

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

23 May 2019 (23.05.2019)



(10) International Publication Number

WO 2019/099982 A1

(51) International Patent Classification:

A61K 48/00 (2006.01) C12N 15/90 (2006.01)

C12N 15/11 (2006.01)

(21) International Application Number:

PCT/US2018/061770

(22) International Filing Date:

19 November 2018 (19.11.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/587,554 17 November 2017 (17.11.2017) US

(71) Applicant: THE JOHNS HOPKINS UNIVERSITY

[US/US]; 3400 N. Charles Street, Baltimore, Maryland 21218 (US).

(72) Inventors: SEYDOUX, Geraldine; 4000 Jones Bridge

Road, Chevy Chase, Maryland 20815 (US). PAIX, Alexandre; 306 E. 32nd, Apt. L, Baltimore, Maryland 21218 (US).

(74) Agent: GALVEZ, Guido J.; 1812 Ashland Avenue, Suite

110, Baltimore, Maryland 21205 (US).

(81) Designated States (unless otherwise indicated, for every

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

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: COMPOSITIONS AND METHODS FOR EFFICIENT GENOME EDITING

(57) Abstract: The present invention relates to the field of genome editing. More specifically, the present invention provides compositions and methods useful in clustered regularly interspaced short palindromic repeats (CRISPR)-based techniques. In one embodiment, the present invention provides a double-stranded, linear donor polynucleotide comprising a template polynucleotide flanked by a first homology arm and a second homology arm, wherein the homology arms are between 30-35 bases in length.



WO 2019/099982 A1

COMPOSITIONS AND METHODS FOR EFFICIENT GENOME EDITING

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 62/587,554,
5 filed November 17, 2017, which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates to the field of genome editing. More specifically, the present invention provides methods and compositions useful in the design of synthetic donor DNAs for efficient genome editing.

10 BACKGROUND OF THE INVENTION

Precision genome editing begins with the creation of a double-strand break (DSB) in the genome near the site of the desired DNA sequence change ("edit") (Jasin, M. & Haber, J.E., 44 DNA REPAIR (AMST.) 6-16 (2016)). Generation of targeted DSBs has been greatly accelerated in recent years by the discovery of CRISPR-Cas9, a programmable DNA
15 endonuclease that can be targeted to a specific DNA sequence by a small "guide" RNA (crRNA) (Doudna, J.A. & Charpentier, E., 346(6213) SCIENCE 1258096 (2014)). DSBs are lethal events that must be repaired by the cell's DNA repair machinery. DSBs can be repaired via imprecise, non-homology-based repair mechanisms, such as non-homologous end-joining (NHEJ), or by precise, homology-dependent repair (HDR) (Danner et al., 28(7-8)
20 MAMMALIAN GENOME 262-74 (2017)). HDR utilizes DNAs that contain homology to sequences flanking the DSB (termed homology arms) to template the repair. If a synthetic "donor" DNA containing the desired edit is available when the DSB is generated, the cellular HDR machinery will use the donor DNA to repair the DSB and the edit will be incorporated at the targeted locus (Jasin & Haber (2016)). Several studies have reported that single-
25 stranded oligonucleotides (ssODNs) can be used to introduce short edits (<50 bases) (Liang et al., 241 J. BIOTECHNOL. 136-46 (2016)). ssODNs that target the DNA strand that is first released by Cas9 after DSB generation have been reported to perform best (Richardson et al., 34(3) NAT. BIOTECHNOL. 339-44 (2016)). This strand preference, however, has only been tested for small edits near the DSB and has not been noticed at all loci (Liang et al. (2016)).
30 Edits at a distance from the DSB (>10 bp) are recovered at lower frequencies (Liang et al. (2016; Paquet et al., 533(7601) NATURE 125-29 (2016)). Recovery of large edits (such as GFP knock-ins) has also been reported to be inefficient, requiring large plasmid donors with long (>500nt) homology arms or selection markers to recover the rare edits (Danner et al. (2017)). Large insertions have been obtained through non-homologous or micro-homology-

mediated end joining reactions (NHEJ and MMEJ), but these approaches require simultaneous Cas9-induced cleavage of donor and target DNAs. *See* Yao et al., 20 EBIO MEDICINE 19-26 (2017); Yao et al., 27(6) CELL RES. 801-14 (2017); Zhang et al., 18(1) GENOME BIOL. 35 (2017); He et al., 44(9) NUCLEIC ACIDS RES. e85 (2016); Suzuki et al., 5 540(7631) NATURE 144-49 (2016); Yamamoto et al., 5(9) G3 (BETHESDA) 1843-47 (2015); Nakade et al., 5 NAT. COMMUN. 5560 (2014). Thus, there exists a need for more efficient genome editing tools and techniques.

SUMMARY OF THE INVENTION

Genome editing, the introduction of precise changes in the genome, is revolutionizing
10 our ability to decode the genome. The present invention is based, at least in part, on the development of compositions and methods for genome editing in mammalian cells that uses linear, double-stranded donor DNAs to introduce precise changes in the genome. As described herein, the present inventors demonstrate that PCR fragments containing edits up to 1kb require only about 35 bp homology arms to initiate Cas9-induced double-strand breaks in
15 human cells and mouse embryos. In addition, the present inventors have developed donor DNA design rules that maximize the recovery of edits without cloning or selection.

Accordingly, in one aspect, the present invention provides compositions useful for more efficient genome editing. In one embodiment, the present invention provides a double-stranded, linear donor polynucleotide comprising a polynucleotide encoding a fluorescent
20 protein flanked by a first homology arm and a second homology arm. In a specific embodiment, the homology arms are 15-60 bases in length. In a more specific embodiment, the homology arms are 25-45 bases in length. In an even more specific embodiment, the homology arms are 30-40 bases in length.

In another embodiment, the present invention provides a double-stranded, linear
25 donor polynucleotide comprising a polynucleotide encoding a fluorescent protein flanked by a first homology arm and a second homology arm, wherein the first and second homology arms are between 30-35 bases in length.

A double-stranded, linear donor polynucleotide can comprise a template polynucleotide encoding an edit flanked by an intervening sequence and two homology arms.
30 In one embodiment, the homology arms are 15-60 bases in length. In another embodiment, the homology arms are 25-45 bases in length. In an alternative embodiment, the homology arms are 30-40 bases in length. In certain embodiments, the template polynucleotide is up to 1 kb in length. The template polynucleotide can comprise a sequence designed to change at least one nucleotide base within 30 bases of a double-stranded break (DSB) of a target

nucleic acid. In another embodiment, the template polynucleotide further comprises a restriction enzyme site.

In an alternative embodiment, a double-stranded, linear donor polynucleotide comprises a template polynucleotide flanked by a first homology arm and a second homology arm, wherein the homology arms are between 30-35 bases in length. In particular
5 embodiments, the template polynucleotide is up to 1 kb in length. In a specific embodiment, the template polynucleotide comprises a sequence designed to change at least one nucleotide base within 30 bases of a DSB of a target nucleic acid. In another embodiment, the template polynucleotide further comprises a restriction enzyme site.

10 In another aspect, the present invention provides methods for more efficient genome editing. In one embodiment, a method comprises the step of performing a clustered regularly interspaced short palindromic repeats (CRISPR)-based technique using a double-stranded, linear donor polynucleotide described herein as the donor polynucleotide. In another embodiment, the present invention provides a method comprising injecting into a target cell a
15 composition comprising (a) an RNA-guided DNA endonuclease; (b) a guide RNA; and (c) a double-stranded, linear donor polynucleotide described herein.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1A-1B. Tagging of the mouse *Adcy3* locus with mCherry using a PCR donor with short homology arms. FIG. 1A: schematic representation of the mouse *Adcy3* locus
20 repair strategy using a PCR donor: mCherry (red), Homology arm Sequences (HS, blue), locus (grey lines), DSB (blue line). FIG. 1B: agarose gel showing representative PCR reactions using primers flanking the DSB at the *Adcy3* locus (primers correspond to sequence outside the HS from the PCR donor). The upper bands ('insert' arrow) correspond to the mCherry insertion.

25 FIG. 2A-2D. PCR fragments with short homology arms are efficient donors to create GFP knock-ins in HEK293T cells. FIG. 2A: diagrams showing PCR donors for GFP insertion at the *Lamin A/C* and *RAB11A* loci. Locus - grey, GFP - green, HS (Homology arm Sequences) - blue, DSB - vertical line. GFP was inserted at the DSB in *Lamin A/C* and 11 bp upstream of the DSB in *RAB11A*. FIG. 2B: graphs showing % of GFP+ cells obtained with
30 PCR donors with HS of the indicated lengths (33/33 refers to a right HS and a left HS, each 33 bp long). Insert size in all cases was 714 bp. Each bar represents the average insertion efficiency from two or more independent experiments (Table 1). Error bars represent the +/- SD. PCR fragments were nucleofected in HEK293T cells at the concentration indicated and were counted by flow cytometer 3 days later. See Table 1. FIG. 2C: graphs showing % of

GFP⁺ cells obtained with PCR or plasmid donors with HS of the indicated lengths. Insert size in all cases was 714 bp. Each bar represents the average insertion efficiency from two or more independent experiments (Table 1). Error bars represent the +/- SD. PCR fragments were nucleofected in HEK293T cells at the concentration indicated and cells were counted by flow cytometer 3 days later. FIG. 2D: confocal images of cells 3 days after nucleofection. GFP: green, DNA: blue. The GFP subcellular localizations are as expected for in frame translational fusions.

FIG. 3A-3C. Editing efficiency increases with decreasing insert size. The graphs show % of GFP⁺ cells obtained with PCR donors with HS and inserts of the indicated lengths. Each bar represents the average insertion efficiency from two or more independent experiments (Table 1). Error bars represent the +/- SD. FIG. 3A: knock-in of donors containing full-length GFP at the *Lamin A/C* locus. PCR fragments were nucleofected in HEK293T cells at the concentration indicated and cells were counted by microscopy 3 days later. FIG. 3B: knock-in of donors containing full-length GFP or GFP11 at the *Lamin A/C* locus. PCR fragments were nucleofected at the concentration indicated in HEK293T (expressing GFP1-10) and cells were counted by microscopy 3 days later. FIG. 3C: knock-in of donors containing full-length GFP or GFP11 at the *RAB11A* locus (11 bp upstream of DSB). PCR fragments were nucleofected at the concentration indicated in HEK293T (expressing GFP1-10) and cells were counted by flow cytometer 3 days later.

FIG. 4A-4C. Repair is a polarity-sensitive process. FIG. 4A: synthesis-dependent strand annealing (SDSA) model for gene conversion. In this, and all other schematics, each line corresponds to a DNA strand. Locus DNA is in grey, donor homology arms are in blue, donor insert is in green, and arrows indicate 3' ends. Donor DNA strands of opposite polarity are shown above and below the locus for clarity. PCR donors contain both strands, ssODNs donors would contain either a sense or antisense strand. Dotted lines represent DNA synthesized during the repair process. Resection of DSB: DSB is resected creating 3' overhangs on each side of the DSB. Strand invasion and DNA synthesis: The overhangs pair with complementary strands in the donor and are extended by DNA synthesis. Annealing: The newly synthesized strands withdraw from the donor and anneal back at the locus. Ligation (not shown) seals the break. FIG. 4B: diagrams showing donor ssODNs with only one HS (same conventions as in A). The ssODNs contain a 126 bp insert (green) coding for 3xFlag and GFP11 and HS targeting either the right or left side of the DSB (Table 2). FIG. 4C: normalized editing efficiency of ssODNs containing only one HS at the *Lamin A/C* and *RAB11A* loci. The polarity that allows pairing between the ssODN and resected ends (as

shown in diagram in A) is favored. Sense and antisense ssODNs were tested in parallel experiments and their efficiency were normalized as follows: normalized efficiency of sense ssODN (light blue) = % GFP+ cells with sense ssODN / [% GFP+ cells with sense ssODN + % GFP+ cells with antisense ssODN]. Normalized efficiency of antisense ssODN (dark blue) = % GFP+ cells with antisense ssODN / [% GFP+ cells with sense ssODN + % GFP+ cells with antisense ssODN]. Numbers on top of each column indicate the non-normalized % of GFP+ cells for each ssODN determined by microscopy (*Lamin A/C*) or flow cytometer (*RAB11A*).

FIG. 5A-5B. Polarity of ssODNs affects incorporation of distal edits. FIG. 5A: schematics showing possible pairing interactions between resected locus (grey) and ssODNs (light or dark blue for sense and antisense ssODN respectively, arrows – 3' ends) coding for a distal insert (green). Sequences between the DSB and insert were recoded to help integration of the distal insert and prevent cutting of edited locus by Cas9. FIG. 5B: normalized efficiency of sense vs antisense ssODNs calculated as in FIG. 4 (see Tables 3-4 for detailed results). Distance from the DSB, locus, and guide RNA polarity are indicated under each experiment. ssODN polarity has little effect on editing efficiency for proximal edits, but has a larger effect for distal edits. The favored polarity changes depending on whether the distal edit is positioned to the left or right of the DSB. Note that the favored ssODN polarity does not correlate with crRNA polarity (for example, first two columns in the graph show crRNAs 1776 and 1777 which cut at the same position but have opposite polarity). Experiments involving the *PYMI* locus were done on HEK293T that were cloned out and genotyped by PCR genotyping (size shift) for 3xFlag insertion (see FIG. 6). All other experiments were performed on HEK293T (GFP1-10) cells that were directly scored for GFP+ by flow cytometer or microscopy 3 days after nucleofection. Numbers on top of each column indicate the overall % of edits. Note that overall frequency decreases with increasing distance from the DSB (also see FIG. 14).

FIG. 6A-6D. Recoding of sequences between the DSB and the edit increases recovery of distal edits. FIG. 6A: schematics showing resected locus (grey with arrow at the 3' ends, *PYMI* locus) and ssODN donor (blue with arrow at the 3' end) coding for a proximal edit (green, restriction enzyme site, 1bp to the right of the DSB) and a distal edit (red, 3x Flag, 23 bp to the left of the DSB). Double arrows represent the region between the proximal and distal edits that is recoded (silent mutations). FIG. 6B: graphs showing % of edited cells containing proximal + distal edits (purple), proximal only (green) or distal only (red), using a ssODN donor with or without a recoded region. >50 cell clones were analyzed by PCR

genotyping (size shift) and RE digestion. FIG. 6C: schematics showing resected locus (grey with arrow at the 3' ends, *Lamin A/C* locus) and PCR donor (blue, thick bar) coding for a proximal edit (green, GFP11 inserted at the DSB) and a distal edit (red, tagRFP, 33 bp to the right of the DSB). Double arrows represent the region between proximal and distal edits that is recoded (silent mutations). FIG. 6D: graphs showing % of edited cells containing proximal + distal edits (purple), proximal only (green) or distal only (red), using a PCR donor with or without a recoded region. Edits were determined by direct examination of >1000 cells by microscopy.

FIG. 7A-7F. Repair is prone to template switching between donors. FIG. 7A: schematics showing repair of a DSB at the *RAB11A* locus (grey) with two ssODN donors. Arrows indicate 3' ends. Donor 1 contains GFP11 (green) with a STOP codon (red cross) and two HS (blue). Donor 2 contains GFP11 with no STOP codon and no HS. Double arrows indicate identical sequence shared between the donors. FIG. 7B: graphs showing the percent of GFP+ cells (Y axis, as determined by flow cytometer) for each donor combination (X axis). Each bar represents the average insertion efficiency from two independent experiments (Table 4). Error bars represent the +/- SD. For comparison, an ssODN identical to donor 1 but without the STOP codon gives 17.2% edits (discontinuous right most bar). FIG. 7C: schematics showing repair of a DSB at the *RAB11A* locus as in diagram A but with two PCR donors (thick bars). FIG. 7D: graphs showing the percent of GFP+ cells as in graphs B but with two PCR donors. Each bar represents the average insertion efficiency from two independent experiments (Table 4). Error bars represent the +/- SD. FIG. 7E: schematics showing repair of a DSB at the *Lamin A/C* locus (grey) with two ssODN donors. Arrows represents 3' ends. Donor 1 contains GFP11 (green) and two HS (blue). Donor 2 contains a recoded GFP11 (stars) with no HS. Double arrows indicate identical sequence shared between the donors. In this experiment, the edits were amplified en masse by PCR using a locus-specific primer and an insert-specific primer and sequenced by Illumina sequencing. FIG. 7F: graph showing the % of reads with evidence of template switching (Y axis) for each donor combination (X axis). Donor 1 + donor 2 without mutations and donor 1 + donor 2 with 1 mutations every 3 nucleotides (1/3) show no evidence of template switching (0%), whereas donor 1 + donor 2 (1/6) and donor 1 + donor 2 (1/12) show evidence of template switching (0.5% and 1.4% respectively). See FIG. 15 and Table 15.

FIG. 8. Guidelines for donor design. FIG. 8A: schematic showing a typical editing experiment using a PCR fragment (thick line) with two homology arms (blue) to introduce an edit (green) at a distance from the DSB (stippled line). FIG. 8B: recommendations based on

results presented in this study. For additional recommendations for ssODNs designed to insert edits at the DSB, see DeWitt et al., 121-11 METHODS 9-15 (2017) and Richardson et al., 34(3) NAT. BIOTECHNOL. 339-44 (2016).

FIG. 9. crRNAs used in this study. Schematics showing guide RNAs (arrows) used in this study mapped on *Lamin A/C*, *RAB11A*, *SMC3*, *PYMI* (human) and *Adcy3* (mouse) loci. Grey boxes indicate coding exons, only the first and last exons are shown for *Lamin A/C*, *RAB11A*, *SMC3*, and mouse *Adcy3*. For each guide, arrows indicate the 3' end. Numbers indicate position of the DSB relative to the ATG or STOP codon. Chemically synthesized crRNAs were used at all loci, except for *PYMI* where we used a plasmid-encoded sgRNA. Guide RNA sequences are in Table 14.

FIG. 10A-10C. Tagging with GFP of the *SMC3* locus using PCR repair template with short homology arms. FIG. 10A: diagram showing PCR donor for GFP insertion at the *SMC3* locus. Locus - grey, GFP - green, HS (Homology arm Sequences) - blue. GFP was inserted 5 bp to the right of the DSB. FIG. 10B: graphs showing % of GFP+ cells obtained with PCR fragments with HS of the indicated lengths. Insert size in all cases was 714 bp. Each bar represents the average insertion efficiency from two or more independent experiments (Table 1). Error bars represent the +/- SD. PCR fragments were nucleofected in HEK293T cells at the concentration indicated and cells were counted by flow cytometer 3 days later. FIG. 10C: confocal images of cells 3 days after nucleofection. GFP: green, DNA: blue. The GFP subcellular localization is as expected for in-frame translational fusion to *SMC3*, a nuclear protein.

FIG. 11A-11B. Flow cytometer plots of cells tagged with PCR repair templates. Flow cytometer plots showing the number of cells (Y axis) and their GFP intensity (X axis). FIG. 11A: *Lamin A/C*, *RAB11A* and *SMC3* were targeted in HEK293T cells with an eGFP containing PCR fragment with or without ~35 bp Homology arm Sequences (HS). Green double arrows indicate the % of GFP+ cells. For every experiment, non-nucleofected cells were also run through the flow cytometer to determine background fluorescence (<0.5% cells). Note that donors without HS yield GFP+ values slightly above background, consistent with a low level of integration by NHEJ or MMEJ. FIG. 11B: *RAB11A* was targeted in HEK293T (GFP1-10) cells using a GFP11-containing repair template with or without ~35 bp Homology arm Sequences (HS). Green double arrows indicate the % of GFP+ cells. Non-nucleofected cells were also run through the flow cytometer to determine background fluorescence (<0.5% cells). Note that HEK293T cells that express GFP1-10 cells have a higher intrinsic fluorescence than HEK293T cells.

FIG. 12A-12B. Derivation of GFP+ and GFP- clones from a single editing experiment targeting the *Lamin A/C* locus with a GFP-containing PCR fragment. FIG. 12A: schematic showing the donor (green with blue Homology arm Sequences - HS) and targeted locus (grey). HEK293T cells were edited at the *Lamin A/C* locus with an eGFP PCR donor with 33/33 HS, and FACS-sorted as GFP+ and GFP- cells. The clones were amplified and examined by confocal microscopy. All GFP+ cells exhibit the expected nuclear membrane localization expected from a GFP translation fusion with *Lamin A/C*. FIG. 12B: statistics of genotyping results for GFP+ and GFP- single clones. See also FIG. 13.

FIG. 13A-13E. Structure of imprecise GFP knock-in edits. Schematics showing the GFP inserts obtained in the experiment described in FIG. 12. *Lamin A/C* locus (grey line), Full-length left HS (L, 33 bp) and right HS (R, 33 bp) (blue), GFP (green, with length of GFP sequence indicated), Indel (red). GFP+ indicates cells with *Lamin A/C* GFP signal. FIG. 13A: precise edit for reference. FIG. 13B: edits with imprecise right junctions—(b1) Contain an 11 bp duplication of the *Lamin A/C* locus sequence just downstream the right HS; (b2) Contain a 6 bp deletion of the *Lamin A/C* locus sequence just downstream the right HS; (b3) Contain a deletion of the last 19 bp of the right HS and of the 8 bp just downstream the right HS; (b4) Contain an 11 bp deletion inside the right HS; (b5) Contain only the 363 first bp of GFP sequence; (b6) Contain only the 70 first bp of GFP sequence followed by a 4 bp insertion and a full deletion of the right HS together with a 4 bp deletion of the *Lamin A/C* locus sequence just downstream the right HS sequence. Sequencing from wild-type size allele from Het GFP+ cell; and (b7) Contain only the 22 first bp of GFP sequence followed by a 5 bp insertion and a deletion of the first 13 bp of the right HS. Sequencing from wild-type size allele from Het GFP+ cell. FIG. 13C: edits with imprecise left junctions—(c1) Contain a 23 bp duplication of the left HS just upstream the GFP sequence; (c2) Contain on the left side the 8 first bp of GFP, followed by the 25 bp of the left HS sequence upstream of GFP, and followed by full-length GFP sequence; (c3) Contain a 52 bp insertion followed by the last 469 bp of GFP sequence; and (c4) Contain a deletion of the last 7 bp of the left HS followed by the last 68 bp of GFP sequence. Sequencing from wild-type size allele from Het GFP+ cell. FIG. 13D: edit with internal deletion—(d1) Contain the 556 first bp of GFP sequence followed by a 12 bp insertion and the last 13 bp of GFP sequence. FIG. 13E: eEdit with inverted insertion—(e1) Contain the left HS and first 501 bp of GFP sequence inverted.

FIG. 14. Insertion efficiency relative to distance from the DSB. Graph showing the efficiency % of editing (Y axis) vs distance from the DSB (X axis) for PCR donors (*RAB11A*, cr1777, GFP11 insertion) or ssODNs (GFP11 insertion (see FIG. 5) except for *PYMI* were

3xFlag insertion was monitored by genotyping single cell colonies (see FIG. 6)). Each line links editing experiments performed with the same guide RNA. ssODNs (optimal polarity, FIG. 5) were designed to insert the edit at varying distances from the DSB as indicated. The sequence between the edit and the DSB was partially recoded to improve insertion efficiency (FIG. 6) and Cas9 re-cutting of the edited locus while preserving coding potential.

FIG. 15A-15B. Illumina sequencing to monitor template switching. FIG. 15A: schematic representation of the experimental design (see FIG. 7). Stars in color represent silent mutations (A, C, G or T) used to monitor template switching. FIG. 15B: the probability of a mutation (relative to the “No mutation” template) at each nucleotide position in the region of the ssODN repair template, after removal of incompletely mapped and low-quality reads. Bars are color-coded by identity of the incorporated nucleotide. Green: A, blue: C, black: G, red: T. PCR control: Two cell populations that received separately a wild-type ssODN or a mutant ssODN (1/6 mutations) were combined for PCR amplification. This control was used to determine basal levels of template switching that might occur during PCR amplification. These levels are 25-fold lower than observed in cells co-transfected with the wild-type ssODN (donor 1) and a 1/6 mutations ssODN (donor 2) (0.02% versus 0.50%).

FIG. 16A-16E. Schematics showing repair of a DSB (grey) using two ssODN donors. Donor 1 contains an insert (green) and two homology arms (blue). Donor 2 contains the same insert with mutations (red stars) and no homology arms. Arrows indicate 3' ends and dotted lines represented newly synthesized DNA. FIG. 16A: strand invasion—the DSB is resected creating two 3' overhangs on each side of the DSB. The right overhang pairs with Donor 1 and is extended by DNA synthesis. FIGS. 16B and 16C: template switching—the newly synthesized strand withdraws from Donor 1 and anneals to Donor 2 (B), and withdraws from Donor 2 and anneals back to Donor 1 (C). FIG. 16D: annealing—the newly synthesized strand withdraws from Donor 1 and anneals back to the locus. FIG. 16E: second strand synthesis and ligation—the newly synthesized strand is used as a template for second strand synthesis. The resulting edit is a hybrid insertion containing sequences from Donor 1 and Donor 2.

FIG. 17A-17B. Comparison between nucleofection (FIG. 17A) and Lipofection (FIG. 17B) in HEK293T cells.

FIG. 18A-18C. Various gene tagging in HEK293T cells using PCR or gBlock donors. FIG. 18A—Nucleofection. FIG. 18B—Lipofection. FIG. 18C—Expression patterns.

FIG. 19A-19B. GFP/RFP co-tagging in HEK293T cells. FIG. 19A—two genes editing (nucleofection). FIG. 19B—two colors editing (lipofection).

FIG. 20. Isolation of HEK293T cells with stress granules proteins tagged with GFP.

FIG. 21A-21B. U2OS and DLD1 genes tagging. FIG. 21A—U2OS (nucleofection).

FIG. 21B—DLD1 (nucleofection).

DETAILED DESCRIPTION OF THE INVENTION

5 It is understood that the present invention is not limited to the particular methods and components, etc., described herein, as these may vary. It is also to be understood that the terminology used herein is used for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention. It must be noted that as used herein and in the appended claims, the singular forms “a,” “an,” and “the” include the plural
10 reference unless the context clearly dictates otherwise. Thus, for example, a reference to a “protein” is a reference to one or more proteins, and includes equivalents thereof known to those skilled in the art and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention
15 belongs. Specific methods, devices, and materials are described, although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention.

All publications cited herein are hereby incorporated by reference including all journal articles, books, manuals, published patent applications, and issued patents. In
20 addition, the meaning of certain terms and phrases employed in the specification, examples, and appended claims are provided. The definitions are not meant to be limiting in nature and serve to provide a clearer understanding of certain aspects of the present invention.

I. Definitions

Unless otherwise indicated, the terms “polynucleotide” and “nucleic acid” refer to a
25 deoxyribonucleotide or ribonucleotide polymer, in linear or circular conformation, and in either single- or double-stranded form. The terms can encompass known analogs of natural nucleotides, as well as nucleotides that are modified in the base, sugar and/or phosphate moieties (e.g., phosphorothioate backbones). In general, an analog of a particular nucleotide has the same base-pairing specificity; i.e., an analog of A will base-pair with T.

30 The term “nucleotide” refers to deoxyribonucleotides or ribonucleotides. The nucleotides may be standard nucleotides (i.e., adenosine, guanosine, cytidine, thymidine, and uridine) or nucleotide analogs. A nucleotide analog refers to a nucleotide having a modified purine or pyrimidine base or a modified ribose moiety. A nucleotide analog may be a naturally occurring nucleotide (e.g., inosine) or a non-naturally occurring nucleotide. Non-

limiting examples of modifications on the sugar or base moieties of a nucleotide include the addition (or removal) of acetyl groups, amino groups, carboxyl groups, carboxymethyl groups, hydroxyl groups, methyl groups, phosphoryl groups, and thiol groups, as well as the substitution of the carbon and nitrogen atoms of the bases with other atoms (e.g., 7-deaza purines). Nucleotide analogs also include dideoxy nucleotides, 2'-O-methyl nucleotides, locked nucleic acids (LNA), peptide nucleic acids (PNA), and morpholinos.

A "gene," as used herein, refers to a DNA region (including exons and introns) encoding a gene product, as well as all DNA regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. Accordingly, a gene includes, but is not necessarily limited to, promoter sequences, terminators, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites, and locus control regions.

As used herein, an "edit" is the desired modification to be introduced into the genome. In other words, an edit is any change in the genomic sequence that is included in the repair template polynucleotide. Edits can include, for example, base pair insertions, deletions or changes.

The term "intervening sequence" refers to a sequence between the edit and the double-stranded break (DSB). An intervening sequence can be unmodified (identical to genome sequence) or can be modified (for example, see FIG. 8.).

As used herein, a "homology arm," "homology sequence" or "sequence homologous" to a reference or target gene/sequence describes a polynucleotide sequence that has substantial sequence identity to a corresponding segment of the reference or target gene/sequence, e.g., at least 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% identical or even 100% identical, to the nucleotide sequence of the reference or target gene/sequence, such that, when placed under appropriate conditions, homologous recombination can take place between a pair of "homologous sequences" and their reference or target gene/sequence. The homology arms have substantial sequence identity to the sequence upstream and downstream of the targeted site in the target nucleic acid molecule.

For edits inserted to the right of a DSB: the right homology arm corresponds to the genomic sequence immediately to the right of the insertion point of the edit and the left homology arm corresponds to the genomic sequence immediately on the left side of the DSB.

