ATP1A1 sgRNA dose (ng)
Indels at AAVS1 (Fold change)
0
1
2
3
4

FIG. 22A

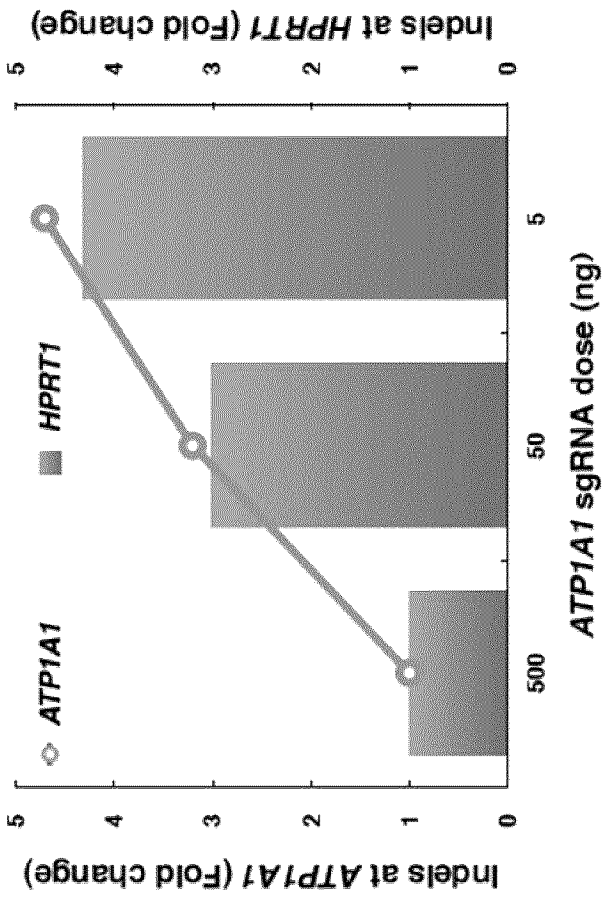

Indels at HPRT1 (Fold change)
0
1
2
3
4
5
ATP1A1
HPRT1
500
50
5
ATP1A1 sgRNA dose (ng)
Indels at ATP1A1 (Fold change)
0
1
2
3
4
5

FIG. 22B

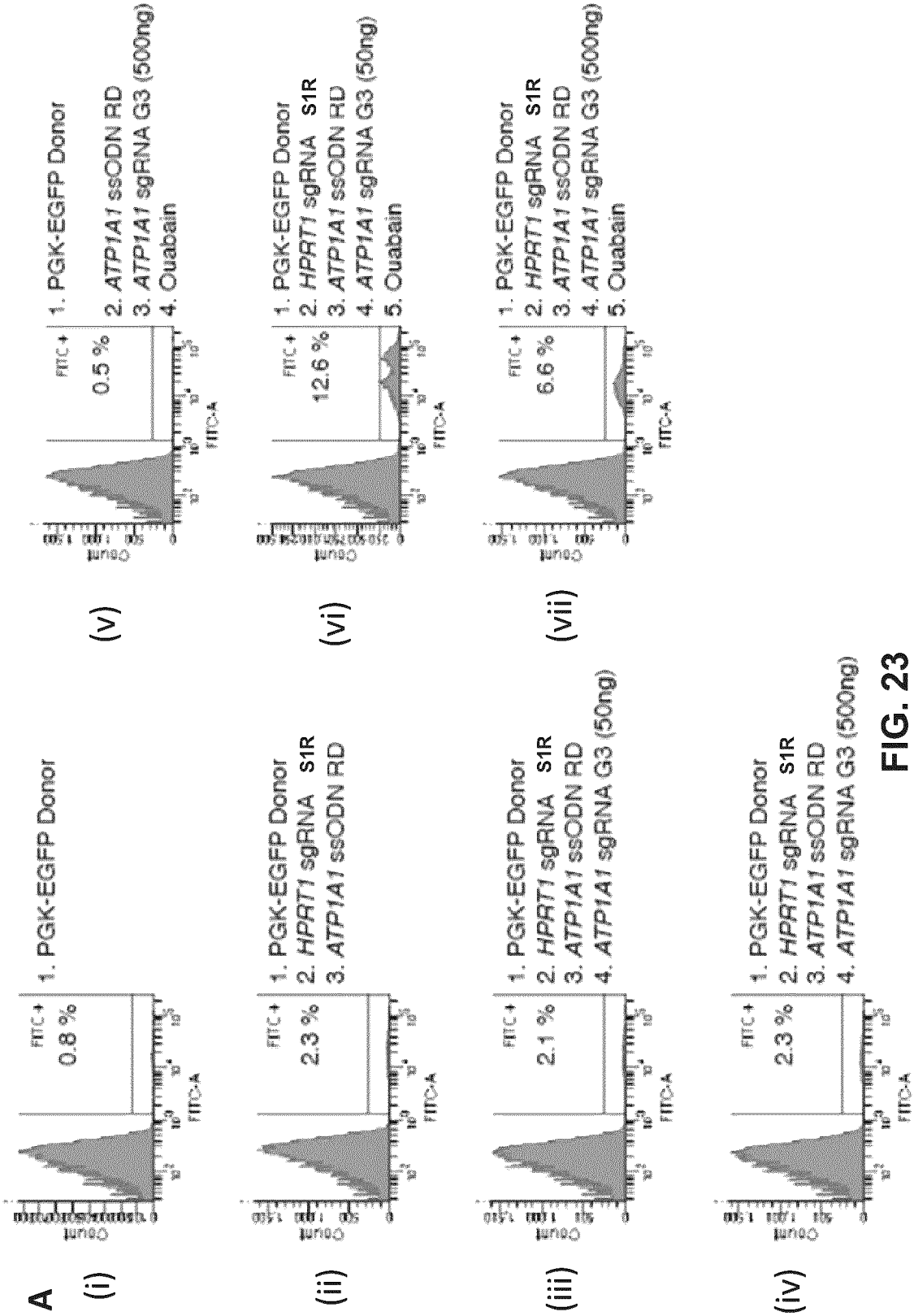

A
(i)
1. PGK-EGFP Donor
FITC +
0.8 %
Count
FITC-A
(ii)
1. PGK-EGFP Donor
2. HPRT1 sgRNA S1R
3. ATP1A1 ssODN RD
FITC +
2.3 %
Count
FITC-A
(iii)
1. PGK-EGFP Donor
2. HPRT1 sgRNA S1R
3. ATP1A1 ssODN RD
4. ATP1A1 sgRNA G3 (50ng)
FITC +
2.1 %
Count
FITC-A
(iv)
1. PGK-EGFP Donor
2. HPRT1 sgRNA S1R
3. ATP1A1 ssODN RD
4. ATP1A1 sgRNA G3 (500ng)
FITC +
2.3 %
Count
FITC-A
(v)
1. PGK-EGFP Donor
2. ATP1A1 ssODN RD
3. ATP1A1 sgRNA G3 (500ng)
4. Ouabain
FITC +
0.5 %
Count
FITC-A
(vi)
1. PGK-EGFP Donor
2. HPRT1 sgRNA S1R
3. ATP1A1 ssODN RD
4. ATP1A1 sgRNA G3 (50ng)
5. Ouabain
FITC +
12.6 %
Count
FITC-A
(vii)
1. PGK-EGFP Donor
2. HPRT1 sgRNA S1R
3. ATP1A1 ssODN RD
4. ATP1A1 sgRNA G3 (500ng)
5. Ouabain
FITC +
6.6 %
Count
FITC-A

FIG. 23

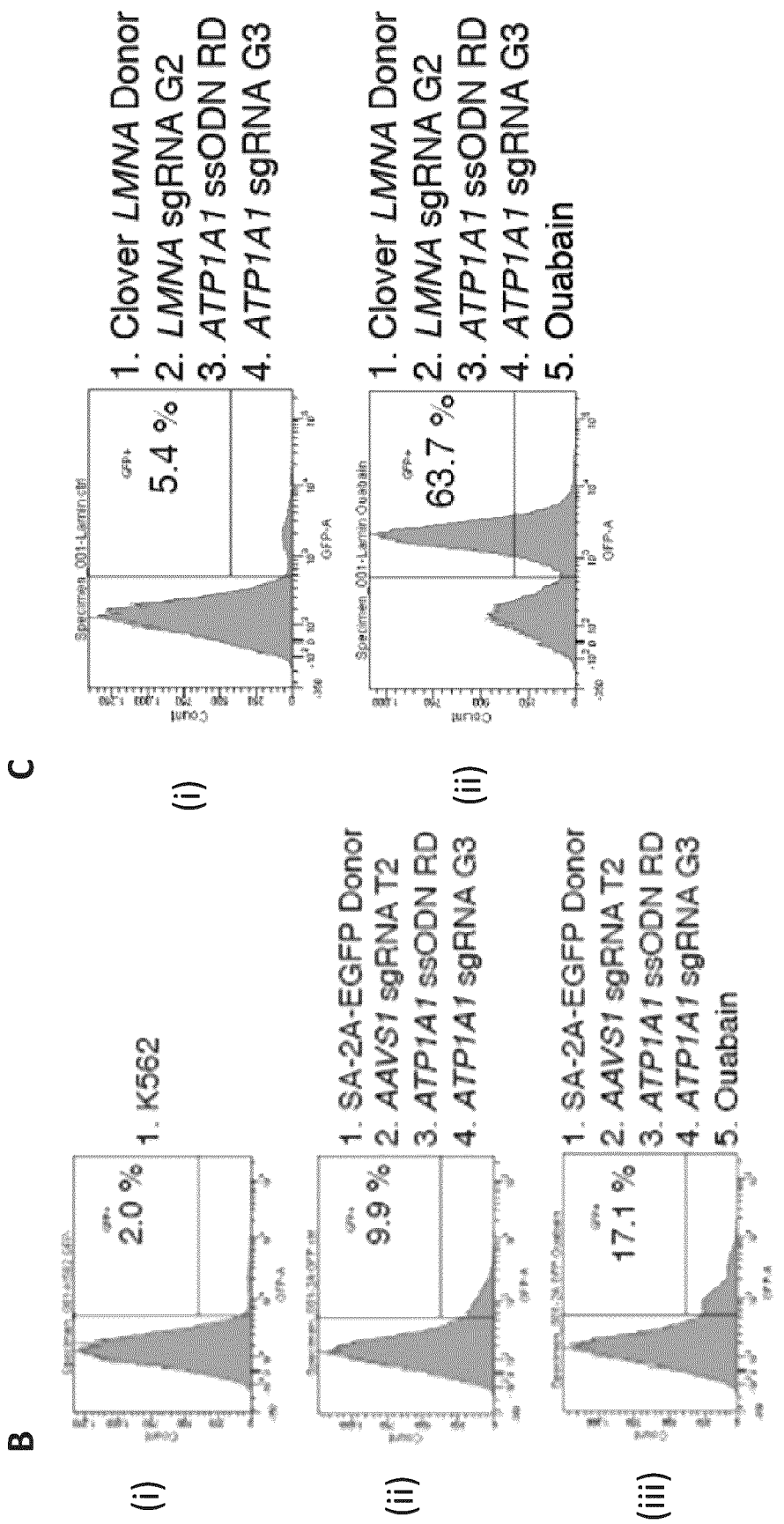

B
(i)
2.0 %
1. K562
(ii)
9.9 %
1. SA-2A-EGFP Donor
2. AAVS1 sgRNA T2
3. ATP1A1 ssODN RD
4. ATP1A1 sgRNA G3
(iii)
17.1 %
1. SA-2A-EGFP Donor
2. AAVS1 sgRNA T2
3. ATP1A1 ssODN RD
4. ATP1A1 sgRNA G3
5. Ouabain
C
(i)
5.4 %
1. Clover LMNA Donor
2. LMNA sgRNA G2
3. ATP1A1 ssODN RD
4. ATP1A1 sgRNA G3
(ii)
63.7 %
1. Clover LMNA Donor
2. LMNA sgRNA G2
3. ATP1A1 ssODN RD
4. ATP1A1 sgRNA G3
5. Ouabain

