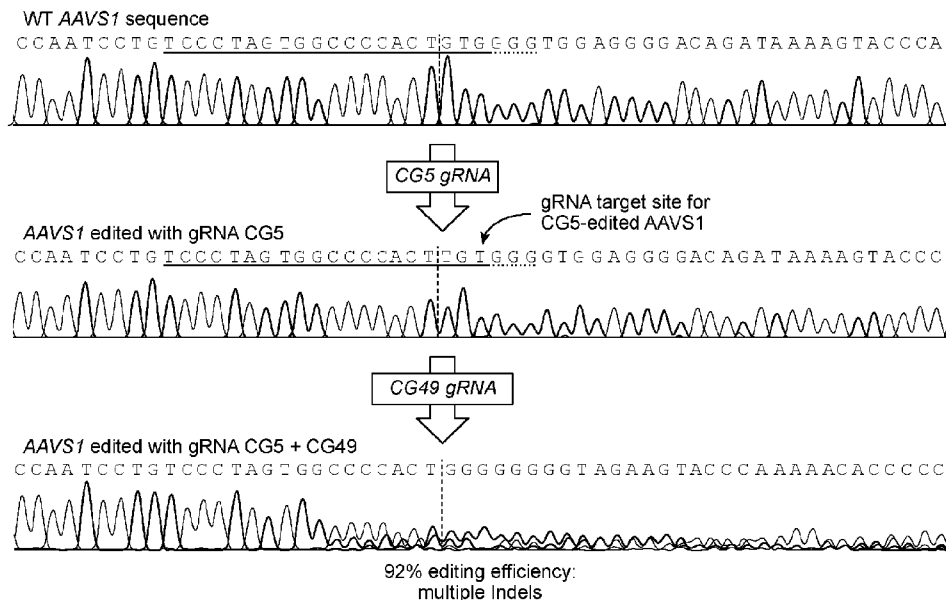




- (51) International Patent Classification:  
C12N 15/11 (2006.01)
- (21) International Application Number:  
PCT/IB2023/057589
- (22) International Filing Date:  
26 July 2023 (26.07.2023)
- (25) Filing Language:  
English
- (26) Publication Language:  
English
- (30) Priority Data:  
63/392,382 26 July 2022 (26.07.2022) US
- (71) Applicant: **BIT BIO LIMITED** [GB/GB]; The Dorothy Hodgkin Building, Babraham Research Campus, Cambridge Cambridgeshire CB22 3FH (GB).
- (72) Inventor: **KLAPHOLZ, Benjamin**; c/o Bit Bio Limited, The Dorothy Hodgkin Building, Babraham Research Campus, Cambridge Cambridgeshire CB22 3FH (GB).

- (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,

(54) Title: MULTI-gRNA GENOME EDITING



"Second Chance" for knock-in

FIG. 1A

(57) Abstract: Described herein are methods and compositions for performing genomic editing with multiple gRNAs, also called second-chance CRISPR/Cas9-mediated genome editing (second-chance editing, or scEditing). In certain aspects, the methods described herein comprise use of CRISPR/Cas9 to target genomic sites of interest at least twice sequentially, providing at least one additional opportunity or "second-chance" at each target polynucleotide sequence for the integration of a DNA cargo into by homology-directed repair (HDR) and are useful for increasing the accuracy of editing target polynucleotide sequences by CRISPR-mediated gene editing.



SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

**Published:**

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

## MULTI-gRNA GENOME EDITING

### CROSS REFERENCE

[0001] This application claims the benefit of and priority to U.S. Provisional Patent Application No. 63/392,382, filed July 26, 2022, which is incorporated by reference in its entirety herein.

### SEQUENCE LISTING

[0002] Not applicable

### BACKGROUND

[0003] Genetic engineering has been revolutionized and democratized by the application of clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins by making genome editing efficient and scalable. CRISPR-Cas mediated gene editing is a powerful and practical tool with potential for discovering new genetic regulatory networks, correcting clinically relevant mutations and engineering new cell-based immunotherapies. The efficiency of CRISPR-Cas mediated gene editing, which harnesses the natural mechanisms of DNA double-strand break repair (DSB), has been iteratively optimized, and, coupled with the design of adapted therapeutic strategies has enabled the scientific community to explore the consequences of genetic variation and develop therapeutic strategies to correct pathogenic genetic variants. However, the ability of CRISPR to mediate the targeted integration of large transgenes and genetic cargos remains limited, as it can be prone to error and a significant percentage of target genomes can remain un-edited. The future of cell therapy and high throughput functional genomic screening will require higher integration efficiencies in order to generate engineered cells with more complex synthetic biological circuitry. Integration of transgenes and genetic cargos rely on targeted CRISPR-mediated DSBs being repaired by homology-directed repair (HDR), a mechanism that competes with other DNA repair mechanisms in the cell, and therefore is not used efficiently at all targeted sites. Therefore, to advance the capabilities of cell engineering, novel methods and techniques that improve the frequency of HDR-mediated transgene integration will be critical.

### SUMMARY

[0004] Disclosed herein is a method of increasing frequency of gene editing at a target polynucleotide/nucleotide sequence in a cell, the method comprising: (i) contacting the target polynucleotide sequence with a first ribonucleoprotein (RNP) comprising a first ribonucleic

acid molecule and a clustered regularly interspaced short palindromic repeat-associated (Cas) protein; sequence to produce an edited target polynucleotide sequence; (ii) contacting the edited target polynucleotide sequence with a second or more RNP(s) comprising a second or more ribonucleic acid molecule(s) and a Cas protein to induce a second gene editing event; wherein the second or more ribonucleic acid molecule(s) comprises a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid molecule and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; wherein the method results in increased gene editing frequency compared to the same method but consisting of only step (i). In certain embodiments, (i) results in an error in one or more bases at the target polynucleotide sequence relative to the first ribonucleic acid molecule sequence. In certain embodiments, wherein (ii) occurs after (i) results in an error in one or more bases at the target polynucleotide sequence relative to the first ribonucleic acid molecule sequence.

**[0005]** In certain embodiments, the error in the one or more bases at the target polynucleotide sequence is an error induced by double strand break repair. In certain embodiments, the error in the one or more bases at the target polynucleotide sequence is an alteration in one base relative to the first ribonucleic acid molecule sequence. In certain embodiments, the error in the one or more bases at the target polynucleotide sequence comprises an insertion of one or more bases. In certain embodiments, the error in the one or more bases at the target polynucleotide sequence comprises a deletion of one or more bases. In certain embodiments, the error in the one or more bases at the target polynucleotide sequence comprises an insertion at the cut site created by double strand break repair.

**[0006]** In certain embodiments, the methods further comprise integrating a donor polynucleotide sequence into the target polynucleotide sequence. In certain embodiments, steps (i) and (ii) occur simultaneously. In certain embodiments, steps (i) and (ii) occur sequentially.

**[0007]** In certain embodiments, the target polynucleotide sequence is a genomic safe harbor site (GSH) locus. In certain embodiments, the GSH locus is selected from the group consisting of AAVS1 (PPP1R12C gene), ROSA26 (THUMP3-AS1 gene), CLYBL, DBI, PCSK9 and ZC3H3. In certain embodiments, the GSH locus is ROSA26. In certain embodiments, the GSH locus is AAVS1. In certain embodiments, the AAVS1 locus is altered by an insertion of one base; and wherein the insertion is an adenine or thymidine. In certain embodiments, the GSH locus is CLYBL. In certain embodiments, the CLYBL locus is altered

by an insertion of one base; and wherein the insertion is an adenine or thymidine. In certain embodiments, the GSH locus is DBI. In certain embodiments, the DBI locus is altered by an insertion of one base; and wherein the insertion is an adenine or thymidine. In certain embodiments, the GSH locus is PCSK9. In certain embodiments, the PCSK9 locus is altered by an insertion of one base; and wherein the insertion is a adenine or thymidine .

**[0008]** In certain embodiments, the increased accuracy of integration of the donor polynucleotide sequence by Homology-Directed Repair (HDR)-mediated integration ranges from about 25% to about 75%. In certain embodiments, the increased gene editing frequency is measured by detecting expression of a gene that is encoded by a donor polynucleotide sequence. In certain embodiments, the gene editing results in the insertion of at least one exogenous gene.

**[0009]** In certain embodiments, the gene editing results in the insertion of one or more non-protein coding sequences. In certain embodiments, the non-protein coding sequence comprises a non-coding RNA sequence. In certain embodiments, the non-coding RNA sequence comprises a sequence selected from the group consisting of one or more microRNAs, siRNAs, piRNAs, snoRNAs, snRNAs, exRNAs, scaRNAs and long ncRNAs.

**[0010]** In certain embodiments, the contacting comprises introducing one or more of: the first ribonucleic acid, second ribonucleic acid, donor polynucleotide, and polynucleotide encoding Cas protein, to a cell. In certain embodiments, the introducing step is performed by at least one of transfection, transduction, electroporation, and microinjection. In certain embodiments, transcription of the first and/or second ribonucleic acid sequence is transiently induced. In certain embodiments, the transcription is transiently induced by activating a regulatable promoter controlling the transcription of the first and/or second ribonucleic acid sequence.

**[0011]** In certain embodiments, the at least one exogenous gene comprises a transcription factor. In certain embodiments, the donor polynucleotide sequence comprises a sequence encoding at least one transcription factor.

**[0012]** In certain embodiments, a donor plasmid DNA comprises the donor polynucleotide sequence. In certain embodiments, the donor plasmid DNA further comprises one or more polynucleotide sequences encoding, a bacterial resistance gene, an origin of replication, a transcriptional promoter for driving expression of an exogenous gene, a selectable marker, a fluorescent protein, a 5' homology arm, and a 3' homology arm.

**[0013]** In an aspect, described herein is a cell comprising the first and second ribonucleic acids of any one of the above claims; and optionally, the donor polynucleotide of any one of

the above claims. In certain embodiments, the cell further comprises Cas protein. In certain embodiments, the cell is a stem cell. In certain embodiments, the cell is an induced pluripotent stem (iPS) cell. In certain embodiments, the iPS cell is a human iPS (hiPS) cell. In certain embodiments, the cell is a somatic cell. In certain embodiments, the cell is an immune cell, optionally a T cell, a B cell, a dendritic cell, a macrophage or an NK cell. In certain embodiments, the cell is a neuronal cell, optionally a microglial cell, a motor neuron, a dopaminergic neuron, a GABAergic neuron, or a glutamatergic neuron. In certain embodiments, the cell is an adipocyte, a hepatocyte, a pancreatic cell, an epithelial cell, a muscle cell, a bone cell, a skin cell or a blood cell. In certain embodiments, the cell is an *ex-vivo* patient-derived cell. In certain embodiments, the cell is reprogrammed to a differentiated cell after the second gene editing event.

**[0014]** In an aspect, described herein is a plurality of polynucleotides comprising: i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence; ii) a second or more ribonucleic acid molecule(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; and optionally, iii) a donor DNA comprising the target polynucleotide sequence; and optionally, iv) a polynucleotide comprising a sequence encoding Cas protein.

**[0015]** In an aspect, described herein is a kit comprising: i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence; ii) a second or more ribonucleic acid molecule(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; and optionally, iii) a donor DNA comprising the target polynucleotide sequence; and optionally, iv) a polynucleotide comprising a sequence encoding Cas protein; and v) instructions for use.

**[0016]** In an aspect, described herein is a population of cells comprising: i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence; ii) a second or more ribonucleic acid molecule(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; and optionally, iii) a donor DNA comprising the target polynucleotide sequence; and optionally, iv) a polynucleotide comprising a sequence encoding Cas protein. In certain embodiments, the cells comprise an increased percentage of the donor polynucleotide integrated into the target polynucleotide sequence compared to a population of the same type of cells comprising the first ribonucleic acid, the target polynucleotide sequence, and the donor polynucleotide, but not the second ribonucleic acid. In certain embodiments, the cells further comprise Cas protein. In certain embodiments, the cells comprise induced pluripotent stem (iPS) cells, optionally human iPS (hiPS) cells. In certain embodiments, the cells comprise immune cells, optionally T cells, B cells, dendritic cell, macrophages, NK cells, or combinations thereof. In certain embodiments, the cells comprise neuronal cells, optionally microglial cells, motor neurons, dopaminergic neurons, GABAergic neurons, glutamatergic neurons or combinations thereof. In certain embodiments, the cells comprise adipocytes, hepatocytes, pancreatic cells, epithelial cells, skeletal muscle cells, smooth muscle cells, cardiomyocytes, bone cells, skin cells or blood cells. In certain embodiments, the cells comprise *ex-vivo* patient derived cells. In certain embodiments, the donor polynucleotide encodes at least one transcription factor. In certain embodiments, the cells are reprogrammed to a differentiated cell after the first or second gene editing event mediated by the first or second ribonucleic acid, Cas protein and the donor polynucleotide.

**[0017]** In an aspect, described herein is a method of generating a cell with at least one exogenous polynucleotide sequence integrated into the cell genome, the method comprising: (i) contacting a target polynucleotide sequence within the cells with a first ribonucleic acid molecule and a Cas protein; wherein the first ribonucleic acid molecule comprises a sequence complementary to a nucleic acid sequence in the target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence; (ii) contacting the target polynucleotide sequence with a second or

more ribonucleic acid molecule(s) molecule(s) and a Cas protein to induce a second gene editing event at the target polynucleotide sequence; wherein the second or more ribonucleic acid molecule(s) comprises a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; wherein the target polynucleotide sequence is a GSH site; wherein the method results in increased gene editing accuracy compared to a method comprising the first ribonucleic acid molecule that is complementary to a nucleic acid sequence in the target polynucleotide sequence but not the second or ribonucleic acid(s); and wherein the method results in integration of at least one exogenous polynucleotide sequence into the genome of the cell.

**[0018]** In an aspect, described herein is a method of generating a differentiated cell from iPS cells, the method comprising: (i) contacting a target polynucleotide sequence within the iPS cells with a first ribonucleic acid molecule and a Cas protein; wherein the first ribonucleic acid molecule is complementary to a nucleic acid sequence in the target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence; (ii) contacting the target polynucleotide sequence with a second or more ribonucleic acid molecule(s) molecule(s) and a Cas protein to induce a second gene editing event at the target polynucleotide sequence; wherein the second or more ribonucleic acid molecule(s) comprises a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; wherein the target polynucleotide sequence is a GSH site; wherein the method results in increased gene editing accuracy compared to a method comprising the first ribonucleic acid molecule that is complementary to a nucleic acid sequence in the target polynucleotide sequence but not the second or more ribonucleic acid molecule(s); wherein the method results in integration of at least one exogenous gene ; and wherein the method results in the generation of one or more differentiated cells.



**BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS**

[0019] These and other features, aspects, and advantages of the present disclosure will become better understood with regard to the following description, and accompanying drawings, where:

[0020] **Figure 1A** is a diagram illustrating the sequence at the AAVS1 target site for the gRNAs CG5 and CG49 for second chance editing.

[0021] **Figure 1B** presents two dot plots depicting the percentage of cells expressing EGFP corresponding to the percent of cells that harbour the inserted donor plasmid DNA encoding the reporter gene when introduced with either a single CG5 gRNA or both the CG5 and CG49 gRNAs.

[0022] **Figure 2A** is a diagram illustrating the sequence at the CLYBL target site for the gRNAs CG65 and CG70 for second chance editing.

[0023] **Figure 2B** presents two dot plots depicting the percentage of cells expressing EGFP corresponding to the percent of cells that harbour the inserted donor plasmid DNA encoding the reporter gene when introduced with either a single CG65 gRNA or both the CG65 and CG70 gRNAs.

[0024] **Figure 3A** is a diagram illustrating the sequence at the DBI target site for the gRNAs CG63 and CG68 for second chance editing.

[0025] **Figure 3B** presents two dot plots depicting the percentage of cells expressing EGFP corresponding to the percent of cells that harbour the inserted donor plasmid DNA encoding the reporter gene when introduced with either a single CG63 gRNA or both the CG63 and CG68 gRNAs.

[0026] **Figure 4A** is a diagram illustrating the sequence at the PCSK9 target site for the gRNAs CG55 and CG69 for second chance editing.

[0027] **Figure 4B** presents two dot plots depicting the percentage of cells expressing EGFP corresponding to the percent of cells that harbour the inserted donor plasmid DNA encoding the reporter gene when introduced with either a single CG55 gRNA or both the CG55 and CG69 gRNAs.

**DETAILED DESCRIPTION****Definitions**

[0028] Certain terms used in the claims and specification are defined as set forth below unless otherwise specified.

**[0029]** As used herein, the “CRISPR-Cas” system refers to a class of bacterial systems for defense against foreign nucleic acids. CRISPR-Cas systems are found in a wide range of eubacterial and archaeal organisms. CRISPR-Cas systems include type I, II, and III subtypes. Type II CRISPR-Cas systems generally utilize an RNA-mediated nuclease, for example, Cas9 protein, in complex with guide and activating RNAs or single-guide RNA (sgRNA) to recognize and cleave foreign nucleic acids, *e.g.*, foreign nucleic acids including natural or modified nucleotides.

**[0030]** As used herein, the term “targetable nuclease” refers to a protein that can recognize a sequence of a cognate nucleic acid sequence (*e.g.*, a target gene within a genome), bind to the cognate nucleic acid sequence, and modify the cognate nucleic acid sequence. In some embodiments, a targetable nuclease is an RNA-guided nuclease, *e.g.*, a Cas protein. In other embodiments, a targetable nuclease is a fusion protein that includes a protein that can bind to a cognate nucleic acid sequence (*e.g.*, a transcription activator-like (TAL) effector DNA-binding protein or a zinc finger DNA-binding protein) and a protein that can modify a cognate nucleic acid sequence (*e.g.*, a nuclease, a transcription activator or repressor). In some embodiments, the targetable nuclease is a chimeric DNA-RNA-guided nuclease. In some embodiments, the targetable nuclease has nuclease activity. In some embodiments, the targetable nuclease can modify a cognate nucleic acid sequence by cleaving the target nucleic acid. The cleaved target nucleic acid can then undergo homologous recombination with a nearby a homology directed repair (HDR) template, such as through homology directed repair or homology mediated end joining (HMEJ).

**[0031]** As used herein, the term “donor DNA” or “donor template” refers to a polynucleotide that comprises a target polynucleotide sequence. The donor DNA can be a single-stranded oligonucleotide donor (ssODN) or a double-strand donor DNA (dsODN). The double-strand donor DNA can be with or without homology regions (homologous to the target polynucleotide sequence) flanking the sequence to integrate donor DNA at the target polynucleotide sequence that is cut by the RNA-guided nuclease (*e.g.*, a Cas protein). In certain embodiments, the donor DNA comprises homology regions that enable the use of homology-directed repair (HDR) by the cell. The donor DNA can include a homology directed repair (HDR) template. An HDR template can include a 5' homology arm, a nucleotide insert (*e.g.*, an exogenous sequence, a transgene, and/or a sequence that encodes a heterologous protein or fragment thereof), and a 3' homology arm. In certain embodiments, the donor DNA lacks homology arms, and the gene editing event with the donor DNA comprises the DNA repair mechanism, Non-Homologous End Joining (NHEJ).

**[0032]** As used herein, the term “target polynucleotide sequence” or “target sequence” refers to a nucleotide sequence that is recognized and bound by a targetable nuclease. In some embodiments, a targetable nuclease, *e.g.*, a transcription activator-like (TAL) effector DNA-binding protein or zinc finger DNA-binding protein, can directly recognize and bind a target sequence. In other embodiments, a targetable nuclease, *e.g.*, an RNA-guided nuclease, can indirectly recognize and bind a target sequence via a donor gRNA. An RNA-guided nuclease binds to the donor gRNA, while the donor gRNA hybridizes to a target sequence. In some embodiments, a target sequence is a portion of genomic nucleic acid targeted by the donor gRNA.

**[0033]** As used herein, the “RNA-guided nuclease” refers to a nuclease that binds or forms a complex with a guide RNA (gRNA) and utilizes the gRNA to selectively bind regions within a DNA polynucleotide. In general, an RNA-guided nuclease can selectively bind nearly any sequence within a DNA polynucleotide that is complementary to the gRNA. In some embodiments, a RNA-guided nuclease has nuclease activity and can cleave the linkage (*e.g.*, phosphodiester bonds) between nucleotides in the DNA polynucleotide. In other embodiments, an RNA-guided nuclease does not have nuclease activity and can be used to selectively bind and/or localize other proteins (*e.g.*, transcriptional activator or repressors) that are fused to the RNA-guided nuclease to the region of interest within the DNA polynucleotide.