For edits inserted to the left of a DSB: the left homology arm corresponds to the genomic sequence immediately to the left of the insertion point of the edit and the right

homology arm corresponds to the genomic sequence immediately on the right side of the DSB.

The terms “target sequence,” “target nucleic acid” or “target DNA sequence,” when used to refer to a pre-determined segment of a genomic sequence or polynucleotide is similarly defined in regard to the percentage sequence identity between the target sequence and its corresponding guide RNA. On the other hand, a “homology arm” or “target sequence” is of the appropriate length that ensures its purpose. Typically, a “homology arm” is in the size range of about 10-100, 10-90, 10-80, 15-75, 15-70, 15-65, 15-60, 15-55, 15-50, 15-45, 15-40, 15-35, 20-50, 20-45, 20-40, 20-35, 25-40, 25-35 or 30-35 nucleotides (e.g., about 30, 35, 40, 45, 50, 55 or 60 nucleotides in length); whereas a “target sequence” may vary in the size range of about 10-50, 15-45, or 20-40 (e.g., about 20, 25, or 30) nucleotides. In some embodiments, the target sequence contains a sequence that is suitable as a substrate for an RNA-guided DNA endonuclease (e.g., a Cas9 nuclease) (i.e., a nuclease target sequence site). In some embodiments, the target sequence contains a sequence that is suitable as a substrate for Cfp1 endonuclease (i.e., an endonuclease target sequence site).

Techniques for determining nucleic acid and amino acid sequence identity are known in the art. Typically, such techniques include determining the nucleotide sequence of the mRNA for a gene and/or determining the amino acid sequence encoded thereby, and comparing these sequences to a second nucleotide or amino acid sequence. Genomic sequences can also be determined and compared in this fashion. In general, identity refers to an exact nucleotide-to-nucleotide or amino acid-to-amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. Two or more sequences (polynucleotide or amino acid) can be compared by determining their percent identity. The percent identity of two sequences, whether nucleic acid or amino acid sequences, is the number of exact matches between two aligned sequences divided by the length of the shorter sequences and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, 2 *ADVANCES IN APPLIED MATHEMATICS* 482-89 (1981). This algorithm can be applied to amino acid sequences by using the scoring matrix developed by Dayhoff, 3 (5 Suppl.) *ATLAS OF PROTEIN SEQUENCES AND STRUCTURE*, M. O. Dayhoff ed. 353-58, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov et al., 14(6) *NUCL. ACIDS RES.* 6745-63 (1986). An exemplary implementation of this algorithm to determine percent identity of a sequence is provided by the Genetics Computer Group (Madison, Wis.) in the “BestFit” utility application. Other suitable programs for calculating the percent identity or similarity

between sequences are generally known in the art, for example, another alignment program is BLAST, used with default parameters. For example, BLASTN and BLASTP can be used using the following default parameters: genetic code=standard; filter=none; strand=both; cutoff=60; expect=10; Matrix=BLOSUM62; Descriptions=50 sequences; sort by=HIGH
 5 SCORE; Databases=non-redundant, GenBank+EMBL+DDBJ+PDB+GenBank CDS translations+Swiss protein+Spupdate+PIR. Details of these programs can be found on the GenBank website.

“Cas9” or (CRISPR associated protein 9) is an RNA-guided DNA endonuclease enzyme associated with the CRISPR (Clustered Regularly Interspersed Palindromic Repeats)
 10 adaptive immunity system in *Streptococcus pyogenes*, among other bacteria. *S. pyogenes* utilizes the CRISPR system to memorize and later interrogate and cleave foreign DNA, such as the DNA of an invading bacteriophage. Cas9, complexed with a guide RNA, performs this interrogation by unwinding foreign DNA and checking whether the DNA contains any sequence segment complementary to a spacer region of the guide RNA. If the guide RNA
 15 finds sequence complementarity in the DNA, it is cleaved by Cas9.

“Cpf1” or “CRISPR/Cpf1” is a DNA editing technology analogous to the CRISPR/Cas9 system. Cpf1 is an RNA-guided DNA endonuclease enzyme associated with the CRISPR adaptive immunity system in *Prevotella* and *Francisella*, among other bacteria. Cpf1 is a smaller and simpler endonuclease as compared to Cas9 because Cpf1 only requires
 20 one RNA molecule to cut DNA while Cas9 requires two. Cpf1 is a Type V CRISPR/Cas system containing a 1,300 amino acid protein.

As used herein, “sgRNA” or “small guide RNA” refers to a short RNA molecule that is capable of forming a complex with Cas9 protein and contains a segment of about 20 nucleotides complementary to a target DNA sequence, such that the Cas9-sgRNA complex
 25 directs Cas9 cleavage of a target DNA sequence upon the sgRNA recognizing the complementary sequence in the target DNA sequence. Accordingly, a sgRNA is approximately a 20-base sequence (ranging from about 10-50, 15-45, or 20-40, for example, 15, 20, 25, or 30 bases) specific to the target DNA 5’ of a non-variable scaffold sequence.

As used herein, the term “endogenous sequence” refers to a chromosomal sequence
 30 that is native to the cell.

The term “exogenous,” as used herein, refers to a sequence that is not native to the cell, or a chromosomal sequence whose native location in the genome of the cell is in a different chromosomal location.

The term “heterologous” refers to an entity that is not endogenous or native to the cell of interest. For example, a heterologous protein refers to a protein that is derived from or was originally derived from an exogenous source, such as an exogenously introduced nucleic acid sequence. In some instances, the heterologous protein is not normally produced by the cell of interest.

II. RNA-Guided Endonucleases

In particular embodiments, the compositions and methods of the present invention utilize RNA-guided endonucleases. In some embodiments, the endonuclease comprises at least one nuclear localization signal, which permits entry of the endonuclease into the nuclei of eukaryotic cells and embryos such as, for example, non-human one-cell embryos. In other embodiments, RNA-guided endonucleases comprise at least one nuclease domain and at least one domain that interacts with a guide RNA. An RNA-guided endonuclease is directed to a specific nucleic acid sequence (or target sequence/site) by a guide RNA. The guide RNA interacts with the RNA-guided endonuclease as well as the target site such that, once directed to the target site, the RNA-guided endonuclease is able to introduce a double-stranded break into the target site nucleic acid sequence. Since the guide RNA provides the specificity for the targeted cleavage, the endonuclease of the RNA-guided endonuclease is universal and can be used with different guide RNAs to cleave different target nucleic acid sequences.

The RNA-guided endonuclease can be derived from a clustered regularly interspersed short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system. The CRISPR/Cas system can be a type I, a type II, or a type III system. Non-limiting examples of suitable CRISPR/Cas proteins include Cas3, Cas4, Cas5, Cas5e (or CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9, Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (or CasA), Cse2 (or CasB), Cse3 (or CasE), Cse4 (or CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csz1, Csx15, Csf1, Csf2, Csf3, Csf4, and Cul966.

In one embodiment, the RNA-guided endonuclease is derived from a type II CRISPR/Cas system. In specific embodiments, the RNA-guided endonuclease is derived from a Cas9 protein. The Cas9 protein can be from *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus* sp., *Nocardiopsis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycooides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus*

delbrueckii, *Lactobacillus salivarius*, *Microscilla marina*, *Burkholderiales bacterium*,
Polaromonas naphthalenivorans, *Polaromonas* sp., *Crocospaera watsonii*, *Cyanothece* sp.,
Microcystis aeruginosa, *Synechococcus* sp., *Acetohalobium arabaticum*, *Ammonifex degensii*,
Caldicelulosiruptor beccsii, *Candidatus Desulforudis*, *Clostridium botulinum*, *Clostridium*
5 *difficile*, *Finegoldia magna*, *Natranaerobius thermophilus*, *Pelotomaculum the*
mopropionicum, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochromatium*
vinosum, *Marinobacter* sp., *Nitrosococcus halophilus*, *Nitrosococcus watsoni*,
Pseudoalteromonas haloplanktis, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*,
Anabaena variabilis, *Nodularia spumigena*, *Nostoc* sp., *Arthrospira maxima*, *Arthrospira*
10 *platensis*, *Arthrospira* sp., *Lyngbya* sp., *Microcoleus chthonoplastes*, *Oscillatoria* sp.,
Petrotoga mobilis, *Thermosipho africanus*, or *Acaryochloris marina*.

In other embodiments, the RNA-guided endonuclease is derived from another Cas
nuclease including, but not limited to, Cpf1, C2c1, C2c2, and C2c3 proteins. Cpf1 is similar
to Cas9, and contains a RuvC-like nuclease domain. See Zetsche et al., 163 CELL 1-13
15 (2015).

III. Guide RNA

In some embodiments of the present disclosure, a CRISPR/Cas nuclease system
includes at least one guide RNA. In some embodiments, the guide RNA and the Cas protein
may form a ribonucleoprotein (RNP), e.g., a CRISPR/Cas complex. The guide RNA may
20 guide the Cas protein to a target sequence on a target nucleic acid molecule, where the guide
RNA hybridizes with, and the Cas protein cleaves, the target sequence. In some
embodiments, the CRISPR/Cas complex may be a Cpf1/guide RNA complex. In some
embodiments, the CRISPR complex may be a Type-II CRISPR/Cas9 complex. In some
embodiments, the Cas protein may be a Cas9 protein. In some embodiments, the
25 CRISPR/Cas9 complex may be a Cas9/guide RNA complex.

A guide RNA for a CRISPR/Cas9 nuclease system comprises a CRISPR RNA
(crRNA) and a tracr RNA (tracr). In another embodiment, a single guide RNA (sgRNA)—a
chimer of cr/tracrRNA—can be used. See Doudna, J.A. & Charpentier, E., 346(6213)
SCIENCE 1258096 (2014). A guide RNA for a CRISPR/Cpf1 nuclease system comprises a
30 crRNA. In some embodiments, the crRNA may comprise a targeting sequence that is
complementary to and hybridizes with the target sequence on the target nucleic acid
molecule. The crRNA may also comprise a flagpole that is complementary to and hybridizes
with a portion of the tracrRNA. In some embodiments, the crRNA may parallel the structure
of a naturally occurring crRNA transcribed from a CRISPR locus of a bacteria, where the

targeting sequence acts as the spacer of the CRISPR/Cas9 system, and the flagpole corresponds to a portion of a repeat sequence flanking the spacers on the CRISPR locus.

The guide RNA may target any sequence of interest via the targeting sequence of the crRNA. In some embodiments, the degree of complementarity between the targeting
 5 sequence of the guide RNA and the target sequence on the target nucleic acid molecule may be about 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%. In some
 embodiments, the targeting sequence of the guide RNA and the target sequence on the target
 nucleic acid molecule may be 100% complementary. In other embodiments, the targeting
 sequence of the guide RNA and the target sequence on the target nucleic acid molecule may
 10 contain at least one mismatch. For example, the targeting sequence of the guide RNA and the
 target sequence on the target nucleic acid molecule may contain 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10
 mismatches. In some embodiments, the targeting sequence of the guide RNA and the target
 sequence on the target nucleic acid molecule may contain 1-6 mismatches. In some
 embodiments, the targeting sequence of the guide RNA and the target sequence on the target
 15 nucleic acid molecule may contain 5 or 6 mismatches.

The length of the targeting sequence of the guide RNA may depend on the CRISPR/Cas9 system and components used. For example, different Cas9 proteins from
 different bacterial species have varying optimal targeting sequence lengths. Accordingly, the
 targeting sequence may comprise 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,
 20 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or more than 50 nucleotides in length. In
 some embodiments, the targeting sequence may comprise 18-24 nucleotides in length. In
 some embodiments, the targeting sequence may comprise 19-21 nucleotides in length. In
 some embodiments, the targeting sequence may comprise 20 nucleotides in length.

IV. Target Site/Sequence of the Target Nucleic Acid Molecule

25 An RNA-guided endonuclease in conjunction with a guide RNA is directed to a target
 site in the chromosomal sequence, wherein the RNA-guided endonuclease introduces a
 double-stranded break in the chromosomal sequence. The target site has no sequence
 limitation except that the sequence is immediately followed (downstream) by a consensus
 sequence. This consensus sequence is also known as a protospacer adjacent motif (PAM).
 30 Examples of PAMs include, but are not limited to, NGG, NGGNG, and NNAGAAW
 (wherein N is defined as any nucleotide and W is defined as either A or T). In particular
 embodiments, the first region (at the 5' end) of the guide RNA is complementary to the
 protospacer of the target sequence. Typically, the first region of the guide RNA is about 19

to 21 nucleotides in length. Thus, in certain aspects, the sequence of the target site in the chromosomal sequence is 5'-N₁₉₋₂₁-*NGG*-3'. The PAM is in italics.

The target site can be in the coding region of a gene, in an intron of a gene, in a control region of a gene, in a non-coding region between genes, etc. The gene can be a protein coding gene or an RNA coding gene. The gene can be any gene of interest.

V. Linear Donor Polynucleotides & Design Parameters Thereof

In certain embodiments, the present invention provides a double-stranded, linear donor polynucleotide comprising a template polynucleotide encoding an edit flanked by an intervening sequence and two homology arms. In other embodiments, the donor polynucleotide comprises a template polynucleotide encoding an edit flanked by two homology arms.

In some embodiments, the template polynucleotide of the double-stranded, linear donor polynucleotide may correspond to an endogenous sequence of a target cell. In some embodiments, the endogenous sequence may be a genomic sequence of the cell. In some embodiments, the endogenous sequence may be a chromosomal or extrachromosomal sequence. In some embodiments, the endogenous sequence may be a plasmid sequence of the cell. In some embodiments, the template sequence may be substantially identical to a portion of the endogenous sequence in a cell at or near the cleavage site, but comprise at least one nucleotide change (i.e., an "edit" as defined herein). In some embodiments, the repair of the cleaved target nucleic acid molecule with the template may result in an edit comprising an insertion, deletion, or substitution of one or more nucleotides of the target nucleic acid molecule. In some embodiments, the edit may result in one or more amino acid changes in a protein expressed from a gene comprising the target sequence. In some embodiments, the edit or mutation may result in one or more nucleotide changes in an RNA expressed from the target gene. In some embodiments, the edit may alter the expression level of the target gene. In some embodiments, the edit may result in increased or decreased expression of the target gene. In some embodiments, the edit may result in gene knockdown. In some embodiments, the edit may result in gene knockout. In some embodiments, the repair of the cleaved target nucleic acid molecule with the template may result in replacement of an exon sequence, an intron sequence, a transcriptional control sequence, a translational control sequence, or a non-coding sequence of the target gene.

In other embodiments, the double-stranded, linear donor polynucleotide encoding an edit may comprise an exogenous sequence. In some embodiments, the exogenous sequence may comprise a protein or RNA coding sequence operably linked to an exogenous promoter

sequence such that, upon integration of the exogenous sequence into the target nucleic acid molecule, the cell is capable of expressing the protein or RNA encoded by the integrated sequence. In other embodiments, upon integration of the exogenous sequence into the target nucleic acid molecule, the expression of the integrated sequence may be regulated by an endogenous promoter sequence. In some embodiments, the exogenous sequence may be a chromosomal or extrachromosomal sequence. In some embodiments, the exogenous sequence may provide a cDNA sequence encoding a protein or a portion of the protein. In yet other embodiments, the exogenous sequence may comprise an exon sequence, an intron sequence, a transcriptional control sequence, a translational control sequence, or a non-coding sequence. In some embodiments, the integration of the exogenous sequence may result in gene knock-in.

In the double-stranded, linear donor polynucleotide, the template polynucleotide is flanked by a first homology arm and a second homology arm, e.g., a left homology arm and a right homology arm. These sequences to the left and right of the template polynucleotide have substantial sequence identity to sequences located to the left and right, respectively, of the target site of the RNA-guided endonuclease in the target nucleic acid molecule. Because of these sequence similarities, homology arms permit homologous recombination between the donor polynucleotide and the targeted sequence such that the template polynucleotide can serve as a template for DNA synthesis. In certain embodiments, the linear donor polynucleotide comprises a template polynucleotide encoding an edit flanked by an intervening sequence and two homology arms.

In certain embodiments, specifically, for edits inserted to the right of a DSB, the right homology arm corresponds to the genomic sequence immediately to the right of the insertion point of the edit and the left homology arm corresponds to the genomic sequence immediately on the left side of the DSB. In other embodiments, specifically, for edits inserted to the left of a DSB, the left homology arm corresponds to the genomic sequence immediately to the left of the insertion point of the edit and the right homology arm corresponds to the genomic sequence immediately on the right side of the DSB.

In particular embodiments, each homology arm can range in length from about 10 nucleotides to about 100 nucleotides. The recited range includes ranges within the recited range including, but not limited to, 10-100, 10-90, 10-80, 15-75, 15-70, 15-65, 15-60, 15-55, 15-50, 15-45, 15-40, 15-35, 20-50, 20-45, 20-40, 20-35, 25-40, 25-35 or 30-35 nucleotides (e.g., about 30, 35, 40, 45, 50, 55 or 60 nucleotides in length). In a specific embodiment, a homology arm is 15-60 nucleotides in length. In another embodiment, a homology arm is 25-

45 nucleotides in length. In yet another embodiment, a homology arm is 30-40 nucleotides in length. In a further embodiment, a homology arm is 35 nucleotides in length. In certain embodiments, homology arms can comprise different lengths within the range.

5 VI. Introducing Genome Editing Compositions into the Cell or Embryo

The RNA-guided endonuclease(s) (or encoding nucleic acid), the guide RNA(s) (or encoding DNA), and the double-stranded, linear donor polynucleotide can be introduced into a cell or embryo by a variety of means. In some embodiments, the cell or embryo is transfected. Suitable transfection methods include calcium phosphate-mediated transfection, nucleofection (or electroporation), cationic polymer transfection (e.g., DEAE-dextran or polyethylenimine), viral transduction, virosome transfection, virion transfection, liposome transfection, cationic liposome transfection, immunoliposome transfection, nonliposomal lipid transfection, dendrimer transfection, heat shock transfection, magnetofection, lipofection, gene gun delivery, impalefection, sonoporation, optical transfection, and
10 proprietary agent-enhanced uptake of nucleic acids. Transfection methods are well known in the art. In other embodiments, the molecules are introduced into the cell or embryo by microinjection. In certain embodiments, the embryo is a fertilized one-cell stage embryo of the species of interest. In such embodiments, the molecules can be injected into the pronuclei of one-cell embryos.
15

The RNA-guided endonuclease(s) (or encoding nucleic acid), the guide RNA(s) (or DNAs encoding the guide RNA), and the double-stranded, linear donor polynucleotide(s) can be introduced into the cell or embryo simultaneously or sequentially. The ratio of the RNA-guided endonuclease(s) (or encoding nucleic acid) to the guide RNA(s) (or encoding DNA) generally will be about stoichiometric such that they can form an RNA-protein complex. In
20 one embodiment, DNA encoding an RNA-guided endonuclease and DNA encoding a guide RNA are delivered together within a plasmid vector.
25

In further embodiments, the method comprises maintaining the cell or embryo under appropriate conditions such that the guide RNA(s) directs the RNA-guided endonuclease(s) to the targeted site(s) in the chromosomal sequence, and the RNA-guided endonuclease(s)
30 introduce at least one double-stranded break in the chromosomal sequence. A double-stranded break can be repaired by a DNA repair process such that the chromosomal sequence is modified by a deletion of at least one nucleotide, an insertion of at least one nucleotide, a substitution of at least one nucleotide, or a combination thereof.

In general, the cell is maintained under conditions appropriate for cell growth and/or maintenance. Suitable cell culture conditions are well known in the art. Those of skill in the art appreciate that methods for culturing cells can and will vary depending on the cell type. Routine optimization may be used, in all cases, to determine the best techniques for a particular cell type.

An embryo can be cultured in vitro (e.g., in cell culture). Typically, the embryo is cultured at an appropriate temperature and in appropriate media with the necessary O₂/CO₂ ratio to allow the expression of the RNA endonuclease and guide RNA, if necessary. Suitable non-limiting examples of media include M2, M16, KSOM, BMOC, and HTF media. A skilled artisan will appreciate that culture conditions can and will vary depending on the species of embryo. Routine optimization may be used, in all cases, to determine the best culture conditions for a particular species of embryo. In some cases, a cell line may be derived from an in vitro-cultured embryo (e.g., an embryonic stem cell line).

Alternatively, an embryo may be cultured in vivo by transferring the embryo into the uterus of a female host. Generally speaking the female host is from the same or similar species as the embryo. In certain embodiments, the female host is pseudo-pregnant. Methods of preparing pseudo-pregnant female hosts are known in the art. Additionally, methods of transferring an embryo into a female host are known. Culturing an embryo in vivo permits the embryo to develop and can result in a live birth of an animal derived from the embryo. Such an animal would comprise the modified chromosomal sequence in every cell of the body.

VII. Cell and Embryo Types

A variety of eukaryotic cells and embryos are suitable for use in the method. For example, the cell can be a human cell, a non-human mammalian cell, a non-mammalian vertebrate cell, an invertebrate cell, an insect cell, a plant cell, a yeast cell, or a single cell eukaryotic organism. In general, the embryo is non-human mammalian embryo. In specific embodiments, the embryos can be a one-cell non-human mammalian embryo. Exemplary mammalian embryos, including one-cell embryos, include without limit mouse, rat, hamster, rodent, rabbit, feline, canine, ovine, porcine, bovine, equine, and primate embryos. In still other embodiments, the cell can be a stem cell. Suitable stem cells include without limit embryonic stem cells, ES-like stem cells, fetal stem cells, adult stem cells, pluripotent stem cells, induced pluripotent stem cells, multipotent stem cells, oligopotent stem cells, unipotent stem cells and others. In exemplary embodiments, the cell is a mammalian cell.

Non-limiting examples of suitable mammalian cells include Chinese hamster ovary (CHO) cells, baby hamster kidney (BHK) cells; mouse myeloma NSO cells, mouse embryonic fibroblast 3T3 cells (NIH3T3), mouse B lymphoma A20 cells; mouse melanoma B16 cells; mouse myoblast C2C12 cells; mouse myeloma SP2/0 cells; mouse embryonic mesenchymal C3H-10T1/2 cells; mouse carcinoma CT26 cells, mouse prostate DuCuP cells; mouse breast EMT6 cells; mouse hepatoma Hepa1c1c7 cells; mouse myeloma J5582 cells; mouse epithelial MTD-1A cells; mouse myocardial MyEnd cells; mouse renal RenCa cells; mouse pancreatic RIN-5F cells; mouse melanoma X64 cells; mouse lymphoma YAC-1 cells; rat glioblastoma 9L cells; rat B lymphoma RBL cells; rat neuroblastoma B35 cells; rat hepatoma cells (HTC); buffalo rat liver BRL 3A cells; canine kidney cells (MDCK); canine mammary (CMT) cells; rat osteosarcoma D17 cells; rat monocyte/macrophage DH82 cells; monkey kidney SV-40 transformed fibroblast (COS7) cells; monkey kidney CVI-76 cells; African green monkey kidney (VERO-76) cells; human embryonic kidney cells (HEK293, HEK293T); human cervical carcinoma cells (HELA); human lung cells (W138); human liver cells (Hep G2); human U2-OS osteosarcoma cells, human A549 cells, human A-431 cells, and human K562 cells. An extensive list of mammalian cell lines may be found in the American Type Culture Collection catalog (ATCC, Manassas, Va.).

Without further elaboration, it is believed that one skilled in the art, using the preceding description, can utilize the present invention to the fullest extent. The following examples are illustrative only, and not limiting of the remainder of the disclosure in any way whatsoever.

EXAMPLES

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how the compounds, compositions, articles, devices, and/or methods described and claimed herein are made and evaluated, and are intended to be purely illustrative and are not intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers (e.g., amounts, temperature, etc.) but some errors and deviations should be accounted for herein. Unless indicated otherwise, parts are parts by weight, temperature is in degrees Celsius or is at ambient temperature, and pressure is at or near atmospheric. There are numerous variations and combinations of reaction conditions, e.g., component concentrations, desired solvents, solvent mixtures, temperatures, pressures and other reaction ranges and conditions that can be used to optimize the product purity and yield obtained from

the described process. Only reasonable and routine experimentation will be required to optimize such process conditions.

Precision Genome Editing Using Synthesis-Dependent Repair of Cas9-Induced DNA Breaks

The RNA-guided DNA endonuclease Cas9 has emerged as a powerful new tool for genome engineering. Cas9 creates targeted double-strand breaks (DSBs) in the genome. Knock-in of specific mutations (precision genome editing) requires homology-directed repair (HDR) of the DSB by synthetic donor DNAs containing the desired edits, but HDR has been reported to be variably efficient. Here, we report that linear DNAs (single and double-stranded) engage in a high-efficiency HDR mechanism that requires only about 35 nucleotides of homology with the targeted locus to introduce edits ranging from about 1 to 1000 nucleotides. We demonstrate the utility of linear donors by introducing fluorescent protein tags in human cells and mouse embryos using PCR fragments. We find that repair is local, polarity-sensitive, and prone to template switching, characteristics that are consistent with gene conversion by synthesis-dependent strand-annealing (SDSA). Our findings enable rational design of synthetic donor DNAs for efficient genome editing.

We documented previously that, in *C. elegans*, HDR can be very efficient provided that the donor DNAs are linear (Paix et al., 44(15) NUCLEIC ACIDS RES. e128 (2016)). Linear donors do not appear to integrate at the DSB, but instead are used as templates for DNA synthesis, as in the synthesis-dependent strand annealing (SDSA) model for gene conversion (Mehta et al., 65(3) MOL. CELL 515-26 e513 (2017); Jasin et al. (2016); Paques, F. & Haber, J.E., 63(2) MICROBIOL. MOL. BIOL. REV. 349-404 (1999)). In *C. elegans*, donors for SDSA can be single (ssODNs) or double-stranded (PCR fragments), and require only short homology arms (~35 bases) to engage the DSB. The repair process is sensitive to insert size and prone to template switching, where synthesis can “jump” between two overlapping donors (Paix et al. (2016)). In human cells, SDSA has been proposed as a repair mechanism for ssODNs (Kan et al., 27(7) GENOME RES. 1099-1111 (2017); Liang et al. (2016)), but not for double-stranded donors, which are thought to participate in a different HDR pathway (Kan et al., 10(4) PLOS GENET. E1004251 (2014)) (Bothmer et al., 8 NAT. COMMUN. 13905 (2017)). Here, we investigate how linear donors engage the DSB repair machinery in mammalian cells. First, we demonstrate that, as in *C. elegans*, PCR fragments with 35bp homology arms function as efficient donors for genome editing in mouse embryos and human cells. Using PCR fragments and ssODNs, we investigate the sequence requirements for efficient repair by linear donors in human cells. Our findings are consistent with SDSA and suggest simple donor DNA design principles to maximize editing efficiency.

Materials and Methods

Detailed results, sequences and solutions. Tables 1-3 lists all experiments, including detailed conditions and results of experimental replicates. Table 5-14 lists sequences of linear donors, plasmids, PCR primers and cr/sgRNAs, respectively. Position of the cr/sgRNAs on the loci targeted in this study can be found in FIG. 9. Results presented in FIGS. 2, 3, 7B/D and 10 are the average of at least two independent experiments and the error bars represent the standard deviation (SD).

Repair templates, Cas9, cr/tracrRNAs and plasmids for cell culture. ssODNs (ultramers) and PCR primers were ordered from IDT and reconstituted at 50 μ M and 100 μ M respectively in water. For the Illumina sequencing experiment shown in FIG. 7F, ssODNs and primers were ordered PAGE purified. PCR fragment donors were synthesized as described in Paix et al., 121-122 METHODS 86-93 (2017).

Cas9 protein was purified as described in Paix et al., 201(1) GENETICS 47-54 (2015). crRNAs and tracrRNA were ordered from IDT and reconstituted in 5mM Tris-HCl pH7.5 at 130 μ M. Plasmids containing repair templates were made using gBlock gene fragments (IDT) and InFusion cloning kit (Clontech), and purified using Qiagen mini-prep kit and eluted in H₂O. For experiments at the *PYMI* locus, the sgRNA was cloned as described in Moyer et al., 129 METHODS CELL BIOL. 19-36 (2015).

Cas9 RNP nucleofection. With the exception of experiments at the *PYMI* locus (see below), all experiments in this study used Cas9 RNP delivery (DeWitt et al., 121-122 METHODS 9-15 (2017)). Nucleofections using Cas9 RNP were performed as described (Leonetti et al., 113(25) PROC. NATL. ACAD. SCI. USA E3501-08 (2016)). HEK293T cells or HEK293T cells expressing a truncated GFP (GFP1-10) (Kamiyama et al., 7 NAT. COMMUN. 11046 (2016)) were grown to 50-75% confluency, trypsinized, pelleted and resuspended at 800000 cells / 80 μ l of PBS. Just before nucleofection, PBS was replaced with 80 μ l of Nucleofection kit V (Lonza). 40 μ l of Cas9 RNP mix (see below) was added to the cells in suspension in Nucleofector kit V and processed using an Amaxa Nucleofector 2b machine (Lonza) using the A023 program. Cells were transferred to culture media and analyzed for fluorescence 3 days after.