FIG. 23

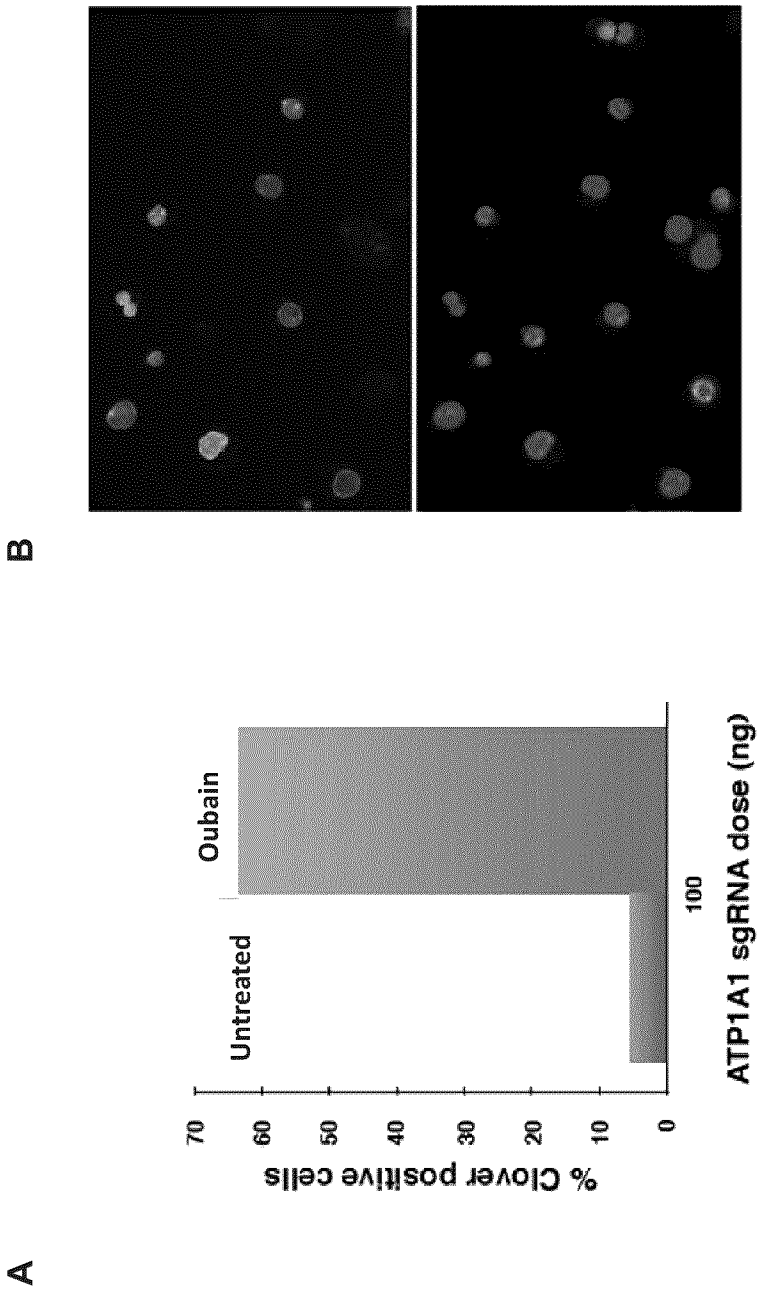

A
B
% Clover positive cells
70
60
50
40
30
20
10
0
Untreated
Oubain
100
ATP1A1 sgRNA dose (ng)

FIG. 24

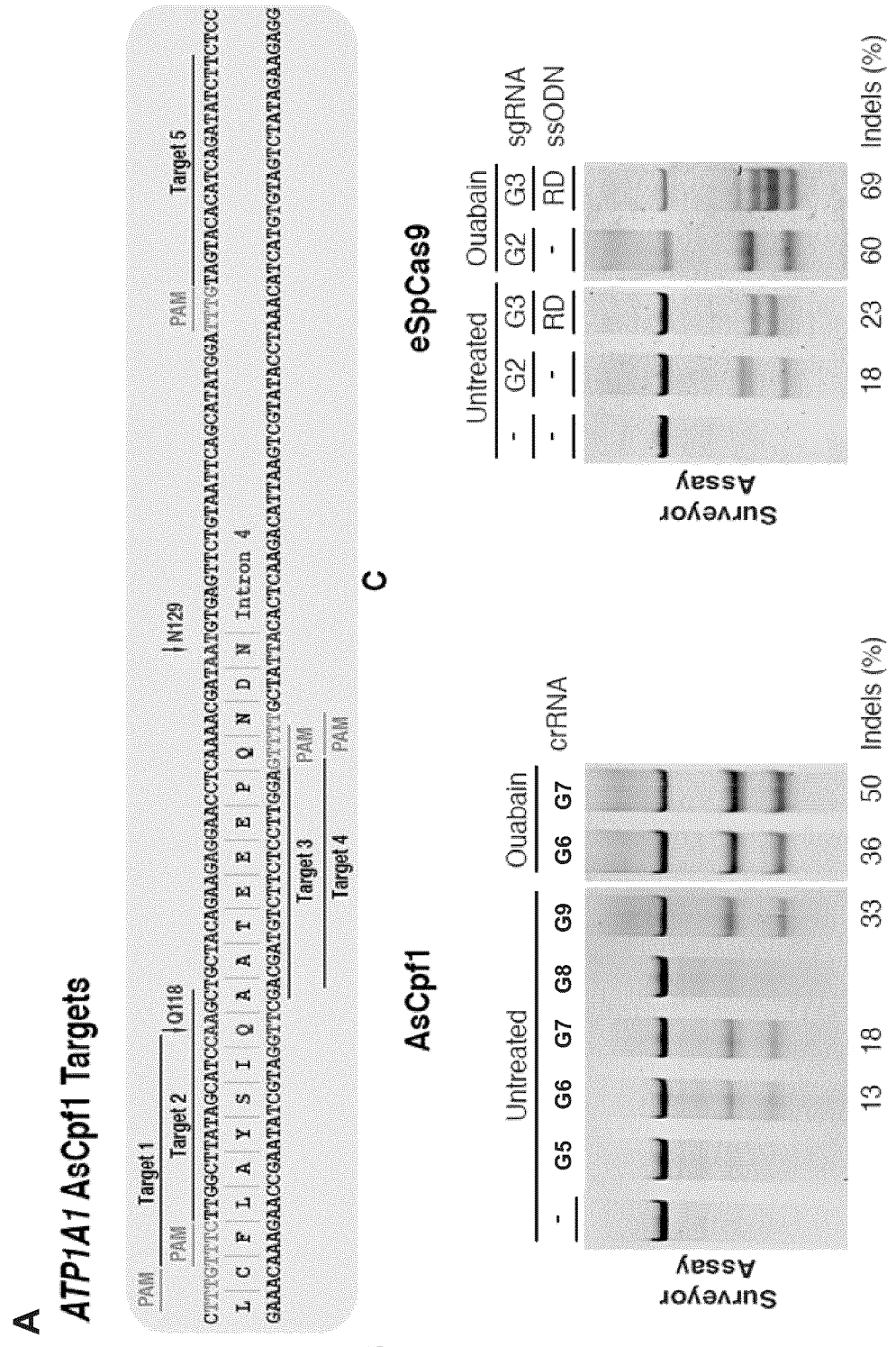

A
ATP1A1 AsCpf1 Targets
PAM
Target 1
Target 2
Q118
N129
PAM
Target 5
CTTTGTTTCTTGGCTTATAGCATCCAAGCTGCTACAGAAGAGGAACCTCAAAACGATAATGTGAGTTCTGTAATTCAGCATATGGATTTTGTAGTACACATCAGATATCTTCTCC
L C F L A Y S I Q A A T E E E P Q N D N Intron 4
GAAACAAAGAACCGAATATCGTAGGTTCGACGATGTCTTCTCCTTGGAGTTTTTGCTATTACACTCAAGACATTAAGTCGTATACCTAAACATCATGTGTAGTCTATAGAAGAGG
Target 3
PAM
Target 4
PAM
B
AsCpf1
Untreated
Ouabain
crRNA
- G5 G6 G7 G8 G9 G6 G7
Surveyor Assay
13 18 33 36 50
Indels (%)
C
eSpCas9
Untreated
Ouabain
sgRNA
- G2 G3 G2 G3
ssODN
- - RD - RD
Surveyor Assay
18 23 60 69
Indels (%)

FIG. 25

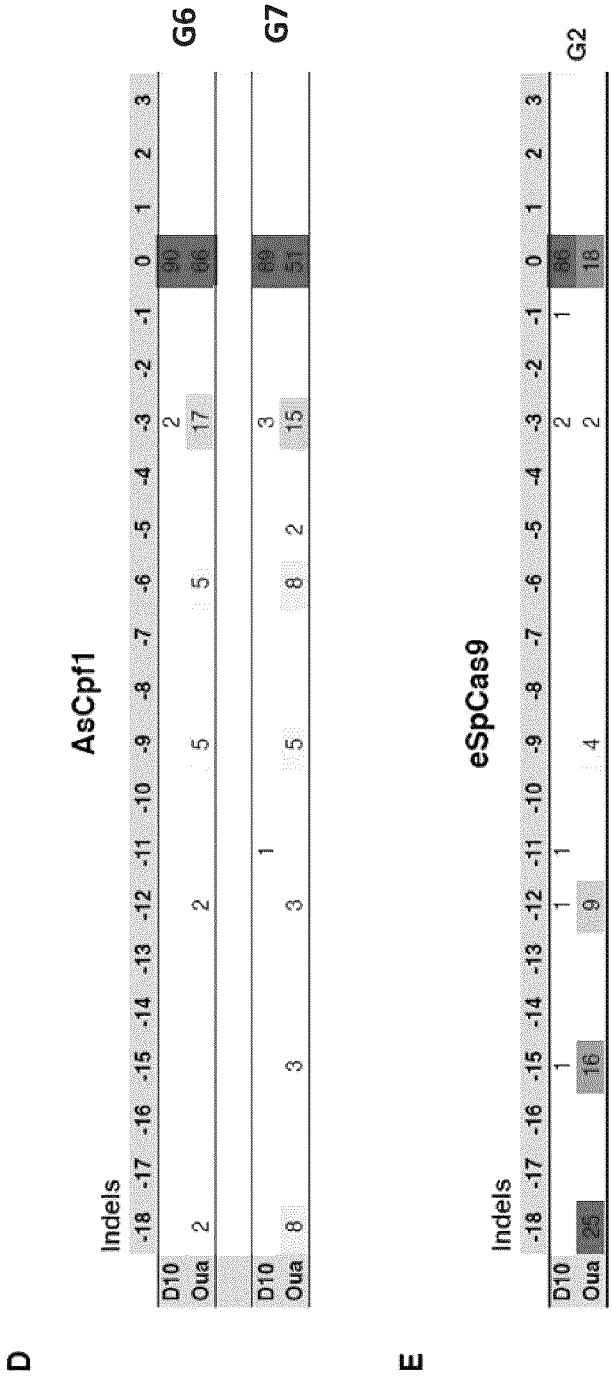

D
AsCpf1
Indels
-18 -17 -16 -15 -14 -13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3
D10 2 2 5 5 2 17 90 66
Oua 8 3 3 1 5 8 2 3 15 89 51
G6
G7
E
eSpCas9
Indels
-18 -17 -16 -15 -14 -13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3
D10 1 1 1 4 2 2 1 86 18
Oua 25 16 9
G2