**[0034]** As used herein, the term “guide RNA” or “gRNA” refers to a DNA-targeting RNA that can guide an RNA-guided nuclease (*e.g.*, a Cas protein) to a cognate nucleic acid sequence by hybridizing to the cognate nucleic acid sequence. In some embodiments, a guide RNA can be a single-guide RNA (sgRNA), which contains (1) a guide sequence (*e.g.*, crRNA equivalent portion of the single-guide RNA) that guides an RNA-guided nuclease to a cognate nucleic acid sequence and (2) a scaffold sequence (*e.g.*, tracrRNA equivalent portion of the single-guide RNA) that interacts with the RNA-guided nuclease. In other embodiments, a guide RNA can contain two components, (1) a guide sequence (*e.g.*, crRNA equivalent portion of the single-guide RNA) that guides an RNA-guided nuclease to cognate nucleic acid sequence and (2) a scaffold sequence (*e.g.*, tracrRNA equivalent portion of the single-guide RNA) that interacts with the RNA-guided nuclease. A portion of the guide sequence can hybridize to a portion of the scaffold sequence to form the two-component guide RNA.

**[0035]** As used herein, the term “target guide RNA” or “target gRNA” refers to a gRNA that can hybridize to a cognate nucleic acid sequence to be modified, *e.g.*, at a location in a DNA

polynucleotide where integration of an HDR template is desired, such as a chromosome of a T cell and/or safe-harbor genomic locations.

**[0036]** As used herein, the term “donor guide RNA” or “donor gRNA” refers to a gRNA that can hybridize to a target polynucleotide sequence within a plasmid donor template. In some embodiments, a target polynucleotide sequence is partially complementary or completely complementary to an equal length portion of the sequence of a donor gRNA.

**[0037]** As used herein, the term “single-guide RNA” or “sgRNA” refers to a DNA-targeting RNA including (1) a guide sequence (*e.g.*, crRNA equivalent portion of the single-guide RNA) that targets a Cas protein to a cognate nucleic acid sequence and (2) a scaffold sequence (*e.g.*, a tracrRNA-equivalent portion of the single-guide RNA) that interacts with a Cas protein.

**[0038]** As used herein, the term “complex” refers to a joining of at least two components. The two components may each retain the properties/activities they had prior to forming the complex or gain properties as a result of forming the complex. The joining includes, but is not limited to, covalent bonding, non-covalent bonding (*i.e.*, hydrogen bonding, ionic interactions, Van der Waals interactions, and hydrophobic bond), use of a linker, fusion, or any other suitable method. Contemplated components of the complex include polynucleotides, polypeptides, or combinations thereof. For example, a complex comprises an endonuclease and a guide RNA.

**[0039]** As used herein, the term “complementary” or “complementarity” refers to the capacity for base pairing between nucleobases, nucleosides, or nucleotides, as well as the capacity for base pairing between one polynucleotide to another polynucleotide. In some embodiments, one polynucleotide can have “complete complementarity,” or be “completely complementary,” to another polynucleotide, which means that when the two polynucleotides are optionally aligned, each nucleotide in one polynucleotide can engage in Watson-Crick base pairing with its corresponding nucleotide in the other polynucleotide. In other embodiments, one polynucleotide can have “partial complementarity,” or be “partially complementary,” to another polynucleotide, which means that when the two polynucleotides are optionally aligned, at least 60% (*e.g.*, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 97%) but less than 100% of the nucleotides in one polynucleotide can engage in Watson-Crick base pairing with their corresponding nucleotides in the other polynucleotide. In other words, there is at least one (*e.g.*, one, two, three, four, five, six, seven, eight, nine, ten, or more) mismatched nucleotide base pairs when the two polynucleotides are hybridized. Pairs of nucleotides that engage in Watson-Crick base pairing include, *e.g.*, adenine and thymine,

cytosine and guanine, and adenine and uracil, which all pair through the formation of hydrogen bonds. Examples of mismatched bases include guanine and uracil, guanine and thymine, and adenine and cytosine hydrogen bonding.

**[0040]** As used herein, the term “Cas protein” or “Cas” refers to a Clustered Regularly Interspaced Short Palindromic Repeats-associated protein or nuclease. A Cas protein can be a wild-type Cas protein or a Cas protein variant. Cas9 protein is an example of a Cas protein that belongs in the type II CRISPR-Cas system (*e.g.*, Rath et al., *Biochimie* 117:119, 2015). Other examples of Cas proteins are described in more detail herein. A naturally-occurring type II Cas protein generally requires both a crRNA (“crRNA”) and a trans-activating crRNA (“tracrRNA”) for site-specific DNA recognition and cleavage. The crRNA associates with the tracrRNA through a region of partial complementarity to guide the Cas protein to a region homologous to the crRNA in the target DNA called a “protospacer”. A naturally-occurring type II Cas protein cleaves DNA to generate blunt ends at the double-strand break at sites specified by a guide sequence contained within a crRNA transcript. In some embodiments of the compositions and methods described herein, a Cas protein associates with a target gRNA or a donor gRNA to form a ribonucleoprotein (RNP) complex. In some embodiments of the compositions and methods described herein, the Cas protein has nuclease activity. In other embodiments, the Cas protein does not have nuclease activity.

**[0041]** As used herein, the term “Cas protein variant” refers to a Cas protein that has at least one amino acid substitution (*e.g.*, one, two, three, four, five, six, seven, eight, nine, ten, or more amino acid substitutions) relative to the sequence of a wild-type Cas protein and/or is a truncated version or fragment of a wild-type Cas protein. In some embodiments, a Cas protein variant has at least 75% sequence identity (*e.g.*, at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity) to the sequence of a wild-type Cas protein. In some embodiments, a Cas protein variant is a fragment of a wild-type Cas protein and has at least one amino acid substitution relative to the sequence of the wild-type Cas protein. A Cas protein variant can be a Cas9 protein variant. In some embodiments, a Cas protein variant has nuclease activity. In other embodiments, a Cas protein variant does not have nuclease activity.

**[0042]** As used herein, the term “ribonucleoprotein complex” or “RNP complex” refers to a complex comprising a Cas protein or variant (*e.g.*, a Cas9 protein or variant) and at least one gRNA.

**[0043]** As used herein, the term “modifying” in the context of modifying a target nucleic acid in the genome of a cell refers to inducing a change (*e.g.*, cleavage) in the target nucleic acid.

In some embodiments, the change can be a structural change in the sequence of the target nucleic acid. For example, the modifying can take the form of inserting a nucleotide sequence into the target nucleic acid. For example, an exogenous nucleotide sequence can be inserted into the target nucleic acid. In certain embodiments, the exogenous nucleotide sequence encodes a transgene. The target nucleic acid can also be excised and replaced with an exogenous nucleotide sequence. In another example, the modifying can take the form of cleaving the target nucleic acid without inserting a nucleotide sequence into the target nucleic acid. For example, the target nucleic acid can be cleaved and excised. Such modifying can be performed, for example, by inducing a double stranded break within the target nucleic acid, or a pair of single stranded nicks on opposite strands and flanking the target nucleic acid. Methods for inducing single or double stranded breaks at or within a target nucleic acid include the use of a targetable nuclease (*e.g.*, a Cas protein) as described herein directed to the target nucleic acid by a gRNA/sgRNA. In other embodiments, modifying a target nucleic acid includes targeting another protein to the target nucleic acid and does not include cleaving the target nucleic acid.

**[0044]** As used herein, the term “first gene editing event” refers to modification of a target polynucleotide sequence, and includes DNA repair of double stranded breaks that leads to at least one base alteration (*e.g.*, insertion, deletion or substitution) in the target polynucleotide sequence, but does not lead to an insertion of donor DNA.

**[0045]** As used herein, the term “second gene editing event” refers to modification of a target polynucleotide sequence by a DNA repair mechanism and can involve a donor DNA.

**[0046]** As used herein, the term “frequency of gene editing” refers to the frequency that a desired gene editing event (*e.g.*, integration of donor DNA) occurs at a target polynucleotide sequence.

**[0047]** As used herein, the term genomic safe harbor (GSH) refers to chromosomal locations where transgenes can integrate and function in a predictable manner (*e.g.*, are less prone to silencing), without perturbing endogenous gene activity. In certain embodiments, a GSH is a genomic locus 50 kb away from a known gene, 300 kb away from a known oncogene, 300 kb away from a miRNA, 150 kb away from a lncRNA or tRNA, 300 kb away from a telomere or centromere, and 20 kb away from a known enhancer region (Aznauryan E, Yermanos A, Kinzina E, et al. Discovery and validation of human genomic safe harbor sites for gene and cell therapies. *Cell Rep Methods*. 2022;2(1):100154).

**[0048]** Abbreviations used in this application include the following: “CAS” (Clustered Regularly Interspaced Short Palindromic Repeats-associated protein or nuclease), “CRISPR”

(clustered regularly interspaced short palindromic repeat), “ssODN” (single-stranded oligonucleotide donor), “dsODN” (double-stranded oligonucleotide donor), “NHEJ” (Non-Homologous End Joining), “HDR” (homology-directed repair), “RNP” (ribonucleoprotein), “gRNA” (guide RNA), “sgRNA” (single guide RNA), “crRNA” (crispr RNA), and “tracrRNA” (trans-activating crispr RNA).

**[0049]** It must be noted that, as used in the specification and the appended claims, the singular forms “a,” “an” and “the” include plural referents unless the context clearly dictates otherwise.

### **Methods**

**[0050]** Described herein are methods and compositions for performing genomic editing with multiple gRNAs, also called second-chance CRISPR/Cas9-mediated genome editing (second-chance editing, or scEditing). In certain aspects, the methods described herein comprise use of CRISPR/Cas9 to target genomic sites of interest at least twice sequentially, providing at least one additional opportunity or a “second-chance” at each target polynucleotide sequence for the integration of a transgene by homology-directed repair (HDR). This method can increase the probability of target site transgene integration. In some embodiments, the methods described herein utilize the observation that double-strand DNA break (DSB) repair at certain genomic loci can be very consistent, reflected by the presence of a very predictable DNA sequence after repair. This can allow for the design of gRNAs that recognize target polynucleotide sequences that have undergone DSB, but have not had an integration of donor DNA. In certain aspects, the use of the gRNAs that recognize the unmodified target polynucleotide sequence in combination with gRNAs that recognize sequences that have undergone DSB repair allows for increased overall accuracy (*e.g.*, frequency) of integration of donor DNA sequence at target polynucleotide sequences by CRISPR-mediated gene editing. Thus, this disclosure is useful for increasing the accuracy of editing target polynucleotide sequences by CRISPR-mediated gene editing.

**[0051]** In certain aspects, described herein are methods of increasing the frequency of gene editing at a target polynucleotide sequence in a cell leading to insertion of a donor polynucleotide sequence of interest. Any method of making specific, targeted double strand breaks in the genome in order to affect the insertion of a donor polynucleotide sequence (*e.g.*, a gene/inducible cassette) may be used in the method of the disclosure. It may be preferred that the method for inserting the gene/inducible cassette utilizes any one or more of zinc finger nucleases, TALENs and/or CRISPR/Cas9 systems or any derivatives thereof.

**[0052]** In certain aspects, the gene editing is performed by a CRISPR mechanism of gene editing. Three types of CRISPR mechanisms for gene editing have been identified, of which type II is the most studied. The type II CRISPR/Cas9 system utilizes the Cas9 nuclease to make a double-stranded break in DNA at a site determined by a short guide RNA. The CRISPR/Cas system is a prokaryotic immune system that confers resistance to foreign genetic elements. CRISPR are segments of prokaryotic DNA containing short repetitions of base sequences. Each repetition is followed by short segments of “protospacer DNA” from previous exposures to foreign genetic elements. CRISPR spacers recognize and cut the exogenous genetic elements using RNA interference. The CRISPR immune response occurs through two steps: CRISPR-RNA (crRNA) biogenesis and crRNA-guided interference. CrRNA molecules are composed of a variable sequence transcribed from the protospacer DNA and a CRISPR repeat. Each crRNA molecule then hybridizes with a second RNA, known as the trans-activating CRISPR RNA (tracrRNA) and together these two eventually form a complex with the nuclease Cas9. The protospacer DNA encoded section of the crRNA directs Cas9 to cleave complementary target DNA sequences if they are adjacent to short sequences known as protospacer adjacent motifs (PAMs). This natural system has been engineered and exploited to introduce double stranded breaks (DSBs) in specific sites in genomic DNA, amongst many other applications. In particular, the CRISPR type II system from *Streptococcus pyogenes* (*S. pyogenes* or *Sp*) may be used. At its simplest, the CRISPR/Cas9 system comprises two components that are delivered to the cell to provide genome editing: the Cas9 nuclease itself and a sgRNA. The sgRNA is a fusion of a customized, site-specific crRNA (directed to the target polynucleotide sequence) and a standardized tracrRNA. Once a DSB has been made, if a donor DNA template with homology to the targeted locus is supplied; the DSB may be repaired by the homology-directed repair (HDR) pathway allowing for precise insertions to be made.

**[0053]** Once the DSB has been made by any appropriate means, the donor polynucleotide sequence (*e.g.*, an exogenous gene) for insertion may be supplied in any suitable fashion as described below. The donor polynucleotide sequence and associated genetic material form the donor DNA for repair of the DNA at the DSB and are inserted using standard cellular repair machinery/pathways. How the break is initiated will alter which pathway is used to repair the damage, as noted above.

**[0054]** In certain aspects, the methods for increasing the accuracy of gene editing described herein comprise: (i) contacting the target polynucleotide sequence with a first ribonucleoprotein (RNP) comprising a first ribonucleic acid molecule and a clustered



regularly interspaced short palindromic repeat-associated (Cas) protein; wherein the first ribonucleic acid molecule comprises a sequence complementary to a nucleic acid sequence in the target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence; (ii) contacting the edited target polynucleotide sequence with a second or more RNP(s) comprising a second or more ribonucleic acid molecule(s) and a Cas protein to induce a second gene editing event; wherein the second or more ribonucleic acid molecule(s) comprises a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid molecule and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; wherein the method results in increased gene editing accuracy compared to the same method but consisting of only step (i). In certain embodiments, step (i) results in an error in one or more bases at the target polynucleotide sequence relative to the first ribonucleic acid molecule sequence. In certain embodiments, step (ii) is performed after step (i) results in an error in one or more bases at the target polynucleotide sequence relative to the first ribonucleic acid molecule sequence. In certain embodiments, steps (i) and (ii) occur simultaneously. In certain embodiments, steps (i) and (ii) occur sequentially.

**[0055]** In certain embodiments, the contacting comprises introducing one or more of the first ribonucleic acid molecule, the second (or more) ribonucleic acid molecule(s), the donor polynucleotide, and a polynucleotide encoding a Cas protein, to a cell. In certain embodiments, 2, 3, 4, 5, 6, 7, 8, 9, or 10 of the second or more ribonucleic acids are introduced. In certain embodiments, the introducing step is performed by at least one of transfection, transduction, electroporation, and microinjection. In certain embodiments, one or more ribonucleoprotein (RNP) complexes of Cas protein and ribonucleic acid (*e.g.*, sgRNA) are first generated and the RNP complexes are introduced to the cell. The one or more RNP complexes are introduced to the cell either simultaneously or sequentially.

**Alterations in target polynucleotide sequences by the first gene editing event**

**[0056]** In certain embodiments, the error in the one or more bases at the target polynucleotide sequence as a consequence of the first gene editing event is an error induced by double strand break repair. In certain embodiments, the error in the one or more bases at the target polynucleotide sequence is an alteration in one base relative to the first ribonucleic acid molecule sequence. In certain embodiments, the error in the one or more bases at the target

polynucleotide sequence comprises an insertion of one or more bases. In certain embodiments, the insertion of one or more bases comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more bases. In certain embodiments, the error in the one or more bases at the target polynucleotide sequence comprises a deletion of one or more bases. In certain embodiments, the deletion of one or more bases comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more bases. In certain embodiments, the error in the one or more bases at the target polynucleotide sequence comprises an insertion at the cut site created by double strand break repair.

**[0057]** In certain embodiments, the error in the one or more bases at the target polynucleotide sequence as a consequence of the first gene editing event occurs in 0.01-0.1%, 0.1-1.0%, 1.0-10%, 10-20%, 20-30%, 30-40% or 40-50% of the populations of cells subjected to the first gene editing event. In certain embodiments, more than one error or alteration type occurs in the population of cells after the first gene editing event.

### **Donor DNA**

**[0058]** In certain aspects, the method comprises integrating a donor polynucleotide sequence from the donor DNA into the target polynucleotide sequence. In certain embodiments, the donor polynucleotide sequence is configured for insertion into the genomic target sequence of a cell.

**[0059]** In certain embodiments, the donor DNA comprises a single-stranded oligonucleotide donor DNA (ssODN) sequence. In certain embodiments, the donor DNA comprises a double-stranded donor polynucleotide sequence.

**[0060]** In some embodiments, the donor DNA comprises AAVS1 (PPP1R12C gene), ROSA26 (THUMPD3-AS1 gene), CLYBL, DBI, PCSK9, or ZC3H3 gene sequences.

**[0061]** In some embodiments, the donor comprises a nucleic acid sequence having at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 70% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 75% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 80% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 85% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic

acid sequence having at least about 90% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 95% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 96% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 97% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 98% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having at least about 99% identity to SEQ ID NO: 1. In some embodiments, the donor comprises a nucleic acid sequence having 100% identity to SEQ ID NO: 1.

**[0062]** In some embodiments, the donor comprises a nucleic acid sequence having at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 70% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 75% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 80% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 85% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 90% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 95% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 96% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 97% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 98% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having at least about 99% identity to SEQ ID NO: 2. In some embodiments, the donor comprises a nucleic acid sequence having 100% identity to SEQ ID NO: 2.

**[0063]** In some embodiments, the donor comprises a nucleic acid sequence having at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ

ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 70% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 75% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 80% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 85% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 90% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 95% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 96% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 97% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 98% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having at least about 99% identity to SEQ ID NO: 3. In some embodiments, the donor comprises a nucleic acid sequence having 100% identity to SEQ ID NO: 3.

**[0064]** In certain embodiments, donor polynucleotide sequences from donor DNA comprising homology arms are integrated into the target polynucleotide sequence by homology-directed repair (HDR). Double-stranded donor DNA can comprise homology regions comprising one or more homology arms flanking the donor polynucleotide sequence to be integrated into the target polynucleotide sequence. Any design of donor DNA sequences known in the art for integration of donor DNA by homology directed repair can be used. For example, in certain embodiments, each of the homology regions is 0.8-1 kilobase pair (Kb), 15 bases – 1 Kb, 100-200 bases, 200-300 bases, 300-400 bases, 400-500 bases, 500-600 bases, 600-700 bases, 700-800 bases, 800-900 bases, 900-1000, or 1 Kb-2 Kb in length. In certain embodiments, the homology regions are complementary to the genomic target polynucleotide sequence, and the homology arms are complementary to nucleic acid sequences flanking the genomic target polynucleotide sequence of the cell.

**[0065]** In certain aspects, the donor DNA lacks homology arms flanking the sequence to be integrated into the target polynucleotide sequence. In certain embodiments, donor polynucleotide sequences from donor DNA without homology arms is integrated into the target polynucleotide sequence by a DNA repair mechanism (*e.g.*, non-homologous end joining).

**[0066]** In certain embodiments, the donor DNA is a plasmid that comprises: 1) a plasmid “backbone”, containing an antibiotic resistance gene and a bacterial origin of replication, and

2) a transgene comprising a coding sequence to be inserted in the target polynucleotide sequence. In certain embodiments, the transgene comprises 5' and 3' homology arms and a promoter driving the expression of the coding sequence. In certain embodiments, the coding sequence comprises a sequence that codes for one or more selectable markers. In certain embodiments, the coding sequence comprises a sequence that encodes a fluorescent marker (*e.g.*, EGFP).