The Cas9 RNP mix contains: 6.5 μ M of crRNA and tracrRNA, 9.8 μ M of Cas9 (1.6 μ g/ μ l), a variable concentration of repair templates (see Tables 1-3 for details), 10.4% Glycerol, 131mM KCl, 5.2mM Hepes, 1mM MgCl₂, 0.5mM Tris-HCl, pH7.5.

For sequencing of GFP edits at the *Lamin A/C* locus, cells were sorted (at the JHU Ross Flow Cytometry Core Facility) for GFP signal and cloned in 96 wells plates for

genotyping or pooled in a 6-well plate for microscopy analysis. Single cell clones were lysed using QuickExtract DNA Extraction Solution (Epicentre) and genotyped by PCR using Phusion taq (NEB) with genomic primers outside of the HDR fragment. PCR products were analyzed on agarose gel and sequenced (see FIGS. 12 and 13).

5 Cas9 plasmid transfections. For experiments at the *PYMI* locus, Cas9 and the sgRNA were delivered on plasmids. HEK293T cells were grown to 50-75% confluency in 6 wells plate (with 2ml of culture media per wells). 10.8µl of Cas9 plasmid mix (containing 3.6µl of X-tremeGENE 9 DNA Transfection Reagent from Roche, 892ng of plasmid pX458 containing *PYMI* sgRNA and 3.24pmol of repair template) was added to 120µl of optiMEM
10 glutaMAX media (ThermoFisher), incubated for 15min at room temperature, and next added to the cells. 48h after transfection, cells were sorted for GFP signal (to select for cells that received pX458) and grown out as single cell clones. The single cell clones were lysed and genotyped by PCR. PCR products were directly analyzed on agarose gel or mix with *EcoRI* (NEB) and the corresponding Restriction Enzyme (RE) buffer, digested over-night and
15 analyzed on agarose gel.

Cytometer analysis. For each experiment, 5000 to 10000 cells were analyzed using a Guava EasyCyte 6/2L (Millipore) cytometer. Cells were scored as GFP+ if they exhibited a higher signal than 99.5% of non-transfected control cells. HEK293T (GFP1-10) cells exhibit a higher basal green fluorescence than wild-type HEK293T cells. Cytometer analysis could
20 not be performed on these cells for GFP11-tagged Lamin A/C and SMC3. For those experiments, as well as for RFP tagging, cells were analyzed by fluorescence microscopy and scored manually.

Microscopy. Cells were fixed in 4% PFA and mounted with DAPI. Cells were imaged using a confocal microscope with a 63X objective. > 50 fields of cells (>1000 cells)
25 were selected in the DAPI channel, photographed, and analyzed for GFP or RFP expression manually.

PCR amplicons for Illumina sequencing. HEK293T (GFP1-10) were nucleofected with different combinations of repair ssODNs (FIG. 7E, Tables 1-3). To control for possible template-switching during PCR amplification, we also introduced single donors (wild-type or
30 mutant) in two separate cell populations and combined the cells during PCR amplification. 60h after nucleofection, cells were trypsinized, washed in PBS, and 500000 cells were lysed in 40µl of QuickExtract DNA Extraction Solution. 40µl of H2O was added to each lysis. A total of 6µl of DNA from each experiments were PCR amplified using Phusion Taq and the primer 390 (Forward, in the left end of the insert) and the primer 1849 (Reverse, in the *Lamin*

A/C locus downstream of the right HS of the ssODN used for repair) for 10 cycles at 68.5°C (see Tables 10-13 for primer sequences). After 10 PCR cycles, no band could be detected on agarose gel and ethidium bromide staining. Each PCR reaction was purified using Qiagen Minelute columns and eluted in 10µl of H₂O. 2µl of each PCR were amplified using
5 Phusion taq at 65°C for 20 cycles. PCR reactions did not reach an amplification plateau with this number of cycles. The PCR reactions were performed using primers 1928 (Forward, containing the Illumina sequence and annealing in the same region than primer 390) and Reverse primers containing the Illumina sequence and a specific barcode. The Illumina reverse primers anneal with the *Lamin A/C* locus just upstream of primer 1849 and
10 downstream the right HS of the ssODN used for repair.

PCR amplicons were purified on a 10% non-denaturing TBE/PAGE gel and the band corresponding to the PCR product was cut from the gel, eluted over-night, and precipitated with isopropanol. After resuspension, sample concentrations were quantified on a bioanalyzer, and the barcoded samples were pooled to a concentration of 0.4µM per sample
15 in 10µl. This sample was submitted to the Johns Hopkins School of Medicine Genetics Resources Core Facility for 250 cycle paired-end sequencing on an Illumina MiSeq instrument.

Illumina Sequencing analysis. After de-multiplexing of barcoded samples, the 3' adaptor and all downstream nucleotides were trimmed from the forward reads using Cutadapt
20 (Martin, M., 17(1) EMBNET.JOURNAL 10-12 (2011), and the resulting sequences were mapped to the insert + *Lamin A/C* locus using Bowtie 2 (Langmead, B. & Salzberg, S.L., 9(4) NAT. METHODS 357-59 (2012)). After removing reads that did not fully map to the template and low-quality reads (Q score less than 35; error probability of 0.00032), sequences were parsed for template switching. To score template switches, we evaluated sequencing reads at
25 diagnostic positions and determined whether each position matched the sequence of the wild-type or mutated template. Reads with a diagnostic nucleotide that did not match either the wild-type or mutated template were discarded. Because the PCR control sample contained a mixture of the fully wild-type and fully mutated templates, we used the first diagnostic position (from the right side of the insert) only as an "anchor" to determine the initial identity
30 of the template; this position was not used to score switching. Thereafter, whenever two or more contiguous diagnostic nucleotides indicated a switch in template identity, we scored this as a switch. For the control sample in which both templates were wild-type, we used the "1/6" mutated template for comparison, to determine the rate of false-positive switches in the assay. Because the PCR control experiment was performed with the wild-type and "1/6"

mutated template (FIG. 15 and Tables 1-3), we also used the “1/6” mutated template for scoring switches in this sample. See Table 15 for details.

Cas9 RNP injection in mouse zygotes. All mouse experiments were carried out under protocols approved by the JHU animal care and use committee. The PCR fragment donor was synthesized as described in Paix et al. (2017). The plasmid donor was generated using a gBlock and restriction enzyme cloning, and purified by Qiagen midi-prep kit and eluted in injection buffer (10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA). Pronuclear injections of zygotes (from B6SJLF1/J parents (Jackson labs)) was performed by the JHU Transgenic facility at a final concentration: 30ng/μl Cas9 protein (PNABio), 0.6μM each of crRNA/TracrRNA (Dharmacon) and PCR donor (3ng/μl or 5ng/μl) or plasmid donor (10ng/μl). The Cas9 protein, crRNA, tracrRNA were combined from stocks at 1000ng/μl, 20μM, 20μM respectively and incubated at 4°C for 10 minutes. Then injection buffer was added to dilute to the final working concentrations above (Tables 1-3) along with repair vector or fragment. The solution was microcentrifuged 5 min at 13000xg and the solution used for injection.

Pups were genotyped using genomic primers immediately outside of the PCR donor sequence, or using one primer in mCherry and one upstream of the 483 bp homology arms in the case of the plasmid donor. Genomic DNA from all pups was also subjected to PCR amplification with internal mCherry specific primers to identify random insertions of the donor template (locus-specific mCherry negative/internal mCherry product positive).

We identified 7 pups (11%, out of 60 pups without mCherry insertion at the *Adcy3* locus) with potential transgenic insertions of the PCR fragment at other undetermined loci. In contrast, we identified no transgenics (0%, out of 20 pups without mCherry insertion at the *Adcy3* locus) when using the plasmid donor.

Results

mCherry-tagging of a mouse locus using a PCR donor with short homology arms. In mammalian systems, ssODNs and plasmids are most commonly used as donors for genome editing (Danner et al., 28(708) MAMMALIAN GENOME 262-74 (2017)). To test whether PCR fragments with short homology arms can also function as donors, we designed a PCR fragment to insert mCherry near the C-terminus of the mouse adenylyl cyclase 3 (*Adcy3*) locus. The mCherry open reading frame (739 bp) flanked by 36bp homology arm sequences (HS) for the *Adcy3* locus was amplified by PCR. The purified PCR fragment and *in vitro* assembled Cas9 complexes were co-injected into mouse zygotes, and the resulting pups were genotyped by PCR and Sanger sequencing (FIG. 1). We identified 27/87 pups with a correct size insertion at the *Adcy3* locus (31% editing efficiency). Sequencing of 10 full-size

mCherry edits revealed them all to be precise (no indels). A parallel editing experiment using an mCherry supercoiled plasmid with 500 bp HS yielded 5 edits from 25 pups (20% editing efficiency). Similar knock-in efficiencies have also been reported using long single-stranded donors (Quadros et al., 18(1) GENOME BIOL. 92 (2017)). These results suggest that single-stranded DNAs, plasmids and PCR fragments function with similar efficiency for genome editing in mouse embryos. Unlike single-strand DNAs and plasmids, PCR fragments have the added convenience of ease of synthesis especially for long inserts.

GFP-tagging of human loci using PCR donors with short homology arms. To determine whether PCR fragments can also function for genome editing in human cells, we attempted to knock-in GFP at three loci in HEK293T cells. We designed the HS to insert GFP 0, 11 and 5 bp away from a Cas9 cleavage site in the *Lamin A/C*, *RAB11A*, and *SMC3* ORFs, respectively (FIGS. 2 and S2). The PCR fragments (0.33-0.21 μ M) and *in vitro*-assembled Cas9-guide RNA complexes were introduced by nucleofection into HEK293T cells without selection as in Leonetti et al. (2016). The efficiency of GFP integration was examined 3 days later by cytometer or fluorescence microscopy. These methods permit the scoring of > 5000 cells (cytometer) and > 1000 cells (fluorescence) per each nucleofection experiment, and we performed at least two independent experiment for each condition. We obtained an average of 14.9%, 17.5% and 14.0% GFP+ positive cells for the *Lamin A/C*, *RAB11A* and *SMC3* loci, respectively (FIGS. 2B and 10B). In each case, the cells expressed GFP in a pattern consistent for the targeted ORF (FIGS. 2D and 10C).

Reducing the molarity of the PCR fragments by 10-fold reduced efficiency by $\sim 1/2$ (Compare FIGS. 2B and 2C). Increasing the length of the homology arms to 500 bp did not increase editing efficiency, even when controlling for the reduced molarity of the longer PCR fragments (FIG. 2C). Reducing the length of the homology arms to ~ 15 bp, however, decreased efficiency (FIG. 2B). PCR fragments with no homology sequence or homology arms for a locus not targeted by Cas9 yielded GFP+ positives in the range of the background levels obtained with cells that did not receive any repair template (FIGS. 2, 10, 11 and Table 1). Plasmid donors with ~ 500 bp homology arms also performed poorly (FIG. 2C) as reported previously (He et al., 44(9) NUCL. ACIDS RES. E85 (2016)). We conclude that PCR fragments function as efficient donors in HEK293T cells, performing better than plasmids with much longer homology arms. Because ~ 35 bp homology arms are convenient to introduce by PCR amplification, we used that length for subsequent experiments. 30-40 nt homology arms have also been reported to be optimal for ssODNs (Liang et al., 241 J. BIOTECHNOL. 136-46 (2017)).

Editing efficiency is sensitive to insert size. To test the effect of insert size on editing efficiency, we added varied sizes of DNA sequence to the GFP insert. For ease of synthesis and to maintain equimolar amounts of donor DNAs, we introduced donor fragments at the same low molarity (0.12 μ M). We found that inserts beyond 1kb performed very poorly, yielding less than 0.5% edits (FIG. 3A). By varying the size of the homology arms, we found that the size of the insert, and not the overall size of the donor DNA, determines editing efficiency. An 1188 bp donor (714 bp insert with two 237 bp HS) performed as well as a 780 bp donor with the same size insert and 33 bp HS (8.5% versus 9.8% edits, FIG. 3A). The 1188 bp donor, however, performed much better than an 1188 bp donor with a longer insert (1122 bp) and 33 bp HS (8.5% versus 0.3% edits, FIG. 3A).

To test whether decreasing insert size below the size of GFP would increase editing efficiency, we took advantage of the split-GFP system (Kamiyama et al. (2016; Leonetti et al. (2016)). In this system, the 11th beta-strand of GFP (57 bp, GFP11) is knocked-in in cells expressing a complementary GFP fragment (GFP1-10). We generated PCR products containing the GFP11 insert and ~35 bp HS and introduced these at 0.33 μ M. We obtain 45.4% edits at the *Lamin A/C* locus (FIG. 3B) and 32.8% at the *RAB11A* locus (FIG. 3C). A donor with no homology arm yielded only 1.3% edits (FIG. 11B). Again, we found that increasing insert size reduced efficiency, down to 17.9% for a 993 bp insert (FIG. 3B). We conclude that dsDNAs engage in an efficient repair process that requires only 35 bp homology arms, but favors relatively short inserts (<1kb at the molarities tested here).

Accuracy of repair is asymmetric. To investigate the accuracy of repair with PCR fragments, we isolated GFP+ and GFP- cells by fluorescence-activated cell sorting from a single editing experiment targeting the *Lamin A/C* locus with a GFP-containing PCR fragment under optimal conditions (FIG. 2B, 33/33 HS, 0.33 μ M molarity). Each cell was grown out as a clone and the *Lamin A/C* locus was amplified using two primers flanking the insertion site. As expected, all 48 GFP+ clones contained at least one *Lamin A/C* allele with a full-size insert (4 were homozygous with two edited alleles). We sequenced the GFP insert in 23 of the 48 GFP+ clones and identified 20 precise insertions and 3 imprecise insertions containing small in-frame indels at the left or right junction (FIGS. 12 and 13). We also sequenced the wild-type-sized allele in 11 of the 44 heterozygous GFP+ clones, and identified 2 with wild-type sequence, 6 with indels at the DSB, and 3 with small inserts (<100 bp) corresponding to either the N-terminus or C-terminus of GFP (FIG. 13). We also screened 37 GFP- clones by PCR and, surprisingly, identified 10 that contained inserts at the

Lamin A/C locus. We sequenced 7 of the 10 inserts and identified 3 with a full-size GFP insert with out-of-frame indels at one junction and 4 with smaller GFP inserts (FIG. 13).

In total, we sequenced 13 imprecise GFP edits and found only one internal deletion and one insertion in the wrong orientation (FIG. 13). All other imprecise edits were full-size or truncated GFP fragments inserted in the correct orientation. All had one precise junction on the non-truncated terminus of GFP. The other junction was imprecise and contained indels (FIG. 13). These observations are consistent with an asymmetric repair process that uses mechanisms with different homology requirements to initiate and resolve repair.

Repair is a polarity-sensitive process. In the SDSA model, initiation and resolution of repair proceeds via distinct steps. First, the DSB is resected to yield 3' overhangs on both sides of the DSB (FIG. 4A). The 3' overhangs pair with the donor and are extended by DNA synthesis copying donor sequences (FIG. 4A). Bridging of the DSB is completed when the newly synthesized strands withdraw from the donor and anneal back at the locus (FIG. 4A). To determine whether initiation and resolution might have different homology requirements, we tested the editing efficiency of single-stranded donors (ssODNs) bearing only one HS. We designed ssODNs with a GFP11 insert and only one HS at either the 3' or 5' end of the ssODN (5' or 3' HS). The HS targeted sequences on the left or right side of the Cas9-induced DSB in *Lamin A/C* and *RAB11A* (FIG. 4B). At both loci, we found that editing efficiency was highest with ssODNs that had a 3' HS that could anneal to a complementary 3' end at the DSB (FIG. 4C). ssODNs of the opposite polarity yielded only background-level edits. These observations are consistent with a replicative repair process that requires pairing between a 3' HS on the donor and sequences on at least one side of the DSB. Apparently, a different, less stringent mechanism can be used to bridge the donor to the other side. One possibility is that NHEJ was used to repair the gap on the side with no HS. Coupling of homologous and non-homologous repair mechanisms has already been documented in mammalian cells (Richardson, C. & Jasin, M., 20(23) MOL. CELL. BIOL. 9068-75 (2000)).

Polarity of single-stranded donors affects incorporation of distal edits. We wondered whether the different requirements for homology on the 3' and 5' ends of single-stranded donors might also apply to donors that contain two HS at different distances from the DSB. Such HS are found in donors designed to insert an edit at a distance from the DSB. In these donors, one HS (proximal HS) matches sequences immediately next to the DSB and the other HS (recessed HS) matches sequences at a distance from the DSB on the distal side of the edit (FIG. 5A). We tested whether proximal and recessed HS function equivalently on the 5' and 3' ends of ssODNs using a series of 23 pairs of sense and antisense ssODNs with inserts

ranging from 0 to 41 nucleotides from the DSB at four loci (FIG. 5B and Table 3). (In all ssODNs, the sequence between the DSB and edit was partially recoded to promote edit incorporation as described in the next section). Strikingly, we observed an increasing bias for a particular polarity with increasing edit-to-DSB distance (FIG. 5B). The favored ssODN polarity changed whether the edit (and recessed homology arm) was positioned to the left or right of the DSB (sense polarity when the edit is on the left side of the DSB, and antisense when the edit is on the right side). ssODNs with inserts close to the DSB did not show much polarity bias (FIG. 5B). These findings demonstrate that repair favors ssODNs with 3' HS that directly abut the DSB (proximal HS) and suggest that initiation of repair synthesis is enhanced by donors that can pair with sequences directly flanking the DSB. These experiments also showed that, in contrast to ssODN polarity, the polarity of the guide RNA used to create the DSB had no discernable effect on editing efficiency (FIG. 5B). We conclude that, under the conditions used here, the requirements for replicative repair have a greater impact on editing efficiency than the strand-bias imposed by asymmetric Cas9 release of the DSB (Richardson et al. (2016)).

Recoding of sequences between the DSB and the edit increases recovery of distal edits. Editing efficiency has been observed to decrease with increasing distance between the edit and the DSB (Paquet et al., 533(7601) NATURE 125-29 (2016)). This observation is also consistent with replicative repair, which predicts that synthesis that generates sequence complementary to the other side of the DSB will promote annealing back to the locus, potentially even before the edit is copied (FIG. 6). To test this prediction directly, we designed an ssODN donor with two inserts: a proximal insert (restriction enzyme site) one base away from the DSB in the *PYMI* locus and a distal insert (3xFlag) 23 bases away from the DSB. Each insert was flanked by an HS targeting the *PYMI* locus (FIG. 6A). We generated 63 single cell clones and genotyped the *PYMI* locus by PCR (see Material and Methods). 46% of the clones contained only the proximal edit and 12.6% contained both the proximal and distal edits (FIG. 6B). The finding that ~80% of the edits contained only the proximal edit is consistent with annealing using sequence between the two edits. To test this hypothesis, we mutated 7 bases in the 23 bases region separating the proximal and distal edit. The mutations were designed to reduce homology with the locus while preserving coding potential (FIG. 6A). This partial recoding reduced the frequency of proximal edit-only clones to 10.3% and increased the frequency of proximal+distal edits to 25.8% (FIG. 6B). We conclude that sequences on the donor that span the DSB can prevent incorporation of distal edits. We note that, although recoding enhances the recovery of distal edits, recoding does

not eliminate the preference for proximal edits, which are still recovered at higher frequency than distal edits even when using recoded templates (FIG. 14).

To test whether internal homologies can also participate in the repair process when using double-stranded donors, we performed a similar experiment with a PCR fragment
5 designed to incorporate GFP11 at the DSB, and tagRFP 33 bases from the DSB in the *Lamin A/C* locus (FIG. 6C). We recovered 10.8% GFP-only edits and 8.6% GFP-RFP double positives (FIG. 6D). Partial recoding of the sequence between GFP11 and tagRFP (by introducing 10 silent mutations) reduced the percent of GFP-only edits to 4.4% and raised the percent of GFP-RFP double positives to 17.6% (FIG. 6D). We conclude that internal
10 homologies on double-stranded templates can also interact with the targeted locus. Since both polarities are present in double-stranded templates, internal sequences could participate in principle in both the initial invasion step and the annealing step back to the locus.

Repair is prone to template switching between donors. Another characteristic of SDSA first observed in yeast is the ability of the repair process to undergo sequential rounds
15 of invasion and synthesis (29, 30). “Template switching” can create edits that combine sequences from overlapping donors (14). To test whether template switching also occurs in human cells, we used two donors to correct a single DSB. The first donor was an ssODN with two HS and a GFP11-coding insert containing a STOP codon to prevent translation of the full-length fusion (FIG. 7A). The second donor was an ssODN with the same GFP11
20 insert but without the STOP codon and without any HS. Consistent with template switching, we obtained 3.2% GFP+ edits when using both donors, compared to 0.3% and 0.4% GFP+ edits when using only the first or second ssODN, respectively (FIG. 7B). We repeated this experiment with double-stranded donors and obtained similar results (FIG. 7C/D). We
conclude that template switching between donors can occur in human cells (FIG. 16).

25 To visualize template switching more directly, we combined wild-type donors with recoded donors where the GFP11 insert contained several silent mutations and used Illumina sequencing to sequence the insertional edits en masse (FIG. 7E). Using recoded donors with silent mutations every 12 bases in the GFP11 insert, we identified evidence of template switching in 1.4% of edits (“chimeric edits”). Interestingly, the same experiment performed
30 with donors that contained silent mutations every 6 or every 3 nucleotides resulted in only 0.5% and 0% chimeric edits, respectively (FIGS. 7F and 15, Table 15). The chimeric edits could not have resulted from sequential rounds of Cas9 cleavage and repair, since the edit destroyed the crRNA pairing sequence. The chimeric edits also could not have arisen during PCR amplification, since we observed no chimeric edits in a control experiment mixing two

different cell populations (FIG. 15). We conclude that template switching occurs between donors in human cells and is sensitive to the degree of homology between donors (FIG. 16), as reported previously in yeast (Anand et al., 28(21) GENES DEV. 2394-2406 (2014) and Tsaponina, O. & Haber, J.E., 55(4) MOL. CELL. 615-25 (2014)).

Discussion

In this report, we demonstrate that PCR fragments are efficient donors for genome editing in mouse embryos and human cells. PCR fragments with short homology arms (HS ~35 bp) can be used to integrate edits up to 1kb, long enough to encode fluorescent reporters such as GFP. Experiments using single and double-stranded DNAs suggest that linear donors participate in a replicative repair mechanism that broadly conforms to the SDSA model for gene conversion. Our findings suggest simple guidelines to streamline donor design and maximize editing efficiency (FIG. 8).

Linear DNAs repair Cas9-induced DSBs by templating repair synthesis. In principle, linear donors could repair Cas9-induced breaks by integrating directly at the DSB. For example, microhomology-mediated end-joining (MMEJ) could cause donor ends to become ligated to each side of the DSB (Yao et al., 20 EBIO MEDICINE 19-26 (2017)). Alternatively, HS on the donor could form holiday junctions with sequences on each side of the DSB. Cross-over resolution of the two holiday junctions could cause donor sequences to become integrated at the DSB. This type of HDR has been proposed to underlie genome editing with plasmid and viral donors (Kan et al., 27(7) GENOME RES. 1099-1111 (2017)). In these models, repair is symmetric: the same mechanism (MMEJ or recombination) is used to ligate donor sequences to each side of the break. In contrast, our observations suggest that repair with linear donors proceeds by an asymmetric, likely replicative, process. First, ssODNs with only one HS show strong polarity specificity (FIG. 4C), consistent with a specific requirement for pairing with 3' ends at the DSB (FIG. 4A). Second, recessed HS (HS at a distance from the DSB) are rarely used to initiate repair synthesis, but can be used to resolve a repair event (FIGS. 5 and 6). Third, internal homologies on the donor can bypass integration of distal edits (FIG. 6). Fourth, most imprecise edits have asymmetric junctional signatures (FIG. 13). These observations suggest that the repair process is polar, like DNA synthesis, and has different requirements to initiate and resolve repair. These findings are consistent with the SDSA model for gene conversion (Paques, F. & Haber, J.E., 63(2) MICROBIOL. MOL. BIOL. REV. 349-404 (1999)) (FIG. 4A). SDSA initiates with DNA synthesis templated by the donor to extend 3' ends at the DSB, and resolves by annealing of the newly replicated strand(s) back to the locus. Our observations suggest that initiation of

DNA synthesis is the most homology-stringent step, requiring a ~35 base HS on the donor complementary to sequences directly adjacent to one side of the DSB. Either side of the DSB can initiate repair and, contrary to an earlier report (Liang et al. (2016)), we did not observe a preference consistent with biased strand-release by Cas9. The observations that HS longer than 35 bases do not perform significantly better, and that distal HS perform more poorly, also suggest that resection exposes only short regions of ssDNA on either side of the DSB. In contrast to the initiation step, the resolution step has more relaxed homology requirements. Recessed homology arms can be used for that step, and in fact repair can proceed with no HS on the “annealing side” (FIG. 4C). In that case, NHEJ (or MHEJ) may be used to fuse the newly replicated strand to the other side of the DSB. One possibility is that NHEJ or MHEJ competes with annealing during resolution, especially in the case of long edits where synthesis has a higher chance of stalling before reaching the distal HS or before synthesis of a complementary strand primed from the other side of the DSB (FIG. 4A). Consistent with this view, we recovered several partial GFP insertions that were integrated in the correct orientation but contained one imprecise junction on the truncated side of GFP, consistent with premature withdrawal from the donor. We cannot exclude the possibility, however, that in these partial edits, the non-homologous joint was made first using a broken donor.

If partial edits are due to premature withdrawal of the newly replicated strand from the donor, partial edits should be less frequent when using donors with shorter inserts.

Consistent with this prediction, we found that editing efficiency is inversely proportional to insert size. At the *Lamin A/C* locus, we obtained 45.4% edits for a 57 bp insert, 23.5% edits for 714 bp insert (GFP) and 17.9% edits for a 993 bp insert. The size of the insert, and not the overall size of the donor, correlated with efficiency, arguing against the possibility that breakage of longer donors contributes to reduced efficiency (FIG. 3). We suggest that the low processivity of repair polymerases (Parsons et al., 18(8) ANTIOXID. REDOX SIGNAL 851-73 (2013)) increases the chances of aberrant dissociation/annealing events on long inserts.

We also obtained evidence for dissociation and invasion events between donors.

Such “template switching” was also observed in yeast and *C. elegans* and can cause sequences from overlapping donors to become incorporated in the same edit (Anand et al. (2014); Tsaponina et al. (2014)) (Paix et al., 44(15) NUCLEIC ACIDS RES. e128 (2016)). We found that template switching is sensitive to the degree of homology between donors and is reduced significantly by mutations every 3 or 6 bases, as was also found in yeast (Anand et al. (2014); Tsaponina et al. (2014)). Similarly, recoding of sequences between the DSB and the edit promotes the incorporation of distal edits, presumably by increasing the rejection rate

of heteroduplexes formed during annealing between the newly replicated strand and sequences flanking the DSB (Sugawara et al., 101(25) PROC. NATL. ACAD. SCI. USA 9315-20 (2004)). Template switching may also explain why editing efficiency is sensitive to donor molarity, since high donor molarity is predicted to lower the frequency of aberrant dissociation/re-annealing events during synthesis. It will be interesting to determine which repair polymerases are responsible for synthesis templated by linear donors and whether their processivity characteristics account for our observations of template switching. In this regard, it is interesting to note that we identified a higher frequency of full-length edits (and lower frequency of partial edits) in mice compared to HEK293T cells. This difference could reflect differences in the properties of the enzymes that mediate SDSA in the two systems. Alternatively, the higher precision in mice could be due to a more efficient method for delivering donors at high molarity (pronuclear injection in mouse zygotes versus nucleofection in HEK293T cells).

SDSA as a repair mechanism for Cas9-induced DSBs: implications for genome editing. The demonstration that ssODNs and PCR fragments engage in a SDSA-like mechanism to repair Cas9-induced DSBs has two important implications for genome editing. First, the SDSA model makes simple predictions for optimal donor design (FIG. 8). These predictions improve editing efficiencies for edits at distance from the DSB, and eliminate the effort and expense used in creating donor DNAs with unnecessarily long homology arms. Linear donors with short homology arms can be chemically synthesized as single-stranded or double-stranded DNA or PCR amplified, avoiding the need for cloning. In this manner, tagging of genes with GFP can be achieved readily, without resorting to split-GFP approaches that also require expression of a complementary GFP1-10 fragment (Leonetti et al., 113(25) PROC. NATL. ACAD. SCI. USA E3501-08 (2016)). Second, because SDSA is thought to be a widespread mechanism for DSB repair among eukaryotes (Iyama, T. & Wilson, D.M., 12(8) DNA REPAIR (AMST) 620-36 (2013)), it is likely that the approaches outlined here will be applicable to other cell types and organisms. We documented previously that PCR fragments with short HS perform well in *C. elegans* (Paix et al. (2016)), and we demonstrate here the same for HEK293T cells and mouse embryos. It will be interesting to investigate whether linear donors with short HS can also be used for genome editing in pluripotent cells and post-mitotic cells.