FIG. 25

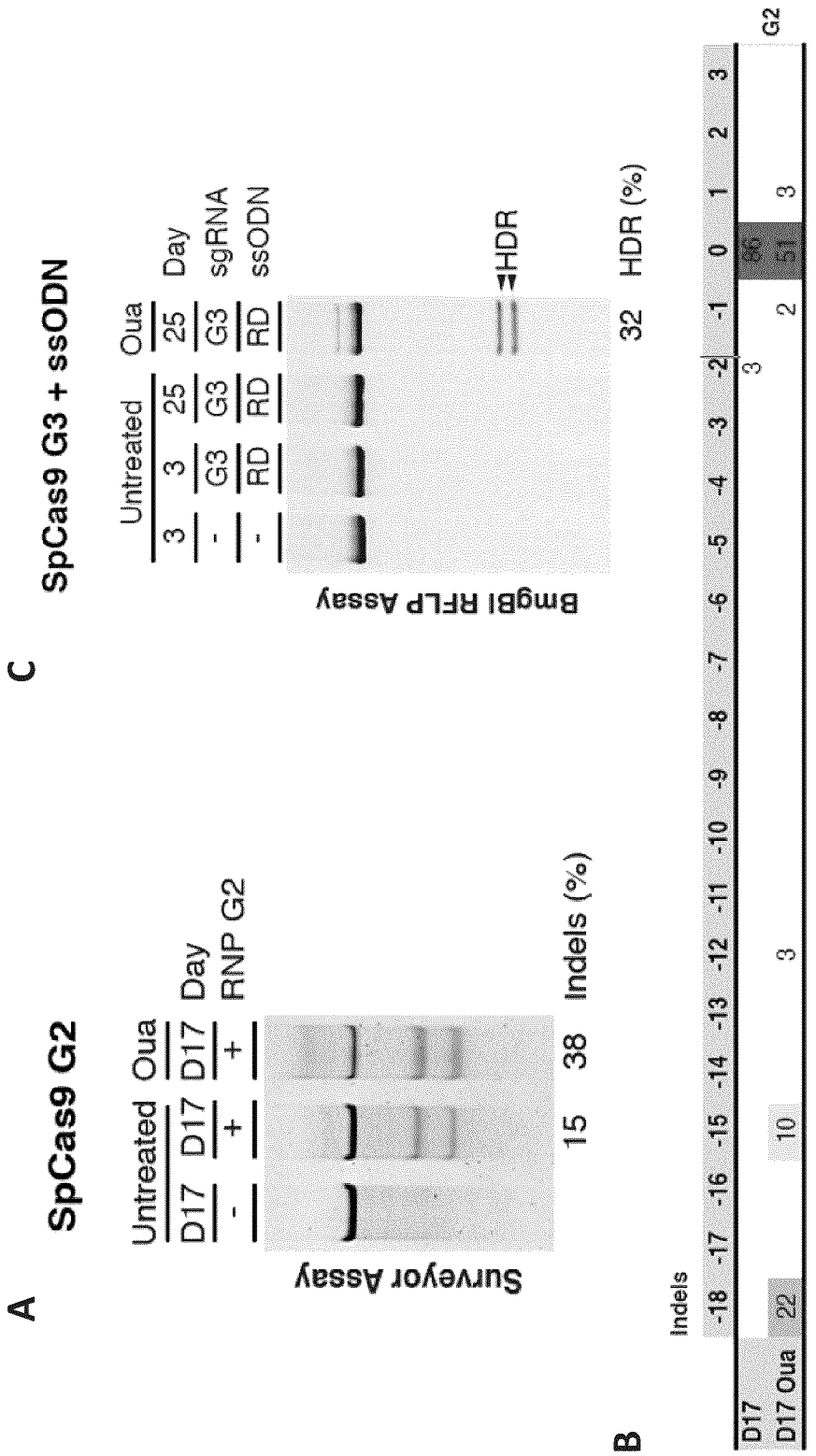

A
SpCas9 G2
Untreated
D17
-
Oua
D17
+
Day
RNP
G2
Surveyor Assay
15
38
Indels (%)
B
Indels
D17
D17 Oua
22
10
3
C
SpCas9 G3 + ssODN
Untreated
3
-
3
G3
RD
25
G3
RD
Oua
25
G3
RD
Day
sgRNA
ssODN
BmgBI RFLP Assay
HDR
32
HDR (%)
G2

FIG. 26

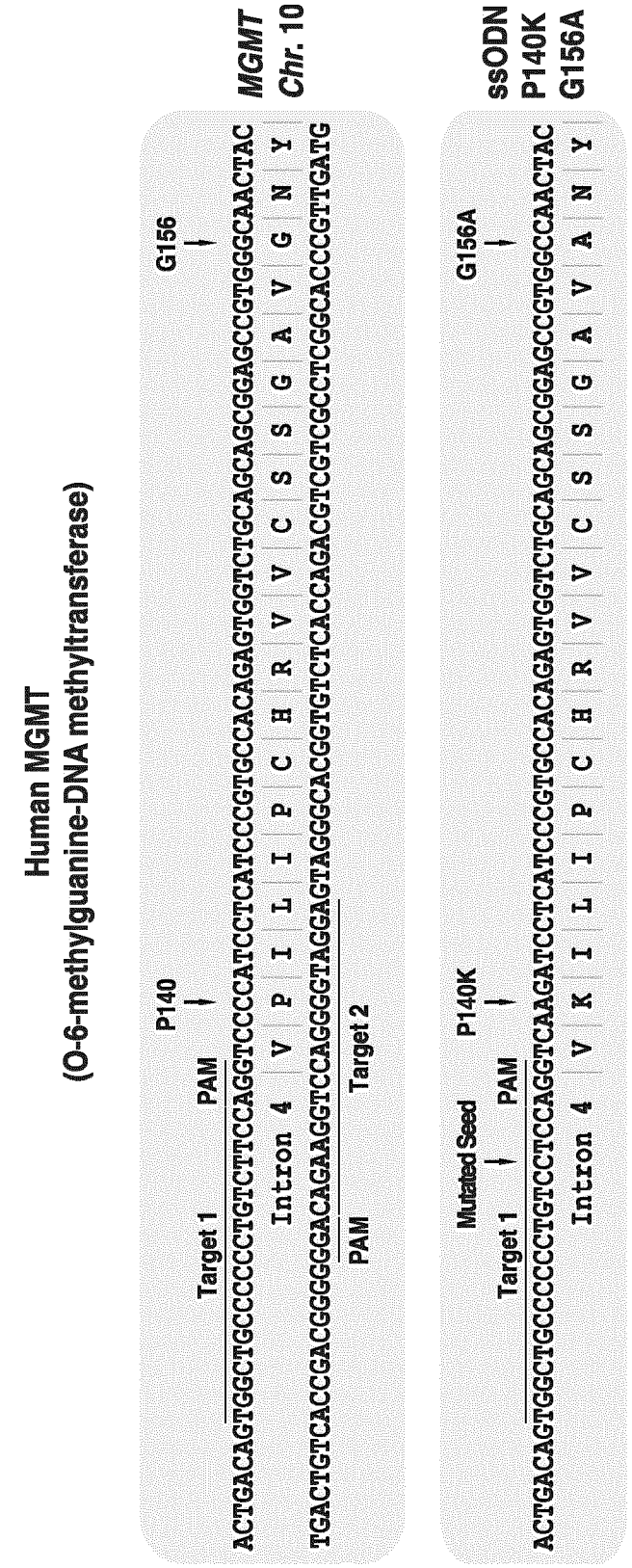

Human MGMT
(O-6-methylguanine-DNA methyltransferase)
P140
G156
Target 1
PAM
ACTGACAGTGGCTGCCCCCCTGTCTTCCAGGTCCCCATCCTCATCCCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGGGCAACTAC
Intron 4 V P I L I P C H R V V C S S G A V G N Y
TGACTGTCACCGACGGGGGGACAGAAGGTCCAGGGGTAGGAGTAGGGCACGGTGTCTCACCAGACGTCGTCGCCTCGGCACCCGTTGATG
PAM
Target 2
MGMT
Chr. 10
Mutated Seed
P140K
G156A
ssODN
P140K
G156A

FIG. 27

sgRNA	Exon 5 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T46	TCAGCCGTTGTAATCATAAC	TGG	90
T47	AAAAGACAAGTTAGGGCCCG	TGG	91
T48	CGTGGCTCTACACAATAATT	AGG	92
T49	GACAAAGCGTCTGGTACCTG	AGG	93
T50	GGGGCTGTCAAGTCATCCAC	TGG	94
T51	GTGGCTCTACACAATAATTA	GGG	95
T52	TAATTATTGTGTAGAGCCAC	GGG	96
T53	GCAACCAGTTATGATTACAA	CGG	97
T54	TTAACTTCACACTAATTCAC	AGG	98
T55	CTGAGGGACCATGTTTTTGA	AGG	99
T56	CTTCAAAAACATGGTCCCTC	AGG	100
T57	AATTCACAGGCTGACATGCT	TGG	101
T58	ACAAAGCGTCTGGTACCTGA	GGG	102
T59	CGGCTGATAGCACCACACCC	AGG	103
T60	GGCTGTCAAGTCATCCACTG	GGG	104
T61	GTCAAGTCATCCACTGGGGA	AGG	105
T62	ACACCCAGGTACAGCTGGAA	GGG	106
T63	TTCACAGGCTGACATGCTTG	GGG	107
T64	TGTGAATTAGTGTGAAGTTA	AGG	108
T65	GGGCTGTCAAGTCATCCACT	GGG	109
T66	ATTCACAGGCTGACATGCTT	GGG	110
T67	TTTCCCTTCCAGCTGTACCT	GGG	111
T68	GACGCTTTGTCCTTCCCCAG	TGG	112
T69	CTGGGGAAGGACAAAGCGTC	TGG	113
T70	GCACCACACCCAGGTACAGC	TGG	114
T71	CACACCCAGGTACAGCTGGA	AGG	115
T72	TTTTCCCTTCCAGCTGTACC	TGG	116
T73	CATGGAATCCTTCAAAAACA	TGG	117
T74	CTTCCAGCTGTACCTGGGTG	TGG	118
T75	ACTTTTAGCTTCTTGATAGT	AGG	119
T76	AGCTAAAAGTTCAAAGATCA	TGG	120
T77	GGAAGGGAAAAGACAAGTTA	GGG	121
T78	TGGAAGGGAAAAGACAAGTT	AGG	122
T79	CTAATTATTGTGTAGAGCCA	CGG	123

FIG. 28A

sgRNA	Exon 8 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T80	TCTTTTTAAGGCACCGCACG	TGG	124
T81	ATCGTAGCCAATGTGCCGGA	AGG	125
T82	TGGCTACGATGATACCGATG	AGG	126
T83	GTAGACAACAATACCACGTG	CGG	127
T84	CGTGGTATTGTTGTCTACAC	TGG	128
T85	TATCATCGTAGCCAATGTGC	CGG	129
T86	CTGGCCACTGTCACGGTAAG	AGG	130
T87	CGTGCGGTGCCTTAAAAAGA	AGG	131
T88	CAATGTGCCGGAAGGTTTGC	TGG	132
T89	TATCCACATCATCACGGGTG	TGG	133
T90	CAGCAAACCTTCCGGCACAT	TGG	134
T103	CCTTGAGTACACCTGGCTTG	AGG	147
T104	AGTTGGCAATACCTTAGAAT	AGG	148
T105	TAAAATCCGTGGCTTCCTTC	AGG	149
T106	CTCTCATCCTTGAGTACACC	TGG	150
T107	TGCGGTGCCTTAAAAAGAAG	GGG	151
T108	ACAGTGGCCAGCAAACCTTC	CGG	152
T109	CCTCAAGCCAGGTGTACTCA	AGG	153
T110	GTGCGGTGCCTTAAAAAGAA	GGG	154
T111	GTGCAAGCTGATCTGAGTCA	GGG	155
T112	AGAATTGCCACACTTGCTTC	TGG	156
T113	TGGTATTGTTGTCTACACTG	GGG	157
T114	GAATTGCCACACTTGCTTCT	GGG	158
T115	TGCCACACTTGCTTCTGGGC	TGG	159
T116	AATACCTTAGAATAGGATAG	CGG	160
T117	GTTCAATTTCTGCAGCAATG	GGG	161
T118	AGGTTTGCTGGCCACTGTCA	CGG	162
T119	TCACGGTAAGAGGCAGGTGA	TGG	163
T120	ACATTTTATCCACATCATCA	CGG	164
T121	TTTCTGCAGCAATGGGGGTC	TGG	165
T122	ATCTAACCTGAAGGAAGCCA	CGG	166
T123	GAGGCTGTCATCTTCCTCAT	CGG	167
T124	TGTTCAATTTCTGCAGCAAT	GGG	168
T125	TTCAATTTCTGCAGCAATGG	GGG	169
T126	GGTGCCTTAAAAAGAAGGGG	AGG	170
T127	ACACTTGCTTCTGGGCTGGA	AGG	171
T128	GAAGATGACAGCCTCAAGCC	AGG	172
T129	GGATGAAGAAAGACACACCC	AGG	173
T130	TTCCAGCCCAGAAGCAAGTG	TGG	174
T131	ACGGGTGTGGCTGTGTTCCT	GGG	175
T132	CACGGGTGTGGCTGTGTTCC	TGG	176
T133	CTTGCTTCTGGGCTGGAAGG	AGG	177
T134	CATTTTATCCACATCATCAC	GGG	178
T136	TCTCCCTCCCCTTCTTTTTA	AGG	180
T137	GTGCCTTAAAAAGAAGGGGA	GGG	181
T138	ATGTTCAATTTCTGCAGCAA	TGG	182