**[0067]** In certain embodiments, the DNA plasmid is approximately 1 Kb - 10 Kb, 10 Kb - 20 Kb, 5 Kb - 10 Kb, 1 Kb - 5 Kb, 1 Kb - 2 Kb, 2 Kb - 3 Kb, 3 Kb - 4 Kb, 4 Kb - 5 Kb, 5 Kb - 6 Kb, 6 Kb - 7 Kb, 7 Kb - 8 Kb, 8 Kb - 9 Kb, 9 Kb - 10 Kb, 10 Kb - 15 Kb, or 15 Kb - 20 Kb or more in length.

**[0068]** In certain embodiments, the transgene is approximately 1 Kb - 10 Kb, 10 Kb - 20 Kb, 5 Kb - 10 Kb, 1 Kb - 5 Kb, 1 Kb - 2 Kb, 2 Kb - 3 Kb, 3 Kb - 4 Kb, 4 Kb - 5 Kb, 6 Kb - 6 Kb, 7 Kb - 8 Kb, 8 Kb - 9 Kb, 9 Kb - 10 Kb, 10 Kb - 15 Kb, or 15 Kb - 20 Kb or more in length.

**[0069]** In certain embodiments, the donor DNA comprises at least one exogenous gene to be integrated into the target polynucleotide sequence. In certain embodiments, the donor DNA comprises 1, 2, 3, 4, 5 or more exogenous genes. In certain embodiments, the donor DNA comprises one or more protein coding sequences. In certain embodiments, the donor polynucleotide encodes at least one transcription factor. In certain embodiments, the donor DNA comprises sequences encoding at least one functional version or variant of a protein (*e.g.*, a heterologous protein, or a T cell receptor), or a chimeric protein (*e.g.*, a chimeric antigen receptor). In some embodiments, a donor DNA includes regulatory sequences, for example, a promoter sequence and/or an enhancer sequence to regulate expression of the exogenous gene or fragment thereof, *e.g.*, after insertion into the genome of a cell.

**[0070]** In certain embodiments, the donor DNA comprises one or more non-protein coding sequences. In certain embodiments, the non-protein coding sequence is a non-coding RNA sequence. In certain embodiments, the non-coding RNA sequence comprises a sequence selected from the group consisting of one or more microRNAs (miRNAs), siRNAs (small interfering RNAs), piRNAs (Piwi-interacting RNAs), snoRNAs (small nucleolar RNAs), snRNAs (small nuclear RNAs), exRNAs (extracellular RNAs), scaRNAs (Small Cajal body-specific RNAs) and long ncRNAs (long non-coding RNAs).

**[0071]** Exogenous gene sequences can be between 100-200 bases in length, between 100-300 bases in length, between 100-400 bases in length, between 100-500 bases in length, between 100-600 bases in length, between 100-700 bases in length, between 100-800 bases in length, between 100-900 bases in length, or between 100-1000 bases in length. Exogenous sequences

can be between 100-2000 bases in length, between 100-3000 bases in length, between 100-4000 bases in length, between 100-5000 bases in length, between 100-6000 bases in length, between 100-7000 bases in length, between 100-8000 bases in length, between 100-9000 bases in length, or between 100-10,000 bases in length. Exogenous sequences can be between 1000-2000 bases in length, between 1000-3000 bases in length, between 1000-4000 bases in length, between 1000-5000 bases in length, between 1000-6000 bases in length, between 1000-7000 bases in length, between 1000-8000 bases in length, between 1000-9000 bases in length, or between 1000-10,000 bases in length.

**[0072]** Exogenous gene sequences can be greater than or equal to 10 bases in length, greater than or equal to 20 bases in length, greater than or equal to 30 bases in length, greater than or equal to 40 bases in length, greater than or equal to 50 bases in length, greater than or equal to 60 bases in length, greater than or equal to 70 bases in length, greater than or equal to 80 bases in length greater than or equal to 90 bases in length, or greater than or equal to 95 bases in length. Exogenous gene sequences can be between 1-100 bases in length, between 1-90 bases in length, between 1-80 bases in length, between 1-70 bases in length, between 1-60 bases in length, between 1-50 bases in length, between 1-40 bases in length, or between 1-30 bases in length. Exogenous gene sequences can be between 1-20 bases in length, between 2-20 bases in length, between 3-20 bases in length, between 5-20 bases in length, between 10-20 bases in length, or between 15-20 bases in length. Exogenous sequences can be between 1-10 bases in length, between 2-10 bases in length, between 3-10 bases in length, between 5-10 bases in length, between 1-5 bases in length, or between 1-15 bases in length. Exogenous gene sequences can be 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 115, 120, 125, 150, 175, 200, 225, 250, or more bases in length. Exogenous gene sequences can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 bases in length. Exogenous gene sequences can be greater than about 200 bp, 250 bp, 300 bp, 350 bp, 400 bp, 450 bp, 500 bp, 550 bp, 600 bp, 650 bp, 700 bp, 750 bp, 800 bp, 850 bp, 900 bp, 1 Kb, 1.1 Kb, 1.2 Kb, 1.3 Kb, 1.4 Kb, 1.5 Kb, 1.6 Kb, 1.7 Kb, 1.8 Kb, 1.9 Kb, 2.0 Kb, 2.1 Kb, 2.2 Kb, 2.3 Kb, 2.4 Kb, 2.5 Kb, 2.6 Kb, 2.7 Kb, 2.8 Kb, 2.9 Kb, 3 Kb, 3.1 Kb, 3.2 Kb, 3.3 Kb, 3.4 Kb, 3.5 Kb, 3.6 Kb, 3.7 Kb, 3.8 Kb, 3.9 Kb, 4.0 Kb, 4.1 Kb, 4.2 Kb, 4.3 Kb, 4.4 Kb, 4.5 Kb, 4.6 Kb, 4.7 Kb, 4.8 Kb, 4.9 Kb, 5.0 Kb, 5.1 Kb, 5.2 Kb, 5.3 Kb, 5.4 Kb, 5.5 Kb, 5.6 Kb, 5.7 Kb, 5.8 Kb, 5.9 Kb, 6.0 Kb, 6.1 Kb, 6.2 Kb, 6.3

Kb, 6.4 Kb, 6.5 Kb, 6.6 Kb, 6.7 Kb, 6.8 Kb, 6.9 Kb, 7.0 Kb or any size of template in between these sizes.

**[0073]** Donor DNA can further contain one or more additional spacer sequences between a donor polynucleotide sequence and an HDR arm or region. In some embodiments, a spacer sequence can have at least 2 nucleotides, *e.g.*, between 2 and 24 nucleotides (*e.g.*, between 2 and 22, between 2 and 20, between 2 and 18, between 2 and 16, between 2 and 14, between 2 and 12, between 2 and 10, between 2 and 8, between 2 and 6, between 2 and 4, between 4 and 24, between 6 and 24, between 8 and 24, between 10 and 24, between 12 and 24, between 14 and 24, between 16 and 24, between 18 and 24, between 20 and 24, between 22 and 24 nucleotides; 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, or 24 nucleotides).

**[0074]** In examples where multiple exogenous gene sequences are introduced, the multiple exogenous gene sequences can be different sizes, *e.g.*, a first exogenous gene sequence can be greater than or equal to 100 base pairs and a second gene exogenous sequence can be greater than or equal to 100 base pairs, or a first exogenous gene sequence can be greater than or equal to 100 base pairs and a second exogenous gene sequence can be less than 100 base pairs (*e.g.*, between 1-100 base pairs in length).

**[0075]** In certain embodiments, the donor DNA is a circular DNA plasmid. In some cases, the donor DNA is a double-stranded circular plasmid. In some cases, donor DNA is a single-stranded circular plasmid. In some cases, a plasmid donor DNA is a mini-circle plasmid. In some cases, a plasmid donor DNA is a nano-plasmid.

**[0076]** In some embodiments, the size or length of the donor DNA is greater than about 200 bp, 250 bp, 300 bp, 350 bp, 400 bp, 450 bp, 500 bp, 550 bp, 600 bp, 650 bp, 700 bp, 750 bp, 800 bp, 850 bp, 900 bp, 1 Kb, 1.1 Kb, 1.2 Kb, 1.3 Kb, 1.4 Kb, 1.5 Kb, 1.6 Kb, 1.7 Kb, 1.8 Kb, 1.9 Kb, 2.0 Kb, 2.1 Kb, 2.2 Kb, 2.3 Kb, 2.4 Kb, 2.5 Kb, 2.6 Kb, 2.7 Kb, 2.8 Kb, 2.9 Kb, 3 Kb, 3.1 Kb, 3.2 Kb, 3.3 Kb, 3.4 Kb, 3.5 Kb, 3.6 Kb, 3.7 Kb, 3.8 Kb, 3.9 Kb, 4.0 Kb, 4.1 Kb, 4.2 Kb, 4.3 Kb, 4.4 Kb, 4.5 Kb, 4.6 Kb, 4.7 Kb, 4.8 Kb, 4.9 Kb, 5.0 Kb, 5.1 Kb, 5.2 Kb, 5.3 Kb, 5.4 Kb, 5.5 Kb, 5.6 Kb, 5.7 Kb, 5.8 Kb, 5.9 Kb, 6.0 Kb, 6.1 Kb, 6.2 Kb, 6.3 Kb, 6.4 Kb, 6.5 Kb, 6.6 Kb, 6.7 Kb, 6.8 Kb, 6.9 Kb, 7.0 Kb, 7.1 Kb, 7.2 Kb, 7.3 Kb, 7.4 Kb, 7.5 Kb, 7.6 Kb, 7.7 Kb, 7.8 Kb, 7.9 Kb, 8.0 Kb, 8.1 Kb, 8.2 Kb, 8.3 Kb, 8.4 Kb, 8.5 Kb, 8.6 Kb, 8.7 Kb, 8.8 Kb, 8.9 Kb, 9.0 Kb, 9.1 Kb, 9.2 Kb, 9.3 Kb, 9.4 Kb, 9.5 Kb, 9.6 Kb, 9.7 Kb, 9.8 Kb, 9.9 Kb, 10.0 Kb, 15 Kb, 20 Kb, any length in between these sizes, or greater than 10 Kb. For example, the size of the donor DNA can be about 200 bp to about 500 bp, about 200 bp to about 750 bp, about 200 bp to about 1 Kb, about 200 bp to about 1.5 Kb, about 200 bp to

about 2.0 Kb, about 200 bp to about 2.5 Kb, about 200 bp to about 3.0 Kb, about 200 bp to about 3.5 Kb, about 200 bp to about 4.0 Kb, about 200 bp to about 4.5 Kb, about 200 bp to about 5.0 Kb, about 200 bp to about 10.0 kb, about 200 bp to about 15.0 Kb, or about 200 bp to about 20.0 Kb.

### Cas Protein

[0077] A Cas nuclease can direct cleavage of one or both strands at a location in a target polynucleotide sequence. Non-limiting examples of Cas nucleases include Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, Cpf1, homologs thereof, variants thereof, mutants thereof, and derivatives thereof. There are three main types of Cas nucleases (type I, type II, and type III), and 10 subtypes including 5 type I, 3 type II, and 2 type III proteins (see, e.g., Hochstrasser and Doudna, *Trends Biochem Sci*, 2015:40(1):58-66). Type II Cas nucleases include Cas1, Cas2, Csn2, Cas9, and Cfp1. These Cas nucleases are known to those skilled in the art. For example, the amino acid sequence of the *Streptococcus pyogenes* wild-type Cas9 polypeptide is set forth, e.g., in NCBI Ref. Seq. No. NP\_269215, and the amino acid sequence of *Streptococcus thermophilus* wild-type Cas9 polypeptide is set forth, e.g., in NCBI Ref. Seq. No. WP\_011681470.

[0078] Cas nucleases, e.g., Cas9 nucleases, can be derived from a variety of bacterial species including, but not limited to, *Veillonella atypical*, *Fusobacterium nucleatum*, *Filifactor alocis*, *Solobacterium moorei*, *Coprococcus catus*, *Treponema denticola*, *Peptoniphilus duerdenii*, *Catenibacterium mitsuokai*, *Streptococcus mutans*, *Listeria innocua*, *Staphylococcus pseudintermedius*, *Acidaminococcus intestine*, *Olsenella uli*, *Oenococcus kitaharae*, *Bifidobacterium bifidum*, *Lactobacillus rhamnosus*, *Lactobacillus gasseri*, *Finegoldia magna*, *Mycoplasma mobile*, *Mycoplasma gallisepticum*, *Mycoplasma ovipneumoniae*, *Mycoplasma canis*, *Mycoplasma synoviae*, *Eubacterium rectale*, *Streptococcus thermophilus*, *Eubacterium dolichum*, *Lactobacillus coryniformis* subsp. *Torquens*, *Ilyobacter polytropus*, *Ruminococcus albus*, *Akkermansia muciniphila*, *Acidothermus cellulolyticus*, *Bifidobacterium longum*, *Bifidobacterium dentium*, *Corynebacterium diphtheria*, *Elusimicrobium minutum*, *Nitratifactor salsuginis*, *Sphaerochaeta globus*, *Fibrobacter succinogenes* subsp. *Succinogenes*, *Bacteroides fragilis*, *Capnocytophaga ochracea*, *Rhodopseudomonas palustris*, *Prevotella micans*, *Prevotella*



*ruminicola*, *Flavobacterium columnare*, *Aminomonas paucivorans*, *Rhodospirillum rubrum*, *Candidatus Puniceispirillum marinum*, *Verminephrobacter eiseniae*, *Ralstonia syzygii*, *Dinoroseobacter shibae*, *Azospirillum*, *Nitrobacter hamburgensis*, *Bradyrhizobium*, *Wolinella succinogenes*, *Campylobacter jejuni* subsp. *Jejuni*, *Helicobacter mustelae*, *Bacillus cereus*, *Acidovorax ebreus*, *Clostridium perfringens*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria meningitidis*, *Pasteurella multocida* subsp. *Multocida*, *Sutterella wadsworthensis*, *proteobacterium*, *Legionella pneumophila*, *Parasutterella excrementihominis*, *Wolinella succinogenes*, and *Francisella novicida*.

**[0079]** Cas9 protein refers to an RNA-guided double-stranded DNA-binding nuclease protein or nickase protein. Wild-type Cas9 nuclease has two functional domains, *e.g.*, RuvC and HNH, that cut different DNA strands. Cas9 can induce double-strand breaks in genomic DNA (target DNA) when both functional domains are active. The Cas9 enzyme can comprise one or more catalytic domains of a Cas9 protein derived from bacteria belonging to the group consisting of *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flaviivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*, *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratifactor*, and *Campylobacter*. In some embodiments, a Cas9 protein can be a fusion protein, *e.g.*, the two catalytic domains are derived from different bacterial species.

**[0080]** In some embodiments, a Cas protein can be a Cas protein variant. For example, useful variants of the Cas9 nuclease can include a single inactive catalytic domain, such as a RuvC<sup>-</sup> or HNH<sup>-</sup>enzyme or a nickase. A Cas9 nickase has only one active functional domain and can cut only one strand of a cognate nucleic acid sequence, thereby creating a single strand break or nick. In some embodiments, a Cas9 nuclease can be a mutant Cas9 nuclease having one or more amino acid mutations. For example, a mutant Cas9 having at least a D10A mutation is a Cas9 nickase. In other embodiments, a mutant Cas9 nuclease having at least a H840A mutation is a Cas9 nickase. Other examples of mutations present in a Cas9 nickase include, without limitation, N854A and N863A. A double-strand break can be introduced using a Cas9 nickase if at least two DNA-targeting RNAs that target opposite DNA strands are used. A double-nicked induced double-strand break can be repaired by NHEJ or HDR (Ran *et al.*, 2013, Cell, 154:1380-1389). Non-limiting examples of Cas9 nucleases or nickases are described in, for example, U.S. Patent Nos. 8,895,308; 8,889,418; and 8,865,406 and U.S. Application Publication Nos. 2014/0356959, 2014/0273226 and 2014/0186919. The Cas9 nuclease or nickase can be codon-optimized for the target cell or target organism.

**[0081]** In some embodiments, a Cas protein variant lacks cleavage (*e.g.*, full cleavage or nickase) activity. A Cas protein variant may contain one or more point mutations that eliminates the protein's nickase activity. In some embodiments, Cas protein variants can be fused to other proteins and serve as targeting domains to direct the other proteins to the target nucleic acid. For example, Cas protein variants without cleavage activity may be fused to transcriptional activation (for CRISPR activation, or CRISPRa assays) or repression (for CRISPR inhibition or CRISPRi assays) domains to control gene expression (Ma *et al.*, *Protein and Cell*, 2(11):879-888, 2011; Maeder *et al.*, *Nature Methods*, 10:977-979, 2013; and Konermann *et al.*, *Nature*, 517:583-588, 2014). A Cas protein variant that lacks cleavage activity may be used to target genomic regions, resulting in RNA-directed transcriptional control. In some embodiments, a Cas protein variant without any cleavage activity may be used to target an exogenous protein to the target nucleic acid. An exogenous protein may be fused to the Cas protein variant. An exogenous protein may be an effector protein domain. An exogenous protein may be a transcription activator or repressor. Other examples of exogenous proteins include, but are not limited to, VP64-p65-Rta (VPR), VP64, P65, Krab, Ten-eleven translocation methylcytosine dioxygenase (TET), and DNA methyltransferase (DNMT). Specific Cas protein variants that lack cleavage (*e.g.*, nickase) activity are also described below.

**[0082]** In some embodiments, a Cas nuclease can be a high-fidelity or enhanced specificity Cas9 polypeptide variant with reduced off-target effects and robust on-target cleavage. Non-limiting examples of Cas9 polypeptide variants with improved on-target specificity include the SpCas9 (K855A), SpCas9 (K810A/K1003A/R1060A) (also referred to as eSpCas9(1.0)), and SpCas9 (K848A/K1003A/R1060A) (also referred to as eSpCas9(1.1)) variants described in Slaymaker *et al.*, *Science*, 351(6268):84-8 (2016), and the SpCas9 variants described in Kleinstiver *et al.*, *Nature*, 529(7587):490-5 (2016) containing one, two, three, or four of the following mutations: N497A, R661A, Q695A, and Q926A (*e.g.*, SpCas9-HF1 contains all four mutations).

**[0083]** In some embodiments, the Cas nuclease can also be a fusion of two or more proteins that contains a protein that can bind to a cognate nucleic acid sequence and a protein that can cleave the cognate nucleic acid sequence. For example, a protein that can recognize and bind to a cognate nucleic acid sequence can be a Cas protein variant without any cleavage activity. A Cas protein variant without any cleavage activity can be a Cas9 polypeptide that contains two silencing mutations of the RuvC1 and HNH nuclease domains (D10A and H840A), also referred to as dCas9 (Jinek *et al.*, *Science*, 2012, 337:816-821; Qi *et al.*, *Cell*, 152(5):1173-

1183). In one embodiment, the dCas9 polypeptide from *Streptococcus pyogenes* comprises at least one mutation at position D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, A987 or any combination thereof. Descriptions of such dCas9 polypeptides and variants thereof are provided in, for example, International Patent Publication No. WO 2013/176772. The dCas9 enzyme can contain a mutation at D10, E762, H983, or D986, as well as a mutation at H840 or N863. In some instances, the dCas9 enzyme can contain a D10A or D10N mutation. Also, the dCas9 enzyme can contain a H840A, H840Y, or H840N. In some embodiments, the dCas9 enzyme can contain D10A and H840A; D10A and H840Y; D10A and H840N; D10N and H840A; D10N and H840Y; or D10N and H840N substitutions. The substitutions can be conservative or non-conservative substitutions to render the Cas9 polypeptide catalytically inactive while still able to bind to a cognate nucleic acid sequence.

**[0084]** A Cas nuclease can also be fused with a localization peptide or protein. For example, a targetable nuclease can be fused with one or more nuclear localization signal (NLS) sequences, which can direct a targetable nuclease, and/or an RNP complex it forms, to the nucleus to modify a cognate nucleic acid sequence. Examples of NLS sequences are known in the art, *e.g.*, as described in Lange et al., *J Biol Chem.* 282(8):5101-5, 2007, and also include, but are not limited to, AVKRPAATKKAGQAKKKKLD, KRPAATKKAGQAKKKK, MSRRRKANPTKLSENAKKLAKEVEN, PAAKRVKLD, KLKIKRPVK, PKKKRKV, PKKKRRV and the NLS of nucleoplasmin. Examples of other peptides or proteins that can be fused to a targetable nuclease, such as cell-penetrating peptides and cell-targeting peptides are available in the art and described, *e.g.*, Vivès et al., *Biochim Biophys Acta.* 1786(2):126-38, 2008.