TABLE 1: Detailed experimental conditions and results									
Cell type	CRISPR/Cas9 delivery	CRISPR/Cas9 name	CRISPR/Cas9 polarity (relative to gene coding sequence)	Repair template type/name	Repair template left/right homology arms (per deletion/insertion)	Distance between DSB and edit (bp, relative to DSB)	Repair template concentration (µM for RNP/Nucleofection)	Insert size (predicted/actual)	Efficiency (%) (for cytoskeletal analysis; non-nucleofected cells: <0.5%) (SD: Standard Deviation)
FIGURE 1									
Profounder injection of mouse embryos	RNP/Injection	crAlcy3	AS	Plasmid pBS-ΔC3term/Genomic	483/421	between -4/+2	3.5 nM of plasmid repair template and 30 ng/µl Cas9 protein and 0.6 µM of crAlcy3/rasRNA in injection buffer	739	PCR on mouse tail DNA (predicted size 310): 20.0 (±2.5), 0.6 (0/20) of the negative for Adgyl. crCherry insertion are positive by PCR using crCherry internal primers
Profounder injection of mouse embryos	RNP/Injection	crAlcy3	AS	PCR 1536/1597	346/36	between -4/+2	60 nM 30 ng/µl PCR repair template and 30 ng/µl Cas9 protein and 0.6 µM of crAlcy3/rasRNA in injection buffer	739	PCR on mouse tail DNA (predicted size 310): 31.0 (±7.9), including 3 homozygotes; 10 out of 10 fetuses from crCherry have perfect insertion; 1.06% (1/94) of the negative for Adgyl. crCherry insertion are positive by PCR using crCherry internal primers
FIGURES 2, 10, 11									
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 1630/1632	0/0	0	0.33	714	Cytometer: 1.3 (average of independent experiments: 1.3 ± 1.7) (n=2, SD=0.3)
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 1630/1632	167/16	0	0.33	714	Cytometer: 7.2 (average of independent experiments: 4.9 ± 9.4) (n=2, SD=3.2)
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 1630/1632 or PCR 1743/1744	33/33	0	0.33	714	Cytometer: 14.0 (average of independent experiments: 10.2 ± 14.9 ± 13.2 ± 15.8 ± 36.3 ± 25.5 ± 14.9 ± 12.7 ± 11.8 ± 18.1) (n=10, SD=2.5). Sequencing results can be found in Figure 4 and 5.
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 1743/1744	33/33	0	0.33	714	Cytometer: 7.4 (average of independent experiments: 7.5 ± 8) (n=2, SD=0.4)
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 1741/1742	518/238	0	0.33	714	Cytometer: 8.0 (average of independent experiments: 5.4 ± 8.4) (n=2, SD=0.7)
HEK293T	RNP/Nucleofection	crRNA 1629	S	Plasmid 1716	518/238	0	0.33	714	Cytometer: 1.0 (average of independent experiments: 1.1 ± 1.2 ± 0.8) (n=3, SD=0.2)
HEK293T	RNP/Nucleofection	crRNA 1648	AS	PCR 1630/1632	0/0	-11 (recoded)	0.21	714	Cytometer: 1.2 (average of independent experiments: 1.3 ± 1.1) (n=2, SD=0.1)
HEK293T	RNP/Nucleofection	crRNA 1648	AS	PCR 1630/1632	157/15	-11 (recoded)	0.21	714	Cytometer: 9.0 (average of independent experiments: 11.0 ± 8.7) (n=2, SD=0.8)
HEK293T	RNP/Nucleofection	crRNA 1648	AS	PCR 1640/1641 or PCR 1652/1653	33/33	-11 (recoded)	0.21	714	Cytometer: 17.5 (average of independent experiments: 15.5 ± 22.5 ± 12.8) (n=3, SD=5.0)
HEK293T	RNP/Nucleofection	crRNA 1648	AS	PCR 1640/1641	33/33	-11 (recoded)	0.02	714	Cytometer: 7.6 (average of independent experiments: 7.6 ± 7.5) (n=2, SD=0.1)
HEK293T	RNP/Nucleofection	crRNA 1648	AS	PCR 1646/1647	465/132	-11 (recoded)	0.02	714	Cytometer: 5.3 (average of independent experiments: 6.4 ± 4.1) (n=2, SD=1.6)
HEK293T	RNP/Nucleofection	crRNA 1648	AS	Plasmid 1791	465/132	-11 (recoded)	0.02	714	Cytometer: 2.7 (average of independent experiments: 3.7 ± 1.7) (n=2, SD=1.4)
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 1652/1653	No homology arms. PCR containing 5' and 3' homology arms used with Lamin A/C crRNA	0	0.21	714	Microscope: no Lamin A/C or 646/114 eGFP signal
HEK293T	RNP/Nucleofection	crRNA 1553	AS	PCR 1630/1632	0/0	+5 (recoded)	0.33	714	Cytometer: 1.4 (average of independent experiments: 1.9 ± 0.9) (n=2, SD=0.7)
HEK293T	RNP/Nucleofection	crRNA 1553	AS	PCR 1630/1632	167/17	+5 (recoded)	0.33	714	Cytometer: 12.0 (average of independent experiments: 10.1 ± 13.8) (n=2, SD=2.7)
HEK293T	RNP/Nucleofection	crRNA 1553	AS	PCR 1554/1555	37/38	+5 (recoded)	0.33	714	Cytometer: 14.0 (average of independent experiments: 16.4 ± 15.5 ± 11.4 ± 9.0 ± 33.5 ± 18.4 ± 13.0 ± 10.8) (n=3, SD=3.9)
FIGURES 3, 11, 14									
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 2002/2006	33/33	0	0.12	2229	Microscope: <0.1 (average of independent experiments: 0.07 (SD=0.06)) but few positives can be found when the entire microscope slide was examined: <0.1 (1/1529) (n=2, SD=0.1)
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 2005/2015	33/33	0	0.12	1122	Microscope: 0.3 (average of independent experiments: 0.5 (1/1350) ± 0.1 (1/1878) (n=2, SD=0.3)
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 2040/1610	33/33	0	0.12	993	Microscope: 3.1 (average of independent experiments: 3.2 ± 3.0) (n=2, SD=0.1)
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 1638/1610	33/33	0	0.12	714	Microscope: 9.8 (average of independent experiments: 11.4 ± 8.1) (n=2, SD=1.3)
HEK293T	RNP/Nucleofection	crRNA 1629	S	PCR 2058/2059	237/237	0	0.12	714	Microscope: 8.5 (average of independent experiments: 10.0 ± 6.9) (n=2, SD=2.2)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	S	PCR 2040/1610	33/33	0	0.33	993	Microscope: 17.5 (average of independent experiments: 20.4 ± 15.3) (n=2, SD=3.6)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	S	PCR 1638/1610	33/33	0	0.33	714	Microscope: 2.5 (average of independent experiments: 3.2 ± 2.7 ± 2.2 ± 3.4) (n=5, SD=0.8)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	S	PCR 2051/2052	33/32	0	0.33	336	Microscope: 30.5 (average of independent experiments: 29.0 ± 32.0) (n=2, SD=2.1)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	S	PCR 2030/2034	33/32	0	0.33	57	Microscope: 45.4 (average of independent experiments: 46.9 ± 45.8) (n=2, SD=0.8)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	PCR 1640/1641 or PCR 1652/1653	33/33	-11 (recoded)	0.21	714	Cytometer: 12.8 (average of independent experiments: 14.1 ± 11.5) (n=2, SD=1.8)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	PCR 2028/2029	33/33	-11 (recoded)	0.21	57	Cytometer: 32.8 (average of independent experiments: 27.4 ± 25.2) (n=2, SD=6.5)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1777	S	PCR 2052/2054	33/34	-2	0.33	57	Cytometer: 50.0
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1777	S	PCR 2053/2054	33/34	-32 (recoded)	0.33	57	Cytometer: 20.7
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1777	S	PCR 2032/2034	No homology arms. PCR containing GFP1 with Lamin A/C homology arms used with RAB12A crRNA	0	0.33	57	Cytometer: 1.3

TABLE 2. Detailed experimental conditions and results											
Edit	Cell type	Cas9 delivery	crRNA/gRNA name	crRNA/gRNA polarity (relative to gene coding sequence)	Repair template type/name	Repair template left/right homology arms (in nucleotides/bases/pairs)	Repair template polarity (relative to gene coding sequence)	Distance between DSB and edit (bp, relative to DSB)	Repair template concentration (μM for RNP/Nucleofection)	Insert size (nucleotides/bases/pairs)	Efficiency (%) (for cytometer analysis; non-nucleofected cells: <0.5%) (SD: Standard deviation)
FIGURE 4											
G1551: 5'-ATGG-3' (2000x42)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	5	scODN 1798	33/0	5	0	5	126	Microscope 0.2
	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	5	scODN 1799	33/0	AS	0	5	126	Microscope 20.8
	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	5	scODN 1705	0/32	5	0	5	126	Microscope 9.5
	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	5	scODN 1706	0/32	AS	0	5	126	Microscope 0.4
G1551: 5'-ATGG-3' (2000x42)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	scODN 1816	33/0	5	+1	5	126	Cytometer 1.6
	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	scODN 1817	33/0	AS	+1	5	126	Cytometer 21.9
	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	scODN 1819	0/33	5	-2	5	126	Cytometer 15.4
	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	scODN 1820	0/33	AS	-2	5	126	Cytometer 1.1

FIGURE 4

TABLE 3: Detailed experimental conditions and results										
Cell type	CRISPR/Cas9 delivery	CRISPR/Cas9 name	CRISPR/Cas9 polarity (relative to gene coding sequence)	Repair template type/name	Repair template left/right homology arms (per cleavage/repair)	Repair template polarity (relative to gene coding sequence)	Distance between DSB and edit (bp, relative to DSB)	Repair template concentration (µM for RNP/Nucleofection)	Insert size (per cleavage/repair)	Efficiency (%) (for cytoskeletal analysis; non-nucleofected cells: <0.5%) (SD: Standard Deviation)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	5	siODN 1620	33/32	5	0	5	57	Microscope 11.6
	RNP/Nucleofection	crRNA 1629	5	siODN 1732	33/32	AS	0	5	57	Microscope 11.8
	RNP/Nucleofection	crRNA 1629	5	siODN 1678	38/35	5	-12 (recoded)	5	57	Microscope 17.9
	RNP/Nucleofection	crRNA 1629	5	siODN 1679	38/35	AS	-32 (recoded)	5	57	Microscope 11.2
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	5	siODN 1793	33/33	5	-12 (recoded)	5	57	Microscope 9.7
	RNP/Nucleofection	crRNA 1629	5	siODN 1794	33/33	AS	-42 (recoded)	5	57	Microscope 11.9
	RNP/Nucleofection	crRNA 1629	5	siODN 1795	33/33	5	-43 (recoded)	5	57	Microscope 6.3
	RNP/Nucleofection	crRNA 1629	5	siODN 1796	33/33	AS	-43 (recoded)	5	57	Microscope 12.4
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1729	AS	siODN 1736	38/33	5	-32 (recoded)	5	57	Microscope 1.1
	RNP/Nucleofection	crRNA 1729	AS	siODN 1737	38/33	AS	-32 (recoded)	5	57	Microscope 0.2
	RNP/Nucleofection	crRNA 1728	5	siODN 1736	38/34	5	-31 (recoded)	5	57	Microscope 12.8
	RNP/Nucleofection	crRNA 1728	5	siODN 1737	38/34	AS	-31 (recoded)	5	57	Microscope 1.3
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	siODN 1778	33/33	5	-11 (recoded)	5	57	Cytometer 14.2
	RNP/Nucleofection	crRNA 1648	AS	siODN 1779	33/33	AS	-11 (recoded)	5	57	Cytometer 9.7
	RNP/Nucleofection	crRNA 1648	AS	siODN 1864	33/33	5	-2	5	57	Cytometer 25.7
	RNP/Nucleofection	crRNA 1648	AS	siODN 1865	33/33	AS	-2	5	57	Cytometer 32.1
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	siODN 1851	33/34	5	-13 (recoded)	5	57	Cytometer 2.5
	RNP/Nucleofection	crRNA 1648	AS	siODN 1852	33/34	AS	-13 (recoded)	5	57	Cytometer 14.3
	RNP/Nucleofection	crRNA 1777	5	siODN 1782	33/34	5	-32 (recoded)	5	57	Cytometer 14.8 (experimental replicate 14.4)
	RNP/Nucleofection	crRNA 1777	5	siODN 1783	33/34	AS	-32 (recoded)	5	57	Cytometer 2.3 (experimental replicate 1.3)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1777	5	siODN 1833	33/34	5	-17 (recoded)	5	57	Cytometer 31.0
	RNP/Nucleofection	crRNA 1777	5	siODN 1834	33/34	AS	-17 (recoded)	5	57	Cytometer 7.3
	RNP/Nucleofection	crRNA 1777	5	siODN 1827	33/34	5	-2	5	57	Cytometer 36.6
	RNP/Nucleofection	crRNA 1777	5	siODN 1828	33/34	AS	-2	5	57	Cytometer 31.5
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1776	AS	siODN 1782	33/34	5	-32 (recoded)	5	57	Cytometer 10.5 (experimental replicate 10.1)
	RNP/Nucleofection	crRNA 1776	AS	siODN 1783	33/34	AS	-32 (recoded)	5	57	Cytometer 1.1 (experimental replicate 1.4)
	RNP/Nucleofection	crRNA 1776	AS	siODN 1833	33/34	5	-17 (recoded)	5	57	Cytometer 21.6
	RNP/Nucleofection	crRNA 1776	AS	siODN 1834	33/34	AS	-17 (recoded)	5	57	Cytometer 3.9
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1776	AS	siODN 1827	33/34	5	-2	5	57	Cytometer 22.6
	RNP/Nucleofection	crRNA 1776	AS	siODN 1828	33/34	AS	-2	5	57	Cytometer 24.6
	RNP/Nucleofection	crRNA 1910	AS	siODN 1911	35/32	5	-3	5	57	Cytometer 26.5
	RNP/Nucleofection	crRNA 1910	AS	siODN 1912	35/32	AS	-3	5	57	Cytometer 31.5
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1910	AS	siODN 1924	35/35	5	-42 (recoded)	5	57	Cytometer 1.0
	RNP/Nucleofection	crRNA 1910	AS	siODN 1925	35/35	AS	-42 (recoded)	5	57	Cytometer 5.5
	RNP/Nucleofection	crRNA 1909	5	siODN 1911	33/35	5	-42	5	57	Cytometer 15.5
	RNP/Nucleofection	crRNA 1909	5	siODN 1912	33/35	AS	-42	5	57	Cytometer 18.8
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1909	5	siODN 1922	35/35	5	-42 (recoded)	5	57	Cytometer 0.7
	RNP/Nucleofection	crRNA 1909	5	siODN 1923	35/35	AS	-42 (recoded)	5	57	Cytometer 3.4
	RNP/Nucleofection	crRNA 1748	AS	siODN 1751	34/38	5	-42 (recoded)	5	57	Microscope 1.8
	RNP/Nucleofection	crRNA 1748	AS	siODN 1752	34/38	AS	-42 (recoded)	5	57	Microscope 11.0
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1747	5	siODN 1753	34/38	5	-41 (recoded)	5	57	Microscope 0.4
	RNP/Nucleofection	crRNA 1747	5	siODN 1754	34/38	AS	-41 (recoded)	5	57	Microscope 10.3
	Plasmid (Casp3/7:GFP:GFP sorting)/Transfection	sgRNA1	5	siODN 1583	38/40	5	-41 and -42 (recoded)	3.2 µmol of siODN and 850 µg of Casp3 plasmid in a mix of 10 µl containing 3.0 µl of X-tremeGEM	6 and 66	RFP single cell colonies (predicted size 4.8 and 4.8 for RFP and RFP:GFP, 0.025% for RFP:GFP alone)
	Plasmid (Casp3/7:GFP:GFP sorting)/Transfection	sgRNA1	5	siODN 1584	38/40	AS	-41 and -42 (recoded)	3.2 µmol of siODN and 850 µg of Casp3 plasmid in a mix of 10 µl containing 3.0 µl of X-tremeGEM	6 and 66	RFP single cell colonies (predicted size 4.8 and 4.8 for RFP and RFP:GFP, 0.025% for RFP:GFP alone)
HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1729	AS	siODN 1734	33/33	5	-41	5	57	Microscope 2.0
	RNP/Nucleofection	crRNA 1728	5	siODN 1734	32/33	5	-42	5	57	Microscope 22.8

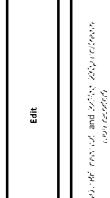
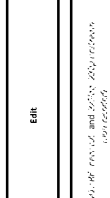
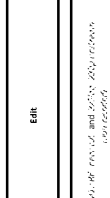
TABLE 2 Detailed experimental conditions and results												
	Cell	Cell type	Cas9 delivery	crRNA/gRNA name	crRNA/gRNA polarity (relative to gene coding sequence)	Repair template type/name	Repair template left/right homology arms (nucleotides/homology)	Repair template polarity (relative to gene coding sequence)	Distance between DSB and edit (bp, relative to DSB)	Repair template concentration (µM for RNP/Nucleofection)	Insert size (predicted/homology)	Efficiency (%) (for cyclometer analysis; non-nucleofected cells: <0.5%) (SD: Standard Deviation)
FIGURES 5 & 6, 14												
	HEK293T	HEK293T	Plasmid (Cas9-T2A-GFP: GFP sorting)/Transfection	gPM1	S	sCDN 1582	38/37	S	+1 and -23	3.2 µmol of sCDN and 850 µg of Cas9 plasmid in a mix of 10.8 µl containing 3.6 µl of X-tremeGENE 9	6 and 66	PCR single cell colonies (predicted size shift and Restriction Enzyme digest): 2.6 (N/3) for RE and 348 bp insertion, 0.063 for 348 bp insertion alone, 46.0 (2/63) for RE insertion alone. Total cells: 36.7 (3/76)
	HEK293T	HEK293T	Plasmid (Cas9-T2A-GFP: GFP sorting)/Transfection	gPM1	S	sCDN 1580	38/37	S	+1 and -23 (recoded)	3.2 µmol of sCDN and 850 µg of Cas9 plasmid in a mix of 10.8 µl containing 3.6 µl of X-tremeGENE 9	6 and 66	PCR single cell colonies (predicted size shift and Restriction Enzyme digest): 25.8 (15/58), including 3 homoygous for RE and 348 bp insertion, 1.7 (2/58) for 348 bp insertion alone, 0.3 (1/58) for RE insertion alone. Total cells: 37.2 (2/58)
	HEK293T	HEK293T	Plasmid (Cas9-T2A-GFP: GFP sorting)/Transfection	gPM1	S	sCDN 1584	38/37	AS	+1 and -23 (recoded)	3.2 µmol of sCDN and 850 µg of Cas9 plasmid in a mix of 10.8 µl containing 3.6 µl of X-tremeGENE 9	6 and 66	PCR single cell colonies (predicted size shift and Restriction Enzyme digest): 3.2 (2/61) for RE and 348 bp insertion, 0.051 for 348 bp insertion alone, 0.051 for RE insertion alone. Total cells: 3.2 (2/61)
	HEK293T	HEK293T	Plasmid (Cas9-T2A-GFP: GFP sorting)/Transfection	gPM1	S	sCDN 1518	46/43	S	+1		3.2 µmol of sCDN and 850 µg of Cas9 plasmid in a mix of 10.8 µl containing 3.6 µl of X-tremeGENE 9	66
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	S	PCR 1388/1389 (on plasmid 1382)	33/33	dsDNA	0 and +33	0.23	57 and 708	Microscope 3.6 (180/1842) for GFP1-1 and tagRFP insertion, 0.713 (382) for tagRFP insertion alone, 0.3 (1/39) for GFP1-1 insertion alone. Total cells: 20.1 (372/1842)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	S	PCR 1388/1389 (on plasmid 1383)	33/33	dsDNA	0 and +33 (recoded)	0.23	57 and 708	Microscope 27.6 (288/1520) for GFP1-1 and tagRFP insertion, 0.117 (620) for tagRFP insertion alone, 4.4 (72/1629) for GFP1-1 insertion alone. Total cells: 22.3 (186/1629)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	sCDN 1597 and sCDN 1379 (unrelated)	33/33 and non-applicable	S and non-applicable	0	5 and 5	133	Cyclometer: 0.3 (average of independent experiments: 0.3 ± 0.3) (n=2, SD=0)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	sCDN 1379 and sCDN 1379 (unrelated)	no homology arm but 65/63 homologous sequences flanking the STOP (frameshift) in GFP1-1 and non-applicable	S and non-applicable	0	5 and 5	129	Cyclometer: 0.4 (average of independent experiments: 0.6 ± 0.2) (n=2, SD=0.3)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	sCDN 1597 and sCDN 1355	33/33 and no homology arm but 65/63 homologous sequences flanking the STOP (frameshift)	S and S	0	5 and 5	133 and 129	Cyclometer: 3.2 (average of independent experiments: 4.0 ± 2.2) (n=2, SD=1.1) (GFP positive cells confirmed by microscopy)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	sCDN 1354 and sCDN 1379 (unrelated)	33/33 and non-applicable	S and non-applicable	0	5 and 5	132	Cyclometer: 17.2 (average of independent experiments: 20.6 ± 14.7) (n=2, SD=3.5)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	sCDN 1356 and sCDN 1379 (unrelated)	33/33 and non-applicable	S and non-applicable	0	5 and 5	132	Cyclometer: 0.6
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	sCDN 1356 and sCDN 1355	33/33 and no homology arm but 65/63 homologous sequences flanking the STOP	S and S	0	5 and 5	132 and 129	Cyclometer: 4.1
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	PCR 2083/2084 (on sCDN 1357) + PCR 2086/2091 (on sCDN 1379, unrelated)	33/33 and non-applicable	dsDNA and dsDNA	0	0.33 and 0.33	133	Cyclometer: 0.5 (average of independent experiments: 0.4 ± 0.2) (n=2, SD=0.1)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	PCR 2086/2087 (on sCDN 1355) + PCR 2090/2091 (on sCDN 1379, unrelated)	no homology arm but 65/63 homologous sequences flanking the STOP (frameshift) in GFP1-1 and non-applicable	dsDNA and dsDNA	0	0.33 and 0.33	129	Cyclometer: 0.6 (average of independent experiments: 0.5 ± 0.7) (n=2, SD=0.1)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	PCR 2083/2084 (on sCDN 1357) + PCR 2086/2087 (on sCDN 1355)	33/33 and no homology arm but 65/63 homologous sequences flanking the STOP (frameshift)	dsDNA and dsDNA	0	0.33 and 0.33	133 and 129	Cyclometer: 2.3 (average of independent experiments: 2.2 ± 2.3) (n=2, SD=0.1) (GFP positive cells confirmed by microscopy)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1648	AS	PCR 2083/2084 (on sCDN 1354) + PCR 2090/2091 (on sCDN 1379, unrelated)	33/33 and non-applicable	dsDNA and dsDNA	0	0.33 and 0.33	132	Cyclometer: 21.7 (average of independent experiments: 20.7 ± 23.6) (n=2, SD=2.8)
	HEK293T (GFP1-10)	HEK293T (GFP1-10)	RNP/Nucleofection	crRNA 1629	S	sCDN 1799	33/32	S	0	5	132	DNA was extracted from cells edited with sCDN 1799 or sCDN 1835. Next the DNA from these experiments was mixed and the insertion was PCR amplified and sequenced by Illumina technology (Barcode 10). This control was performed to ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did not occur during PCR amplification (PCR control ensure that template switching did

TABLE 5: Repair templates used in this study

TABLE 5: Repair templates used in this study						
Gene targeted	crRNA sgRNA used	Repair template coding for	Homology arms (nucleotides/basipairs)	Distance from DSB	Polarity (5-Sense, AS-AntiSense)	Sequence
non-applicable	non-applicable	uncloned sODN (142nt)	non-applicable	non-applicable	non-applicable	genotype
non-applicable	non-applicable	uncloned PCR donor (142nt)	non-applicable	non-applicable	non-applicable	as sODN 1379 but dDNA
60000-AC	1629	5'-TTC-3'	33/32	0	sODN 1620	\$
15000-AC	1629	5'-GTA-3'	33/32	0	sODN 1732	AS
15000-AC	1629	5'-GTA-3'	33/32	0	PCR 2092/2093 (on sODN 1620/1732)	dDNA
50000-AC	1629	5'-GTA-3'	0/0	0	PCR 1630/1632 (on plasmid 1698)	dDNA
40000-AC	1629	5'-TTC-3'	16/16	0	PCR 1685/1686 (on plasmid 1698) or PCR 1858/1859 (on plasmid 1716)	dDNA
15000-AC	1629	5'-GTA-3'	33/33	0	PCR 1618/1619 (on plasmid 1698) or PCR 1743/1744 (on plasmid 1716)	dDNA
15000-AC	1629	5'-GTA-3'	237/237	0	PCR 2058/2059 (on plasmid 1716)	dDNA
15000-AC	1629	5'-GTA-3'	518/518	0	PCR 1741/1742 (on plasmid 1716)	dDNA
15000-AC	1629	5'-GTA-3'	518/518	0	Plasmid 1716	dDNA circular
15000-AC	1629	5'-TTC-3'	33/32	0	PCR 2051/2052 (on plasmid 2050)	dDNA
60000-AC	1629	5'-GTA-3'	33/33	0	PCR 2049/1619 (on plasmid 2042)	dDNA
15000-AC	1629	5'-GTA-3'	33/33	0	PCR 2005/2015 (on plasmid 1894)	dDNA
15000-AC	1629	5'-GTA-3'	33/33	0	PCR 2005/2006 (on plasmid 1894)	dDNA
60000-AC	1629	5'-GTA-3'	33/0	0	sODN 1788	\$
60000-AC	1629	5'-GTA-3'	33/0	0	sODN 1789	AS
15000-AC	1629	5'-GTA-3'	0/32	0	sODN 1705	\$
15000-AC	1629	5'-GTA-3'	0/32	0	sODN 1706	AS
15000-AC	1629	5'-TTC-3'	33/33	0 and +33	PCR 1948/1949 (on plasmid 1892)	dDNA
60000-AC	1629	5'-TTC-3'	33/33	0 and +33 (recoded)	PCR 1948/1949 (on plasmid 1893)	dDNA
15000-AC	1629	5'-GTA-3'	33/32	0	sODN 1799	\$
60000-AC	1629	5'-GTA-3'	33/32	0	sODN 1835	\$
15000-AC	1629	5'-GTA-3'	non-applicable	non-applicable	sODN 1813	\$
60000-AC	1629	5'-GTA-3'	non-applicable	non-applicable	sODN 1804	\$
15000-AC	1629	5'-GTA-3'	non-applicable	non-applicable	sODN 1805	\$
60000-AC	1629	5'-GTA-3'	non-applicable	non-applicable	sODN 1806	\$

TABLE 6: Repair templates used in this study

TABLE 6: Repair templates used in this study						
Gene targeted	crRNA sgRNA used	Repair template coding for	Homology arms (nucleotides/basopins)	Distance from DSB	Type and Name	Polarity (5-Sense, AS-Antisense)
Lamin A/C	1629	CGGCG	38/25	-12 (required)	ssODN 1678	S
	1629	CGGCG	38/25	-12 (required)	ssODN 1679	AS
	1629	CGGCG	33/33	+12 (required)	ssODN 1793	S
	1629	CGGCG	33/33	+12 (required)	ssODN 1794	AS
	1629	CGGCG	33/33	+33 (required)	ssODN 1795	S
Lamin A/C	1629	CGGCG	33/33	+33 (required)	ssODN 1796	AS
	1729	CGGCG	38/33	-32 (required)	ssODN 1736	S
	1729	CGGCG	38/33	-32 (required)	ssODN 1737	AS
	1729	CGGCG	33/33	+1	ssODN 1734	S
	1728	CGGCG	38/34	-31 (required)	ssODN 1736	S
Lamin A/C	1728	CGGCG	38/34	-31 (required)	ssODN 1737	AS
	1728	CGGCG	32/33	-2	ssODN 1734	S
	1648	CGGCG	33/33	-11 (required)	ssODN 1778	S
	1648	CGGCG	33/33	-11 (required)	ssODN 1779	AS
	1648	CGGCG	33/33	-11 (required)	PCR 2008/2009 (on ssODN 1778/1779)	dsDNA
R6521A	1648	CGGCG	0/0	-11 (required)	PCR 1630/1832 (on plasmid 1698)	dsDNA
	1648	CGGCG	15/15	-11 (required)	PCR 1838/1839 (on plasmid 1791)	dsDNA
	1648	CGGCG	33/33	-11 (required)	PCR 1652/1653 (on plasmid 1698) or PCR 1840/1841 (on plasmid 1791)	dsDNA
	1648	CGGCG	63/63	-11 (required)	PCR 1842/1843 (on plasmid 1791)	dsDNA
	1648	CGGCG	461/432	-11 (required)	PCR 1846/1847 (on plasmid 1791)	dsDNA
R6521A	1648	CGGCG	461/432	-11 (required)	Plasmid 1791	dsDNA circular
	1648	CGGCG	33/0	+1	ssODN 1816	S
	1648	CGGCG	33/0	+1	ssODN 1817	AS
	1648	CGGCG	0/33	-2	ssODN 1819	S
	1648	CGGCG	0/33	-2	ssODN 1820	AS
R6521A	1648	CGGCG	33/33	-2	ssODN 1864	S
	1648	CGGCG	33/33	-2	ssODN 1865	AS
	1648	CGGCG	33/34	+19 (required)	ssODN 1831	S
	1648	CGGCG	33/34	+19 (required)	ssODN 1832	AS
	1648	CGGCG	33/33	0	ssODN 1957	S
R6521A	1648	CGGCG	no homology arm but 65/63 homologous sequences flanking the STOP (frameshift) in GFP11	non applicable	ssODN 1955	S
	1648	CGGCG	33/33	0	ssODN 1956	S
	1648	CGGCG	33/33	0	ssODN 1954	S
	1648	CGGCG	33/33	0	ssODN 1954	S
	1648	CGGCG	33/33	0	ssODN 1954	S