FIG. 28B

sgRNA	Exon 9 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T139	AGCCACCTGCTGACTTCGCT	TGG	183
T140	TTACTGCCAAACGCATGGCA	AGG	184
T141	GATTATTACTTCACTCAACC	AGG	185
T142	GACACTTACTGCCAAACGCA	TGG	186
T143	GTTTTTCCTTGCCATGCGTT	TGG	187
T144	GACTCAGAACCGGATGACAG	TGG	188
T145	ATTACTTCACTCAACCAGGC	TGG	189
T146	GATGGTGGACGTGGACCCCA	AGG	190
T147	TCTCTGTCGTATCAGCTTCA	TGG	191
T148	GCTGAGGTGTGTGGTAACCC	TGG	192
T149	AACCAAGCGAAGTCAGCAGG	TGG	193
T150	ACCATCTGCTCTGATAAAAC	TGG	194
T151	AAAAACCAAGCGAAGTCAGC	AGG	195
T152	TTTGATTGTCAAACCACATG	TGG	196
T153	GGATGACAGTGGCCCACATG	TGG	197
T154	TTAGAAGCTGTGGAGACCTT	GGG	198
T155	TAGAAGCTGTGGAGACCTTG	GGG	199
T156	CTGGAACTCTGACTCAGAAC	CGG	200
T157	ATCTGTATAGAATTCCAGCC	TGG	201
T158	ACATGTGGGCCACTGTCATC	CGG	202
T159	GAATCAGAGTGGTAAGGCCA	GGG	203
T160	TCCAGTTTTATCAGAGCAGA	TGG	204
T161	AGTTTTATCAGAGCAGATGG	TGG	205
T162	ATTCTTGCTATATCCTAATG	TGG	206
T163	ATCAGAGCAGATGGTGGACG	TGG	207
T164	AGAATCAGAGTGGTAAGGCC	AGG	208
T165	GACAGAGAATCAGAGTGGTA	AGG	209
T166	AGTGAAGAACTTAGAAGCTG	TGG	210
T167	CTTAGAAGCTGTGGAGACCT	TGG	211
T168	AGACAAATTTCTTCCACATT	AGG	212
T169	GATACGACAGAGAATCAGAG	TGG	213
T170	GCGAAGTCAGCAGGTGGCTG	AGG	214
T171	AGACAGACCTAAAAAATATT	AGG	215
T172	TTGATTGTCAAACCACATGT	GGG	216
T173	TAATCTTCCTAATATTTTTT	AGG	217
T174	CAGCAGGTGGCTGAGGTGTG	TGG	218

FIG. 28C

sgRNA	Exon 17 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T175	GTCACAGATCGATAGTAGTG	AGG	219
T176	TATCAGGAACGGGGTGATCT	CGG	220
T177	CTCGGGAATGTTACTGGTTA	AGG	221
T178	GGTGATCTCGGGAATGTTAC	TGG	222
T179	GCAAACATTCCACTACCACT	GGG	223
T180	ATCAGGAACGGGGTGATCTC	GGG	224
T187	AACGCCAGCCTCATGGATGC	TGG	231
T188	CACCATCCTCTGCATTGACT	TGG	232
T189	CATTGACTTGGGCACTGACA	TGG	233
T190	CAGTGCCCAAGTCAATGCAG	AGG	234
T191	AGCATCCATGAGGCTGGCGT	TGG	235
T192	ACCATCCTCTGCATTGACTT	GGG	236
T193	TGAGCTGTGTCTTCATTCAC	TGG	237
T194	CTATGCTCCCAGCATCCATG	AGG	238
T195	ACTGGTTAAGGTATAAGCAA	TGG	239
T196	ATACAGAAATGGATGCTAAT	GGG	240
T197	GCCCAAGTCAATGCAGAGGA	TGG	241
T198	AGGATGGTGACAGTCCCCAG	TGG	242
T199	CAATAATAAATATCAGGAAC	GGG	243
T200	GCTCCCAGCATCCATGAGGC	TGG	244
T201	TATACAGAAATGGATGCTAA	TGG	245
T202	AATAATAAATATCAGGAACG	GGG	246
T203	ATCCATTTCTGTATACTTCT	TGG	247
T204	TGTTTGCAATAATAAATATC	AGG	248
T205	ATCCAAGAAGTATACAGAAA	TGG	249
T206	GCAATAATAAATATCAGGAA	CGG	250
T207	TAAAAAATCTTGGTTTTCAT	AGG	251
T208	AAGCTAACAGTAAAAAATCT	TGG	252

FIG. 28D

sgRNA	Exon 18 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T209	AAATCAACCTCAACCTGAGG	TGG	253
T210	CGGCTGATCAGCATGGCCTA	TGG	254
T211	ATGGTGTCCACCTCAGGTTG	AGG	255
T212	GTGGGAAGCTGGCACATCTA	AGG	256
T213	GAATGAGCGGCTGATCAGCA	TGG	257
T214	CAGCTTACCAATCTGCCCAT	AGG	258
T215	TCTCAGCCTGCTCATAAGCC	AGG	259
T216	GGCTGATCAGCATGGCCTAT	GGG	260
T217	CTCAGCCTGCTCATAAGCCA	GGG	261
T218	AGCATGGCCTATGGGCAGAT	TGG	262
T219	CTGAAATCAACCTCAACCTG	AGG	263
T220	CAGACAAACTTGTGAATGAG	CGG	264
T221	ATCTAAGGCATCTGAGGTGA	TGG	265
T222	ATTTCAGAGACTGCAAATCC	AGG	266
T223	CATCTCCCTGGCTTATGAGC	AGG	267
T224	TGGCACATCTAAGGCATCTG	AGG	268
T225	CAGATTGGTAAGCTGCAGCC	TGG	269
T243	CAGGAACCTGGGAAGGGGTG	GGG	287
T244	CCTTCTCCCCACCCCTTCCC	AGG	288
T245	CCAGGGAGATGGCAGGAACC	TGG	289
T246	CCTGGGAAGGGGTGGGGAGA	AGG	290
T247	GGCAGGAACCTGGGAAGGGG	TGG	291
T248	CAGGGAGATGGCAGGAACCT	GGG	292
T249	GATGGCAGGAACCTGGGAAG	GGG	293

FIG. 28E

sgRNA	Exons 19-20 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T250	AGCCTTCTTCGTCAGTATCG	TGG	294
T251	TGACCGCTGGATCAACGATG	TGG	295
T252	CACCACGATACTGACGAAGA	AGG	296
T253	AGTCCACTCGGAGGCCCAAC	AGG	297
T254	TGTTGGGCCTCCGAGTGGAC	TGG	298
T255	GTGTGTGAGGTTCTAGTAAT	TGG	299
T256	CGGAGGCCCAACAGGTGAAT	TGG	300
T257	ACGATGTGGAAGACAGCTAC	GGG	301
T258	GGTGGTGCAGTGGGCCGACT	TGG	302
T259	GGCCTGGATCATTCCTAAGT	TGG	303
T260	GAGGAATTCGGTCTTCCAGC	AGG	304
T261	CTTCTTCGTCAGTATCGTGG	TGG	305
T262	CTTCCACATCGTTGATCCAG	CGG	306
T263	GTTGGGCCTCCGAGTGGACT	GGG	307
T264	GGCCCAACAGGTGAATTGGG	AGG	308
T265	TCACCTGTTGGGCCTCCGAG	TGG	309
T266	TCCCAACTTAGGAATGATCC	AGG	310
T267	CTGGAAGACCGAATTCCTCC	TGG	311
T268	AGGAATTCGGTCTTCCAGCA	GGG	312
T269	GGAGGCCCAACAGGTGAATT	GGG	313
T270	GCCTGGATCATTCCTAAGTT	GGG	314
T271	GAGTGGACTGGGATGACCGC	TGG	315
T272	GGTCATCCCAGTCCACTCGG	AGG	316
T273	AGACAGCTACGGGCAGCAGT	GGG	317
T274	CTATGAGCAGAGGAAAATCG	TGG	318
T275	TTCCTCCCAATTCACCTGTT	GGG	319
T276	GAGGAGCTCTGACACTGGTC	AGG	320
T277	TGTGTGAGGTTCTAGTAATT	GGG	321
T278	CAGTATCGTGGTGGTGCAGT	GGG	322
T279	AAGACAGCTACGGGCAGCAG	TGG	323
T280	CTTCCTCCCAATTCACCTGT	TGG	324
T281	CATTTCAGACCTATGAGCAG	AGG	325
T282	TCAGTATCGTGGTGGTGCAG	TGG	326
T283	AGCGGTCATCCCAGTCCACT	CGG	327
T284	TTAGGAATGATCCAGGCCCT	GGG	328
T285	AACGATGTGGAAGACAGCTA	CGG	329
T286	CACGATTTTCCTCTGCTCAT	AGG	330
T287	GTAAGTAAAGAAGCCTCCCA	GGG	331
T288	ACTTGGTCATCTGTAAGACC	AGG	332
T289	TACTGACGAAGAAGGCTGTG	TGG	333
T290	CAGGCCTGGGAGAGCACATA	TGG	334
T291	TGGTCATCTGTAAGACCAGG	AGG	335
T292	TAATCAGCTGAAGGCTATAC	AGG	336
T293	TTACTAGAACCTCACACACA	AGG	337
T294	CTGCTGCAGTAATCAGCTGA	AGG	338
T295	TACTAGAACCTCACACACAA	GGG	339

FIG. 28F

sgRNA	Exons 19-20 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T296	TACAGGTACCCTTGTGTGTG	AGG	340
T297	GGCTTGAGGAGCTCTGACAC	TGG	341
T298	CTTAGGAATGATCCAGGCCC	TGG	342
T299	TGGCTCTTCACCTGCATGGC	TGG	343
T300	AGTAAGTAAAGAAGCCTCCC	AGG	344
T301	CTGTAAGACCAGGAGGAATT	CGG	345
T302	CTGATTACTGCAGCAGGCCT	GGG	346
T303	CTTCAGCTGATTACTGCAGC	AGG	347
T304	TCTGACACTGGTCAGGTCAG	AGG	348
T305	GCTGATTACTGCAGCAGGCC	TGG	349
T306	ACGGGCAGCAGTGGGTGAGT	GGG	350
T307	GGTCTTACAGATGACCAAGT	CGG	351
T308	TTGGGTTGTTTTCCCAACTT	AGG	352
T309	CTCTTCACCTGCATGGCTGG	TGG	353
T310	GCTCTGGCTCTTCACCTGCA	TGG	354
T311	CCTCAAGCCCCAGAAGACTG	AGG	355
T312	GGAATGATCCAGGCCCTGGG	AGG	356
T313	AGTGTGGCCTCAGTCTTCTG	GGG	357
T314	GACGAAGAAGGCTGTGTGGC	AGG	358
T315	TACGGGCAGCAGTGGGTGAG	TGG	359
T316	CTTTACTTACTTTGTGATTC	TGG	360
T317	CACACTGCCACCAGCCATGC	AGG	361
T318	TTTGTGATTCTGGCTGAGAA	CGG	362
T319	CCTCAGTCTTCTGGGGCTTG	AGG	363
T320	TAAAGAAGCCTCCCAGGGCC	TGG	364
T321	AGAAGACAGAAGATGAGCTC	TGG	365
T322	GCAGTGTGGCCTCAGTCTTC	TGG	366
T323	CTGCATGGCTGGTGGCAGTG	TGG	367
T324	CAGTGTGGCCTCAGTCTTCT	GGG	368
T325	AACATATGACACATATAGCC	AGG	369
T326	TTACTTACTTCATCCCCTGC	TGG	370
T327	CTTGGCTCAAAGAAGGGGAC	TGG	371
T328	TGAAGGACATGTCAAGGCCT	TGG	372
T329	CAGTCCCCTTCTTTGAGCCA	AGG	373
T330	AAGTAATGAAGGACATGTCA	AGG	374
T331	TCAAGGCCTTGGCTCAAAGA	AGG	375
T332	AAGGCCTTGGCTCAAAGAAG	GGG	376
T333	CAAGGCCTTGGCTCAAAGAA	GGG	377
T334	ATATGTTTAGTTTTTGGCAA	AGG	378
T335	GGGGATGAAGTAAGTAATGA	AGG	379
T336	TGTGTCATATGTTTAGTTTT	TGG	380
T337	TATGTTTAGTTTTTGGCAAA	GGG	381