**[0085]** In certain aspects, the Cas protein forms a first or a second ribonucleoprotein (RNP) complex with an sgRNA. The RNP can contain the Cas protein nuclease and an sgRNA in a molar ratio of between 1:10 and 2:1 (*e.g.*, between 1:5 and 2:1, between 2:5 and 2:1, between 3:5 and 2:1, between 4:5 and 2:1, between 1:1 and 2:1, between 1:10 and 1:1, between 1:10 and 4:5, between 1:10 and 3:5, between 1:10 and 2:5, or between 1:10 and 1:5), respectively.

**[0086]** In certain embodiments the amount of Cas protein and donor DNA that is added to the cells can be donor in a molar ratio of Cas protein to donor DNA between 10:1 and 1000:1 (*e.g.*, between 50:1 and 1000:1, between 100:1 and 1000:1, between 200:1 and 1000:1, between 300:1 and 1000:1, between 400:1 and 1000:1, between 500:1 and 1000:1, between 600:1 and 1000:1, between 700:1 and 1000:1, between 800:1 and 1000:1, between 900:1 and 1000:1, between 10:1 and 900:1, between 10:1 and 800:1, between 10:1 and 700:1, between

10:1 and 600:1, between 10:1 and 500:1, between 10:1 and 400:1, between 10:1 and 300:1, between 10:1 and 200:1, between 10:1 and 100:1, or between 10:1 and 50:1), respectively.

### **gRNAs**

**[0087]** A Cas protein may be guided to the target polynucleotide nucleotide sequence to be cleaved by a single-guide RNA (sgRNA). An sgRNA is a version of the naturally occurring two-piece guide RNA (crRNA and tracrRNA) engineered into a single, continuous sequence. An sgRNA may contain a guide sequence (*e.g.*, the crRNA-equivalent portion of the sgRNA) that targets the Cas protein to the cognate nucleic acid sequence and a scaffold sequence that interacts with the Cas protein (*e.g.*, the tracrRNA-equivalent portion of the sgRNA). An sgRNA may be selected using a software tool. As a non-limiting example, considerations for selecting an sgRNA can include, *e.g.*, the PAM sequence for the Cas9 protein to be used, and strategies for minimizing off-target modifications. Tools, such as NUPACK® and the CRISPR Design Tool, can provide sequences for preparing the sgRNA, for assessing target modification efficiency, and/or assessing cleavage at off-target sites.

**[0088]** In certain embodiments, prior to performing the methods of this disclosure, the gRNAs (*e.g.*, sgRNAs) are designed to comprise a sequence complementary to a nucleic acid sequence in the target polynucleotide sequence to induce a gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence.

**[0089]** In some embodiments, the gRNA is encoded by any one of SEQ ID NOs: 4-11 or a sequence having at least 90%, 95%, 97%, 98%, or 99% sequence identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA is encoded by a sequence having at least about 80% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA is encoded by a sequence having at least about 85% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA is encoded by a sequence having at least about 90% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA is encoded by a sequence having at least about 95% identity to any one of SEQ ID NOs: 4-11. In some embodiments,

the gRNA is encoded by a sequence having at least about 96% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA is encoded by a sequence having at least about 97% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA is encoded by a sequence having at least about 98% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA is encoded by a sequence having at least about 99% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA is encoded by a sequence having 100% identity to any one of SEQ ID NOs: 4-11.

**[0090]** In some embodiments, the gRNA hybridizes or targets a sequence complementary to a target nucleic acid sequence within the AAVS1 gene or locus within an intron of the AAVS1 gene or locus. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a target nucleic acid sequence within the CLYBL gene or locus or within an intron of the CLYBL gene or locus. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a target nucleic acid sequence within the Diazepam binding inhibitor (DBI) gene or locus or within an intron of the DBI gene or locus. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a target nucleic acid sequence within the proprotein convertase subtilisin/kexin type 9 (PCSK9) gene or locus or within an intron of the PCSK9 gene or locus. In some embodiments, the gRNA hybridizes or targets a sequence complementary to any one of SEQ ID NOs: 4-11 or a sequence having at least 90%, 95%, 97%, 98%, or 99% sequence identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a sequence having at least about 80% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a sequence having at least about 85% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a sequence having at least about 90% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a sequence having at least about 95% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a sequence having at least about 96% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a sequence having at least about 97% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a sequence having at least about 98% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA hybridizes or targets a sequence complementary to a sequence having at least about 99% identity to any one of SEQ ID NOs: 4-11. In some embodiments, the gRNA

hybridizes or targets a sequence complementary to a sequence having 100% identity to any one of SEQ ID NOs: 4-11.

**[0091]** In some embodiments, the guide sequence of a gRNA (*e.g.*, an sgRNA) may comprise about 10 to about 2000 nucleotides, for example, about 10 to about 100 nucleotides, about 10 to about 500 nucleotides, about 10 to about 1000 nucleotides, about 10 to about 1500 nucleotides, about 10 to about 2000 nucleotides, about 50 to about 100 nucleotides, about 50 to about 500 nucleotides, about 50 to about 1000 nucleotides, about 50 to about 1500 nucleotides, about 50 to about 2000 nucleotides, about 100 to about 500 nucleotides, about 100 to about 1000 nucleotides, about 100 to about 1500 nucleotides, about 100 to about 2000 nucleotides, about 500 to about 1000 nucleotides, about 500 to about 1500 nucleotides, about 500 to about 2000 nucleotides, about 1000 to about 1500 nucleotides, about 1000 to about 2000 nucleotides, or about 1500 to about 2000 nucleotides at the 5' end of the gRNA that can direct the Cas protein to a target polynucleotide sequence using RNA-DNA complementarity base pairing. In some embodiments, the guide sequence of a gRNA comprises about 100 nucleotides at the 5' end of the gRNA that can direct the Cas protein to a cognate nucleic acid sequence site using RNA-DNA complementarity base pairing. In some embodiments, the guide sequence comprises 20 nucleic acids at the 5' end of the gRNA that can direct the Cas protein to a site of the target polynucleotide sequence using RNA-DNA complementarity base pairing. In other embodiments, the guide sequence comprises less than 20, *e.g.*, 19, 18, 17 or less, nucleotides that are complementary to a cognate nucleic acid sequence. In some instances, the guide sequence in the sgRNA contains at least one nucleic acid mismatch in the complementarity region of a cognate nucleic acid sequence. In some instances, the guide sequence contains from about 1 to about 10 nucleic acid mismatches in the complementarity region of a cognate nucleic acid sequence.

**[0092]** In certain embodiments, the gRNAs comprise a sequence complementary to the target polynucleotide sequence 10-50 nucleotides in length, 10-20 nucleotides in length, 20-30 nucleotides in length, 10-15 nucleotides in length, 15-20 nucleotides in length, or 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides in length by RNA-DNA complementarity base pairing.

**[0093]** The scaffold sequence in a gRNA may serve as a protein-binding sequence that interacts with the Cas protein or a variant thereof. In some embodiments, the scaffold sequence in an sgRNA can comprise two complementary stretches of nucleotides that hybridize to one another to form a double-stranded RNA duplex (dsRNA duplex). The scaffold sequence may have structures such as lower stem, bulge, upper stem, nexus, and/or

hairpin. In some embodiments, the scaffold sequence in the sgRNA can be between about 90 nucleotides to about 120 nucleotides, *e.g.*, about 90 nucleotides to about 115 nucleotides, about 90 nucleotides to about 110 nucleotides, about 90 nucleotides to about 105 nucleotides, about 90 nucleotides to about 100 nucleotides, about 90 nucleotides to about 95 nucleotides, about 95 nucleotides to about 120 nucleotides, about 100 nucleotides to about 120 nucleotides, about 105 nucleotides to about 120 nucleotides, about 110 nucleotides to about 120 nucleotides, or about 115 nucleotides to about 120 nucleotides.

**[0094]** In certain embodiments, the second or more ribonucleic acid molecule(s) (*e.g.*, second sgRNA) comprises a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid molecule and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence. Thus, in certain embodiments of the methods, the base alterations are determined in the target polynucleotide sequence of the population of target cells after the first gene editing event, and the second or more ribonucleic acid molecule(s) is designed to comprise the one or more base alterations identified.

#### **Genomic Safe Harbor Sites**

**[0095]** In certain aspects, the target polynucleotide sequence is a genomic safe harbor site (GSH) locus. In certain embodiments, the GSH locus is selected from the group consisting of a GSH locus selected from the group consisting of AAVS1 (PPP1R12C gene), ROSA26 (THUMPD3-AS1 gene), CLYBL, DBI, PCSK9 and ZC3H3. In certain embodiments, the AAVS1 locus is altered by an insertion of one base after the first gene editing event. In certain embodiments, the insertion is an adenine or thymidine. In certain embodiments, the CLYBL locus is altered by an insertion of one base. In certain embodiments, the insertion is an adenine or thymidine. In certain embodiments, the DBI locus is altered by an insertion of one base. In certain embodiments, the insertion is an adenine or thymidine. In certain embodiments, the PCSK9 locus is altered by an insertion of one base. In certain embodiments, the insertion is an adenine or thymidine.

#### **Integration of Donor Polynucleotide Sequences**

**[0096]** In certain embodiments, the methods described herein result in integration of at least one exogenous gene from donor DNA into the target polynucleotide sequence. In certain embodiments, the methods result in integration of 1, 2, 3, 4, 5 or more exogenous genes. In

certain embodiments, the methods result in integration of one or more protein coding sequences. In certain embodiments, the methods result in integration of one or more non-protein coding sequences. In certain embodiments, the methods result in integration of a sequence selected from the group consisting of one or more microRNAs, siRNAs, piRNAs, snoRNAs, snRNAs, exRNAs, scaRNAs and long ncRNAs.

**[0097]** In certain aspects, the increased accuracy of gene editing of the methods described herein comprises increased integration of donor DNA by Homology Dependent Repair (HDR)-mediated integration ranges 1- to 5-fold, 1- to 2-fold, 1- to 1.1-fold, 1.1- to 1.2-fold, 1.2- to 1.3-fold, 1.3- to 1.4-fold, 1.4- to 1.5-fold, 1.5- to 1.6-fold, 1.6- to 1.7-fold, 1.8- to 1.9-fold, 1.9- to 2.0-fold, 2- to 3-fold, 3- to 4-fold or 4- to 5-fold compared to the same method but consisting of only step (i). In certain embodiments, the increased accuracy of gene editing of the methods described herein comprises increased integration of donor DNA by Homology Dependent Repair (HDR)-mediated integration from about 47% to about 96% compared to the same method but consisting of only step (i).

**[0098]** In some embodiments of the methods disclosed herein, the integration of the donor polynucleotide sequence into the target polynucleotide sequence is induced in greater than about 5% of the population of cells (*e.g.*, the population of primary cells), *e.g.*, about 6%, about 7%, about 8%, about 9%, about 10%, about 12%, about 14%, about 16%, about 18%, about 20%, about 22%, about 24%, about 26%, about 28%, about 30%, about 32%, about 34%, about 36%, about 38%, about 40% or about 50% of the population of cells. In some embodiments, the integration of the donor polynucleotide sequence into the target polynucleotide sequence is induced in greater than about 50% of the population of cells (*e.g.*, the population of primary cells), *e.g.*, about 51%, about 52%, about 53%, about 54%, about 55%, about 56%, about 57%, about 58%, about 59%, about 60%, about 61%, about 62%, about 63%, about 64%, about 65%, about 66%, about 67%, about 68%, about 69%, about 70%, about 71%, about 72%, about 73%, about 74%, about 75%, about 76%, about 77%, about 78%, about 79%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or about 100% of the population of cells. In other embodiments, the integration of the donor polynucleotide sequence into the target polynucleotide sequence is induced in greater than about 70% of the population of cells, *e.g.*, about 71%, about 72%, about 73%, about 74%, about 75%, about 76%, about 77%, about 78%, about 79%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%,



about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or about 100% of the population of cells. In yet other embodiments, the integration of the donor polynucleotide sequence into the target polynucleotide sequence is induced in greater than about 90% of the population of cells, *e.g.*, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or about 100% of the population of cells.

**[0099]** In certain embodiments, the integration of the donor polynucleotide sequence into the target polynucleotide sequence comprises the replacement of a genetic mutation in the target nucleic acid (*e.g.*, to correct a point mutation or a single nucleotide polymorphism (SNP) in the target nucleic acid that is associated with a disease) or the insertion of an open reading frame (ORF) comprising a normal copy of the target nucleic acid (*e.g.*, to knock in a wild-type cDNA of the target nucleic acid that is associated with a disease).

**[0100]** In certain embodiments, integration of the donor polynucleotide sequences is detected by expression of a gene encoded by the donor polynucleotide sequence that has been integrated into the targeted locus. Detection of gene expression in cells comprising genomes with integrated donor polynucleotide sequences can be performed by any method known in the art to detect gene expression. In certain embodiments, the expression of the genes (*e.g.*, reporter gene) is detected by flow cytometry. For example, flow cytometry can be used to detect the expression of a fluorescent reporter expressed from the targeted locus, or cells stained with antibodies fused to fluorescent tags.

### **Cells**

**[0101]** In certain aspects, the methods and compositions described herein comprise increasing the accuracy of gene editing in a cell or population of cells, *e.g.*, a eukaryotic cell, prokaryotic cell, animal cell, plant cell, fungal cell, and the like. Optionally, the cell is a mammalian cell, for example, a human cell. The cell can be *in vitro*, *ex vivo*, or *in vivo*. The cell can also be a primary cell, a germ cell, a stem cell, or a precursor cell. The precursor cell can be, for example, a pluripotent stem cell, or a hematopoietic stem cell. In some embodiments, the cell is a primary hematopoietic cell, a primary hematopoietic stem cell, or a primary T cell. In certain embodiments, the cell comprises an induced pluripotent stem (iPS) cell. In certain embodiments, the iPS cell is a mammalian iPS cell. In certain embodiments, the cell is a human iPS (hiPS) cell. In certain embodiments, the cell is a differentiated cell. In certain embodiments, the cell is an immune cell, a myeloid cell, a neuronal cell, an adipocyte, a hepatocyte, a pancreatic cell, an epithelial cell, a muscle cell (including skeletal muscle cell

and a smooth muscle cell), a cardiomyocyte, a bone cell, a skin cell or a blood cell. In certain embodiments, the cell is a T cell, a B cell, a dendritic cell, a macrophage or an NK cell. In certain embodiments, the cell comprises a neuronal cell selected from the group consisting of a microglial cell, motor neuron, dopaminergic neuron, GABAergic neuron, and a glutamatergic neuron. In some embodiments, the population of primary cells comprises a heterogeneous population of primary cells. In other embodiments, the population of primary cells comprises a homogeneous population of primary cells.

**[00102]** In some embodiments, the primary cell is isolated from a mammal prior to introducing a composition described herein into the primary cell. For instance, the primary cell can be harvested from a human subject. In some instances, the primary cell or a progeny thereof is returned to the mammal after introducing the composition described herein into the primary cell. In other words, the genetically modified primary cell undergoes autologous transplantation. In other instances, the genetically modified primary cell undergoes allogeneic transplantation. For example, a primary cell that has not undergone stable gene modification is isolated from a donor subject, and then the genetically modified primary cell is transplanted into a recipient subject who is different than the donor subject.

**[00103]** In certain embodiments, the cell is reprogrammed after the second gene editing event. As used herein, a cell is “reprogrammed” when genetic alteration of the cell causes the cell to change into a different cell type. In certain embodiments, reprogramming results in differentiation of a stem cell into a mature cell type. In certain embodiments, reprogramming results in de-differentiation of a mature cell to a pluripotent stem cell or progenitor cell. In certain embodiments, reprogramming involves the forced expression of one or more key lineage transcription factor(s) and/or one or more non-coding RNA(s) in order to convert a stem cell into a particular mature cell type.

**[00104]** In certain embodiments, the cell expresses a therapeutic protein after the second gene editing event. For example, the cell can express a functional version or variant of a protein, a chimeric protein (*e.g.*, a chimeric antigen receptor), or a therapeutic RNA after the second gene editing event.

**[00105]** In certain embodiments, disclosed herein are populations of cells comprising:

i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;

ii) a second or more ribonucleic acid molecule(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence of step (i); and optionally,

iii) a donor DNA comprising the target polynucleotide sequence; and optionally,

iv) a polynucleotide comprising a sequence encoding a Cas protein.

**[00106]** In certain embodiments, the cells comprise an increased percentage of the donor polynucleotide integrated into the target polynucleotide sequence compared to a population of the same type of cells comprising the first ribonucleic acid, the target polynucleotide sequence, and the donor polynucleotide, but not the second ribonucleic acid. In certain embodiments, the cells further comprise Cas protein. In certain embodiments, the cells are reprogrammed to a differentiated cell after the first or second gene editing event mediated by the first or second ribonucleic acid, a Cas protein and the donor polynucleotide.

### **Nucleic acids**

**[00107]** The term percent “identity,” in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that have a specified percentage of nucleotides or amino acid residues that are the same, when compared and aligned for maximum correspondence, as measured using one of the sequence comparison algorithms described below (*e.g.*, BLASTP and BLASTN or other algorithms available to persons of skill) or by visual inspection. Depending on the application, the percent “identity” can exist over a region of the sequence being compared, *e.g.*, over a functional domain, or, alternatively, exist over the full length of the two sequences to be compared.

**[00108]** For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

**[00109]** Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the

homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by visual inspection (*see generally* Ausubel *et al.*, *infra*).

**[00110]** One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the Basic Local Alignment Search Tool ("BLAST") algorithm, which is described in Altschul *et al.*, J. Mol. Biol. 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information ([www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)).

**[00111]** In certain embodiments, described herein are a plurality of polynucleotides comprising:

i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;

ii) a second or more ribonucleic acid molecule(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; and optionally,

iii) a donor DNA comprising the target polynucleotide sequence; and optionally,

iv) a polynucleotide comprising a sequence encoding Cas protein.

**[00112]** The nucleic acids described herein for performing the methods of this disclosure can be in the form of a vector (*e.g.*, a plasmid DNA), genomic DNA, single stranded DNA or double stranded DNA, or any suitable form known in the art to support the induction of a gene editing event. The nucleic acids for inducing the first and/or second gene editing event(s) may be introduced in one or more vectors, such as plasmids, for expression in the cell.

### **Kits**

**[00113]** In certain aspects, described herein are kits comprising:

i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;

ii) a second or more ribonucleic acid(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid(s) comprises a sequence complementary to the edited target polynucleotide sequence; and optionally,

iii) a donor DNA comprising the target polynucleotide sequence; and optionally,

iv) a polynucleotide comprising a sequence encoding Cas protein; and

v) instructions for use.

**[00114]** In certain aspects, the kits are used for performing the methods described herein. In certain aspects, the kits are used to increase accuracy and/or efficiency of integration of one or more donor polynucleotide sequences into the genome of a target cell described herein.

## EXAMPLES

**[00115]** Below are examples of specific embodiments for carrying out the present disclosure. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present disclosure in any way. Efforts have been made to ensure accuracy with respect to numbers used (*e.g.*, amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

**[00116]** The practice of the present disclosure will employ, unless otherwise indicated, conventional methods of protein chemistry, biochemistry, recombinant DNA techniques and pharmacology, within the skill of the art. Such techniques are explained fully in the literature. *See, e.g.*, T.E. Creighton, *Proteins: Structures and Molecular Properties* (W.H. Freeman and Company, 1993); A.L. Lehninger, *Biochemistry* (Worth Publishers, Inc., current addition); Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); *Methods In Enzymology* (S. Colowick and N. Kaplan eds., Academic Press, Inc.); *Remington's Pharmaceutical Sciences*, 18th Edition (Easton, Pennsylvania: Mack Publishing Company, 1990); Carey and Sundberg *Advanced Organic Chemistry 3<sup>rd</sup> Ed.* (Plenum Press) Vols A and B(1992).