TABLE 7: Repair templates used in this study

Gene targeted	crRNA/gRNA used	Repair template coding for	Homology arms (nucleotide/basipair)	Distance from DSB	Type and Name	Polarity (5'-Sense/AS+Antisense)	Sequence
<i>hprt1</i>	1648	5'-GGG-3' / 5'-GGG-3'	33/33	0	PCR 2083/2084 (on sCOON 1957)	dsDNA	as sCOON 1957 but dsDNA
<i>hprt1</i>	1648	5'-GGG-3' / 5'-GGG-3'	no homology arm but 65/63 homologous sequences flanking the STOP (frameshift) in GFP11	non-applicable	PCR 2086/2087 (on sCOON 1955)	dsDNA	as sCOON 1955 but dsDNA
<i>hprt1</i>	1648	5'-GGG-3' / 5'-GGG-3'	33/33	0	PCR 2083/2084 (on sCOON 1954)	dsDNA	as sCOON 1954 but dsDNA
<i>hprt1</i>	1776 and 1777	5'-GGG-3' / 5'-GGG-3'	33/34	-32 (recoded)	sCOON 1782	S	
<i>hprt1</i>	1776 and 1777	5'-GGG-3' / 5'-GGG-3'	33/34	-32 (recoded)	sCOON 1783	AS	as sCOON 1782 but AS
<i>hprt1</i>	1776 and 1777	5'-GGG-3' / 5'-GGG-3'	33/34	-17 (recoded)	sCOON 1833	S	
<i>hprt1</i>	1776 and 1777	5'-GGG-3' / 5'-GGG-3'	33/34	-17 (recoded)	sCOON 1834	AS	as sCOON 1833 but AS
<i>hprt1</i>	1776 and 1777	5'-GGG-3' / 5'-GGG-3'	33/34	-2	sCOON 1827	S	
<i>hprt1</i>	1776 and 1777	5'-GGG-3' / 5'-GGG-3'	33/34	-2	sCOON 1828	AS	as sCOON 1827 but AS
<i>hprt1</i>	1777	5'-GGG-3' / 5'-GGG-3'	33/34	-2	PCR 2055/2054 (on sCOON 1827/1828)	dsDNA	as sCOON 1827 but dsDNA
<i>hprt1</i>	1777	5'-GGG-3' / 5'-GGG-3'	33/34	-32 (recoded)	PCR 2053/2054 (on sCOON 1782/1783)	dsDNA	as sCOON 1782 but dsDNA
<i>hprt1</i>	1910	5'-GGG-3' / 5'-GGG-3'	35/32	-3	sCOON 1911	S	
<i>hprt1</i>	1910	5'-GGG-3' / 5'-GGG-3'	35/32	-3	sCOON 1912	AS	as sCOON 1911 but AS
<i>hprt1</i>	1909	5'-GGG-3' / 5'-GGG-3'	33/35	+2	sCOON 1911	S	
<i>hprt1</i>	1909	5'-GGG-3' / 5'-GGG-3'	33/35	+2	sCOON 1912	AS	see of 1910 -3bp insertion
<i>hprt1</i>	1910	5'-GGG-3' / 5'-GGG-3'	35/35	+21 (recoded)	sCOON 1924	S	see of 1910 -3bp insertion
<i>hprt1</i>	1910	5'-GGG-3' / 5'-GGG-3'	35/35	+21 (recoded)	sCOON 1925	AS	as sCOON 1924 but AS
<i>hprt1</i>	1909	5'-GGG-3' / 5'-GGG-3'	35/35	+26 (recoded)	sCOON 1922	S	
<i>hprt1</i>	1909	5'-GGG-3' / 5'-GGG-3'	35/35	+26 (recoded)	sCOON 1923	AS	as sCOON 1922 but AS
<i>hprt1</i>	1553	5'-GGG-3' / 5'-GGG-3'	0/0	+5 (recoded)	PCR 1630/1632 (on plasmid 1698)	dsDNA	see Lamin A/C with GFP PCR without homology arm
<i>hprt1</i>	1553	5'-GGG-3' / 5'-GGG-3'	16/17	+5 (recoded)	PCR 1604/1605 (on plasmid 1698)	dsDNA	
<i>hprt1</i>	1553	5'-GGG-3' / 5'-GGG-3'	37/38	+5 (recoded)	PCR 1554/1555 (on plasmid 1698)	dsDNA	
<i>hprt1</i>	1748	5'-GGG-3' / 5'-GGG-3'	34/38	+29 (recoded)	sCOON 1751	S	
<i>hprt1</i>	1748	5'-GGG-3' / 5'-GGG-3'	34/38	+29 (recoded)	sCOON 1752	AS	as sCOON 1751 but AS
<i>hprt1</i>	1747	5'-GGG-3' / 5'-GGG-3'	34/38	+41 (recoded)	sCOON 1753	S	
<i>hprt1</i>	1747	5'-GGG-3' / 5'-GGG-3'	34/38	+41 (recoded)	sCOON 1754	AS	as sCOON 1753 but AS
<i>hprt1</i>	sgPM1	5'-GGG-3' / 5'-GGG-3'	38/37	+1 and -23 (recoded)	sCOON 1582	S	
<i>hprt1</i>	sgPM1	5'-GGG-3' / 5'-GGG-3'	38/37	+1 and -23 (recoded)	sCOON 1580	S	
<i>hprt1</i>	sgPM1	5'-GGG-3' / 5'-GGG-3'	38/37	+1 and -23 (recoded)	sCOON 1581	AS	as sCOON 1580 but AS
<i>hprt1</i>	sgPM1	5'-GGG-3' / 5'-GGG-3'	46/43	+1	sCOON 1518	S	
<i>hprt1</i>	sgPM1	5'-GGG-3' / 5'-GGG-3'	38/40	+1 and -25 (recoded)	sCOON 1583	S	
<i>hprt1</i>	sgPM1	5'-GGG-3' / 5'-GGG-3'	38/40	+1 and -25 (recoded)	sCOON 1584	AS	as sCOON 1583 but AS
<i>hprt1</i>	sgPM1	5'-GGG-3' / 5'-GGG-3'	38/40	between -4/+2	PCR 1596/1597 (on plasmid pBS-ACE3/termidemic-mCherry)	dsDNA	
<i>hprt1</i>	sgPM1	5'-GGG-3' / 5'-GGG-3'	483/421	between -4/+2	Plasmid pBS-ACE3/termidemic-mCherry	dsDNA circular	see Table 9

TABLE 8: Plasmids used in this study

[illegible]

[illegible]

[illegible]

TABLE 11: Primers used in this study

[illegible]

TABLE 13: Primers used in this study

Primers name	F/R	Description	Sequence (5' to 3')
1555	R	amplification of <i>SMN2</i> with repair template (cr1553) with ~35bp homology arm	attatctctctcccaacacagctggatgtctctcaactctgtgttgggtgttggatctg
1483	F	<i>PYM1</i> genotyping	tgctcctgggtgtgttgggtgctcttg
1485	R	<i>PYM1</i> genotyping	ggatctgggtgctctctctctctctc
sgPYM1 cloning	F	<i>PYM1</i> sgRNA cloning	caccgcgtcaacacagcgacctga
sgPYM1 cloning	R	<i>PYM1</i> sgRNA cloning	aaactcaggtcgtgtgtgacgc
1596	F	amplification of <i>mouse Adcy3</i> mCherry repair template (crAdcy3) with ~35bp homology arm	ctctgtgtctctctgctctctc
1597	R	amplification of <i>mouse Adcy3</i> mCherry repair template (crAdcy3) with ~35bp homology arm	cctctctctctgtgtggcttc
1760 / HK15	F	<i>mouse Adcy3</i> genotyping and sequencing	ctcttggaggtggtggtcttc
1761 / HK16	R	<i>mouse Adcy3</i> genotyping and sequencing	ggtaattctctgggtgatttggctc
1772	F	<i>mouse Adcy3</i> genotyping and sequencing	tgtgtgaggtgaggtctctctctc
1764 / HK11	F	<i>mouse Adcy3</i> genotyping and sequencing	ggcattctgggtgctctctctctc
HK12	R	<i>mouse Adcy3</i> genotyping and sequencing	gactgttgatgaggtgctctctctctc
1765	R	<i>mouse Adcy3</i> genotyping and sequencing	gtctcagaagctctctctctctctc

TABLE 14: crRNA/sgRNA used in this study

Guide Name	Target	Type	Polarity	Sequence
1629	<i>Lamin A/C</i>	crRNA	S	ccatggagaccccgctcccag
1648	<i>RAB11A</i>	crRNA	AS	ggtagtcgtactcgtcgtcg
1728	<i>Lamin A/C</i>	crRNA	S	gcggcgccaccgccgacg
1729	<i>Lamin A/C</i>	crRNA	AS	agctggcctgcgccccgctg
1776	<i>RAB11A</i>	crRNA	AS	ccatggcctcacctttaag
1777	<i>RAB11A</i>	crRNA	S	gagtaacgactactctttaa
1909	<i>RAB11A</i>	crRNA	S	aaccactgaaaaacaagccaa
1910	<i>RAB11A</i>	crRNA	AS	ttctgacagcactgcacctt
1553	<i>SMC3</i>	crRNA	AS	attttccaatttaaccatgtg
1747	<i>SMC3</i>	crRNA	S	tgaatgatcacacagacagaga
1748	<i>SMC3</i>	crRNA	AS	atcatcttctacaaaagtctt
sgPYM1	<i>PYM1</i>	sgRNA	S	gcgtcaacacagcgacctga
crAdcy3	mouse <i>Adcy3</i>	crRNA	AS	gtggagccagaggtcgctca

TABLE 15: Classification of reads from the Illumina sequencing experiment

Sample	Total sequencing reads	Do not fully map to template	Below quality threshold	Unexpected mutation at diagnostic position	Used in downstream analysis	Reads with switching detected (% of the previous column)
No mutation	3,369,768	21.60%	42.30%	0.20%	36.00%	0.00%
PCR control	3,241,689	20.10%	56.40%	0.20%	23.50%	0.02%
1/3	3,411,796	20.80%	42.30%	0.40%	36.60%	0.02%
1/6	5,680,820	21.20%	42.40%	0.20%	36.20%	0.50%
1/12	6,414,459	21.00%	42.40%	0.10%	36.50%	1.40%

We claim:

1. A double-stranded, linear donor polynucleotide comprising a polynucleotide encoding a fluorescent protein flanked by a first homology arm and a second homology arm.
2. The polynucleotide of claim 1, wherein the homology arms are 15-60 bases in length.
3. The polynucleotide of claim 1, wherein the homology arms are 25-45 bases in length.
4. The polynucleotide of claim 1, wherein the homology arms are 30-40 bases in length.
5. A double-stranded, linear donor polynucleotide comprising a polynucleotide encoding a fluorescent protein flanked by a first homology arm and a second homology arm, wherein the first and second homology arms are between 30-35 bases in length.
6. A double-stranded, linear donor polynucleotide comprising a template polynucleotide encoding an edit flanked by an intervening sequence and two homology arms.
7. The polynucleotide of claim 6, wherein the homology arms are 15-60 bases in length.
8. The polynucleotide of claim 6, wherein the homology arms are 25-45 bases in length.
9. The polynucleotide of claim 6, wherein the homology arms are 30-40 bases in length.
10. The polynucleotide of claim 6, wherein the template polynucleotide is up to 1 kb in length.
11. The polynucleotide of claim 6, wherein the template polynucleotide comprises a sequence designed to change at least one nucleotide base within 30 bases of a double-stranded break (DSB) of a target nucleic acid.
12. The polynucleotide of claim 11, wherein the template polynucleotide further comprises a restriction enzyme site.

13. A double-stranded, linear donor polynucleotide comprising a template polynucleotide flanked by a first homology arm and a second homology arm, wherein the homology arms are between 30-35 bases in length.
14. The polynucleotide of claim 14, wherein the template polynucleotide is up to 1 kb in length.
15. The polynucleotide of claim 14, wherein the template polynucleotide comprises a sequence designed to change at least one nucleotide base within 30 bases of a DSB of a target nucleic acid.
16. The polynucleotide of claim 15, wherein the template polynucleotide further comprises a restriction enzyme site.
17. A method comprising the step of performing a clustered regularly interspaced short palindromic repeats (CRISPR)-based technique using a double-stranded, linear donor polynucleotide of any of claims 1-16 as the donor polynucleotide.
18. A method comprising injecting into a target cell a composition comprising (a) an RNA-guided DNA endonuclease; (b) a guide RNA; and (c) a double-stranded, linear donor polynucleotide of any of claims 1-16.