FIG. 28G

sgRNA	Exon 21 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T338	CCAGAGGATCAACTTACTTG	AGG	382
T339	ATACTTTTTGGAGCTACCAG	AGG	383
T340	CCTCAAGTAAGTTGATCCTC	TGG	384
T341	GCTTTTCCCCCCATTACTTG	TGG	385
T342	AGAGGATCAACTTACTTGAG	GGG	386
T343	ACTATTACCACAAGTAATGG	GGG	387
T344	CTAAGAGCAACACCCATTCC	AGG	388
T345	CAGAGGATCAACTTACTTGA	GGG	389
T346	CAACTATTACCACAAGTAAT	GGG	390
T347	TAAGAGCAACACCCATTCCA	GGG	391
T348	CTGGAATGGGTGTTGCTCTT	AGG	392
T349	ACAACTATTACCACAAGTAA	TGG	393
T350	AGGAACAAGATCTTGATATT	TGG	394
T351	AACTATTACCACAAGTAATG	GGG	395
T352	AACACCCATTCCAGGGCAGT	AGG	396
T353	CCTTTCCTACTGCCCTGGAA	TGG	397
T354	CTTTCCTACTGCCCTGGAAT	GGG	398
T355	CCTCTTTGAAGAGACAGCCC	TGG	399
T356	CCAGGGCTGTCTCTTCAAAG	AGG	400
T357	TGTTCCTAAAATCAAAGTAG	GGG	401
T358	TCTAATGCTATGCATTAAAA	TGG	402
T359	CCATTCCAGGGCAGTAGGAA	AGG	403
T360	TTGTTCCTAAAATCAAAGTA	GGG	404
T361	GCTTTCCTTTCCTACTGCCC	TGG	405
T362	AACTTAAAAGGTGCAAATAA	AGG	406
T363	AAAGTATGAAATAGTATGTG	TGG	407
T364	CTATTACCACAAGTAATGGG	GGG	408
T365	ACATACTATTTCATACTTTT	TGG	409
T366	TTCTCCCCTACTTTGATTTT	AGG	410
T367	TAGGGGAGAAAAAACTTAAA	AGG	411
T368	GTAGGAAAGGAAAGCAGCCA	GGG	412
T369	CTTGTTCCTAAAATCAAAGT	AGG	413
T370	AGTAGGAAAGGAAAGCAGCC	AGG	414

FIG. 28H

sgRNA	Exon 22 sgRNA target site (without PAM)	PAM	SEQ ID NO. (without PAM)
T371	GACGCCCTGGCGGTAATTAT	GGG	415
T372	GAATGCCCATAATTACCGCC	AGG	416
T373	CGACGCCCTGGCGGTAATTA	TGG	417
T374	CTCATCATCAGGCGACGCCC	TGG	418
T375	AATGCCCATAATTACCGCCA	GGG	419
T376	GGGAGGGGCTATAGTTCTAT	TGG	420
T377	ATCATCAGGCGACGCCCTGG	CGG	421
T378	GAGGGAGCCTAGACACCTTG	AGG	422
T379	AGGGAGCCTAGACACCTTGA	GGG	423
T380	ACTCTATTTCCAGACCTACC	TGG	424
T381	ACAGAACCACCAGGTAGGTC	TGG	425
T382	GGCATTCTGACTTTGGTTGG	AGG	426
T383	TTGTTTCCCTCAAGGTGTCT	AGG	427
T384	CTATTTCCAGACCTACCTGG	TGG	428
T385	ATGGGCATTCTGACTTTGGT	TGG	429
T386	AGGTGTCTAGGCTCCCTCAG	TGG	430
T387	AAGGCACAGAACCACCAGGT	AGG	431
T388	AAGTCAGAAAACTCATCATC	AGG	432
T389	GCTGATTATTTCATTGTCAC	TGG	433
T390	GTAGGTCTGGAAATAGAGTA	AGG	434
T391	ATTCTGACTTTGGTTGGAGG	AGG	435
T392	GGGGAAGGCACAGAACCACC	AGG	436
T393	TACGAAGATGAGAAGAGAGT	AGG	437
T394	AATTATGGGCATTCTGACTT	TGG	438
T395	GTGTTCGTGTGCATATAAAT	TGG	439
T396	ACTTTGGTTGGAGGAGGAGT	GGG	440
T397	ACGAAGATGAGAAGAGAGTA	GGG	441
T398	CGAAGATGAGAAGAGAGTAG	GGG	442
T399	GACTTTGGTTGGAGGAGGAG	TGG	443
T400	TTGGTTTTTTGTTTCCCTCA	AGG	444
T401	TTGGTTGGAGGAGGAGTGGG	AGG	445
T402	TGGTTGGAGGAGGAGTGGGA	GGG	446
T403	GATGAGAAGAGAGTAGGGGA	AGG	447
T404	GGTTGGAGGAGGAGTGGGAG	GGG	448

FIG. 28I

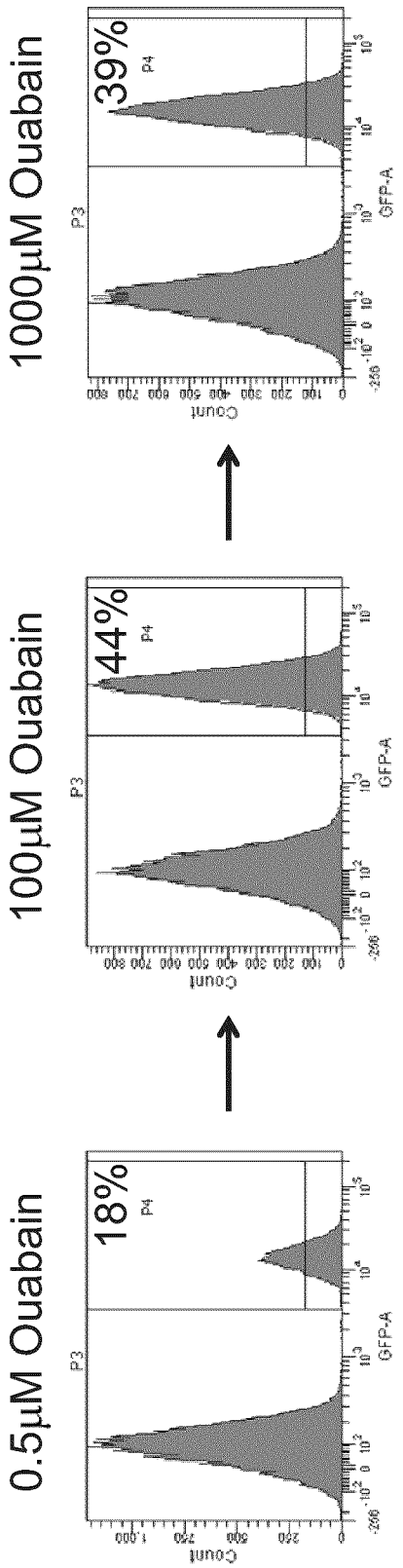

0.5µM Ouabain
18%
P4
P3
GFP-A
Count
100µM Ouabain
44%
P4
P3
GFP-A
Count
1000µM Ouabain
39%
P4
P3
GFP-A
Count

FIG. 29

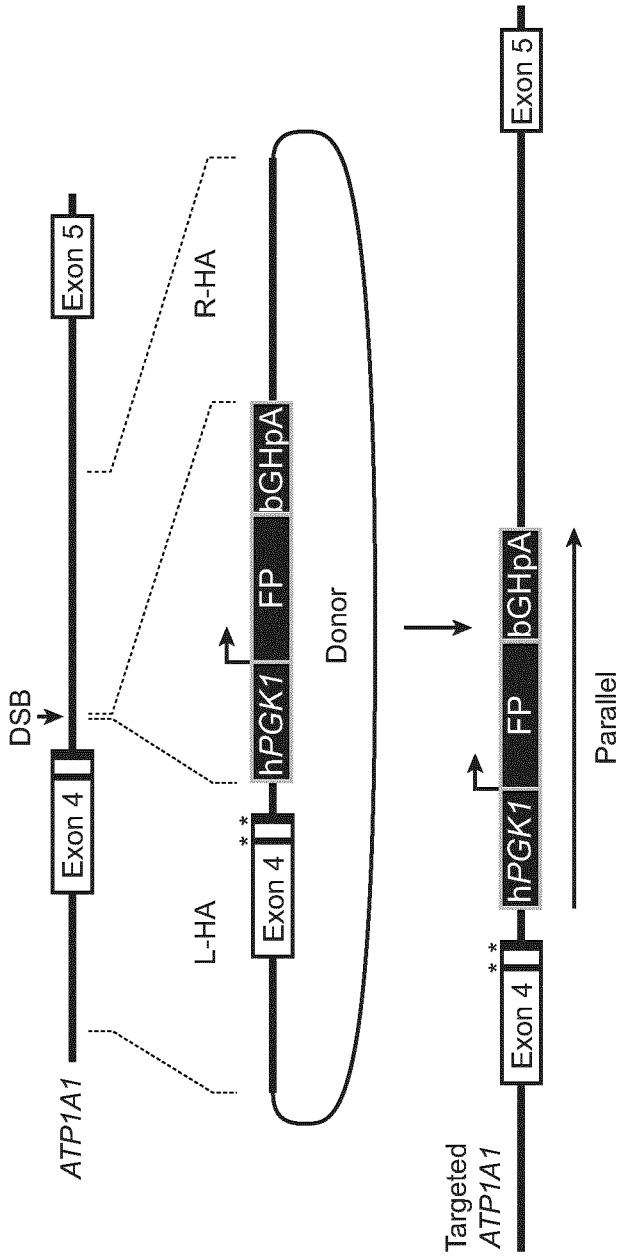

Exon 5
R-HA
DSB
Exon 4
L-HA
Exon 4
hPGK1
FP
bGHpA
Donor
ATP1A1
Exon 5
hPGK1
FP
bGHpA
Parallel
Exon 4
Targeted
ATP1A1