**Example 1: Methods for Genome Engineering**

[00117] Each ribonucleoprotein (RNP) complex was generated by mixing and incubating (10 min at room temperature) 2.5 µg of purified Cas9 protein with 2 µg of synthetic modified single guide RNA (sgRNA). Once formed, RNP combinations for scEditing were mixed at a 1:1 ratio.

[00118] sgRNAs used for the four exemplary GSH sites are shown in Table 1 below:

<b>Table 1:</b> Exemplary gRNA sequences used for four genomic safe harbour sites <b>GSH</b>	<u>gRNA</u>	<u>Sequence</u>	<u>sc- gRNA</u>	<u>Sequence</u>
AAVS1	CG05	UCCCUAGUGGCCCCACUGUG (SEQ ID NO: 4)	CG49	UCCCUAGUGGCCCCACUUGU (SEQ ID NO: 5)
CLYBL	CG65	AAUAUGGAAGUUAACAGUA (SEQ ID NO: 6)	CG70	AAUAUGGAAGUUAACAAGUA (SEQ ID NO: 7)
DBI	CG63	CAUACCCCACCUGAAGAGUA (SEQ ID NO: 8)	CG68	CAUACCCCACCUGAAGAAGU (SEQ ID NO: 9)
PCSK9	CG55	GUGCCCACGGGACUGAUCAG (SEQ ID NO: 10)	CG69	UGCCCACGGGACUGAUUCAG (SEQ ID NO: 11)

[00119] BobC induced pluripotent stem cells (iPSC), expressing the co-transcriptional activator rtTA from the *ROSA26* locus, were dissociated into a single-cell suspension and washed with DPBS. For each condition, 500,000 live cells were prepared in 20 µl of P3 transfection buffer. RNP (one RNP or two mixed RNP for scEditing) and reporter donor DNA plasmid (final concentration of 150 ng/µl, 3 µg per transfection) were added to the cells right before transfection. Transfections were carried out with 16-well strips (one well per condition). Transfected cells recovered in CloneR for 48 hrs before being cultured in standard iPSC conditions.

[00120] *Reporter donor DNA plasmids*

[00121] The structure of the plasmids used to assess scEditing at different genome safe harbour (GSH) sites: [5' homology arm]-[splice-acceptor in frame with a puromycin selection cassette]-[TRE3GV inducible promoter]-[EGFP reporter cassette]-[3' homology arm].

[00122] Each homology arm, mapping to either side (5' and 3') of each GSH CRISPR-Cas9 target site, is approximately 1Kb long. The plasmid backbone originates from pUC18 (Ori, AmpR).

[00123] *Flow cytometry*

[00124] At least 6 days after transfection, the cells were dissociated, washed with DPBS, and stained with a Fixable Live-Dead stain. The cells were analysed by flow cytometry to characterise the proportion of live EGFP-expressing cells.

**Example 2: Increased integration of donor DNA at AAVS1 Genomic Safe Harbor Site by Second Chance Editing**

[00125] Donor plasmid was engineered to target the AAVS1 GSH site which generates an insertion of a single Thymidine base upon repair in the iPSCs when insertion does not occur. Figure 1A shows the sequence at the AAVS1 target site for the gRNAs CG5 and CG49 for second chance editing. The percentage of cells expressing EGFP corresponding to the percent of cells that harbour the inserted donor plasmid DNA encoding the reporter gene when introduced with either a single CG5 gRNA or both the CG5 and CG49 gRNAs is shown in Figure 1B. There was an increase in knock-in efficiency from 25% to 62% both CG5 and CG49 gRNAs were introduced as opposed to only CG5.

**Example 3: Increased integration of donor DNA at CLYBL Genomic Safe Harbor Site by Second Chance Editing**

[00126] Donor plasmid was engineered to target the CLYBL GSH site which generates an insertion of a single thymidine in the iPSCs base upon repair when insertion does not occur. Figure 2A shows the sequence at the CLYBL target site for the gRNAs CG65 and CG70 for second chance editing. The percentage of cells expressing EGFP corresponding to the percent of cells that harbour the inserted donor plasmid DNA encoding the reporter gene when introduced with either a single CG65 gRNA or both the CG65 and CG70 gRNAs is shown in Figure 2B. There was an increase in knock-in efficiency from 21.6% to 37.2% both CG65 and CG70 gRNAs were introduced as opposed to only CG65.

**Example 4: Increased integration of donor DNA at DBI Genomic Safe Harbor Site by Second Chance Editing**

[00127] Donor plasmid was engineered to target the DBI GSH site which generates an insertion of a single thymidine base in the iPSCs upon repair when insertion does not occur. Figure 3A shows the sequence at the DBI target site for the gRNAs CG63 and CG68 for second chance editing. The percentage of cells expressing EGFP corresponding to the percent of cells that harbour the inserted donor plasmid DNA encoding the reporter gene when introduced with either a single CG63 gRNA or both the CG63 and CG68 gRNAs is shown in Figure 3B. There was an increase in knock-in efficiency from 23.9% to 33.7% both CG63 and CG68 gRNAs were introduced as opposed to only CG63.

**Example 5: Increased integration of donor DNA at PCSK9 Genomic Safe Harbor Site by Second Chance Editing**

[00128] Donor plasmid was engineered to target the PCSK9 GSH site which generates an insertion of a single adenine base in the iPSCs upon repair when insertion does not occur. Figure 4A shows the sequence at the DBI target site for the gRNAs CG55 and CG69 for second chance editing. The percentage of cells expressing EGFP corresponding to the percent of cells that harbour the inserted donor plasmid DNA encoding the reporter gene when introduced with either a single CG553 gRNA or both the CG55 and CG69 gRNAs is shown in Figure 4B. There was an increase in knock-in efficiency from 24.0% to 28.9% both CG55 and CG69 gRNAs were introduced as opposed to only CG55.

**Table 2: Summary of increased Kock-in efficiency at different GSH sites**

GSHs	gRNA	Genome Editing	Knock-In (fluorescent reporters)
AAVS1	CG5	98% (+/- 1)	35% (+/- 11) - sc editing (67% +/- 6)
CLYBL	CG22	61% (+/- 10)	21% (+/- 3)- sc editing (37.2%)
DBI	CG63	98% (+/- 1)	21% (+/- 10) - sc editing (38% +/- 5)



<b>PCSK9</b>	CG55	97% (+/- 1)	<b>19% (+/- 7) - sc editing (28% (+/- 1))</b>
--------------	------	-------------	---

\*Values represent at least two replicates per GSH

**[00129] Example 6: Functional genomics screening of transgenes integrated at target GSHs**

A study is conducted to screen libraries of barcoded transgenes (DNA donors) targeted to designated GSHs (e.g., AAVS1 and CLYBL), in order to examine transgene function at a high throughput scale. Increased targeted integration of transgene cassettes is demonstrated in targeted cell pools using scEditing. This increased integration efficiency provides the opportunity to screen less cells and generate more complex cell pools compared to those generated by classical CRISPR-Cas methods. These complex pools harbor one or more transgenes, which can be subsequently used for high throughput functional genomic screening.

**Example 6: Genetic cell engineering with large DNA cargos integrated at target GSHs**

**[00130]** A study is conducted to generate clonal cell lines carrying DNA cargos in designated GSHs. Increased targeted integration frequency of DNA cargos is demonstrated in targeted cell pools using scEditing. This increased integration frequency provides the opportunity to genotype less clonal cell populations and therefore generate a higher number of distinct cell lines in parallel with the same genetic cell engineering process. This increased integration frequency also increases the probability to obtain homozygous integrations (both target alleles with DNA cargo integrated) and more complex clonal populations of cells with multiple DNA cargos integrated at different GSHs, enabling the generation of cells for elaborate functional and cell-based assays.

**[00131]** While the disclosure has been particularly shown and described with reference to a preferred embodiment and various alternate embodiments, it will be understood by persons skilled in the relevant art that various changes in form and details can be made therein without departing from the spirit and scope of the disclosure.

**[00132]** All references, issued patents and patent applications cited within the body of the instant specification are hereby incorporated by reference in their entirety, for all purposes.

SEQUENCE LISTING

SEQ ID NO	DESCRIPTION	SEQUENCE
1	Donor Plasmid DNA-AAVS1	TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACAC ATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGC GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCA GCGGGTGTGGCGGGGTGTCGGGGCTGGCTTAACTATGC GGCATCAGAGCAGATTGTAAGTGTGAGAGTGCACCATATG CGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAT ACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAAC TGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATT ACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGA TTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACG TTGTAACGACGGCCAGTGAATTGTGCTTTCTCTGAC CAGCATTCTCTCCCCTGGGCCTGTGCCGCTTTCTGTCTG TAGCTTGTGGCCTGGGTACCTCTACGGCTGGCCAGA TCCTTCCCTGCCGCTCCTTCAGGTTCCGTCTTCTCCA CTCCCTCTTCCCCTTGCTCTCTGCTGTGTTGCTGCCAA GGATGCTCTTCCGGAGCACTTCCTTCTCGGCCTGCA CCACGTGATGTCCTCTGAGCGGATCCTCCCCGTGTCTG GGTCTCTCCGGGCATCTCTCCTCCCTACCCAACCCC ATGCCGTCTTCACTCGCTGGGTTCCCTTTCTTCTCCT TCTGGGGCCTGTGCCATCTCTCGTTTCTTAGGATGGCCT TCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCTTC TTGTAGGCCTGCATCATCACCGTTTTTCTGGACAACCC CAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCA TCCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGT GGATTCCGGTCACCTCTCACTCCTTTCATTTGGGCAGCT CCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTT CCAGCCCCCTGTCATGGCATCTTCCAGGGGTCCGAGAG CTCAGCTAGTCTTCTTCCCTCCAACCCGGGCCCTATGTC CACTTCAGGACAGCATGTTTGTGCTGCCAGGGATCCT GTGTCCCGAGCTGGGACCACCTTATATTCCAGGGCC GGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCC CCTCCACCCACAGTGGGGCAAGCTTCTGACCTCTTCT CTTCTCCACAGGGCCTCGAGAGATCTGGCAGCGGAG AGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTGGA GGAGAATCCCGGCCCTAGGCTCGAGATGACCGAGTAC AAGCCACGGTGCGCCTCGCCACCCGCGACGACGTCCC CAGGGCCGTACGCACCCTCGCCGCCGCTTCGCCGACT ACCCCGCCACGCGCCACACCGTCGATCCGGACCGCCAC ATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCTCAC GCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTGCGG GACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGG AGAGCGTCGAAGCGGGGGCGGTGTTCCGCCGAGATCGG CCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCG CGCAGCAACAGATGGAAGGCCTCCTGGCGCCGCACCG GCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCG

		<p>TCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGC CGTCGTGCTCCCCGGAGTGGAGGGCGGCCGAGCGCGCC GGGGTGCCCGCCTTCTGGAGACCTCCGCGCCCCGCAA CCTCCCCTTCTACGAGCGGCTCGGCTTACCGTCACCG CCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTG CATGACCCGCAAGCCCGGTGCCTGATCTAGAGGGCCC GTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAG TTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTT CTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCT AATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAG GTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACA GCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGC TGGGGATGCGGTGGGCTCTATGGTCACTCGAGTTTACT CCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCC CTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCT ATCAGTGATAGAGAACGTATAAGGAGTTTACTCCCTAT CAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCA GTGATAGAGAACGTATCTACAGTTTACTCCCTATCAGT GATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGA TAGAGAACGTATGTCGAGGTAGGCGTGTACGGTGGGC GCCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGAT CGCCTGGAGCAATCCACAACACTTTTGTCTTATACTT ACTAGTGCCACCATGGTGAGCAAGGGCGAGGAGCTGT TCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGC GACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGG GCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAA GTTTATCTGCACCACCGGCAAGCTGCCCGTGCCTGGC CCACCCTCGTGACCACCTGACCTACGGCGTGCAGTGC TTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTT CTTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGC GCACCATCTTCTTCAAGGACGACGGCAACTACAAGACC CGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGA ACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA CGGCAACATCCTGGGGCACAAGCTGGAGTACA ACTAC AACAGCCACAACGTCTATATCATGGCCGACAAGCAGA AGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAA CATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTAC CAGCAGAACACCCCATCGGGCAGCGCCCCGTGCTGCT GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGA GCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCT GCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCA TGGACGAGCTGTACAAGTAAGAATTGAGCTCGGTAC CCGGGGATCCTCTAGTCAGCTGACGCGTGCTAGCGCGG CCGCATCGATAAGCTTGTGACGATATCTCTAGAGGAT CATAATCAGCCATACCACATTTGTAGAGGTTTTACTTG CTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAA CATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATT GCAGCTTATAATGGTTACAAATAAAGCAATAGCATCAC AAATTCACAAATAAAGCATTTTTTTTCACTGCCTTGAC AGTACTCTTAAGTCGACTTACTAGGGACAGGATTGGTG</p>
--	--	--

		<p>ACAGAAAAGCCCCATCCTTAGGCCTCCTCCTCCTAGT CTCCTGATATTGGGTCTAACCCCCACCTCCTGTTAGGC AGATTCCTTATCTGGTGACACACCCCCATTCCTGGAG CCATCTCTCCTTGCCAGAACCTCTAAGGTTTGCTTAC GATGGAGCCAGAGAGGATCCTGGGAGGGAGAGCTTGG CAGGGGGTGGGAGGGAAGGGGGGGATGCGTGACCTGC CCGGTTCTCAGTGGCCACCCTGCGCTACCCTCTCCCAG AACCTGAGCTGCTCTGACGCGGCTGTCTGGTGCCTTC ACTGATCCTGGTGCTGCAGCTTCCTTACACTTCCAAG AGGAGAAGCAGTTTGGAAAAACAAAATCAGAATAAGT TGGTCCTGAGTTCTAACTTTGGCTCTTCACCTTTCTAGT CCCCAATTTATATTGTTCCCTCCGTGCGTCAGTTTACCT GTGAGATAAGGCCAGTAGCCAGCCCCGTCCTGGCAGG GCTGTGGTGAGGAGGGGGGTGTCCGTGTGGAAAACCTC CCTTTGTGAGAATGGTGCGTCCTAGGTGTTACCAGGT CGTGGCCGCCTCTACTCCCTTTCTCTTTCTCCATCCTTC TTTCCTTAAAGAGTCCCCAGTGCTATCTGGGACATATT CCTCCGCCAGAGCAGGGTCCCGCTTCCCTAAGGCCCT GCTCTGGGCTTCTGGGTTTGAGTCCTTGGCAAGCCCAG GAGAGGCGCTCAGGCTTCCCTGTCCCCCTTCCCTCGTCC ACCATCTCATGCCCTGGCTCTCCTGCCCTTCCCTACA GGGGTTCCTGGCTCTGCTCTAAGCTTGGCGTAATCATG GTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCAC AATTCCACACAACATACGAGCCGGAAGCATAAAGTGT AAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATT AATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAA ACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC GCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC TTCCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCT GCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATA CGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG TAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCC GCCCCCTGACGAGCATCACAAAATCGACGCTCAAG TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT TCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAC GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCC AAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCCGA CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCA ACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGC CACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGG CTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGC TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCT TGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC TGACGCTCAGTGGAACGAAAACCTACGTTAAGGGATT</p>
--	--	---

		<p>TGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC CTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAG TATATATGAGTAAACTTGGTCTGACAGTTACCAATGCT TAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCA GATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCG AGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAG TTCGCCAGTTAATAGTTTGC GCAACGTTGTTGCCATTG CTACAGGCATCGTGGTGT CACGCTCGTCGTTTGGTATG GCTTCATTCAGCTCCGGT TCCCAACGATCAAGGCGAGT TACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCT CCTTCGGT CCTCCGATCGTTGTCAGAAGTAAGTTGGCC GCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATA CGGGATAATACCGCGCCACATAGCAGAACTTTAAAAG TGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC TC TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA ACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG GCAAAATGCCGCAAAAAGGGAATAAGGGCGACACG GAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGAT ACATATTTGAATGTATTTAGAAAAATAAACAAATAGG GGTTCGCGCACATTTCCCCGAAAAGTGCCACCTGACG TCTAAGAAACCATTATTATCATGACATTAACCTATAAA AATAGGCGTATCACGAGGCCCTTTCGTC</p>
2	Donor Plasmid DNA-DBI	<p>AGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCG GTGGTTTTTTTTGTTT GCAAGCAGCAGATTACGCGCAGA AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTAC GGGGTCTGACGCTCTATTCAACAAAGCCGCCGTCCCGT CAAGTCAGCGTAAATGGGTAGGGGGCTTCAAATCGTC CCCCATACGATATAAGTTGTTACTAGTGCTTGGATTC TCACCAATAAAAAACGCCCGGCGGCAACCGAGCGTTC TGAACAAATCCAGATGGAGTTCTGAGGTCATTACTGGA TCTATCAACAGGAGTCCAAGCGAGCTCGATATCAAATT ACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAAT TCATTAAGCATTCTGCCGACATGGAAGCCATCACAAAC GGCATGATGAACCTGAATCGCCAGCGGCATCAGCACC TTGTCGCCTTGCGTATAATATTTGCCCATGGTGAAAAC GGGGGCGAAGAAGTTGTCCATATTGGCCACGTTTAAAT CAAACCTGGTGAAACTCACCCAGGGATTGGCTGAGAC GAAAAACATATTCTCAATAAACCCTTTAGGGAAATAG GCCAGGTTTTACCGTAACACGCCACATCTTGCGAATA TATGTGTAGAACTGCCGGAATCGTCGTGGTATTCAC TCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAA</p>

		ACGGTGTAACAAGGGTGAACACTATCCCATATCACCA GCTCACCGTCTTTCATTGCCATACGAAATTCCGGATGA GCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCG GATAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAA AGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTA CATTGAGCAACTGACTGAAATGCCTCAAATGTTCTTT ACGATGCCATTGGGATATATCAACGGTGGTATATCCAG TGATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAA ATCTCGATAACTCAAAAATAACGCCCGGTAGTGATCTT ATTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCG ATCAACGTCTCATTGATACCTGAAACAAAACCCATCGT ACGGCCAAGGAAGTCTCCAATAACTGTGATCCACCAC AAGCGCCAGGGTTTTCCAGTCACGACGTTGTAAAACG ACGGCCAGTCATGCATAATCCGCACGCATCTGGAATAA GGAAGTGCCATTCCGCCTGACCTGTCTGAGACCGAGCT ATGTGGGGCGACCTCTGGCTCCTCCCGCCTGCCTCTGC CAATCCGGGCACTGGGACAGAGGTCGGTGTGAAACGC GCGGGCCCCAGGGGGAGGGAGGGGACCAACGGGCTCC GGCGCTGACACCGCGGCACTCATGCCCTGTCCCCTTTC AGCTGTTTCCAGCATACTGTGCCCGTCTGTCTCAGG CCAGGGCTTCGCTGCAGCCCCGGCCACTCCCTAGTGCC TGGCCCGGTGGTGGCCAGGCAGTTGGCCGCGCTGCTTC TCCCGCAGAGGGGACCCCCACTGGGGGCGAAGGCTTG GCCTGCCCTCTTACTGCTGTATTTCCAGACCTGATGCC TGCGTTTGTGAGAGCTCTGGATATATGGTTTTTCGATTG AATGAGTGAACCTGGAGGGGCTTCCCCTTCTTGTGTTGC TGAATCTTTCTAGCTGCCCTGTTGGGGCAGGGAGGGGC AGACACACTTCAGGGGCTGCATTGCCCGAAGGGTGCC ACCTTTCCACCTCTCCATCCCCGTAACCTGGGCTGTGAT CAGGCCACAGTAGGATTCTTACCCTCTCCCACCCAGAG GAGGCCCTCAATCCTCTCCTCTCCCTTCCATTTAGGCTG AGTTTGAGAAAGCTGCAGAGGAGGTTAGGCACCTTAA GACCAAGCCATCGGATGAGGAGATGCTGTTTATCTATG GCCACTACAAACAAGCAACTGTGGGGCGACATAAATAC AGGTATGCAGAGCGGGGGTTGGAAGGGCATCTGCTCA TCAAAGCAGGCTCAGCAGCTCAGACTGGAAGTCCCTG GGAACCTTCACTCTCAAACCTGCCTGAGGCCCTACCGATC GGGCAAGCTTCTGACCTTCTCTTCTCCACAGAAG GCCTCGAGAGATCTGGCAGCGGAGAGGGCAGAGGAAG TCTTCTAACATGCGGTGACGTGGAGGAGAATCCCGGCC CTAGGCTCGAGATGACCGAGTACAAGCCCACGGTGCG CCTCGCCACCCGCGACGACGTCCCAGGGCCGTACGCA CCCTCGCCGCGCGTTCGCCGACTACCCGCCACGCGC CACACCGTCGATCCGGACCGCCACATCGAGCGGGTCA CCGAGCTGCAAGAACTTCTCCTCACGCGCGTCCGGGCTC GACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCG CGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGC GGGGGCGGTGTTCCCGGAGATCGGCCCGCGCATGGCC GAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGA TGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCC
--	--	--