1/37

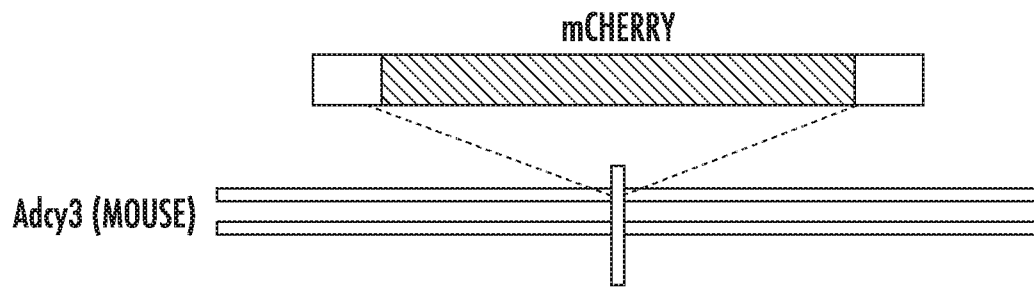


FIG. 1A

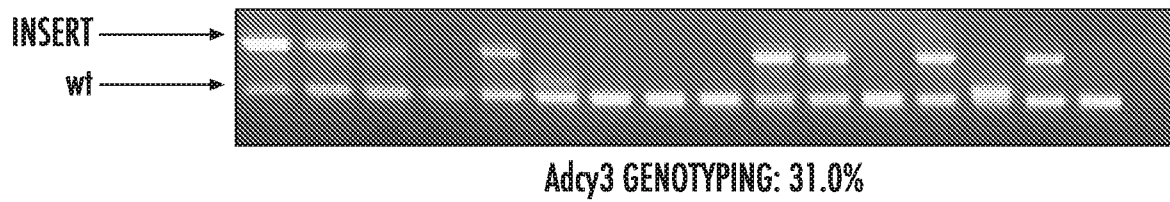


FIG. 1B

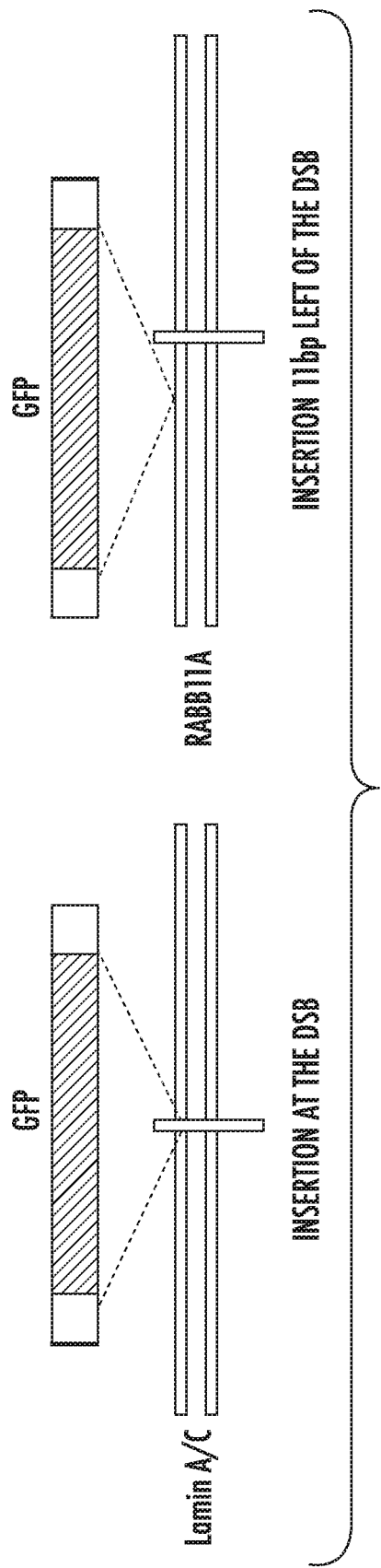


FIG. 2A

2/37

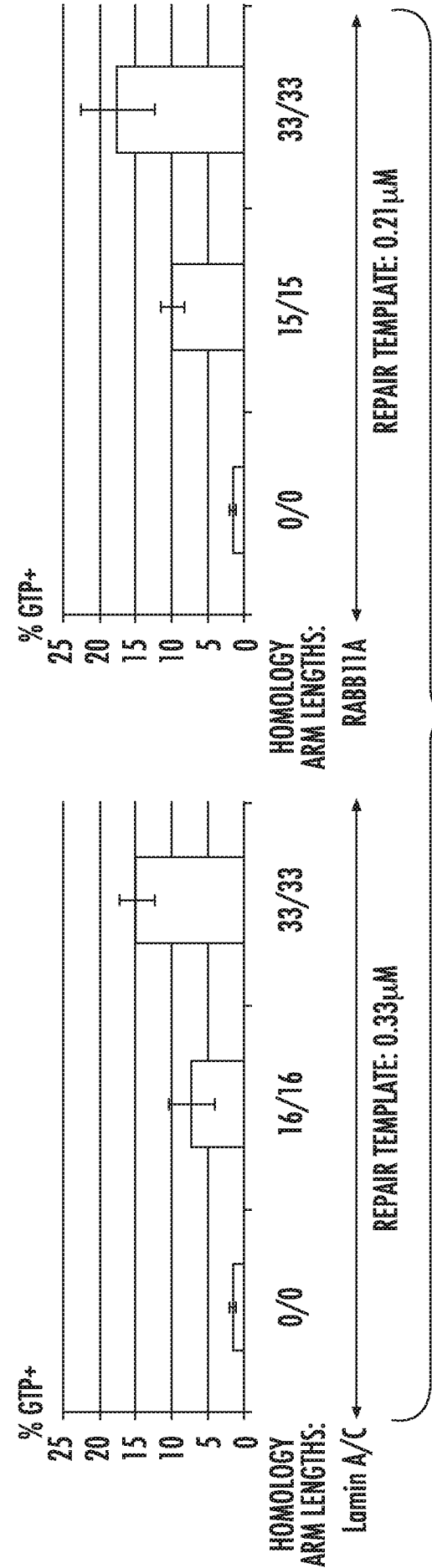


FIG. 2B

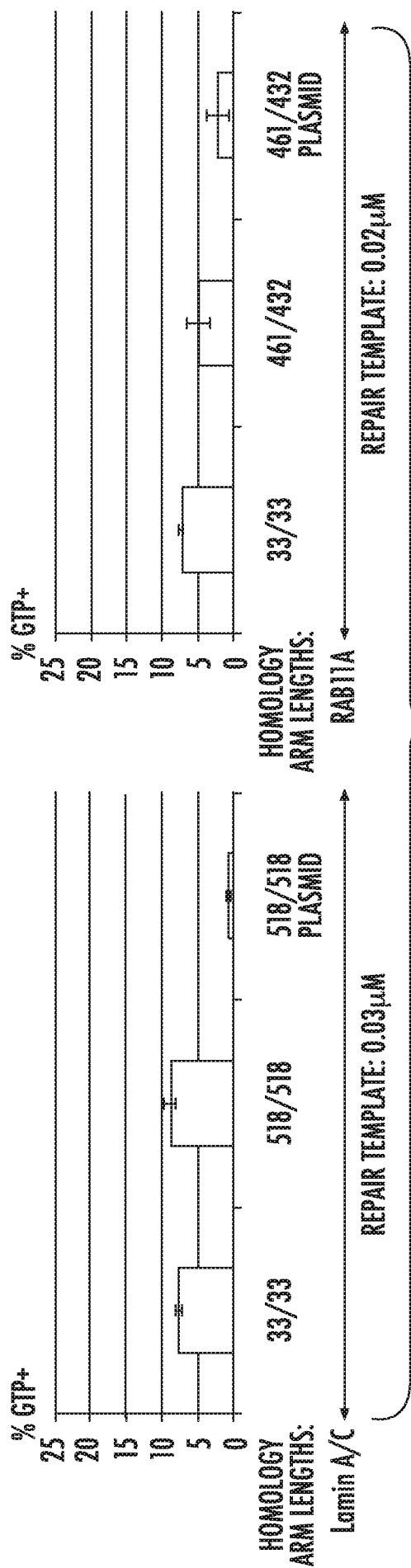


FIG. 2C

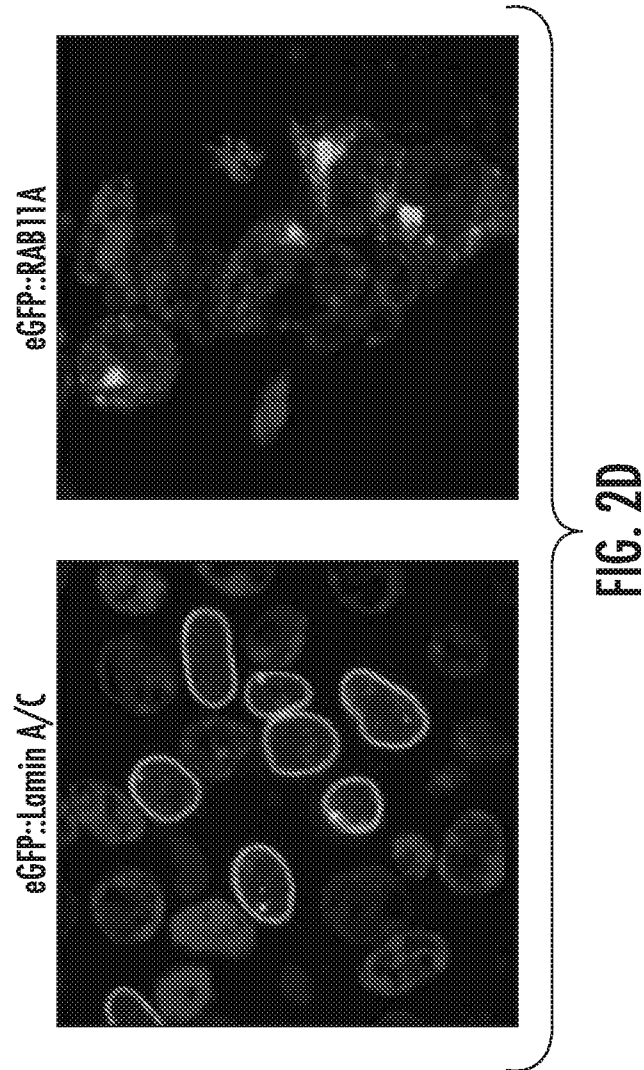


FIG. 2D

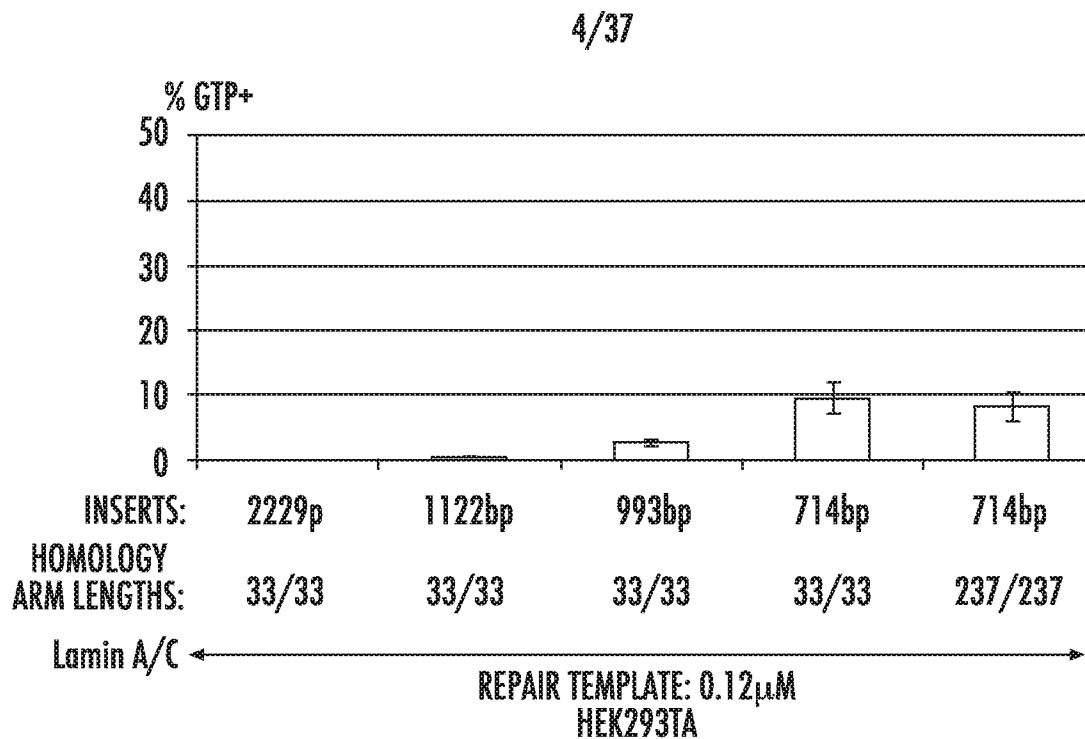


FIG. 3A

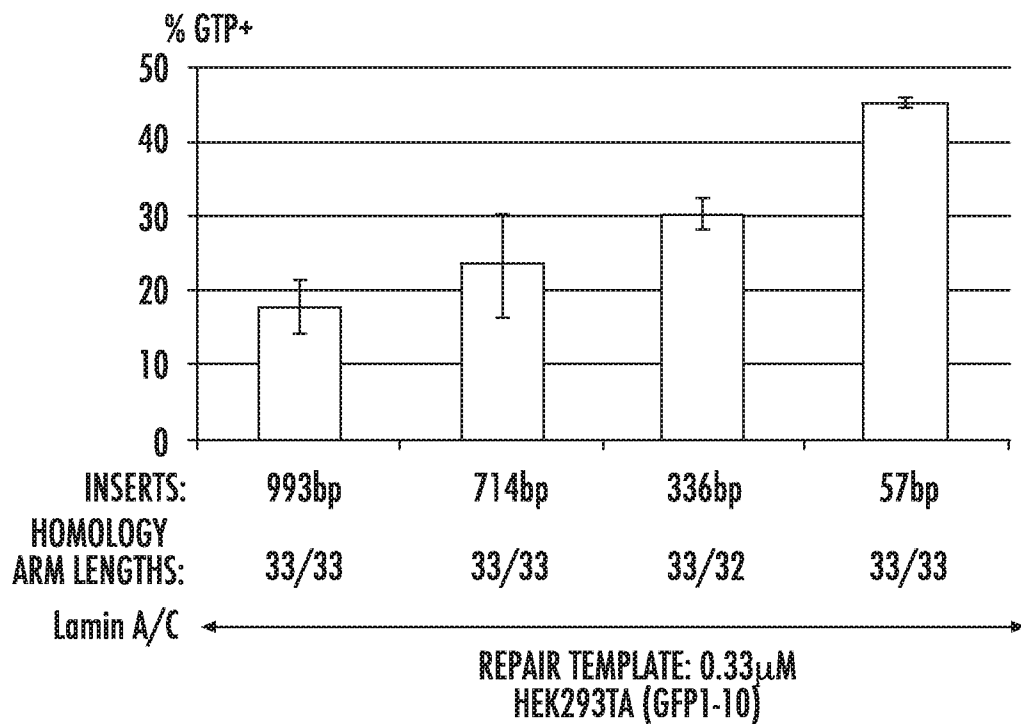


FIG. 3B

5/37

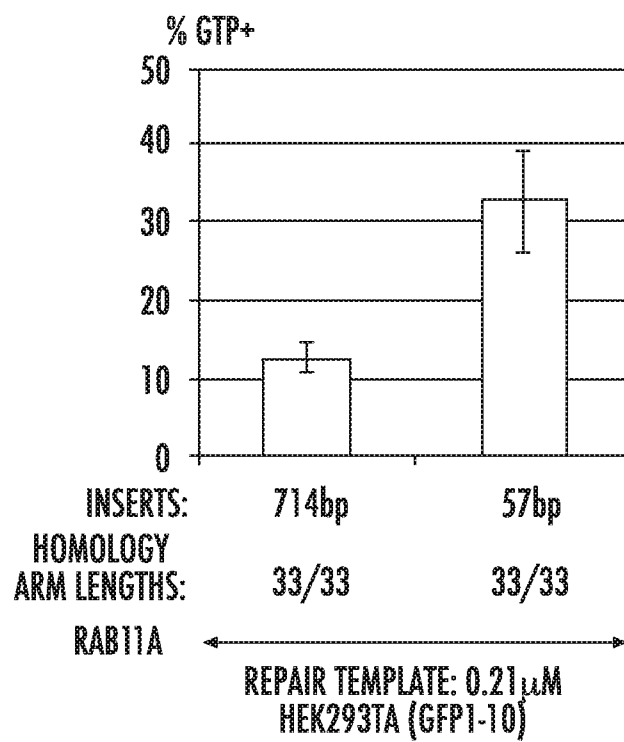


FIG. 3C

6/37

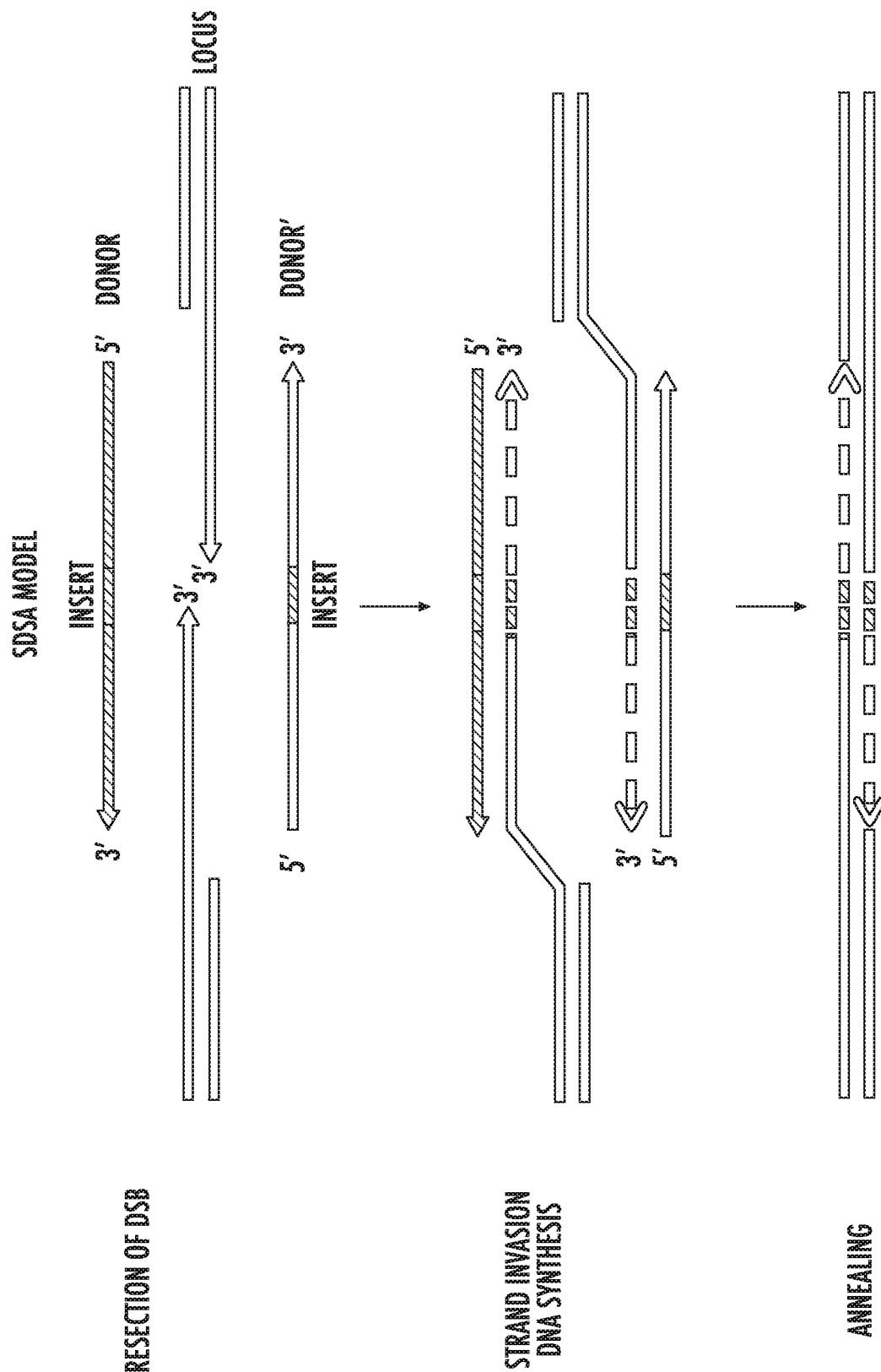


FIG. 4A

7/37

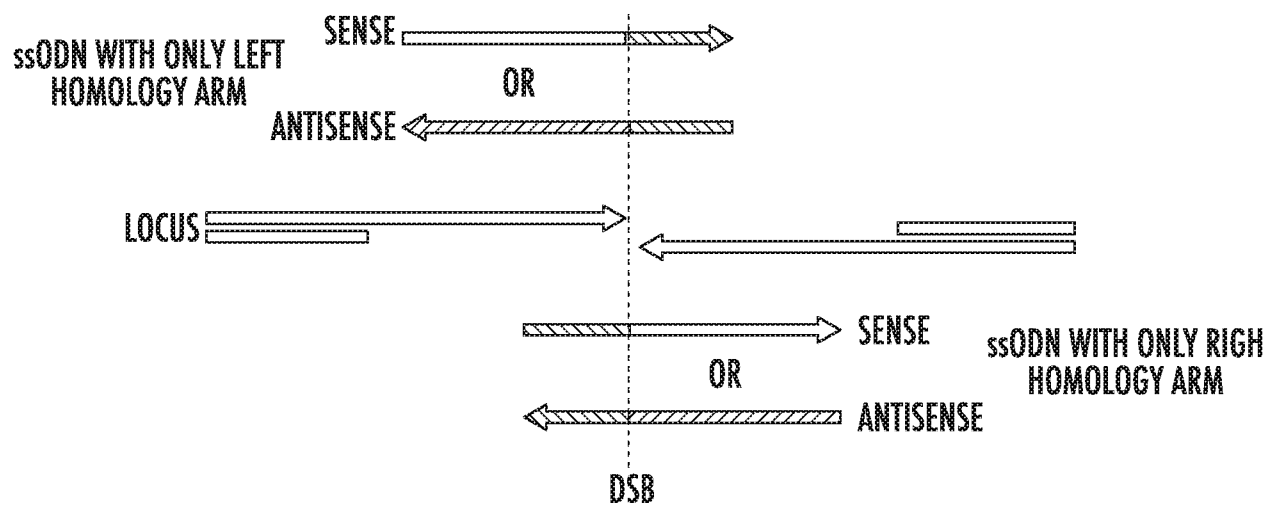


FIG. 4B

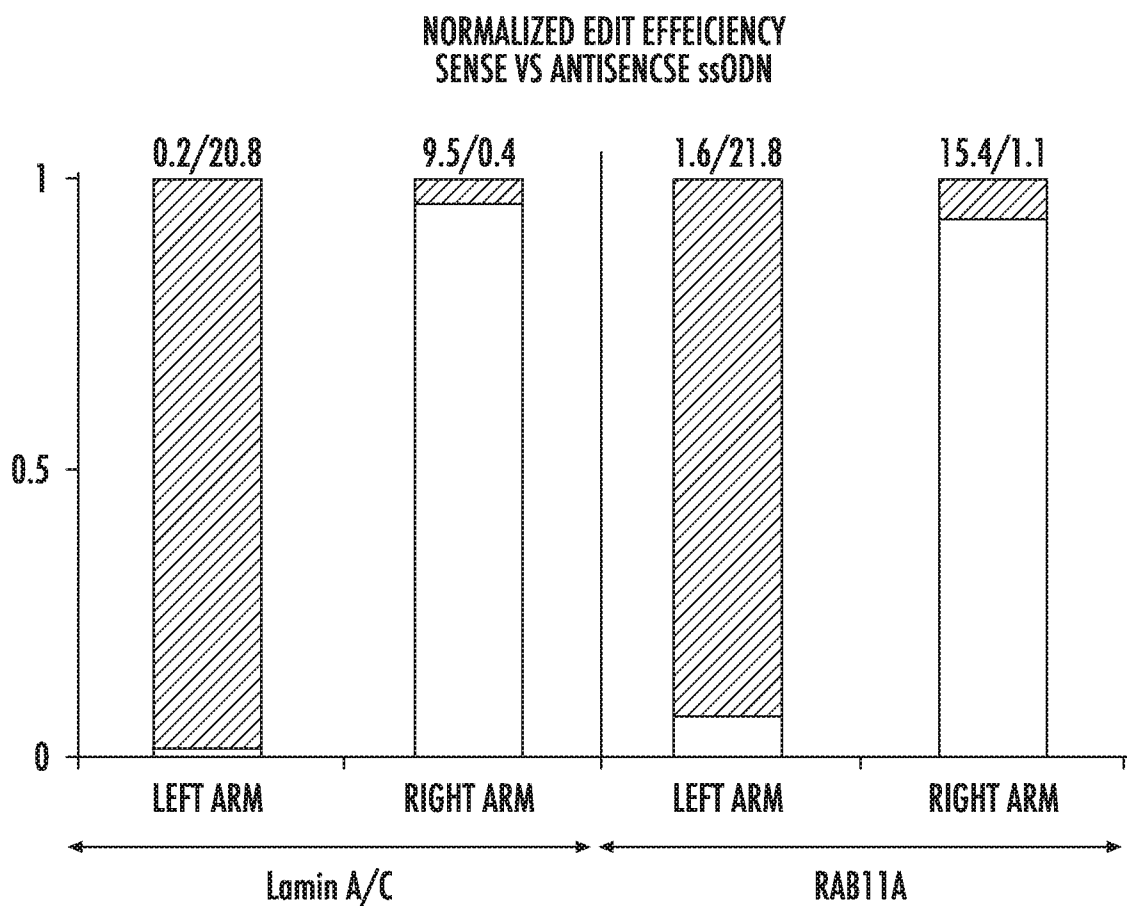
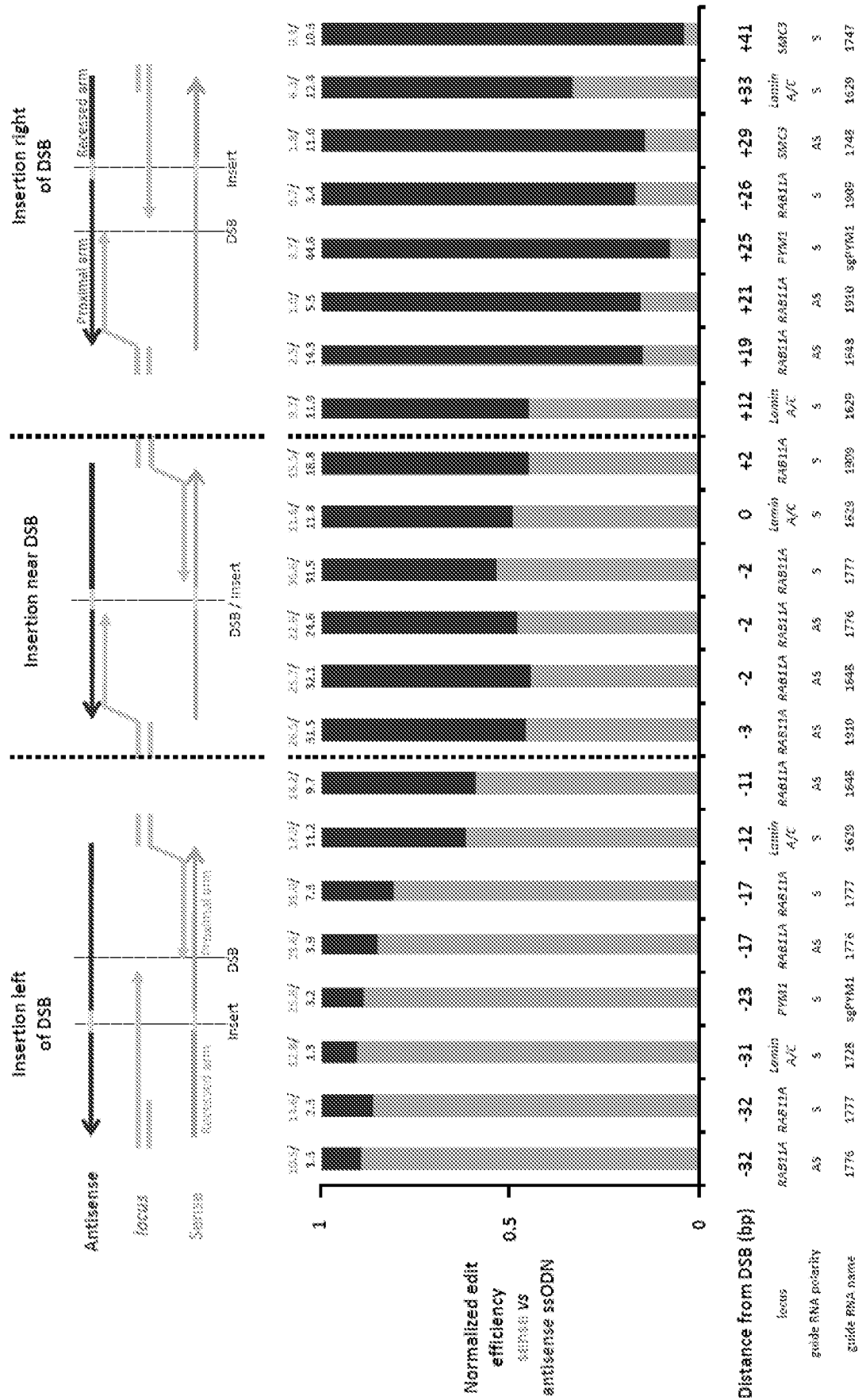