FIG. 30

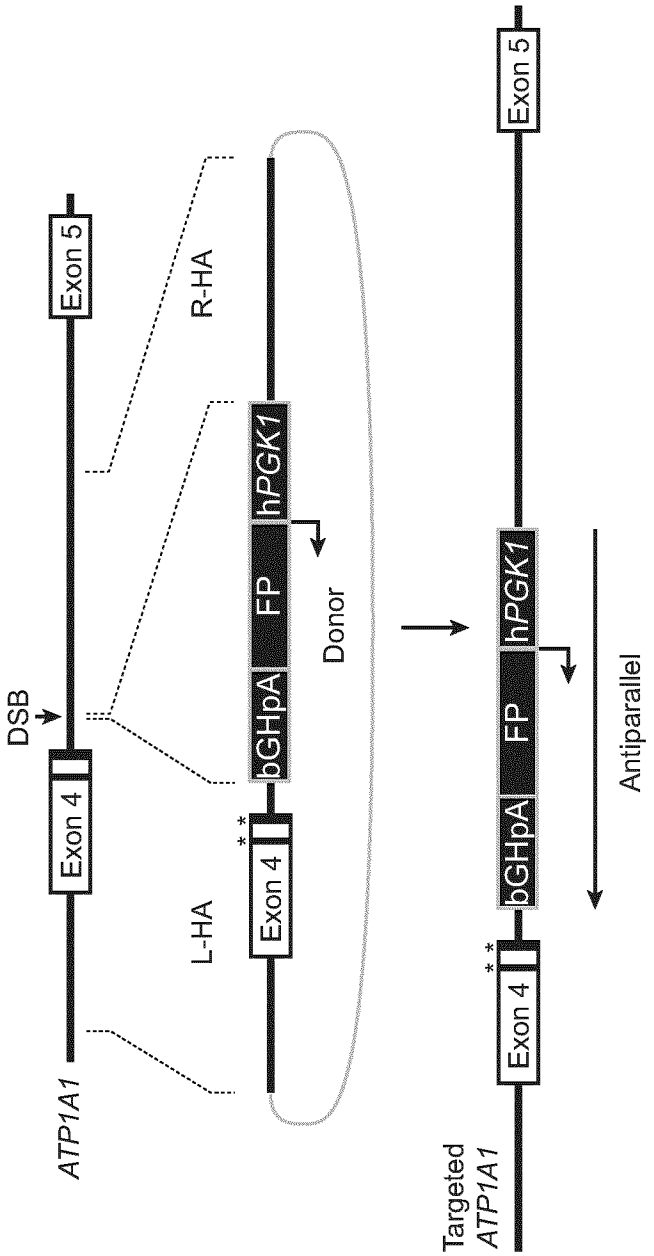

Exon 5
R-HA
DSB
Exon 4
L-HA
Exon 4
bGHpA
FP
hPGK1
Donor
ATP1A1
Exon 5
Exon 4
bGHpA
FP
hPGK1
Antiparallel
Targeted
ATP1A1

FIG. 31

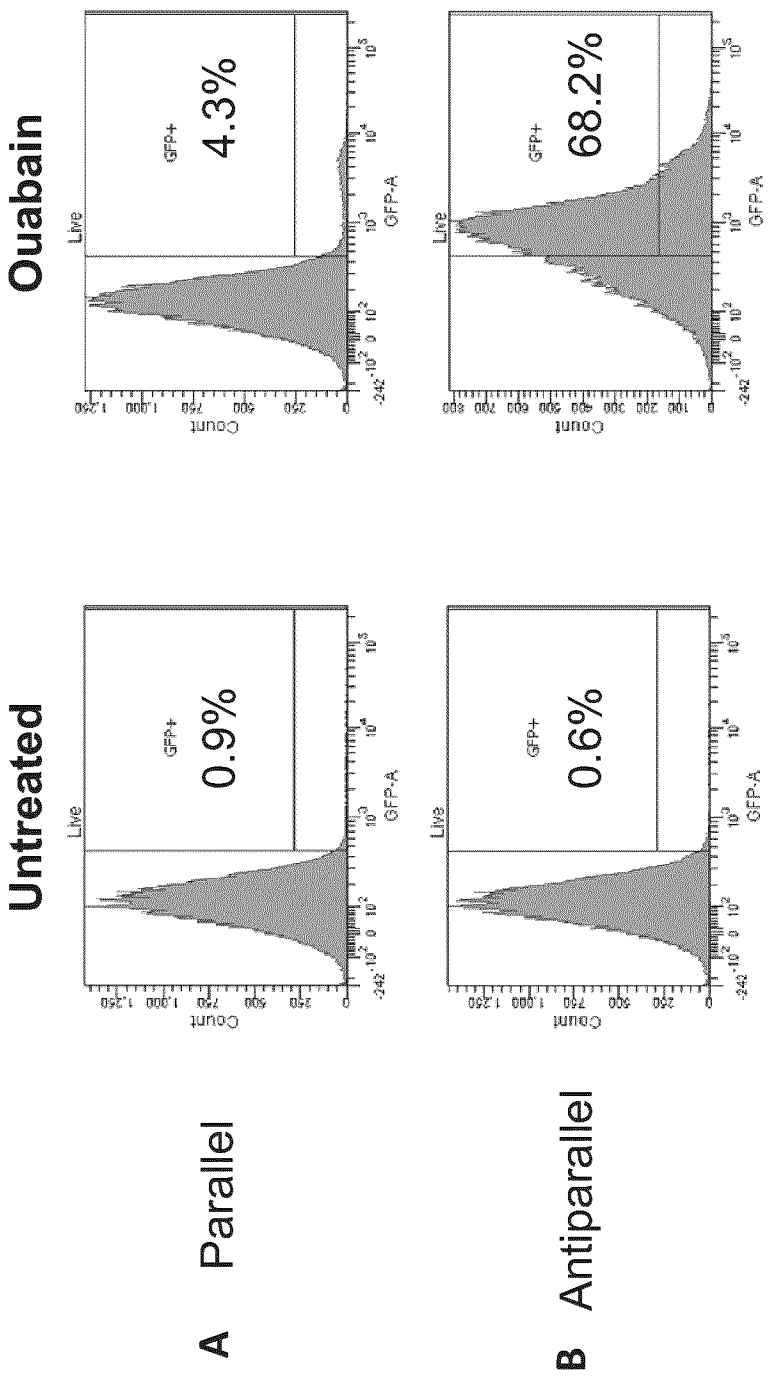

Untreated
Ouabain
A Parallel
B Antiparallel
Live
GFP+
0.9%
4.3%
0.6%
68.2%
Count
GFP-A

FIG. 32

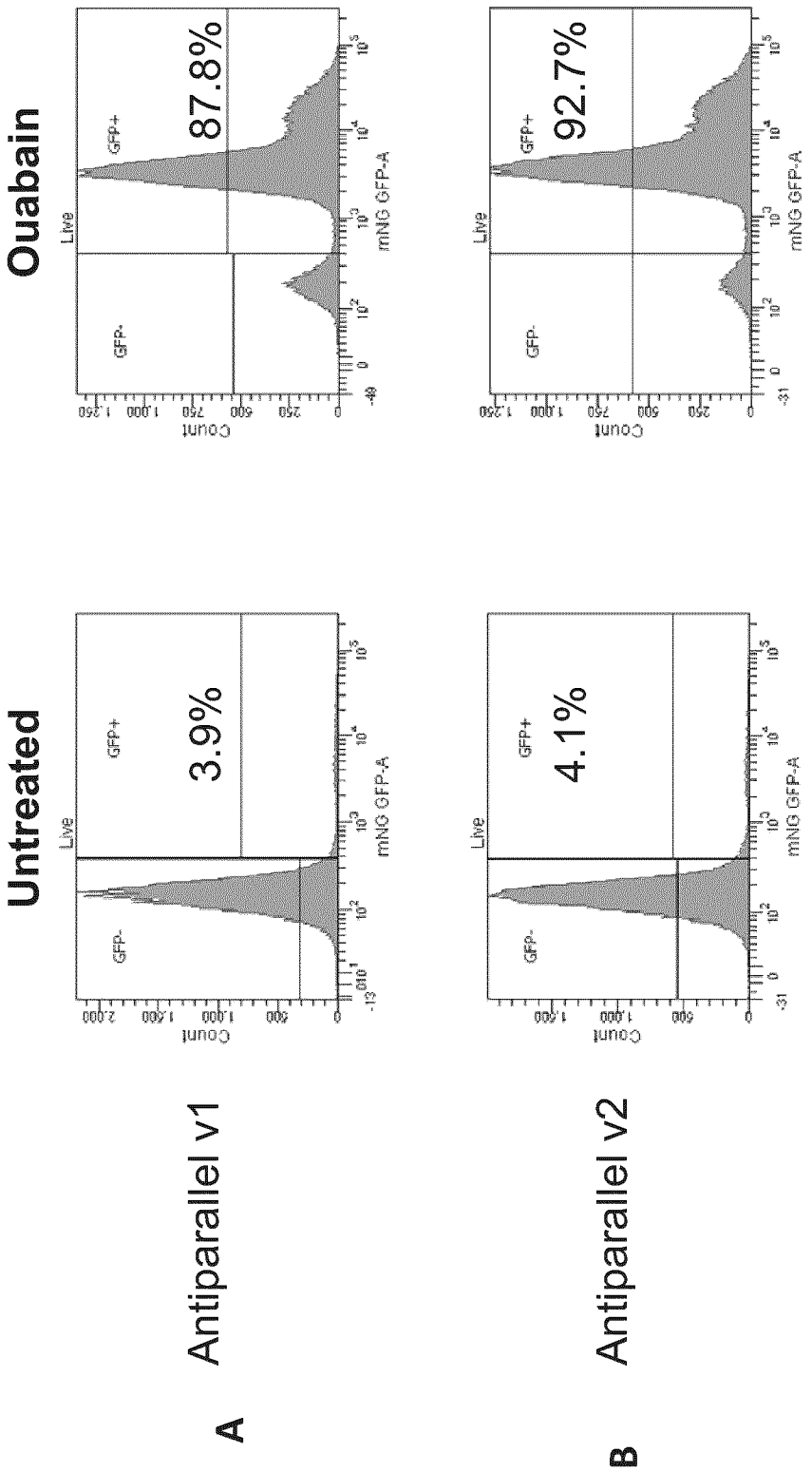

Untreated
Ouabain
Live
GFP-
GFP+
3.9%
87.8%
4.1%
92.7%
mNG GFP-A
Count
A
Antiparallel v1
B
Antiparallel v2

FIG. 33

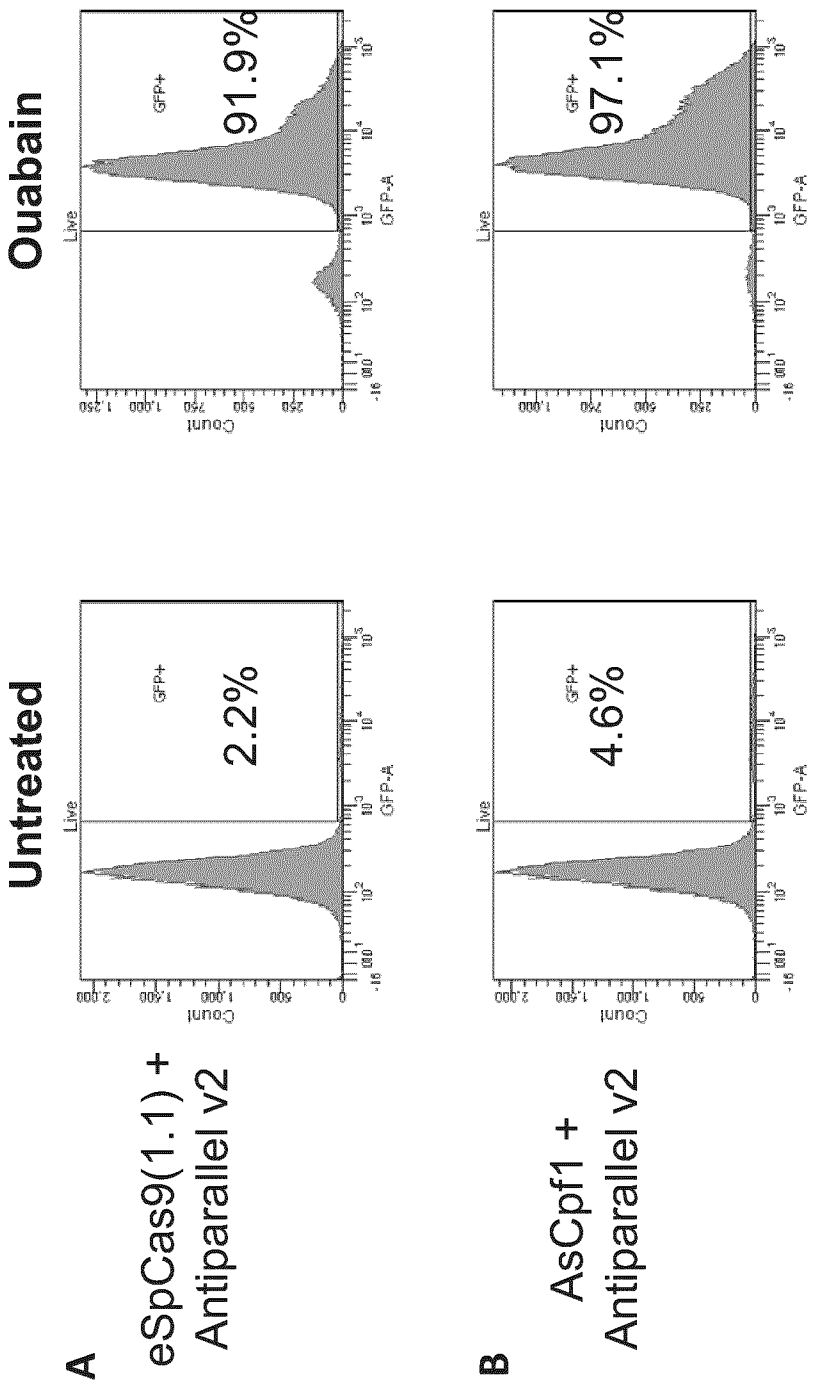

Untreated
Ouabain
A
eSpCas9(1.1) +
Antiparallel v2
B
AsCpf1 +
Antiparallel v2
Live
GFP+
2.2%
91.9%
4.6%
97.1%
Count
GFP-A