		<p>CGCGTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACC ACCAGGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCC CGGAGTGGAGGCGGCCGAGCGCGCCGGGGTGCCCGCC TTCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTA CGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGG TGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAA GCCCCGGTGCCTGATCTAGAGGGCCCGTTTAAACCCGCT GATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCT GTTGTTTGCCCTCCCCCGTGCCTTCTTGACCCTGGAA GGTGCCACTCCCCTGTCTTCTAATAAAAATGAGGA AATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCT GGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA TTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTG GGCTCTATGGTCACTCGAGTTAGACAGACTCATATCCG CGGGAGCTCACGGGGACAGCCCCCCCCCAAAGCCCC CAGGGCTGTAATTGCGTTCCTCTCCCGCTAGGGGGCGG CAGCGAGCCGCCAGGGCTCCGCTCCGGCCCGACGCTC CCCCCGCATCCCCGAGCCGGAGCCCGCAGCGTGCGGG GACAGCCCGGCACGGGGAAGGTGGCACGGGATCGCTT TCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGAC ACCTGGGGGGATACGGGGAAAAGCTTTATCTAGATC CGCGGGAGCTCACGGGGACAGCCCCCCCCCAAAGCCC CCAGGGCTGTAATTGCGTTCCTCTCCCGCTAGGGGGCA GCAGCGAGCCGCCGGGGCTCCGCTCCGGCCCGACGC TCCCCCGCATCCCCGAGCCGGAGCCCGCAGCGTGCGG GGACAGCCCGGCACGGGGAAGGTGGCACGGGATCGCT TTCTCTGAACGCTTCTCACTGCTCTTTGAGCCTGCAGA CACCTGGGGGGATACGGGGAAAAGCTTTATCGCTCT GTCGCAAGTACTGAGTTTACTCCCTATCAGTGATAGAG AACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAA CGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACG TATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTA TGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATC TACAGTTTACTCCCTATCAGTGATAGAGAACGTATATC CAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGA GGTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGC TCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATTCCA CAACACTTTTGTCTTATACTTAGTGCCACCATGGTG AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCA TCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAA GTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC TACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGG CAAGCTGCCCCTGCCCTGGCCCACCCTCGTGACCACC TGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGAC CACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC CGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGG ACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAG GGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGC ACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTA</p>
--	--	--

		TATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTG AACTTCAAGATCCGCCACAACATCGAGGACGGCAGCG TGCAGCTCGCCGACCACTACCAGCAGAACACCCCATC GGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCT GAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG AAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGC CGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGT AAGAATTCGAGCTCGGTACCCGGGGATCCTCTAGTCAG CTGACGCGTGCTAGCGCGGCCGCATCGATAAGCTTGTC GACGATATCTCTAGAGGATCATAATCAGCCATAACCACA TTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCACA CCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTG TTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACA AATAAAGCAATAGCATCACAAATTTACAAATAAAGC ATTTTTTTCCTGACCTTGACAGTACTCCTACTGGCGGAC AGAGCGATAAAGCTTTTTCCCGTATCCCCCAGGTGT CTGCAGGCTCAAAGAGCAGTGAGAAGCGTTCAGAGGA AAGCGATCCCGTGCCACCTTCCCGTGCCGGGCTGTCC CCGCACGCTGCGGGCTCCGGCTCGGGGATGCGGGGGG AGCGTCGGGCCGGAGCGGAGCCCCGGGCGGCTCGCTG CTGCCCCCTAGCGGGAGAGGAACGCAATTACAGCCCT GGGGGCTTTGGGGGGGGGCTGTCCCCGTGAGCTCCCGC GGATCTAGATAAAGCTTTTTCCCGTATCCCCCAGGT GTCTGCAGGCTCAAAGAGCAGCGAGAAGCGTTCAGAG GAAAGCGATCCCGTGCCACCTTCCCGTGCCGGGCTGT CCCCGCACGCTGCGGGCTCCGGCTCGGGGATGCGGGG GGAGCGTCGGGCCGGAGCGGAGCCCCGGGCGGCTCGC TGCTGCCCCCTAGCGGGAGAGGAACGCAATTACAGCC CTGGGGGCTTTGGGGGGGGGCTGTCCCCGTGAGCTCCC GCGGATATGAGTCTGTCTAAGTGCTTAAGTCGACTACC GGTTCTTCAGGTGGGGTATGGTGATGGTTCTGAGGTG GAAAAGACCATGTTCCGGATTCTCAGTGTCTCCAGTAG TAACAGAATTCAAATCCTGGTTTTAGAAGGTCTTTACT GGTTATCACCAGCAGCTACTCTCTACTAGGGAAGAAGC AAAGGCTGCAGCTTGGAAGAAAGACTTGCTGAAGGCTCT CAGCTCAGTAGTATCATTGTTGAGCCGTTGAGCTTCTG CCCTAGATGGGCAGGATCAAAGTTGGAGCACTTTTTGG AGCACTTGACAGCCTGGCCAAGCCTGATGTCAGGAGC AGAGAAGCACCTGGTTTCTTGGGCTAGGTCAGAGCATT TCGCTAACAAGTCTGTGCCTTCCTGATGATAACTTTTTC CCTGCCAGAAATCTTGGTGCAGATTTTGAGGCTGTGC TTTGGACTGTCATGTTCTGTAATAACATCTTTCTGCCT TGGGCAGGTTTCATTCTGTCCCTAAGTCCCTGAAACAT GGGTGGATACTGAGGCAACAGCGCAGTGCATTCTGTG CAAGGACTCAGGGTTATCATGGCAGCACAGAAGGGAG GTCTCCCCTGCCCTGCTGAGGAAGAAGGCGAGCATG GTCCCTATTTCCGCAGTAGCTGGGGTGAAGATGGAGC AGGTGGGCTGGCTGCCAACCAGCTGGAAGCAGGAAAT AGTACCCAGAATGACAGATCACAGGCAGTACCATATC AAACCCTGGGGTTCACATGGAGCACTTAGTTGAAGAA
--	--	---



		<p>GGTCTTATGGCGAAGGTGAGTTTTACAGTGAGTTCGTA AACTCTGTCCTTCCAGGGAGGGGAAGGAAAGGTGAAG TGGGGGAGGCCAGAGGTGCCAAGATGCTTTTCTGACA GGCTAGGTGGAGGCTCAGTGATGATAAGTCTGCGATG GTGGATGCATGTGTCATGGTCATAGCTGTTTCCTGTGT GAAATTGTTATCCGCTCAGAGGGCACAATCCTATTCCG CGCTATCCGACAATCTCCAAGACATTAGGTGGAGTTCA GTTTCGGCGTATGGCATATGTCGCTGGAAAGAACATGTG AGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAA GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAG GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCT TCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAG GTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGG GCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCTGC GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGT AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCC AGTTACCTTCGGAAAAAGAGTTGGT</p>
3	Donor Plasmid DNA-PCSK9	<p>CGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTG GTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC GGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAG AAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA CGGGGTCTGACGCTCTATTCAACAAAGCCGCCGTC TCAAGTCAGCGTAAATGGGTAGGGGGCTTCAAATCGTC CCCCATACGATATAAGTTGTTACTAGTGCTTGGATTC TCACCAATAAAAAACGCCCGGCGGCAACCGAGCGTTC TGAACAAATCCAGATGGAGTTCTGAGGTCATTACTGGA TCTATCAACAGGAGTCCAAGCGAGCTCGATATCAAATT ACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAAT TCATTAAGCATTCTGCCGACATGGAAGCCATCACAAAC GGCATGATGAACCTGAATCGCCAGCGGCATCAGCACC TTGTCGCCTTGCGTATAATATTTGCCCATGGTGAAAAC GGGGGCGAAGAAGTTGTCCATATTGGCCACGTTAAAT CAAACCTGGTGAAACTCACCCAGGGATTGGCTGAGAC GAAAAACATATTCTCAATAAACCTTTAGGGAAATAG GCCAGGTTTTACCGTAACACGCCACATCTTGCGAATA TATGTGTAGAACTGCCGGAATCGTCGTGGTATTCAC TCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAA ACGGTGTAACAAGGGTGAACACTATCCCATATACCA GCTCACCGTCTTTCATTGCCATACGAAATTCCGGATGA GCATTATCAGGCGGGCAAGAATGTGAATAAAGGCCG GATAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAA AGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTA CATTGAGCAACTGACTGAAATGCCTCAAATGTTCTTT ACGATGCCATTGGGATATATCAACGGTGGTATATCCAG</p>

		TGATTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAA ATCTCGATAACTCAAAAAATACGCCCGGTAGTGATCTT ATTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCG ATCAACGTCTCATTGATACCTGAAACAAAACCCATCGT ACGGCCAAGGAAGTCTCCAATAACTGTGATCCACCAC AAGCGCCAGGGTTTTCCAGTCACGACGTTGTAAAACG ACGGCCAGTCATGCATAATCCGCACGCATCTGGAATAA GGAAGTGCCATTCCGCCTGACCTACCCTGTCCCAGCAG GACAGTGGGCTGGTGTGAGCGGGCAGGAACCGCCTGC ACTTAGAAGGTGTGGGGCTGCCTCCCCGAGCTTCCATC TGCCGCTGGGGCCACACCCAGGCCAGGGATGGGAC CCCACAGTGGTCACATCATCTTGCAGCAGAACCCAGGT ACAGCTCCTGGAGCAGATGGTGGTCCCAAGCACGGGT GGGACCAGAAAGGACTCTCACCTGGGCTAACTCAGCT GCAGCCTCAGTTCCTCCTCACACACGACGAGGAACAT GGACTGGAAGCCTGCCAGCAGGCCTTCTGCTCGATGT GCGTTGTGTGGCTTACGTCCAGGGAGGGAAGCAGCCTC TGTGCTGTCTTCTAGATAAGCCTGTATCCCCGGGCTGT CTGCCAATGTATCCAGTTGTCCCGTCAGCCTGGAAGCT CTGAGGGAAAACCTTGGGCTGCTTCTGAGCACCTGTA TCCCCTGCAGCCAGCCGGGGCCTCTGCTAGGAGCAGA CTGAGCATGGCTTATGGGCCTGGCACCATCTGGCCTCT GCCACCTTGCTGGCCTTGTCTTGTGTCTGCCCTTCGA CATTCCATAGCCAGCTCAATATCTAGTGGTTCCTCTA GGGTGGCGAGCACTGTTTGGTCTCCAGATGTCTTCAGG TCGGAGCTCACAGCGCTCTCAGCCACCCCTTCCCAGTG TAGCACCGGGCACATGGTAGATGCCTATTGATGAGTGA AAGCTCCTAACACACTCAGAGAGCAAGGACTCCGCCT CATCCCACAGCCTGGGAGGAGAGGCAGACTGCCAAGG ACCTGCTCAGCATGCTACAGAAGAAACCAAAGTGCCC ACGGCGATCGGGCAAGCTTCTGACCTTCTCTTCTCCTC CCACAGAAGGCCTCGAGAGATCTGGCAGCGGAGAGGG CAGAGGAAGTCTTCTAACATGCGGTGACGTGGAGGAG AATCCCGGCCCTAGGCTCGAGATGACCGAGTACAAGC CCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCAGG GCCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCC CGCCACGCGCCACACCGTCGATCCGGACCGCCACATCG AGCGGGTCACCGAGCTGCAAGAACTTCTCCTCACGCGC GTCGGGCTCGACATCGGCAAGGTGTGGGTGCGCGACG ACGGCGCCGCGGTGGCAGGCTGGACCACGCCGGAGAG CGTCGAAGCGGGGGCGGTGTTCCGCCGAGATCGGCCCG CGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCA GCAACAGATGGAAGGCCTCCTGGCGCCGCACCGGCC AAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTC GCCCGACCACAGGGCAAGGGTCTGGGCAGCGCCGTC GTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCCGGGG TGCCCGCCTTCTGGAGACCTCCGCGCCCCGCAACCTC CCCTTCTACGAGCGGCTCGGCTTACCGTCACCGCCGA CGTCGAGGTGCCGAAGGACCGCGCACCTGGTGCATG ACCCGCAAGCCCGGTGCCTGATCTAGAGGGCCCGTTA
--	--	---

		AACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCC AGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGA CCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAA AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCA TTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAG GGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGG ATGCGGTGGGCTCTATGGTCACTCGAGTTAGACAGACT CATATCCGCGGGAGCTCACGGGGACAGCCCCCCCC AAAGCCCCAGGGCTGTAATTGCGTTCCTCTCCCGCTA GGGGGCGGCAGCGAGCCGCCAGGGCTCCGCTCCGGC CCGACGCTCCCCCGCATCCCCGAGCCGGAGCCCGCAG CGTGCGGGGACAGCCCGGCACGGGGAAGGTGGCACGG GATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGC CTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTA TCTAGATCCGCGGGAGCTCACGGGGACAGCCCCCCCC AAAGCCCCAGGGCTGTAATTGCGTTCCTCTCCCGCTA GGGGGCGGCAGCGAGCCGCCAGGGCTCCGCTCCGGC CCGACGCTCCCCCGCATCCCCGAGCCGGAGCCCGCAG CGTGCGGGGACAGCCCGGCACGGGGAAGGTGGCACGG GATCGCTTTCCTCTGAACGCTTCTACTGCTCTTTGAGC CTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTA TCGCTCTGTGCAAGTACTGAGTTTACTCCCTATCAGT GATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGA TAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATA GAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGA GAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA ACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAAC GTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGT ATGTCGAGGTAGGCGTGTACGGTGGGCGCCTATAAAA GCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGC AATCCACAACACTTTTGTCTTATACTTACTAGTGCCAC CATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGGTG GTGCCATCCTGGTCGAGCTGGACGGCGACGTAACG GCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGA TGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCA CCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTG ACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTA CCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCG CCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGG TGAAGTTCGAGGGGCGACACCCTGGTGAACCGCATCGA GCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATC CTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACA ACGTCTATATCATGGCCGACAAGCAGAAGAACGGCAT CAAGGTGAACTTCAAGATCCGCCACAACATCGAGGAC GGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACA CCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAAC CACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCC CAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCT
--	--	--

		GTACAAGTAAGAATTTCGAGCTCGGTACCCGGGGATCCT CTAGTCAGCTGACGCGTGCTAGCGCGGCCGCATCGATA AGCTTGTCGACGATATCTCTAGAGGATCATAATCAGCC ATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAAC CTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAA TGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAA TGGTTACAAATAAAGCAATAGCATCACAAATTTACAAA ATAAAGCATTTTTTTCACTGCCTTGACAGTACTCCTACT GGCGGACAGAGCGATAAAGCTTTTTTCCCGTATCCCC CAGGTGTCTGCAGGCTCAAAGAGCAGTGAGAAGCGTT CAGAGGAAAGCGATCCCGTGCCACCTTCCCGTGCCGG GCTGTCCCCGCACGCTGCGGGCTCCGGCTCGGGGATGC GGGGGGAGCGTCGGGCCGGAGCGGAGCCCCGGGCGGC TCGCTGCTGCCCCCTAGCGGGAGAGGAACGCAATTAC AGCCCTGGGGGCTTTGGGGGGGGGCTGTCCCCGTGAG CTCCCGCGGATCTAGATAAAGCTTTTTTCCCGTATCCC CCCAGGTGTCTGCAGGCTCAAAGAGCAGCGAGAAGCG TTCAGAGGAAAGCGATCCCGTGCCACCTTCCCGTGCC GGGCTGTCCCCGCACGCTGCGGGCTCCGGCTCGGGGAT GCGGGGGGAGCGTCGGGCCGGAGCGGAGCCCCGGGCG GCTCGCTGCTGCCCCCTAGCGGGAGAGGAACGCAATT ACAGCCCTGGGGGCTTTGGGGGGGGGCTGTCCCCGTG AGCTCCCGCGGATATGAGTCTGTCTAAGTGCTTAAGTC GACTACCGGTGACTGATCAGTGGAGCTTCTTGCCGAGA CTGGAGGCCTTAGGGCAGGGTAGACAGTGTGTGTGCA GGCTGGGGACTCACAGTTCGGACTGTGCCAGACCTAC TAGCATAGTGGGTGGGTGGGAGGATGCGGGACTGGGG GCCGACCTTGCTTCAAATTCATGTGGGATCTCAGAGCA GCCACTGAATTGCTCTGTAGGGGGCTAAATAGTGGCCC CCACAGATACACACACCCAGACAGAGCCTGTGAGCCA GACCTTATTTGGAGAAAAGGTCTTTGTAGATGTAATTA AGCATCTCAAGATGGCATCATCTGGATTATGCGGTGGG CTGTAAGTCCTGTGATGTGTCTTTATGAGAGAAAGGCA GAGGGAGATTTGACACACACAGGAGGGGCCACGTGGA GACAGAGGTGGAGATTGGAGAAATGTGGCCACAAGCC AGGGAACACCAGCAGCCACCAGAAGCCGGAAGACGTG AGGCAGGGTTCTTCCAGAGCCTTCGCTGCTGAGTCTG GGAATTTGTGACCGAAGCCATAAGAAGTGGGTACACG CCCTGAGCCTCCACACTTGCTCACCTGTCCTGAGATG AGAATCTCTACTCTGCAGCATATTTGGAGGATCACTGC GGGGGCCACAGAGGTGCTGTTTCAGATGGCACTTCAGA AGACTCAGGAGACCCTGGGGCAGGAGCAGTTTACTG ACAGCCCAGAGGGCTGCCCTCTGATTCCACCTGAGGCC CTGCTTTTCTGGCTGCAGGGGTTCCAGGGCCAGGCCA TTCCGCTGGCGCAGGACTCTGCTAGCAGCAACCTGCC TGAAGTCTTCCTTTAGGCTAGGTGGAGGCTCAGTGATG ATAAGTCTGCGATGGTGGATGCATGTGTCATGGTCATA GCTGTTTCTGTGTGAAATTGTTATCCGCTCAGAGGGC ACAATCCTATTCCGCGCTATCCGACAATCTCCAAGACA TTAGGTGGAGTTCAGTTCGGCGTATGGCATATGTCGCT
--	--	---

		GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC AGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA TAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGAC GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC GCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGT CCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT CGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCA GCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATG TAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC TACGGCTACACTAGAAGAACAGTATTTGGTATCTG
--	--	---

**CLAIMS**

1. A method of increasing frequency of gene editing at a target polynucleotide/nucleotide sequence in a cell, the method comprising:

(i) contacting the target polynucleotide sequence with a first ribonucleoprotein (RNP) comprising a first ribonucleic acid molecule and a clustered regularly interspaced short palindromic repeat-associated (Cas) protein; wherein the first ribonucleic acid molecule comprises a sequence complementary to a nucleic acid sequence in the target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;

(ii) contacting the edited target polynucleotide sequence with a second or more RNP(s) comprising a second or more ribonucleic acid molecule(s) and a Cas protein to induce a second gene editing event; wherein the second or more ribonucleic acid molecule(s) comprises a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid molecule and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence;

wherein the method results in increased gene editing frequency compared to the same method but consisting of only step (i).

2. The method of claim 1, wherein (i) results in an error in one or more bases at the target polynucleotide sequence relative to the first ribonucleic acid molecule sequence.

3. The method of claim 2, wherein (ii) occurs after (i) results in an error in one or more bases at the target polynucleotide sequence relative to the first ribonucleic acid molecule sequence.