FIG. 4C



9/37

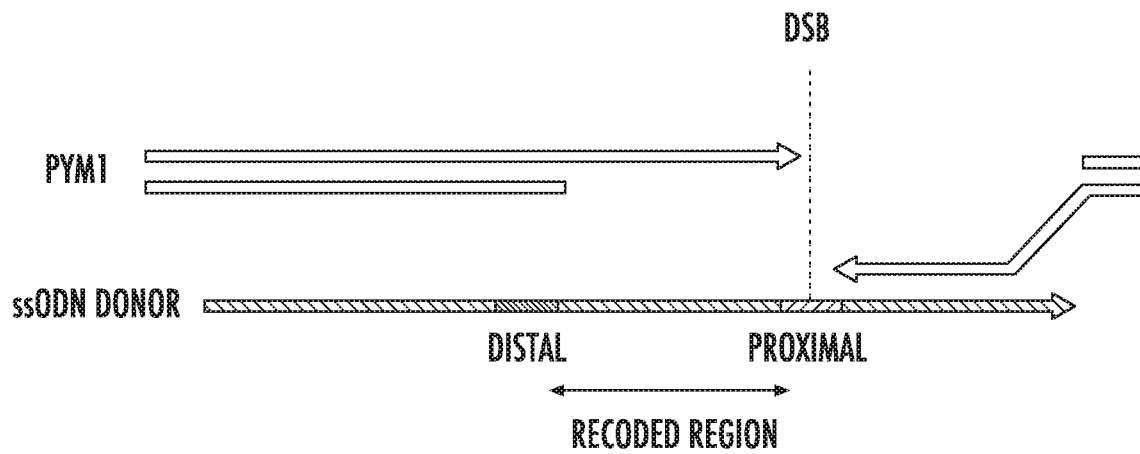


FIG. 6A

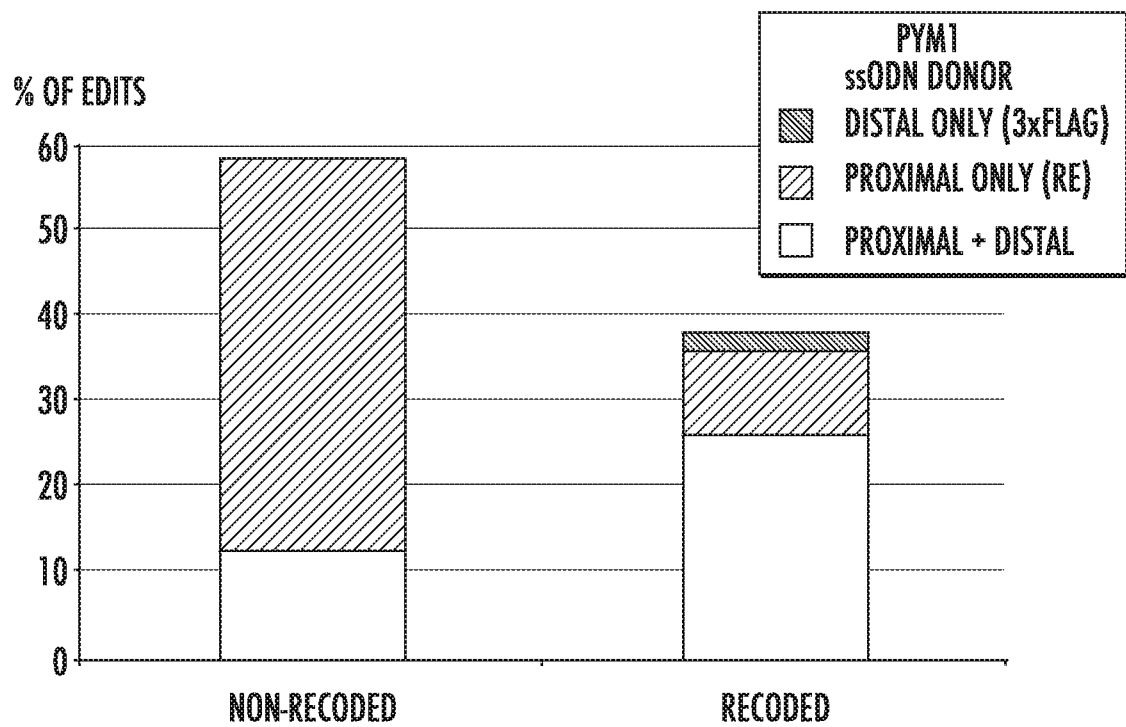


FIG. 6B

10/37

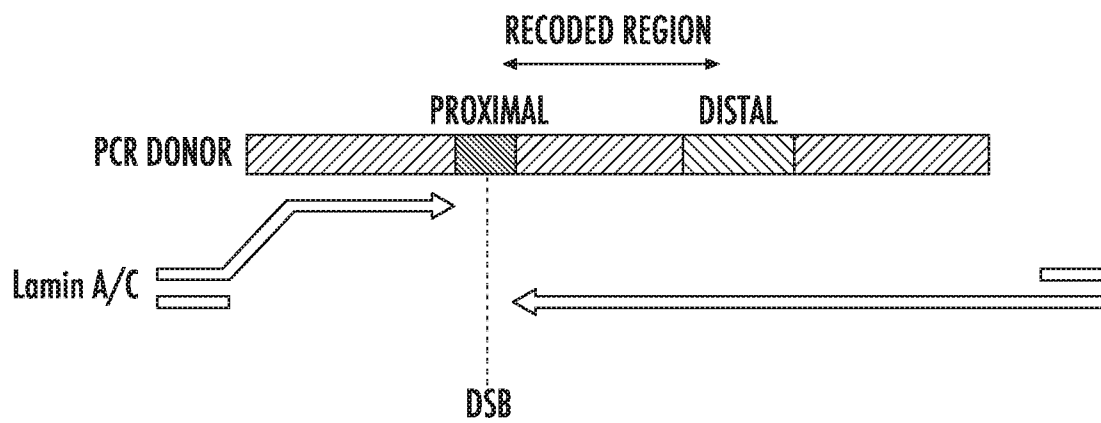


FIG. 6C

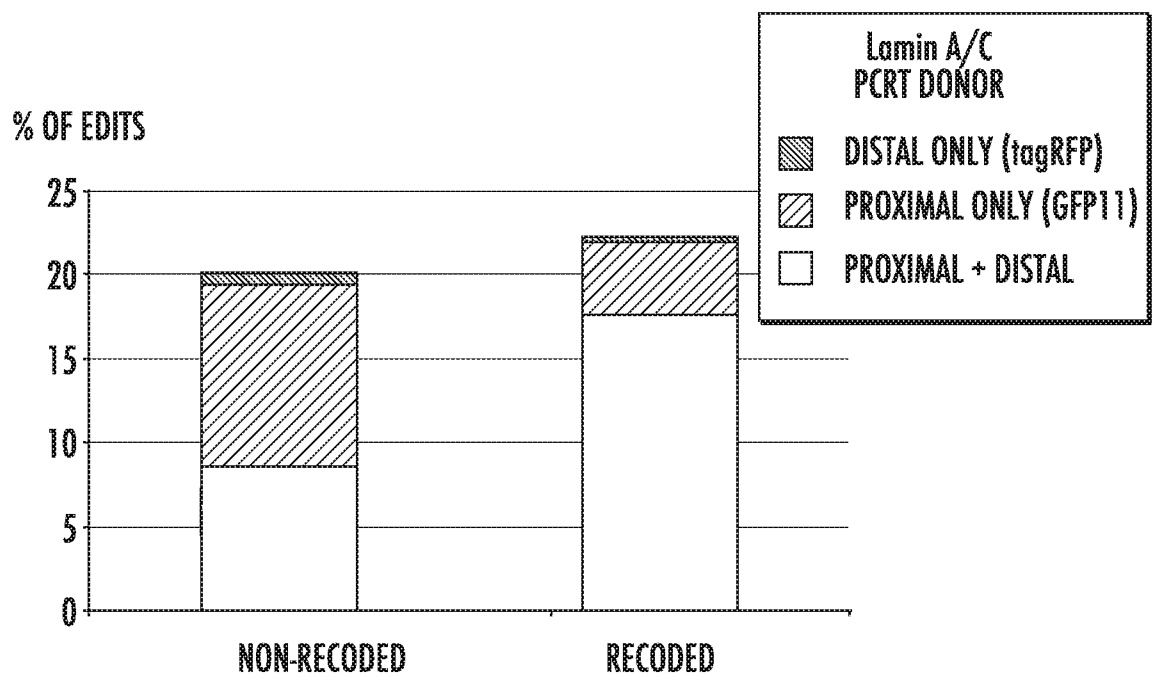


FIG. 6D

11/37

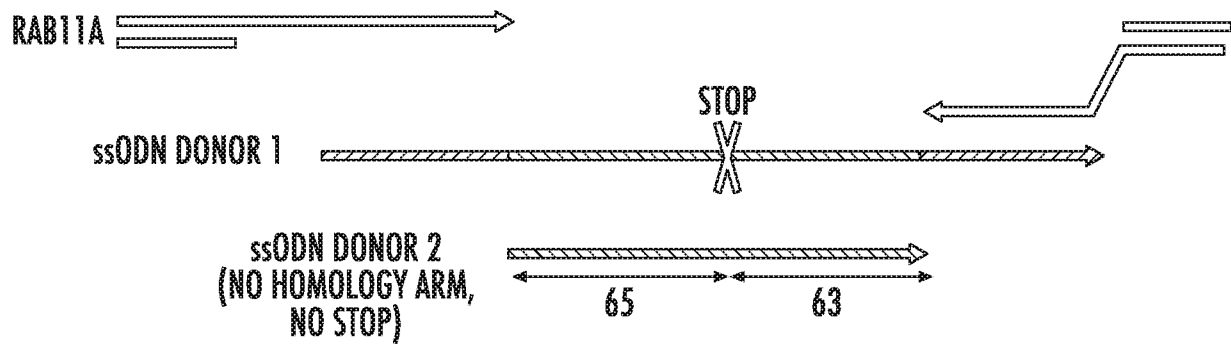


FIG. 7A

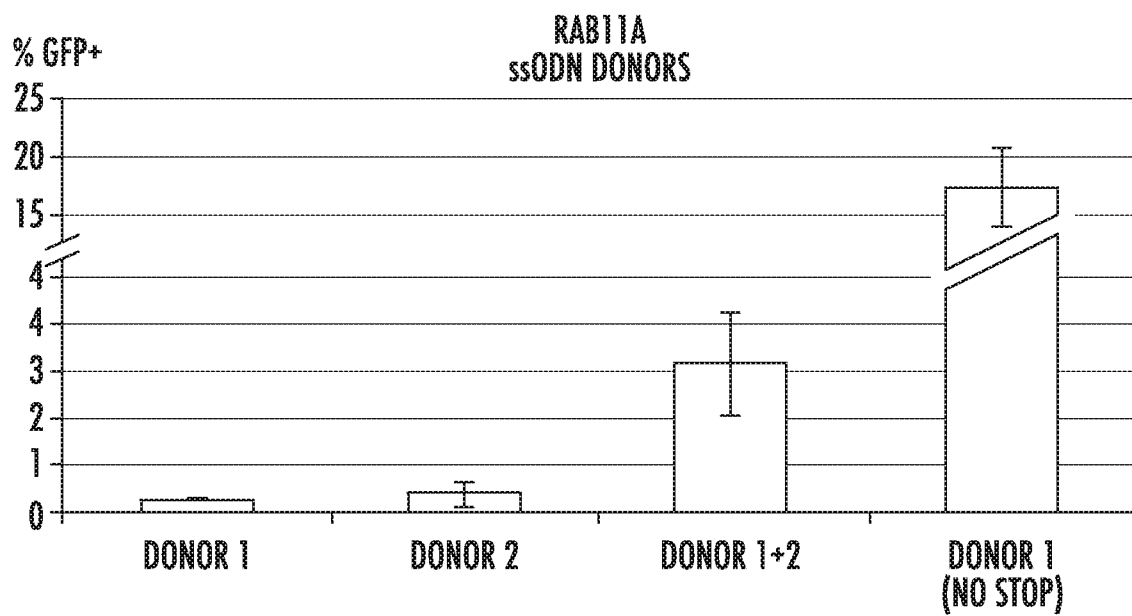


FIG. 7B

12/37

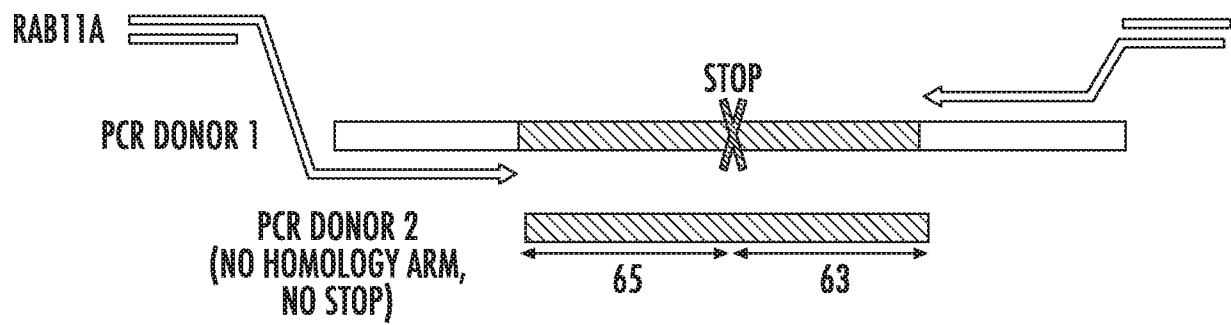


FIG. 7C

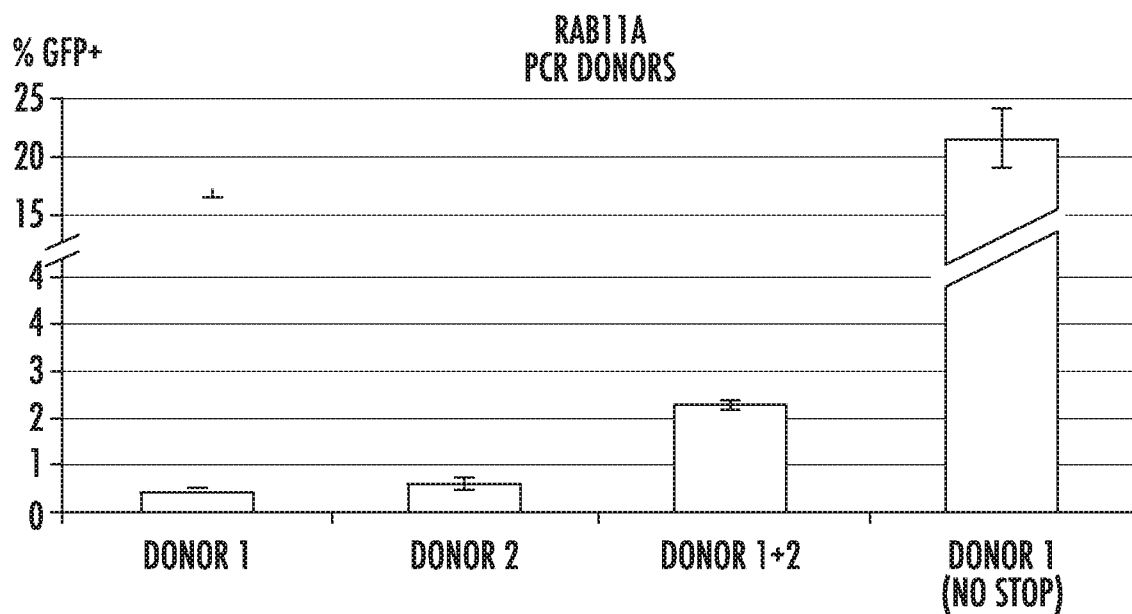


FIG. 7D

13/37

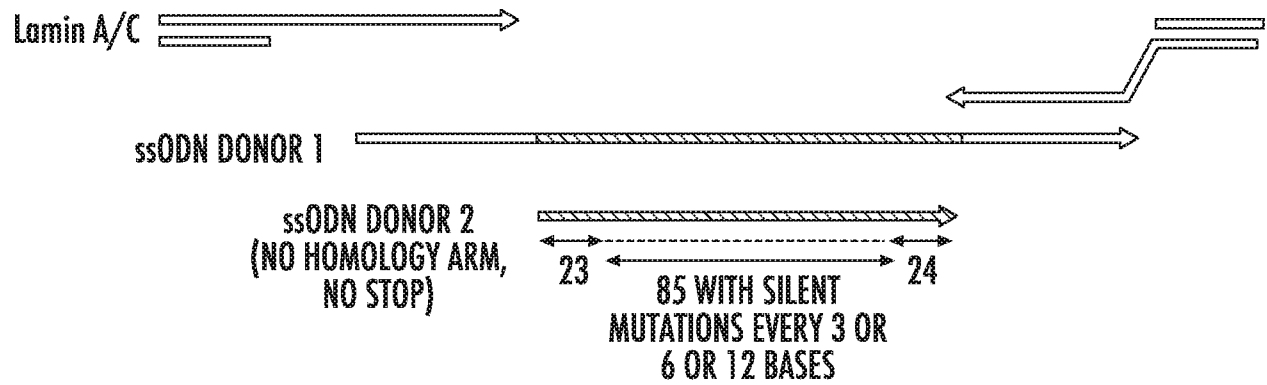


FIG. 7E

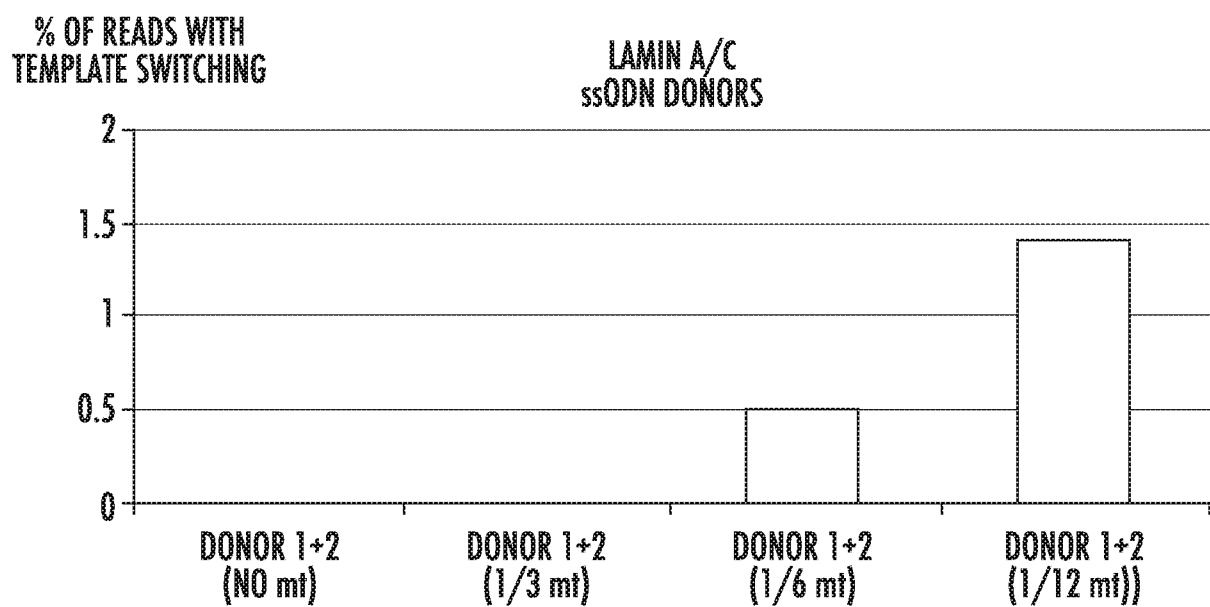


FIG. 7F

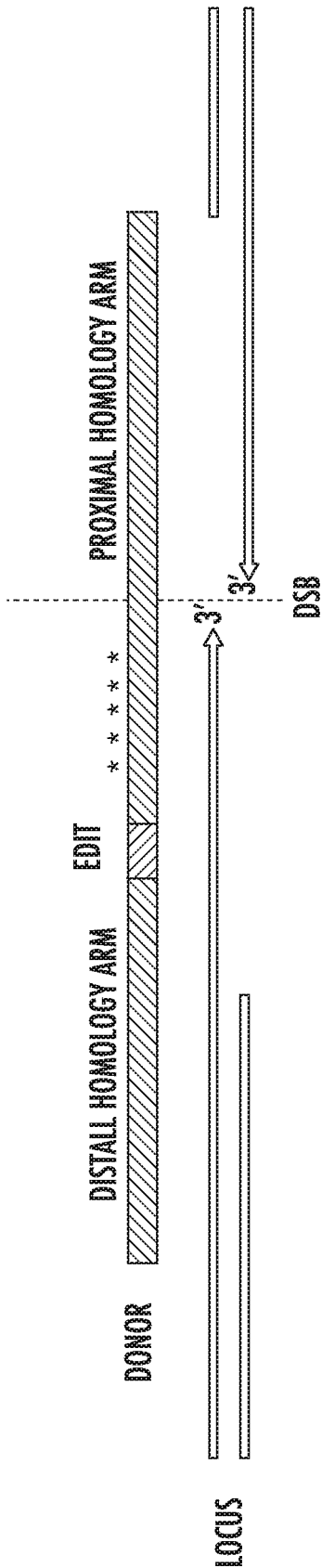


FIG. 8A

1. EDIT: LESS THAN 30 BASES FORM DSB (AND LESS THAN 1kb IN LENGTH IF INSERTION).
2. HOMOLOGY ARMS: ~ 35 BASES.
3. AT LEAST ONE PROXIMAL HOMOLOGY ARM (DIRECTLY ABUTTING DSB) WITH NO/FEW MUTATIONS. IF USING A ssODN, MAKE SURE PROXIMAL HOMOLOGY ARM IS AT 3' END OF THE ssODN.
4. RECODE SEQUENCE BETWEEN EDIT AND DSB TO HELP INTEGRATION OF DISTAL EDITS (ASTERISKS). ALSO INCLUDE HERE ANY MUTATIONS NEEDED TO PREVENT CUTTING OF THE EDITED LOCUS BY Cas9.
5. FOR SMALL EDITS THAT CANNOT BE IDENTIFIED BY SIZE CHANGE, ALSO ADD A RESTRICTION SITE IN EDIT (OR DISTAL TO EDIT) TO FACILITATE DETECTION.

FIG. 8B

15/37

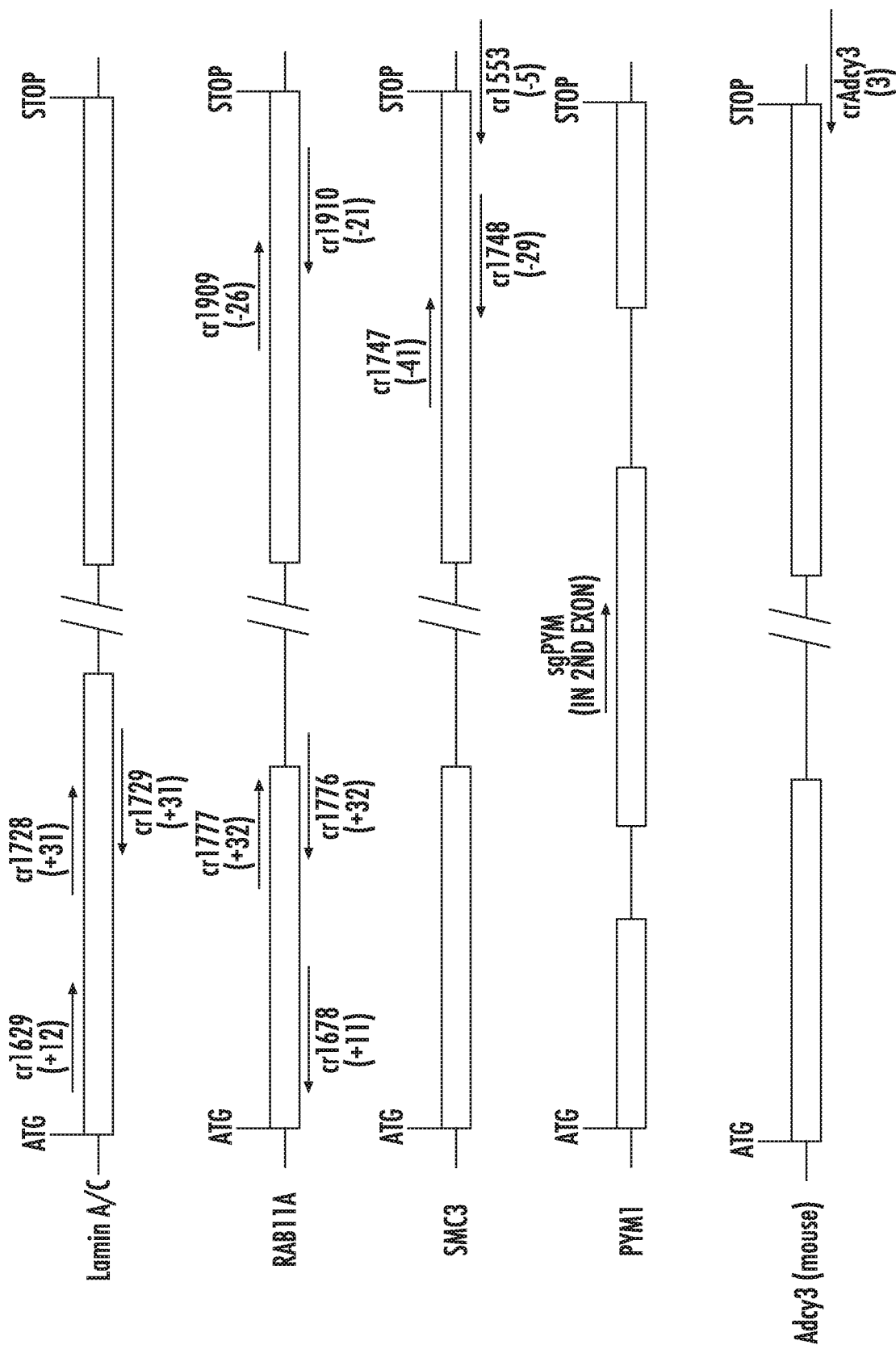


FIG. 9

16/37

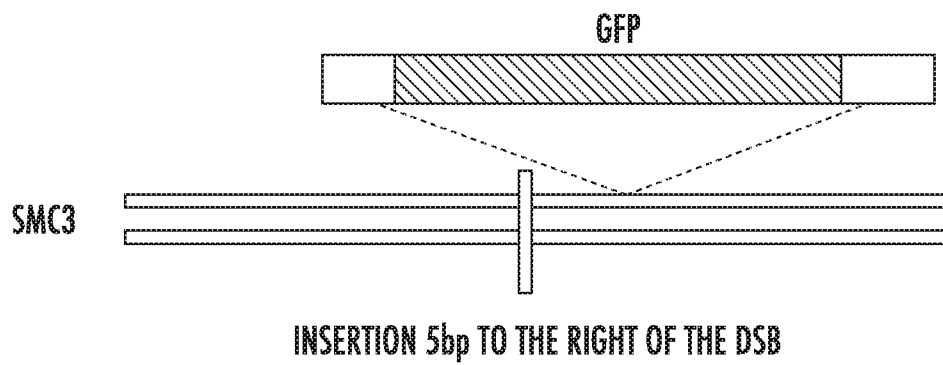


FIG. 10A

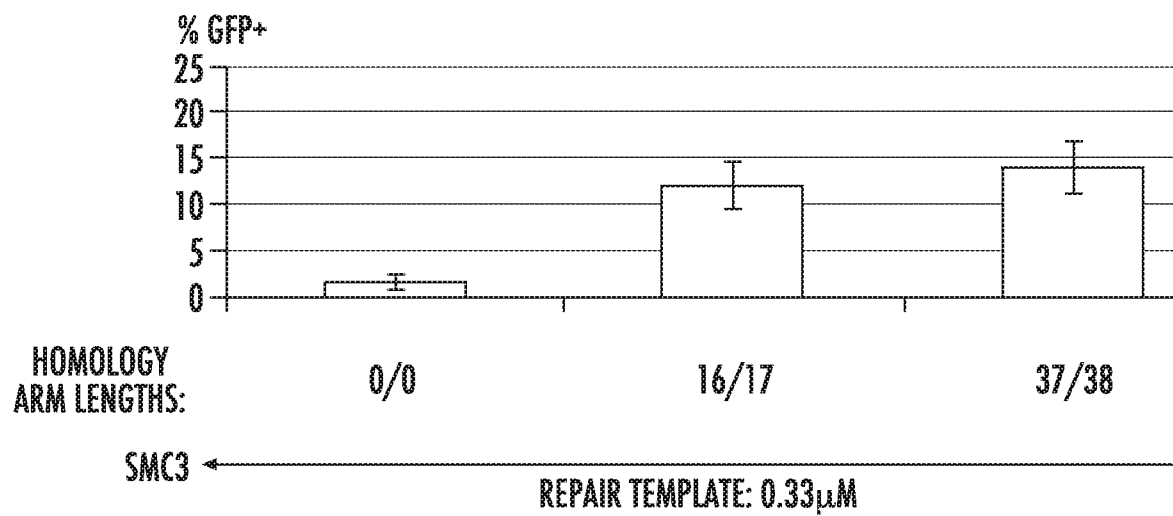


FIG. 10B

17/37

SMC3::eGFP

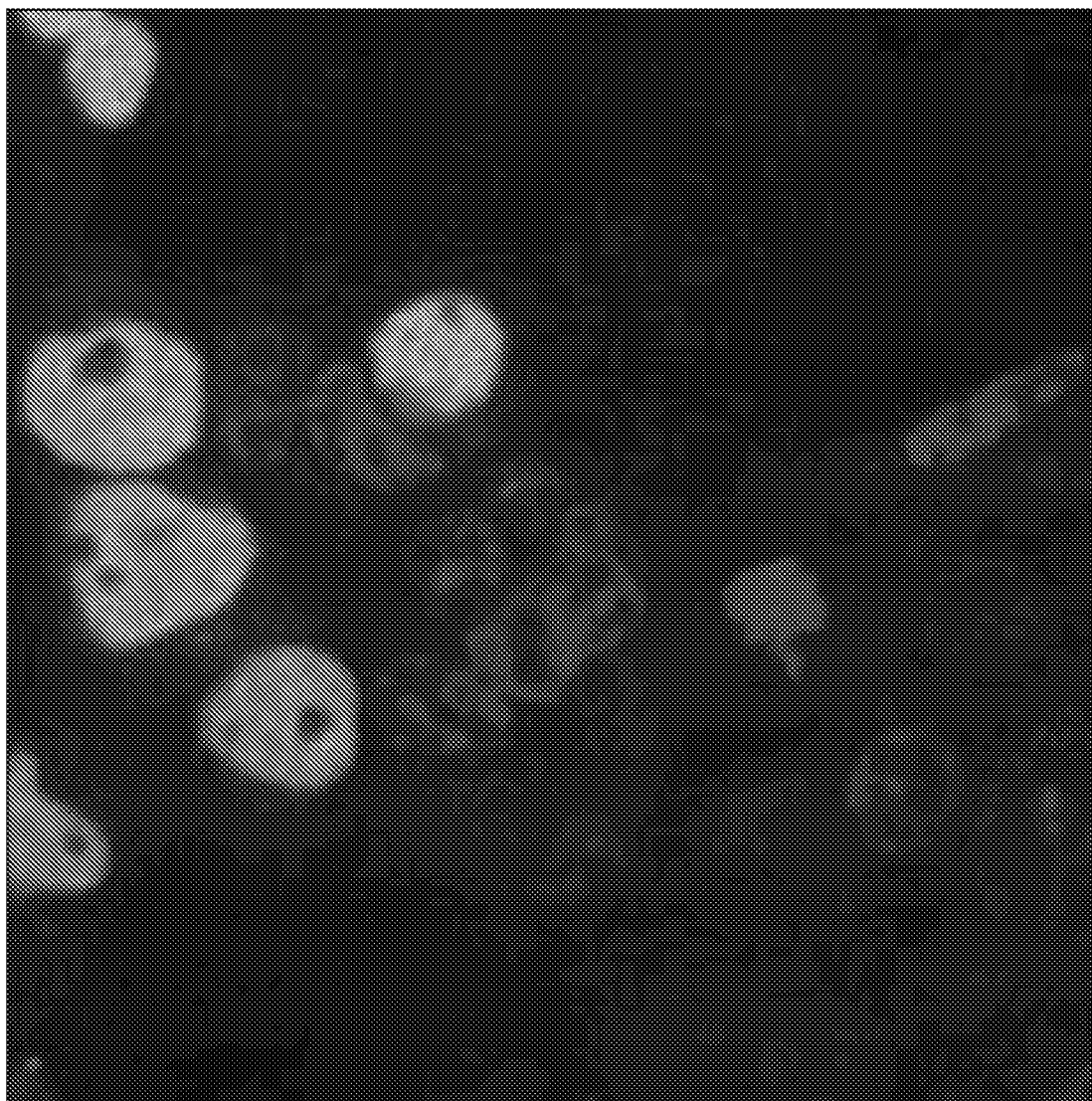


FIG. 10C

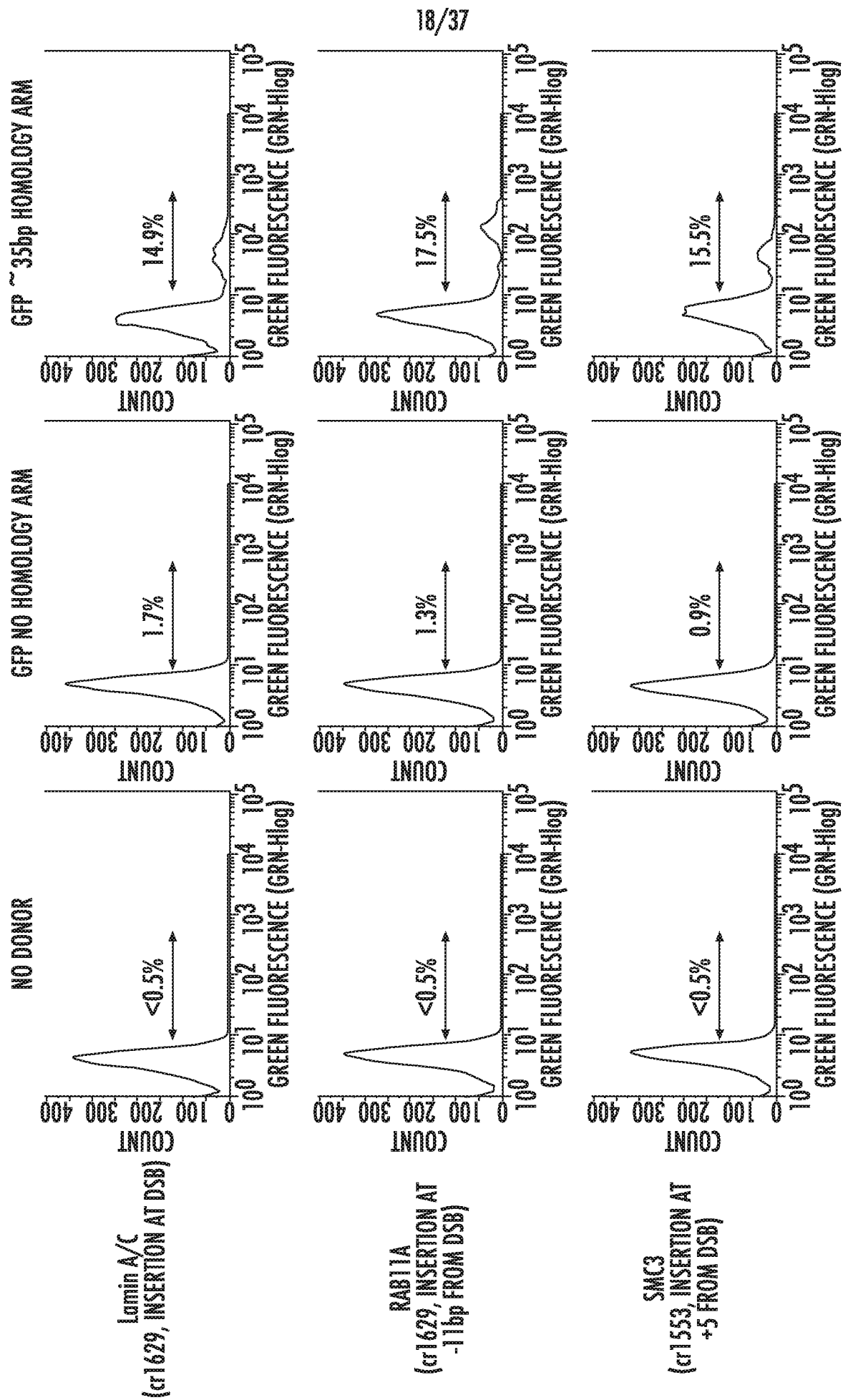


FIG. 11A

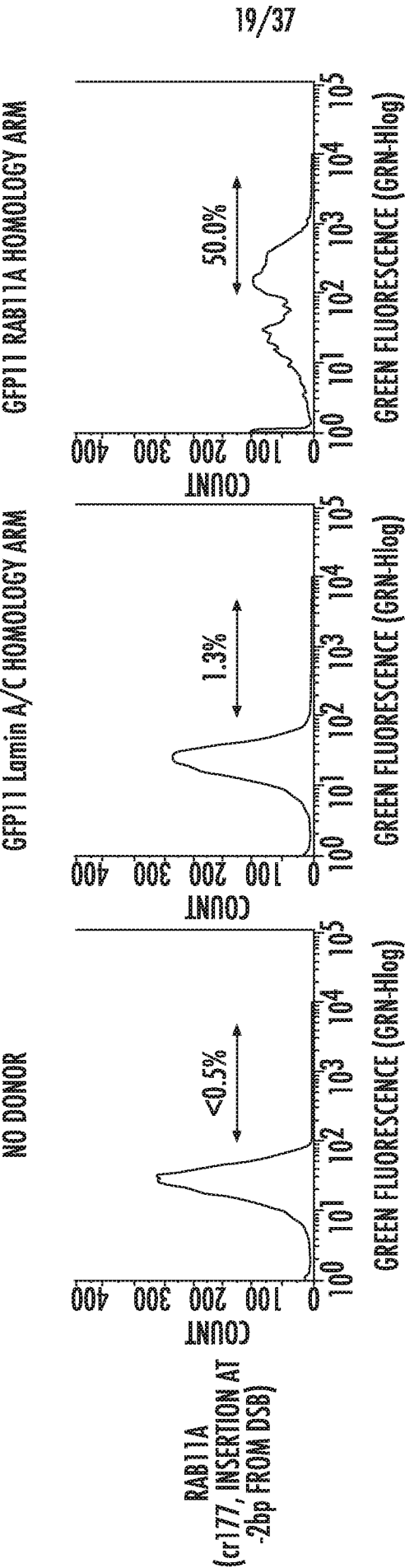


FIG. 11B

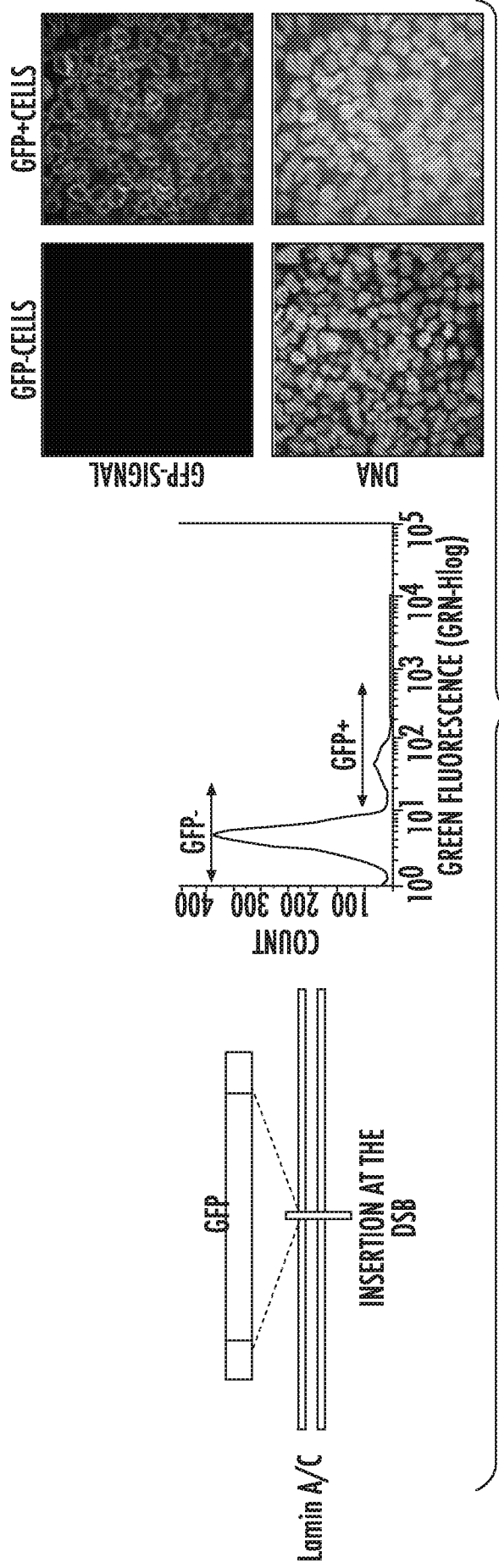


FIG. 12A

	GFP-FRACTION	GFP-FRACTION
% CELLS GFP+(MICROSCOPY)	0% (0/1884)	98% (0/1884)
% CELLS WITH INSERTION (PCR GENOTYPING)	27% (10/37)	91.7% HET (10/37) 8.3% HOMO (4/48)
SEQUENCING RESULTS	6/7 WITH GFP INDEL 1/7 INVERTED GFP INSERTION	6/7 WITH GFP INDEL 1/7 INVERTED GFP INSERTION