FIG. 34

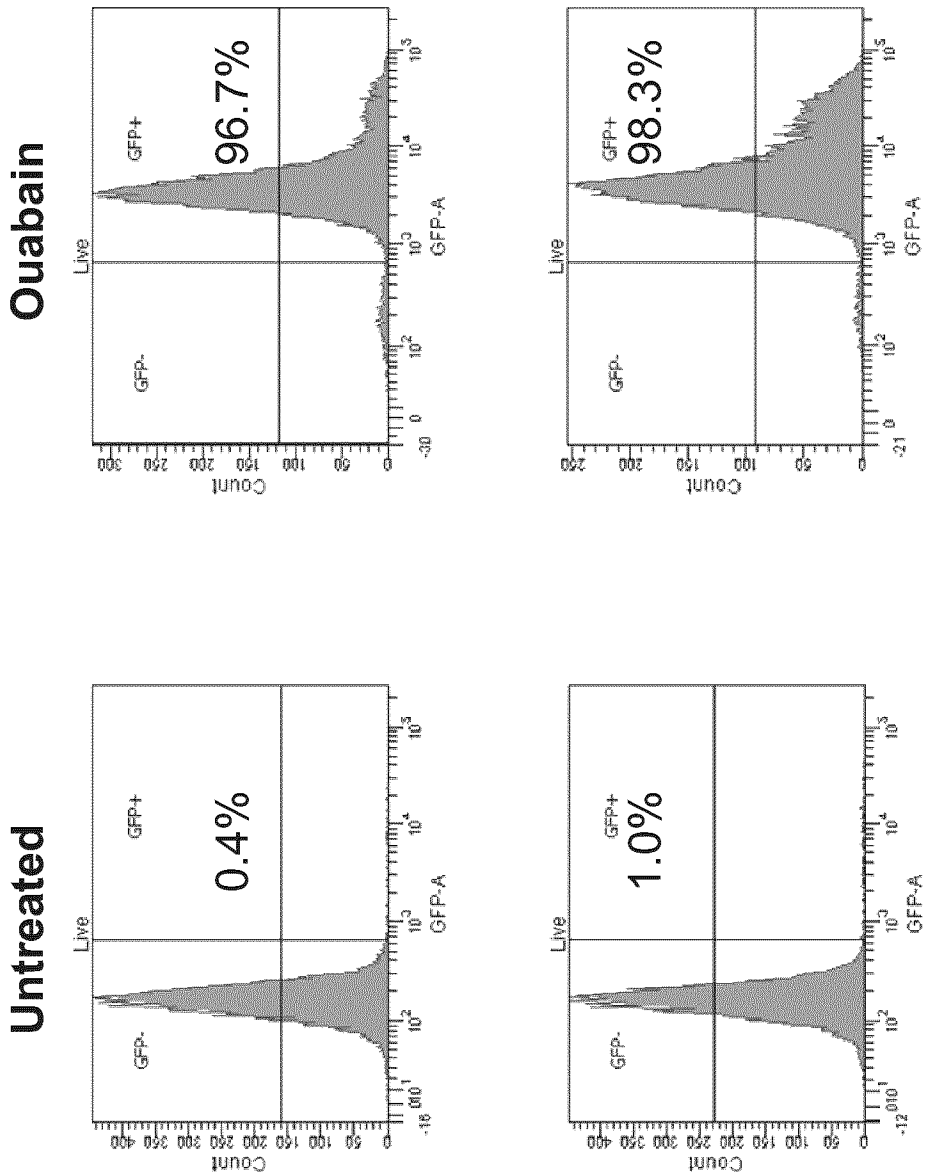

Untreated
Ouabain
A
Antiparallel v2
(500ng)
B
Antiparallel v2
(1000ng)
Live
GFP-
GFP+
0.4%
96.7%
1.0%
98.3%
Count
GFP-A

FIG. 35

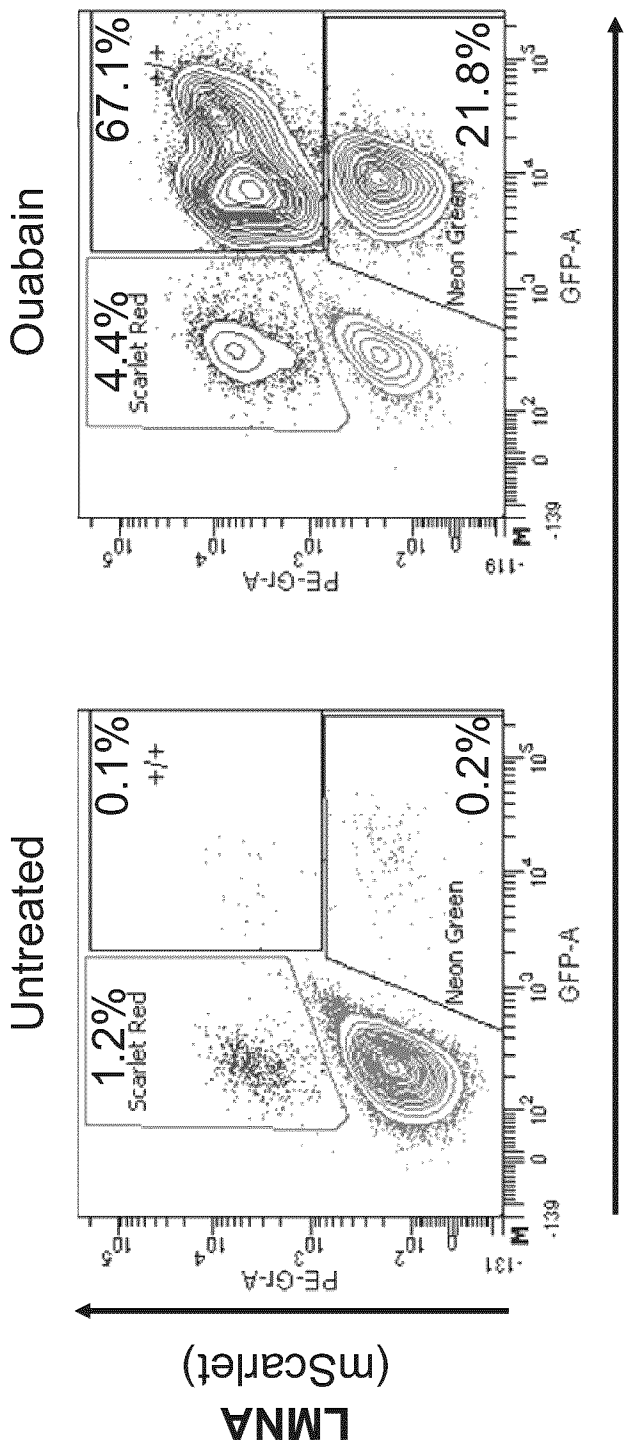

Ouabain
Untreated
67.1%
4.4%
21.8%
0.1%
1.2%
0.2%
Scarlet Red
Neon Green
PE-Gr-A
GFP-A
LMNA
(mScarlet)
ATP1A1
(mNeonGreen)