4. The method of claim 2 or 3, wherein the error in the one or more bases at the target polynucleotide sequence is an error induced by double strand break repair.

5. The method of any one of claims 2-4, wherein the error in the one or more bases at the target polynucleotide sequence is an alteration in one base relative to the first ribonucleic acid molecule sequence.

6. The method of any one of claims 2-5, wherein the error in the one or more bases at the target polynucleotide sequence comprises an insertion of one or more bases.

7. The method of any one of claims 2-5, wherein the error in the one or more bases at the target polynucleotide sequence comprises a deletion of one or more bases.
8. The method of claim 6, wherein the error in the one or more bases at the target polynucleotide sequence comprises an insertion at the cut site created by double strand break repair.
9. The method of any one of the above claims, further comprising integrating a donor polynucleotide sequence into the target polynucleotide sequence.
10. The method of any one of the above claims, wherein steps (i) and (ii) occur simultaneously.
11. The method of any one of claims 1-9, wherein steps (i) and (ii) occur sequentially.
12. The method of any one of the above claims, wherein the target polynucleotide sequence is a genomic safe harbor site (GSH) locus.
13. The method of claim 12 wherein the GSH locus is selected from the group consisting of AAVS1 (PPP1R12C gene), ROSA26 (THUMPD3-AS1 gene), CLYBL, DBI, PCSK9 and ZC3H3.
14. The method of claim 13, wherein the GSH locus is ROSA26.
15. The method of claim 13, wherein the GSH locus is AAVS1.
16. The method of claim 15, wherein the AAVS1 locus is altered by an insertion of one base; and wherein the insertion is an adenine or thymidine.
17. The method of claim 13, wherein the GSH locus is CLYBL.
18. The method of claim 17, wherein the CLYBL locus is altered by an insertion of one base; and wherein the insertion is an adenine or thymidine.
19. The method of claim 13, wherein the GSH locus is DBI.
20. The method of claim 19, wherein the DBI locus is altered by an insertion of one base; and wherein the insertion is an adenine or thymidine.
21. The method of claim 13, wherein the GSH locus is PCSK9.
22. The method of claim 21, wherein the PCSK9 locus is altered by an insertion of one base; and wherein the insertion is an adenine or thymidine.

23. The method of any one of claims 9-22, wherein the increased accuracy of integration of the donor polynucleotide sequence by Homology-Directed Repair (HDR)-mediated integration ranges from about 25% to about 75%.
24. The method of any one of claims 9-23, wherein the increased gene editing frequency is measured by detecting expression of a gene that is encoded by a donor polynucleotide sequence.
25. The method of any one of the above claims, wherein the gene editing results in the insertion of at least one exogenous gene.
26. The method of any one of the above claims, wherein the gene editing results in the insertion of one or more non-protein coding sequences.
27. The method of claim 26, wherein the non-protein coding sequence comprises a non-coding RNA sequence.
28. The method of claim 27, wherein the non-coding RNA sequence comprises a sequence selected from the group consisting of one or more microRNAs, siRNAs, piRNAs, snoRNAs, snRNAs, exRNAs, scaRNAs and long ncRNAs.
29. The method of any one of the above claims, wherein the contacting comprises introducing one or more of: the first ribonucleic acid, second ribonucleic acid, donor polynucleotide, and polynucleotide encoding Cas protein, to a cell.
30. The method of claim 29, wherein the introducing step is performed by at least one of transfection, transduction, electroporation, and microinjection.
31. The method of any one of the above claims, wherein transcription of the first and/or second ribonucleic acid sequence is transiently induced.
32. The method of claim 31, wherein the transcription is transiently induced by activating a regulatable promoter controlling the transcription of the first and/or second ribonucleic acid sequence.
33. The method of claim 25, wherein the at least one exogenous gene comprises a transcription factor.
34. The method of any one of claims 9-33, wherein the donor polynucleotide sequence comprises a sequence encoding at least one transcription factor.



35. The method of any one of claims 9-34, wherein a donor plasmid DNA comprises the donor polynucleotide sequence.
36. The method of claim 35, wherein the donor plasmid DNA further comprises one or more polynucleotide sequences encoding, a bacterial resistance gene, an origin of replication, a transcriptional promoter for driving expression of an exogenous gene, a selectable marker, a fluorescent protein, a 5' homology arm, and a 3' homology arm.
37. A cell comprising the first and second ribonucleic acids of any one of the above claims; and optionally, the donor polynucleotide of any one of the above claims.
38. The cell of claim 37, wherein the cell further comprises Cas protein.
39. The cell of claim 37 or 38, wherein the cell is a stem cell.
40. The cell of claim 37 or 38, wherein the cell is an induced pluripotent stem (iPS) cell.
41. The cell of claim 40, wherein the iPS cell is a human iPS (hiPS) cell.
42. The cell of any one of claims 37 -41, wherein the cell is a somatic cell.
43. The cell of claim 42, wherein the cell is an immune cell, optionally a T cell, a B cell, a dendritic cell, a macrophage or an NK cell.
44. The cell of claim 42, wherein the cell is a neuronal cell, optionally a microglial cell, a motor neuron, a dopaminergic neuron, a GABAergic neuron, or a glutamatergic neuron.
45. The cell of claim 42, wherein the cell is an adipocyte, a hepatocyte, a pancreatic cell, an epithelial cell, a muscle cell, a bone cell, a skin cell or a blood cell.
46. The cell of any one of claims 37-45, wherein the cell is an ex-vivo patient-derived cell.
47. The cell of any one of claims 37-46, wherein the cell is reprogrammed to a differentiated cell after the second gene editing event.
48. A plurality of polynucleotides comprising:
  - i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;
  - ii) a second or more ribonucleic acid molecule(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or

more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; and optionally,

iii) a donor DNA comprising the target polynucleotide sequence; and optionally, iv) a polynucleotide comprising a sequence encoding Cas protein.

49. A kit comprising:

i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;

ii) a second or more ribonucleic acid molecule(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; and optionally,

iii) a donor DNA comprising the target polynucleotide sequence; and optionally,

iv) a polynucleotide comprising a sequence encoding Cas protein; and

v) instructions for use.

50. A population of cells comprising:

i) a first ribonucleic acid molecule comprising a sequence complementary to a nucleic acid sequence in a target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;

ii) a second or more ribonucleic acid molecule(s) comprising a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence; and optionally,

iii) a donor DNA comprising the target polynucleotide sequence; and optionally,

iv) a polynucleotide comprising a sequence encoding Cas protein.

51. The population of cells of claim 50, wherein the cells comprise an increased percentage of the donor polynucleotide integrated into the target polynucleotide sequence compared to a population of the same type of cells comprising the first ribonucleic acid, the target polynucleotide sequence, and the donor polynucleotide, but not the second ribonucleic acid.

52. The population of cells of claim 50 or 51, wherein the cells further comprise Cas protein.

53. The population of cells of any one of claims 50-52, wherein the cells comprise induced pluripotent stem (iPS) cells, optionally human iPS (hiPS) cells.

54. The population of cells of any one of claims 50-53, wherein the cells comprise immune cells, optionally T cells, B cells, dendritic cell, macrophages, NK cells, or combinations thereof.

55. The population of cells of any one of claims 50-53, wherein the cells comprise neuronal cells, optionally microglial cells, motor neurons, dopaminergic neurons, GABAergic neurons, glutamatergic neurons or combinations thereof.

56. The population of cells of any one of claims 50-53, wherein the cells comprise adipocytes, hepatocytes, pancreatic cells, epithelial cells, skeletal muscle cells, smooth muscle cells, cardiomyocytes, bone cells, skin cells or blood cells.

57. The population of cells of claim 53, wherein the cells comprise ex-vivo patient derived cells.

58. The population of cells of any one of claims 50-57, wherein the donor polynucleotide encodes at least one transcription factor.

59. The population of cells of any one of claims 50-58, wherein the cells are reprogrammed to a differentiated cell after the first or second gene editing event mediated by the first or second ribonucleic acid, Cas protein and the donor polynucleotide.

60. A method of generating a cell with at least one exogenous polynucleotide sequence integrated into the cell genome, the method comprising:

(i) contacting a target polynucleotide sequence within the cells with a first ribonucleic acid molecule and a Cas protein; wherein the first ribonucleic acid molecule comprises a

sequence complementary to a nucleic acid sequence in the target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;

(ii) contacting the target polynucleotide sequence with a second or more ribonucleic acid molecule(s) molecule(s) and a Cas protein to induce a second gene editing event at the target polynucleotide sequence;

wherein the second or more ribonucleic acid molecule(s) comprises a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence;

wherein the target polynucleotide sequence is a GSH site;

wherein the method results in increased gene editing accuracy compared to a method comprising the first ribonucleic acid molecule that is complementary to a nucleic acid sequence in the target polynucleotide sequence but not the second or ribonucleic acid(s); and

wherein the method results in integration of at least one exogenous polynucleotide sequence into the genome of the cell.

61. A method of generating a differentiated cell from iPS cells, the method comprising:

(i) contacting a target polynucleotide sequence within the iPS cells with a first ribonucleic acid molecule and a Cas protein; wherein the first ribonucleic acid molecule is complementary to a nucleic acid sequence in the target polynucleotide sequence to induce a first gene editing event at the target polynucleotide sequence to produce an edited target polynucleotide sequence;

(ii) contacting the target polynucleotide sequence with a second or more ribonucleic acid molecule(s) molecule(s) and a Cas protein to induce a second gene editing event at the target polynucleotide sequence;

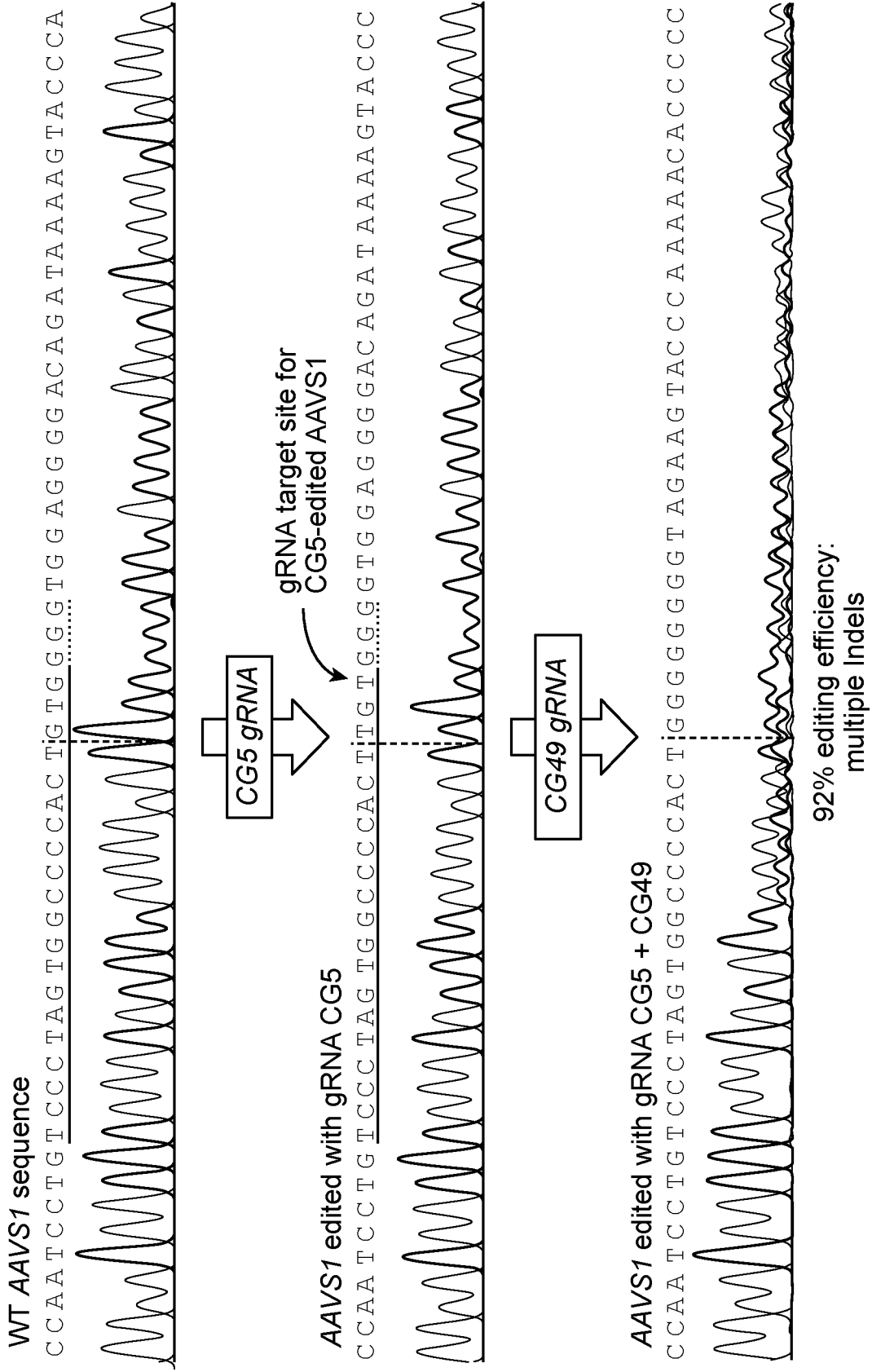
wherein the second or more ribonucleic acid molecule(s) comprises a sequence identical to the sequence of the first ribonucleic acid molecule except that the sequence of the second or more ribonucleic acid molecule(s) comprises one or more base alterations relative to the sequence of the first ribonucleic acid, and the second or more ribonucleic acid

molecule(s) comprises a sequence complementary to the edited target polynucleotide sequence;

wherein the target polynucleotide sequence is a GSH site;

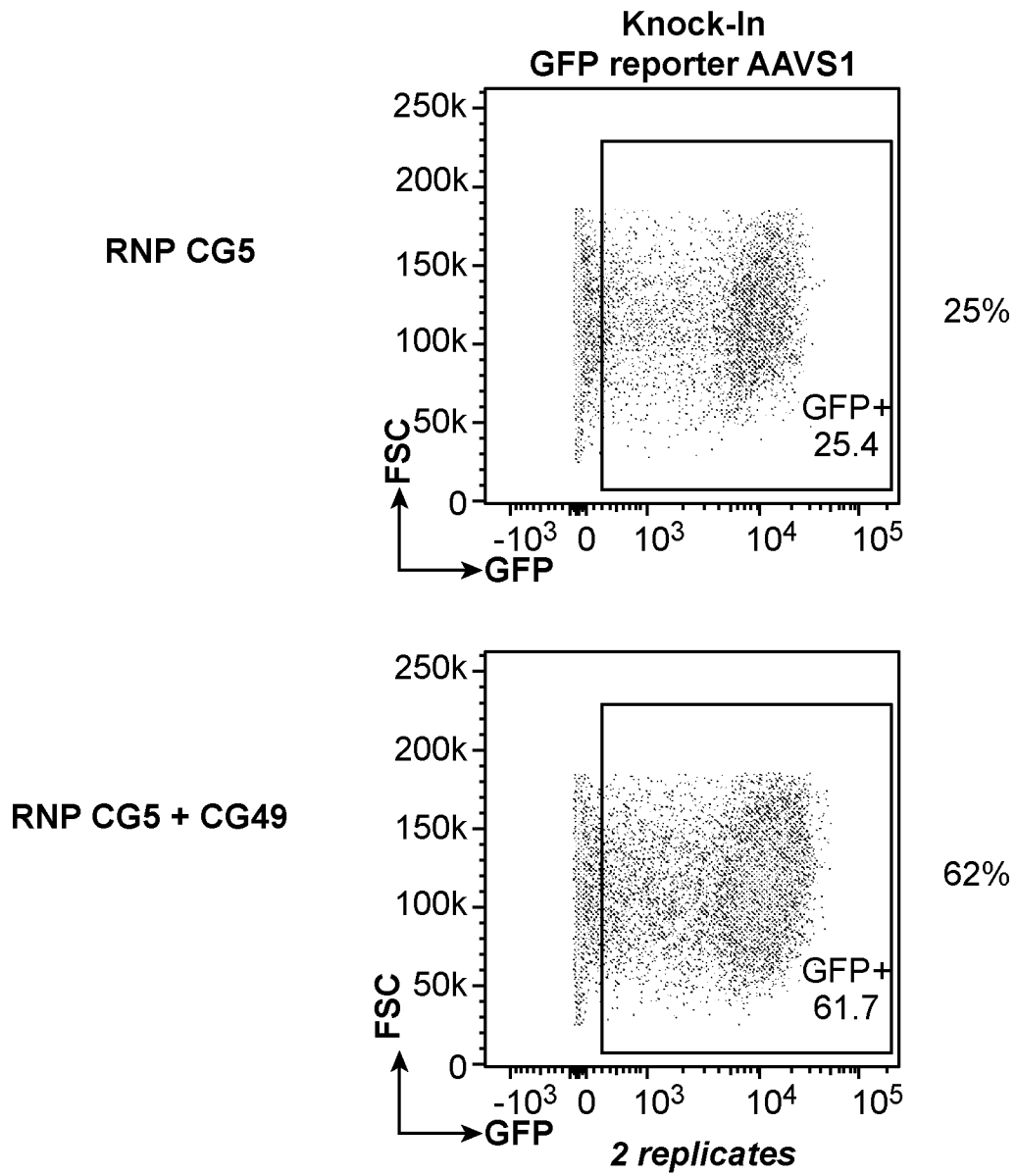
wherein the method results in increased gene editing accuracy compared to a method comprising the first ribonucleic acid molecule that is complementary to a nucleic acid sequence in the target polynucleotide sequence but not the second or more ribonucleic acid molecule(s);

wherein the method results in integration of at least one exogenous gene; and wherein the method results in the generation of one or more differentiated cells.