FIG. 36

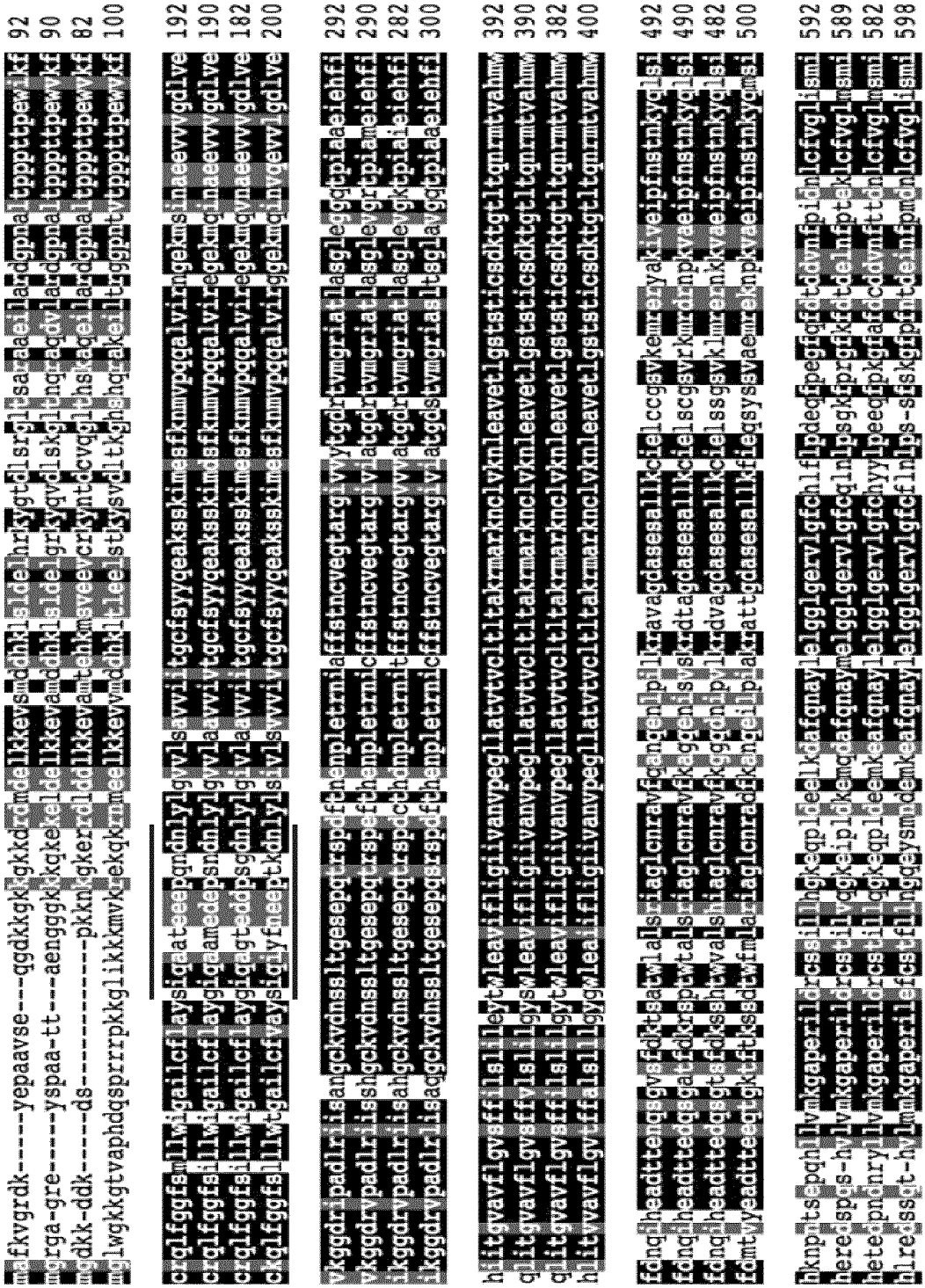


FIG. 37A

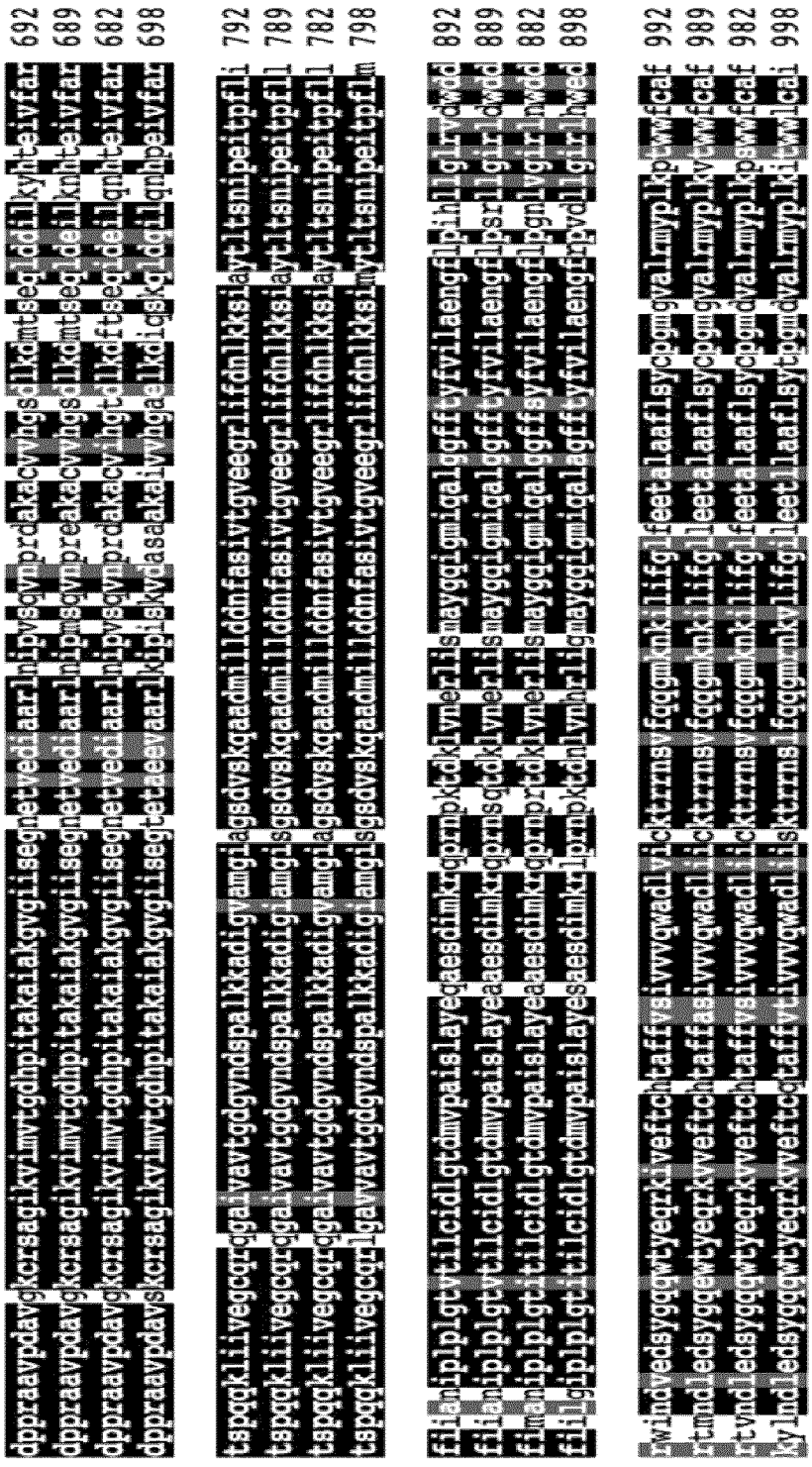


FIG. 37B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2017/051011

A. CLASSIFICATION OF SUBJECT MATTER
IPC: ***C12N 15/11*** (2006.01), ***C07K 14/705*** (2006.01), ***C12N 15/09*** (2006.01), ***C12N 15/55*** (2006.01), ***C12N 15/63*** (2006.01), ***C12N 15/85*** (2006.01), ***C12N 15/90*** (2006.01), ***C12N 9/14*** (2006.01), ***C12N 9/22*** (2006.01), ***C12Q 1/02*** (2006.01), ***C12Q 1/68*** (2006.01)

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: ALL

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)
Databases: Questel-Orbit, NCBI PubMed
Keywords: CRISPR, homology directed repair, non-homologous end joining, NHEJ, drug resistan*, ATP1A1, ouabain

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	D1: PRICE et al., "*Structure-function relationships in the Na,K-ATPase alpha subunit: site-directed mutagenesis of glutamine-111 to arginine and asparagine-122 to aspartic acid generates a ouabain-resistant enzyme.*" Biochemistry, 1988, Vol. 27, No. 22, pp. 8400-8404, ISSN 0006-2960 *whole document*	1-47
A	D2: PRICE et al., "*Structure-function studies of Na,K-ATPase. Site-directed mutagenesis of the border residues from the H1-H2 extracellular domain of the alpha subunit.*" Journal of Biological Chemistry, 1990, Vol. 265, No. 12, pp. 6638-6641, ISSN 0021-9258 *whole document*	1-47
A	D3: TRESCHOW et al., "*OuaSelect, a novel ouabain-resistant human marker gene that allows efficient cell selection within 48h.*" Gene Therapy, 27 September 2007 (27-09-2007), Vol. 14, pp. 1564-1572, ISSN 0969-7128 *whole document*	1-47

☑ 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
23 October 2017 (23-10-2017)	
Name and mailing address of the ISA/CA Canadian Intellectual Property Office Place du Portage I, C114 - 1st Floor, Box PCT 50 Victoria Street Gatineau, Quebec K1A 0C9 Facsimile No.: 819-953-2476	Authorized officer Pascal Lachance (819) 635-3897

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 in 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/CA2017/051011

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of the 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. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim 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. ☐ Claim Nos.:
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:
See extra sheet.

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 claim 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 claim Nos.:
1-30 (completely) and 43-44 and 47 (partially)

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.

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	D4: CROYLE et al. "*Extensive random mutagenesis analysis of the Na+/K+-ATPase alpha subunit identifies known and previously unidentified amino acid residues that alter ouabain sensitivity--implications for ouabain binding.*" European Journal of Biochemistry, 1997, Vol. 248, pp. 488-495, ISSN 0014-2956 *whole document*	1-47
X --- Y	D5: LIAO et al., "*Enriching CRISPR-Cas9 targeted cells by co-targeting the HPRT gene.*" Nucleic Acids Research, 29 June 2015 (29-06-2015), Vol. 43, No. 20, document e134, ISSN 1362-4962 *abstract; introduction; materials and methods; discussion*	1-9 --- 10-11 and 29-30
Y	D6: KIM et al. "*Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells.*" Nature Biotechnology, 06 June 2016 (06-06-2016), Vol. 34, No. 8, pp. 863-868, ISSN 1546-1696 *abstract; Figure 3*	10 and 11
Y	D7: SHY et al. "*Co-incident insertion enables high efficiency genome engineering in mouse embryonic stem cells.*" Nucleic Acids Research, 02 August 2016 (02-08-2016), Vol. 44, No. 16, pp. 7997-8010, ISSN 1362-4962 *abstract; materials and methods; Figure 3*	29 and 30
P, X	D8: AGUDELO et al. "*Marker-free coselection for CRISPR-driven genome editing in human cells.*" Nature Methods, 17 April 2017 (17-04-2017), Vol. 14, No. 6, pp. 615-620, ISSN 1548-7105 *whole document*	1-30, 43-44 and 47

CONTINUATION OF BOX NO. III

The claims are directed to a plurality of inventive concepts as follows:

Group 1 - Claims 1-30 (completely) and claims 43-44 and 47 (partially) are directed to methods featuring the nuclease-based modification of a target polynucleotide in one or more target cells and for enriching the number of said target cells comprising: introduction into said cells by HDR or NHEJ of (a) a nuclease based modification in the target polynucleotide and (b) a second nuclease-based modification in an endogenous gene locus wherein said second modification confers resistance to a drug; contacting said population of cells to the drug; and selecting said one or more target cells by virtue of increased tolerance to the drug, thereby enriching the number of target cells comprising the nuclease-based modification in the target polynucleotide; and to a related kit, use and composition;

Group 2 - Claims 31-32, 34 and 38-47 (partially) are directed to a ouabain-resistant ATP1A1 polypeptide set forth in SEQ ID NO: 5 and to variants thereof; and to encoding nucleic acids, vector, cell, kit, use and composition related thereto;

Groups 3-10 - Claims 31 and 38-47 (partially) are directed to a ouabain-resistant ATP1A1 polypeptide set forth in one of SEQ ID NOs: 8 (Group 3), 10 (Group 4), 12 (Group 5), 14 (Group 6), 16 (Group 7), 18 (Group 8), 20 (Group 9) and 22 (Group 10); and to encoding nucleic acids, vector, cell, kit, use and composition related thereto;

Group 11 - Claims 32, 34 and 38-47 (partially) are directed to ; a ouabain-resistant ATP1A1 polypeptide set forth in part (b) of claim 32, relating to variants of SEQ ID NO: 6; and to encoding nucleic acids, vector, cell, kit, use and composition related thereto;

Groups 12-14 - Claims 32 and 38-47 (partially) are directed to a ouabain-resistant ATP1A2 polypeptide set forth in part (c) of claim 32, relating to variants of SEQ ID NO: 485 (Group 12); a ouabain-resistant ATP1A3 polypeptide set forth in part (d) of claim 32, relating to variants of SEQ ID NO: 486 (Group 13); a ouabain-resistant ATP1A4 polypeptide set forth in part (e) of claim 32, relating to variants of SEQ ID NO: 487 (Group 14); and to encoding nucleic acids, vector, cell, kit, use and composition related thereto;

Groups 15-19 - Claims 33 and 38-47 (partially) are directed to a ouabain-resistant ATP1A1 polypeptide set forth in part (i) of claim 33, relating to variants of one of SEQ ID NOs: 3 (Group 15) and 4 (Group 16); a ouabain-resistant ATP1A2 polypeptide set forth in part (ii) of claim 33, relating to variants of SEQ ID NO: 460 (Group 17); a ouabain-resistant ATP1A3 polypeptide set forth in part (iii) of claim 33, relating to variants of SEQ ID NO: 462 (Group 18); a ouabain-resistant ATP1A4 polypeptide set forth in part (iv) of claim 33, relating to variants of SEQ ID NO: 464 (Group 19); and to encoding nucleic acids, vector, cell, kit, use and composition related thereto; and

Group 20 - Claims 35-37 (completely) and claims 38-47 (partially) are directed to a sgRNA for conferring ouabain resistance in the ATP1A1 gene comprising a target sequence as set forth in Table 3 and Fig. 22 or as defined in claim 36 and 37; and to encoding nucleic acids, vector, cell, kit, use and composition related thereto.

The method of Group 1 features the co-introduction of a modification into a target gene and of a modification into an endogenous gene to confer drug resistance to a cell, to enrich the number of cells comprising said modification into a target gene by selecting for increase in tolerance to the drug. The claims of Groups 2-19 relate to different ouabain-resistant ATP1A1-4 polypeptides. The claims of Group 20 relate to sgRNA molecules for use in conferring ouabain resistance in the ATP1A1 gene. Each of D1, D2, D3 and D4 discloses several variants of ATP1A polypeptides that exhibit resistance to ouabain. It is also known from D3 that ouabain-resistant ATP1A polypeptides have uses beyond the method of the instant claims. The methods of Group 1 and the polypeptides of Groups 2-19 do not share the same general technical features and therefore lack unity. Moreover, the general concept of ouabain-resistant ATP1A1-4 polypeptides is well known in the art, as disclosed for example in each of D1, D2 D3 and D4, and cannot serve to unify the individual polypeptide variants featured in the claims of Groups 2-19. Finally, the sgRNA of Group 20 is for generating ouabain-resistant ATP1A1 polypeptides via NHEJ or HDR but is not limited in use to the method of Group 1, therefore Group 20 also lacks unity with that of Groups 1-19.

The claims must be limited to one inventive concept as set out in PCT Rule 13