“Second Chance” for knock-in

FIG. 1A



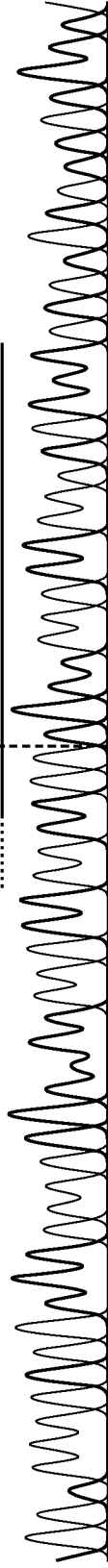
**FIG. 1B**

**CLYBL**

**WT sequence**

GACGAACTACTTTCAACTGTTTTAACTTCCCTTACTGTTAACTTCCAATTTCTAGATAGTATGGT

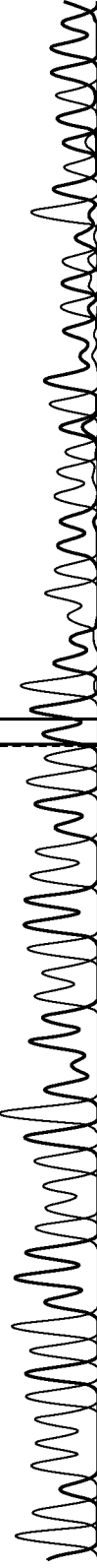
**CG65**



**Sequence after editing with CG65**

GACGAACTACTTTCAACTGTTTTAACTTCCCTTACTTGGTTAACTTCCAATTTCTAGATAGTATGG

**CG70**



Addition of a T

**FIG. 2A**



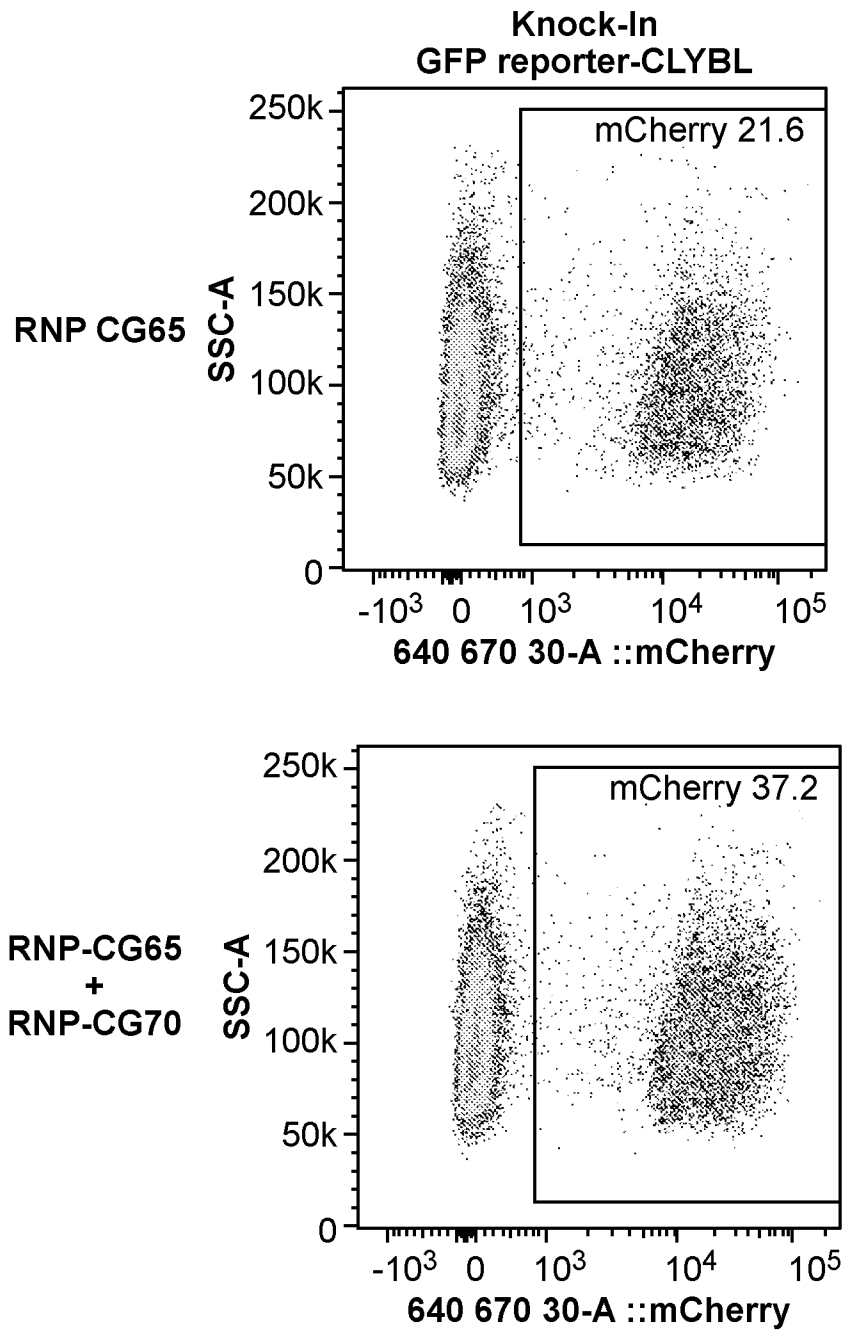


FIG. 2B

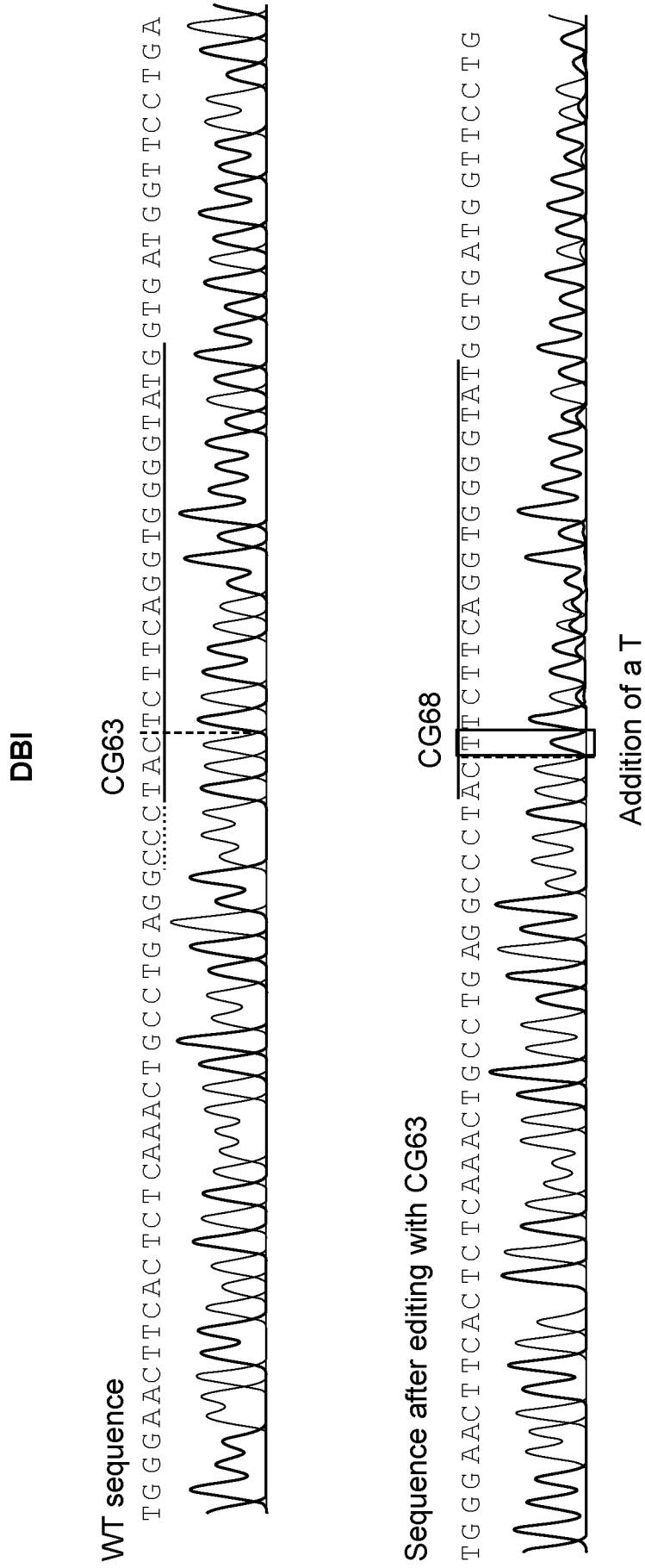


FIG. 3A

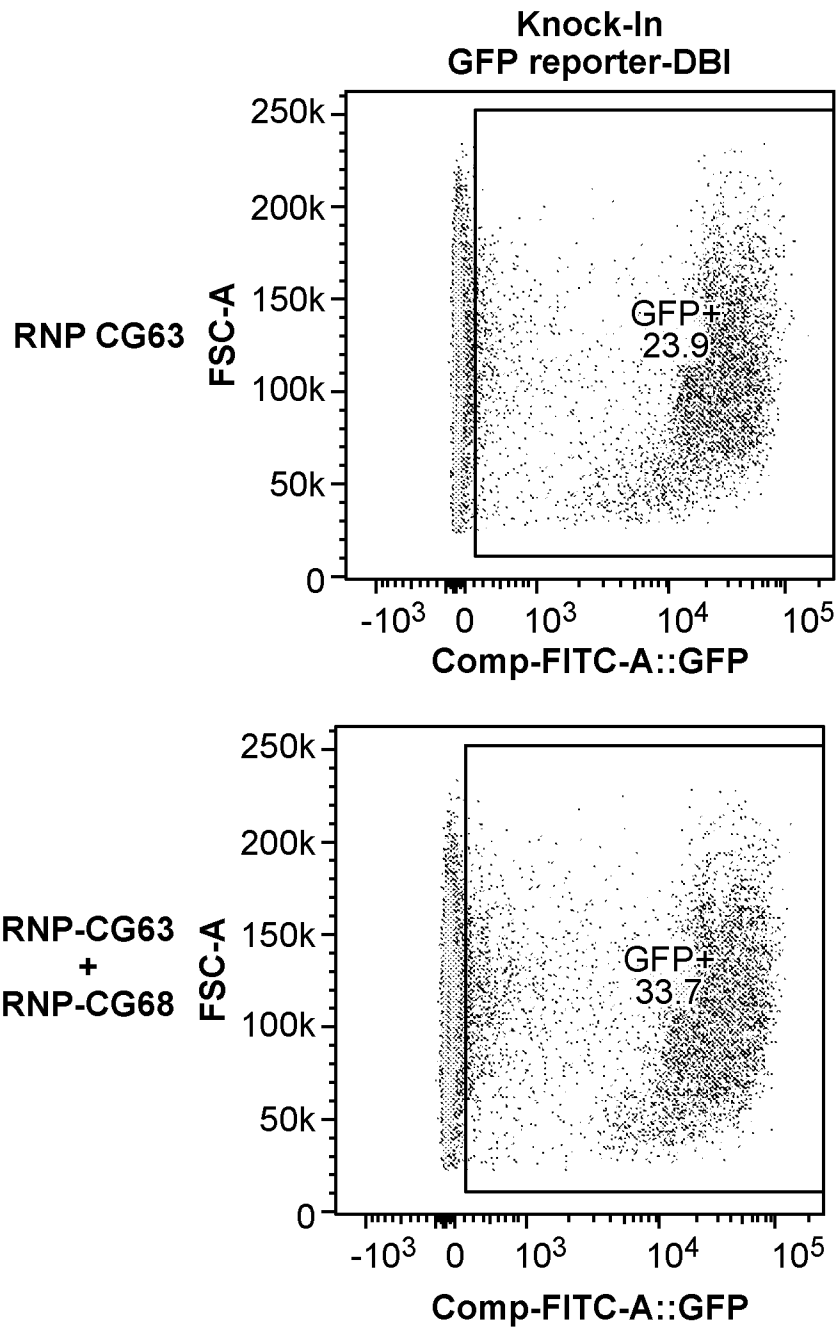
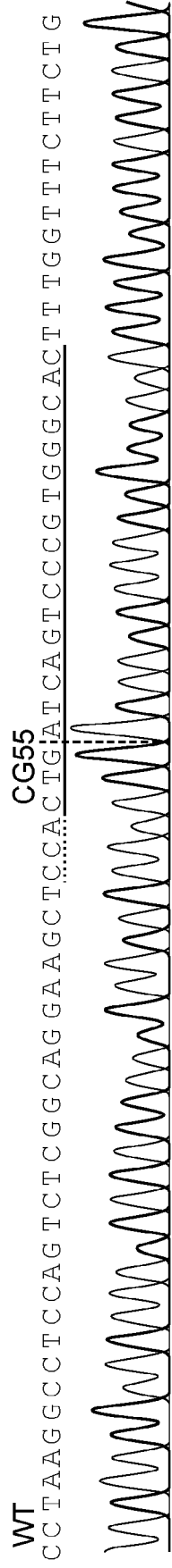
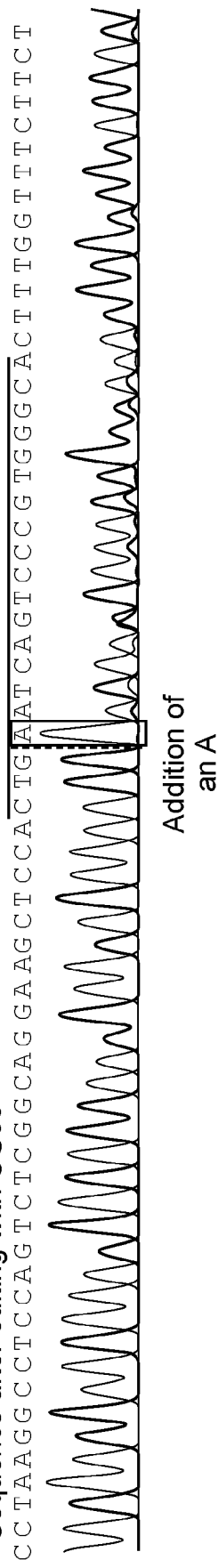


FIG. 3B

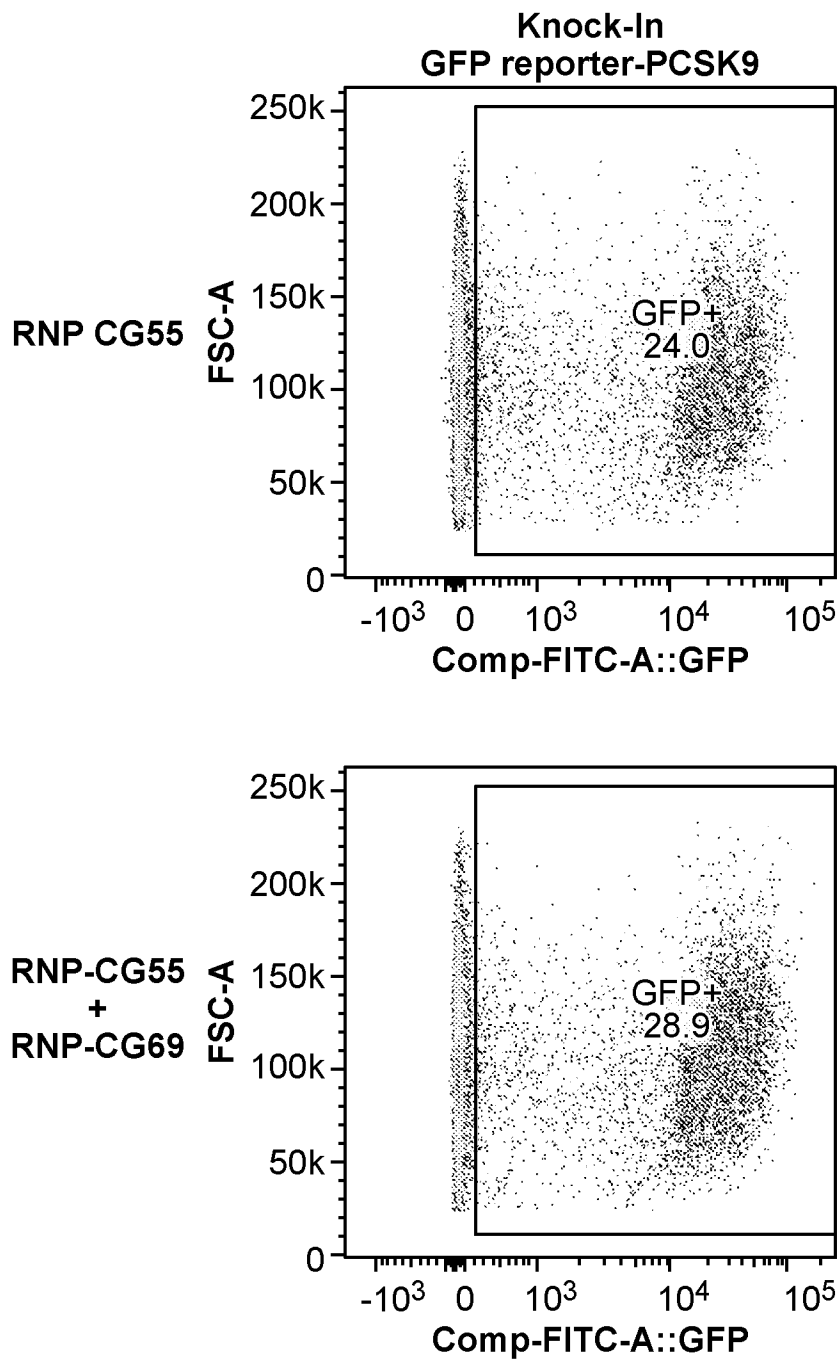
**PCSK9**



**Sequence after editing with CG55**



**FIG. 4A**



**FIG. 4B**

# INTERNATIONAL SEARCH REPORT

International application No  
**PCT/IB2023/057589**

**A. CLASSIFICATION OF SUBJECT MATTER**

**INV. C12N15/11**

**ADD.**

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)

**C12N**

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

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

**EPO-Internal, WPI Data, BIOSIS, EMBASE**

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
<b>X</b>	<p><b>BODAI ZSOLT ET AL.: "Targeting double-strand break indel byproducts with secondary guide RNAs improves Cas9 HDR-mediated genome editing efficiencies", NATURE COMMUNICATIONS, vol. 13, no. 1, 9 May 2022 (2022-05-09), XP093047691, DOI: 10.1038/s41467-022-29989-9 Retrieved from the Internet: URL:https://www.nature.com/articles/s41467-022-29989-9&gt;</b></p>	<p><b>1-9, 11-13, 21-26, 29-32, 34-38, 42, 48-52</b></p>
<b>Y</b>	<p><b>the whole document</b></p>	<p><b>13-20, 27, 28, 33, 39-41, 43-47, 53-61</b></p>
	<p><b>-/--</b></p>	

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

**16 October 2023**

Date of mailing of the international search report

**25/10/2023**

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

**Macchia, Giovanni**

## INTERNATIONAL SEARCH REPORT

International application No

PCT/IB2023/057589

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	-& Bodai Zsolt ET AL.: "supplementary data", , 9 May 2022 (2022-05-09), XP93090847, Retrieved from the Internet: URL:https://www.nature.com/articles/s41467-022-29989-9#Sec20 [retrieved on 2023-10-11]	1-9, 11-13, 21-26, 29-32, 34-38, 42, 48-52
Y	the whole document	13-20, 27, 28, 33, 39-41, 43-47, 53-61
X	----- BISHOP ALENA L. ET AL.: "Double-tap gene drive uses iterative genome targeting to help overcome resistance alleles", NATURE COMMUNICATIONS, vol. 13, no. 1, 9 May 2022 (2022-05-09), XP093047690, DOI: 10.1038/s41467-022-29868-3 Retrieved from the Internet: URL:https://www.nature.com/articles/s41467-022-29868-3> the whole document	1-10, 23, 37, 38, 42
X	-& BISHOP ALENA L. ET AL.: "Double-tap gene drive uses iterative genome targeting to help overcome resistance alleles", NATURE COMMUNICATIONS, vol. 13, no. 1, 9 May 2022 (2022-05-09), XP93091004, DOI: 10.1038/s41467-022-29868-3 the whole document	1-10, 23, 37, 38, 42
Y	----- WO 2021/181110 A1 (BIT BIO LIMITED [GB]) 16 September 2021 (2021-09-16)  page 34 - page 35 the whole document	13-20, 33, 39-41, 43-47, 53-61
Y	----- ALESSANDRO BERTERO ET AL.: "Optimized inducible shRNA and CRISPR/Cas9 platforms for in vitro studies of human development using hPSCs", DEVELOPMENT, vol. 143, no. 23, 29 November 2016 (2016-11-29), pages 4405-4418, XP055421687, GB ISSN: 0950-1991, DOI: 10.1242/dev.138081 the whole document ----- -/--	27, 28

**INTERNATIONAL SEARCH REPORT**

International application No  
**PCT/IB2023/057589**

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
<b>A</b>	<p><b>EIRINI P. PAPAPETROU ET AL. : "Gene insertion into Genomic Safe Harbors for human gene therapy",</b>  <b>MOLECULAR THERAPY,</b>  <b>vol. 24, no. 4, 1 April 2016 (2016-04-01),</b>  <b>pages 678-684, XP055547341,</b>  <b>US</b>  <b>ISSN: 1525-0016, DOI: 10.1038/mt.2016.38</b>  <b>the whole document</b></p> <p align="center">-----</p>	<b>12-22</b>
<b>X,P</b>	<p><b>MÖLLER LUKAS ET AL. : "Recursive editing improves homology-directed repair through retargeting of undesired outcomes",</b>  <b>NATURE COMMUNICATIONS,</b>  <b>vol. 13, no. 1, 5 August 2022 (2022-08-05)</b>  <b>, XP093047689,</b>  <b>DOI: 10.1038/s41467-022-31944-7</b>  <b>Retrieved from the Internet:</b>  <b>URL:https://www.nature.com/articles/s41467-022-31944-7&gt;</b>  <b>the whole document</b></p> <p align="center">-----</p>	<b>1</b>
<b>X,P</b>	<p><b>WO 2023/039135 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US])</b>  <b>16 March 2023 (2023-03-16)</b>  <b>the whole document</b></p> <p align="center">-----</p>	<b>1</b>



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2023/057589

## 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.
  - b.  furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)).  
 accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2.  With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

**PCT/IB2023/057589**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
<b>WO 2021181110</b>	<b>A1</b>	<b>16-09-2021</b>	<b>NONE</b>
-----			
<b>WO 2023039135</b>	<b>A1</b>	<b>16-03-2023</b>	<b>NONE</b>
-----			