FIG. 12B

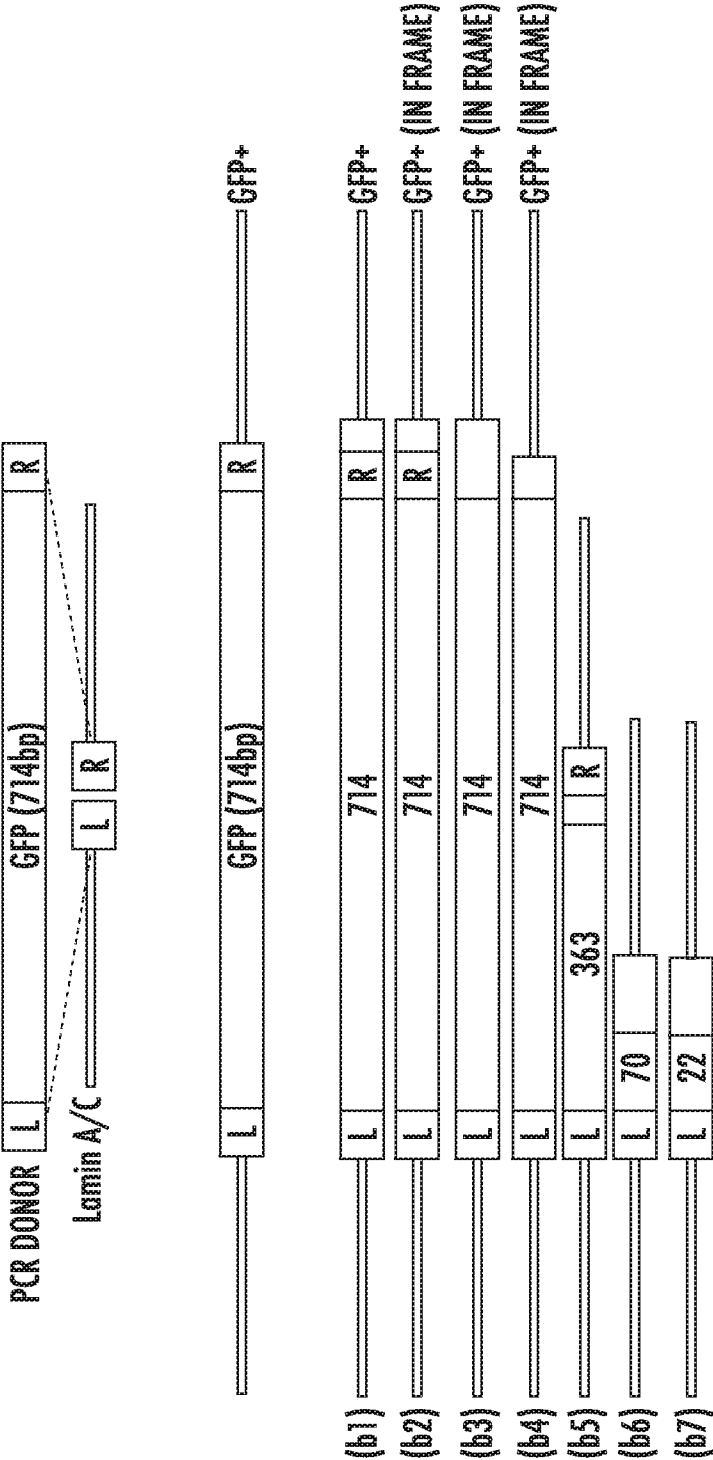


FIG. 13A CORRECT

FIG. 13B LEFT CORRECT

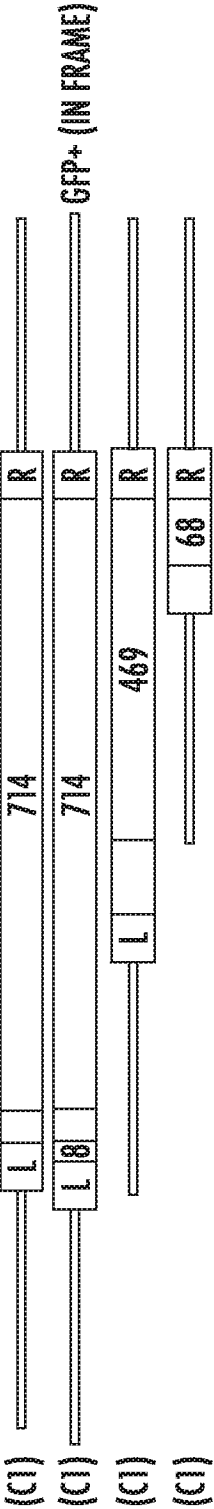


FIG. 13C RIGHT CORRECT



FIG. 13D INTERNAL DELETION



FIG. 13E INVERTED

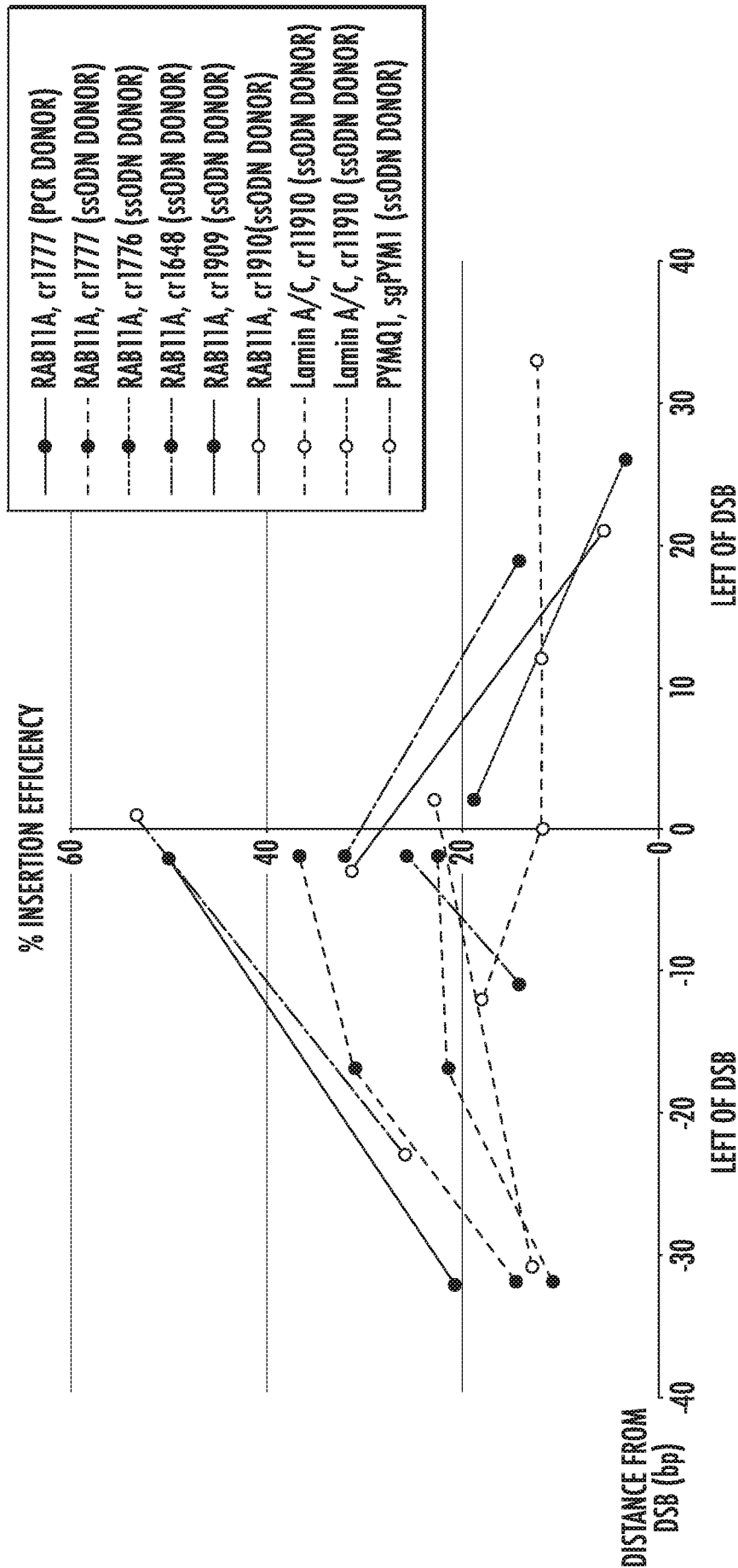


FIG. 14

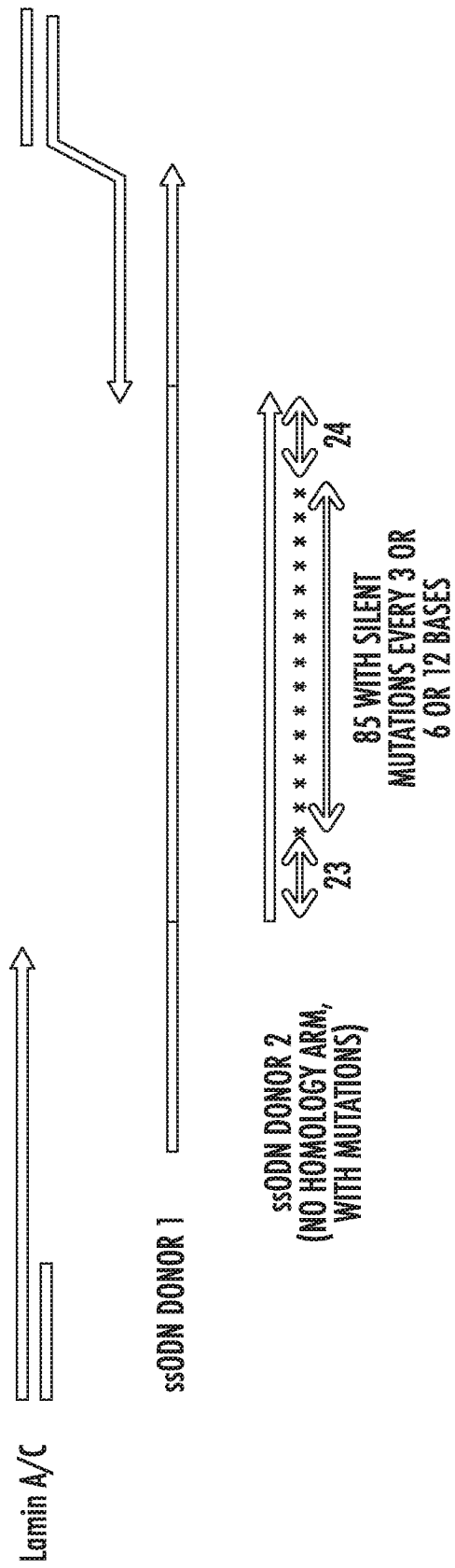


FIG. 15A

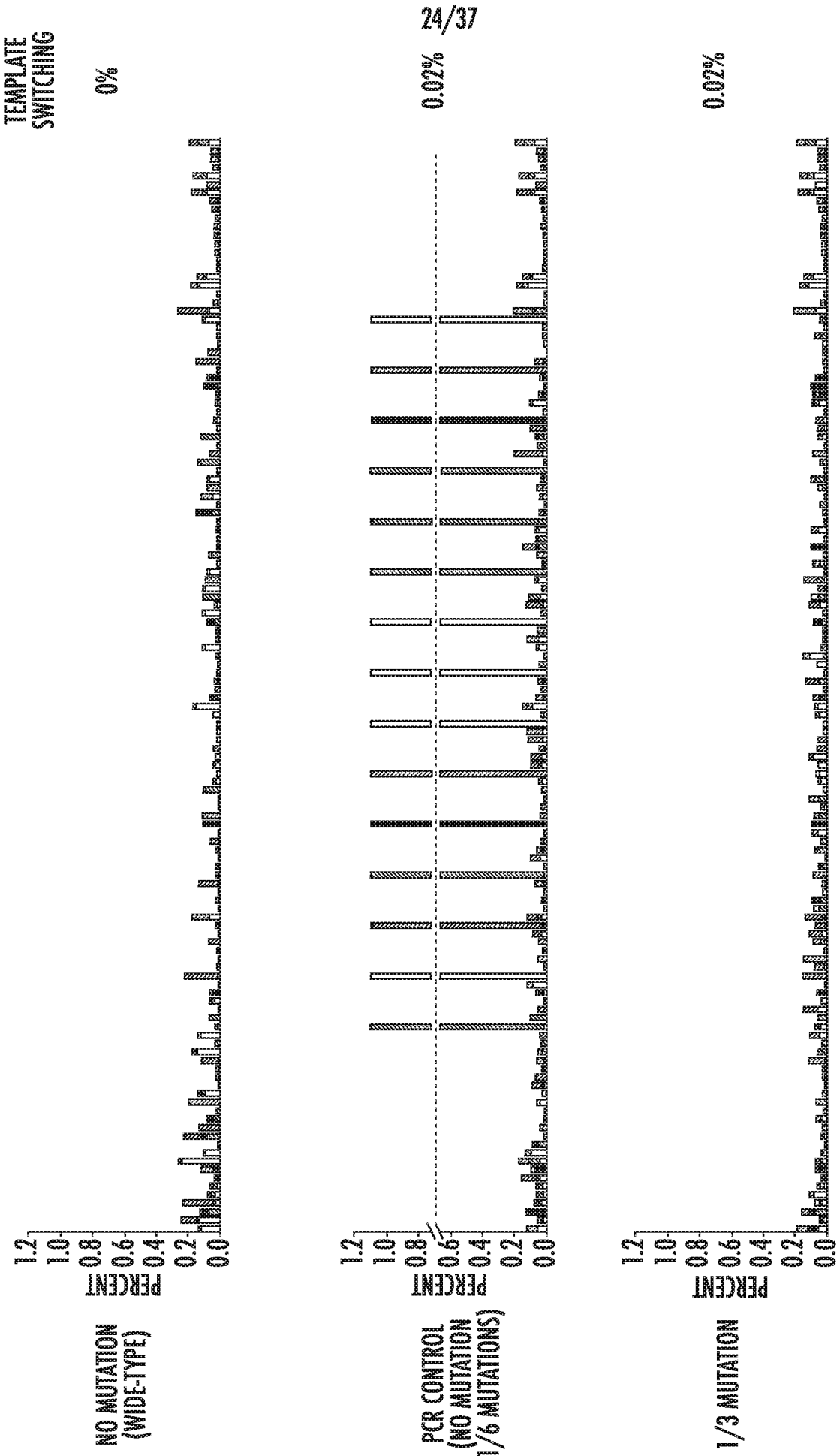


FIG. 15B

25/37

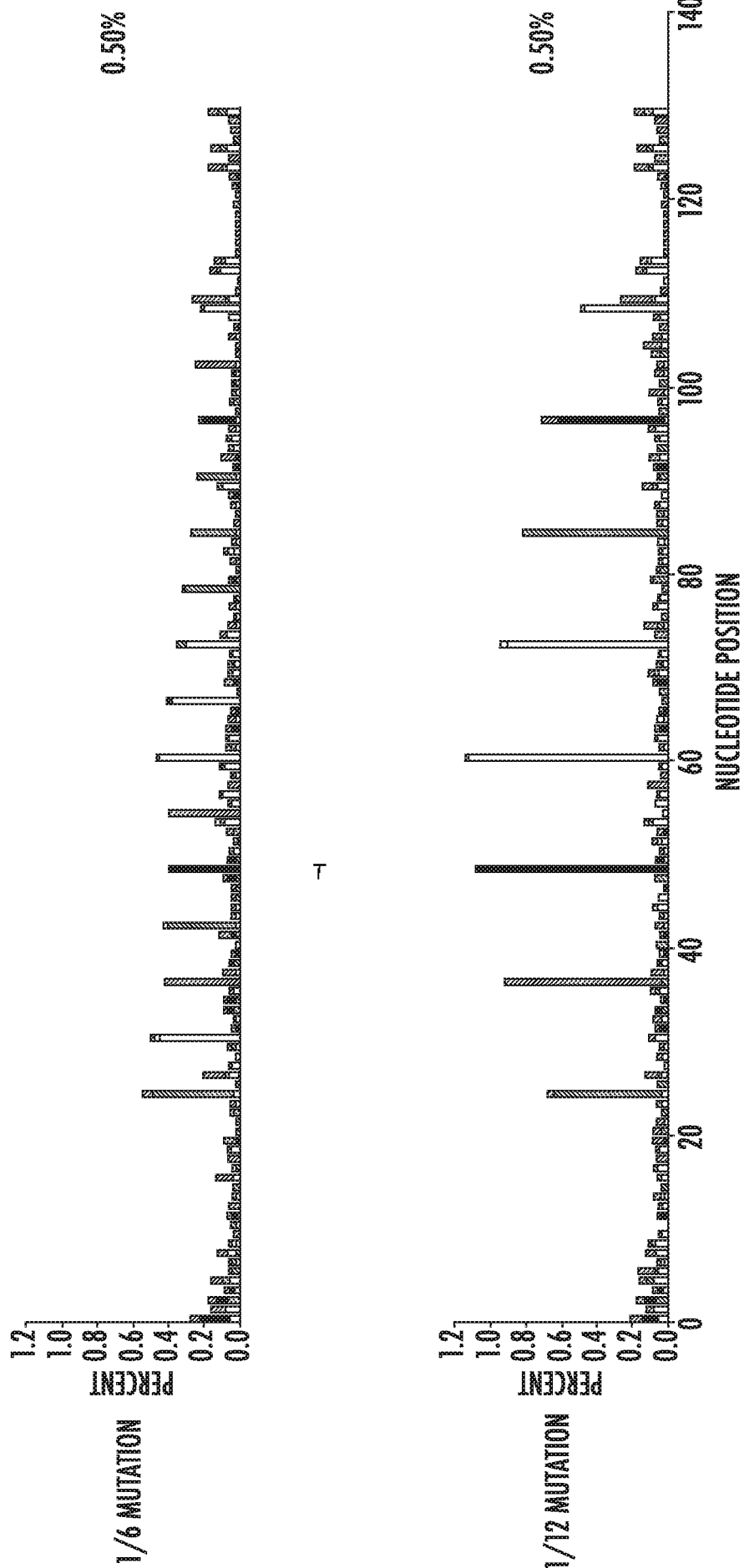
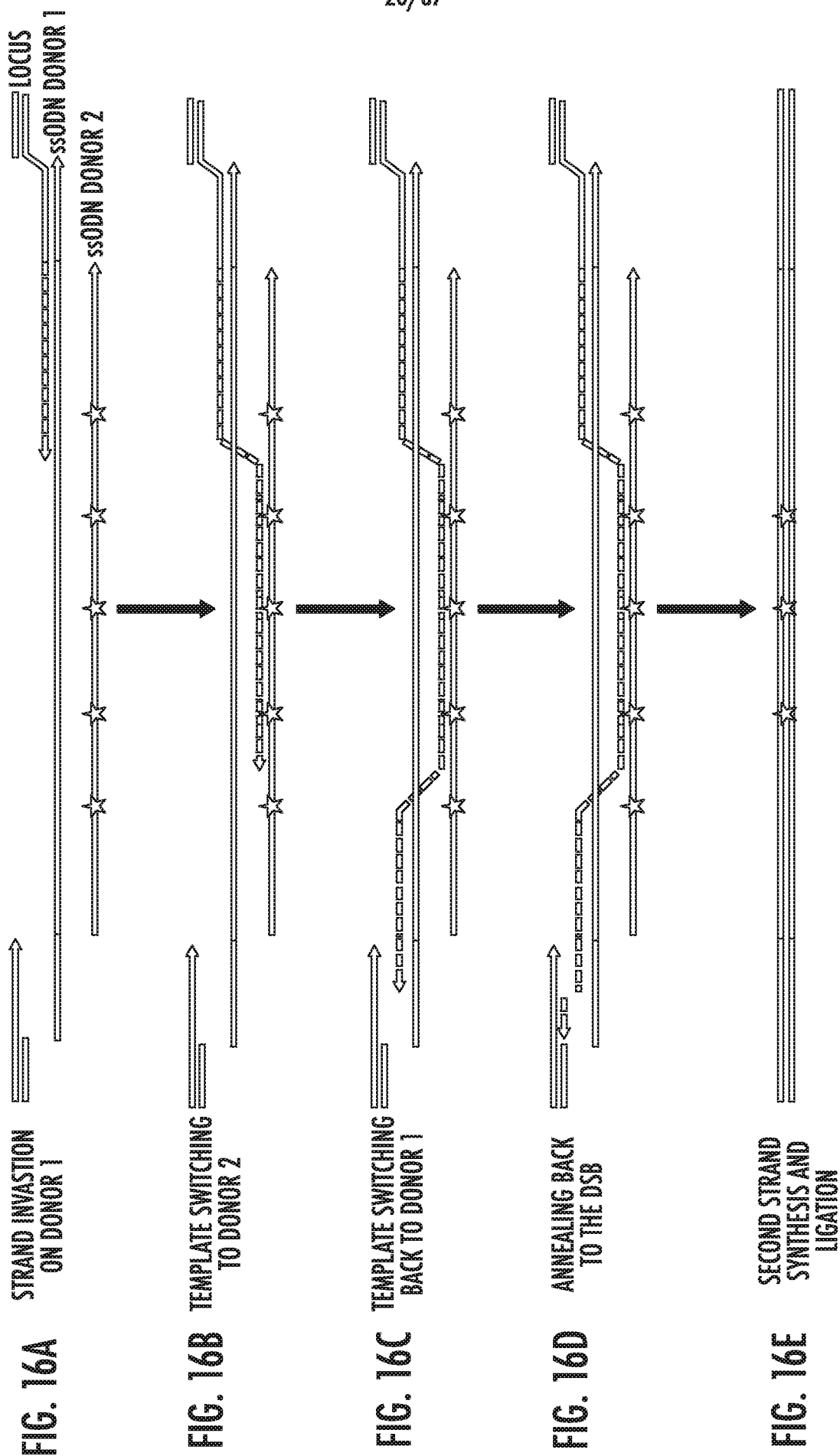
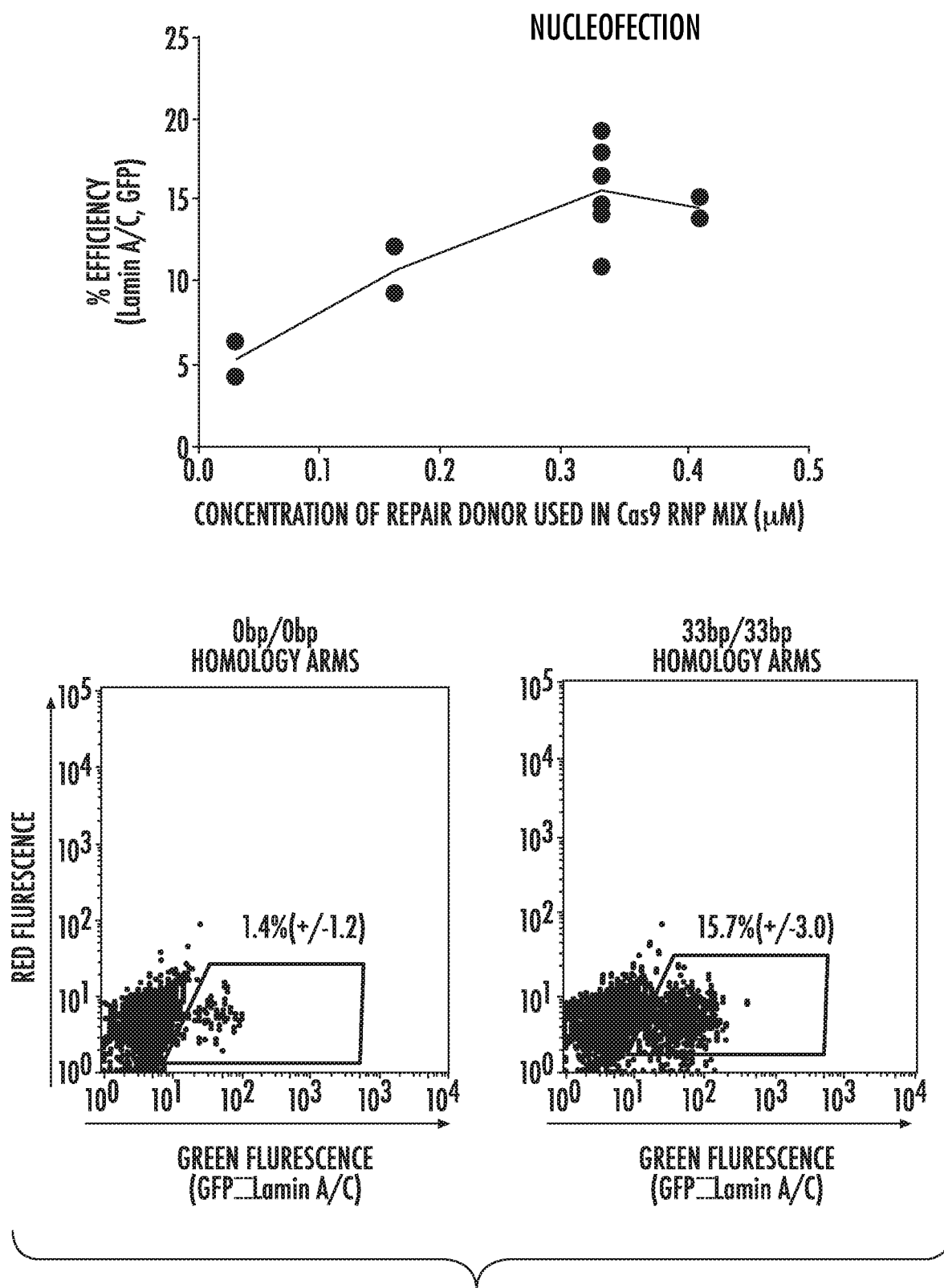


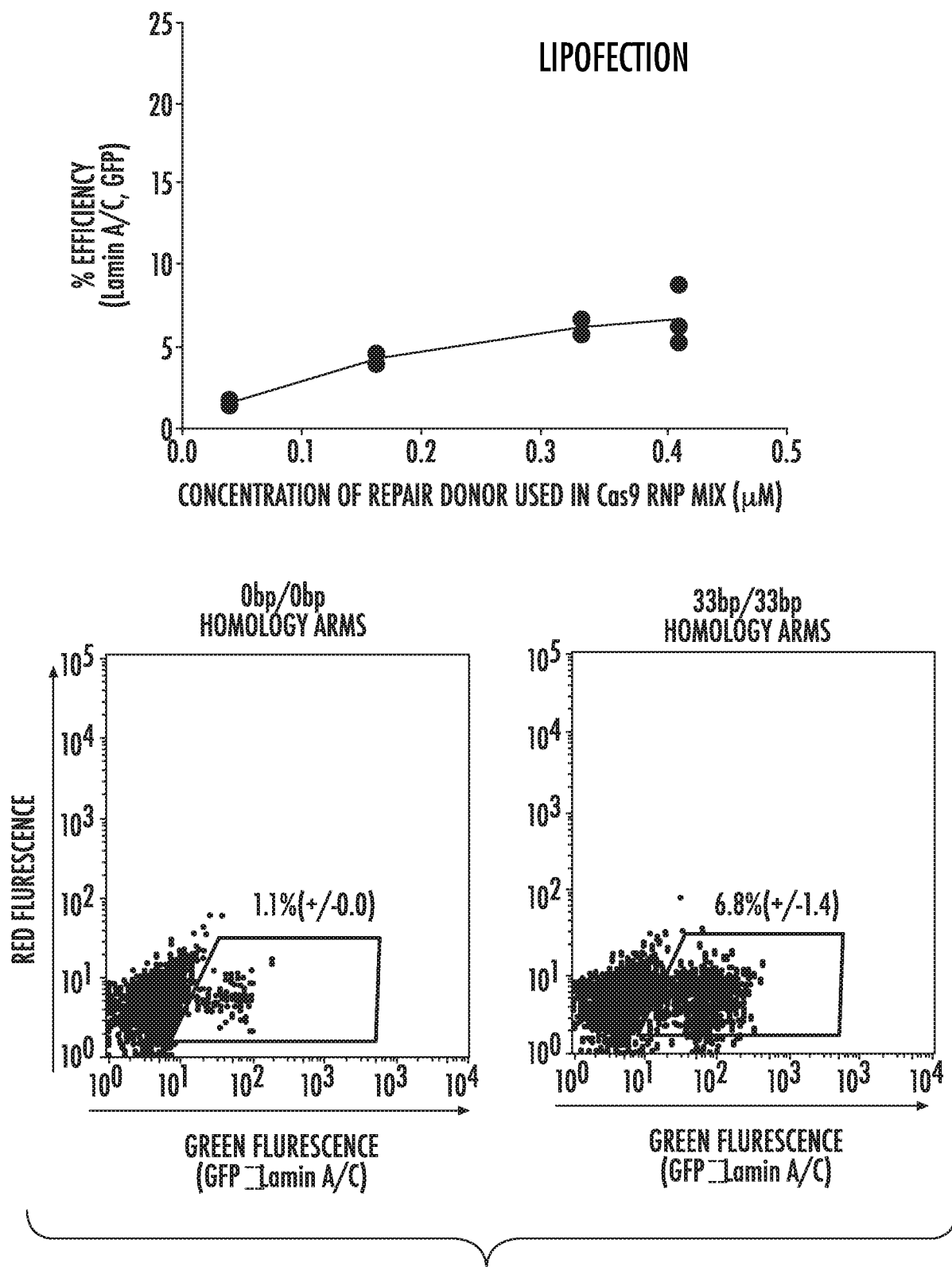
FIG. 15B CONTINUED



27/37



28/37



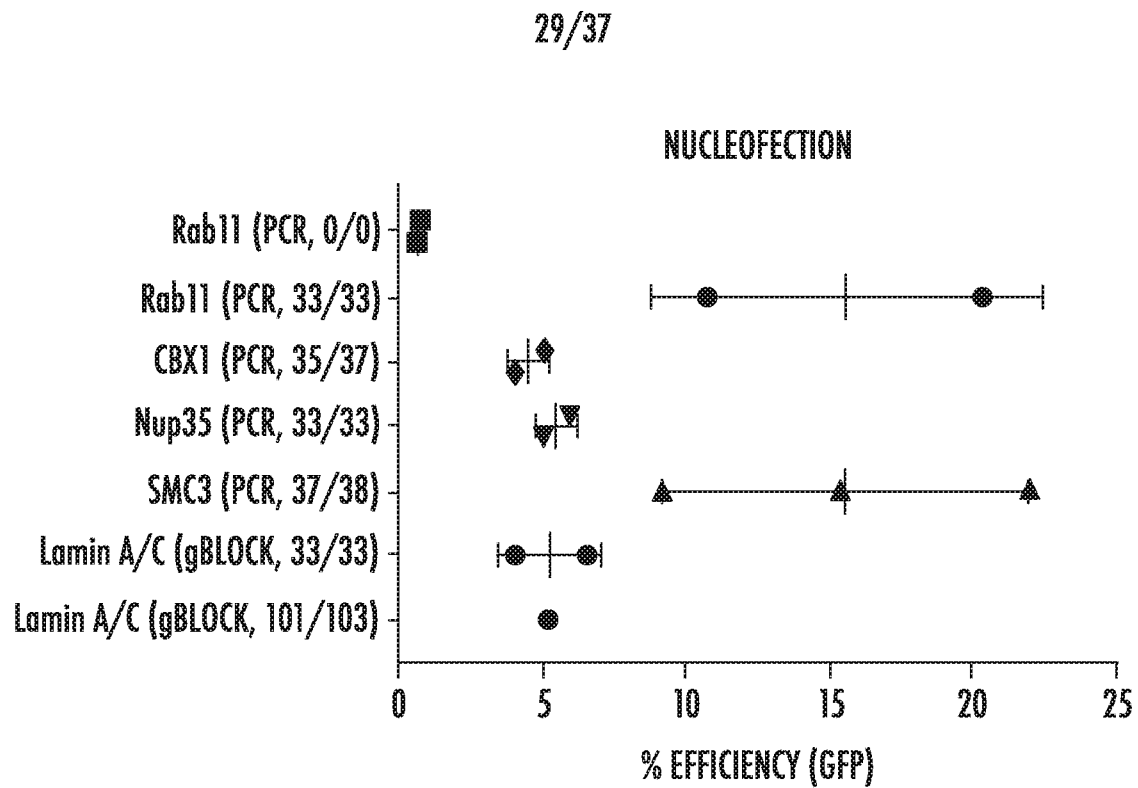


FIG. 18A

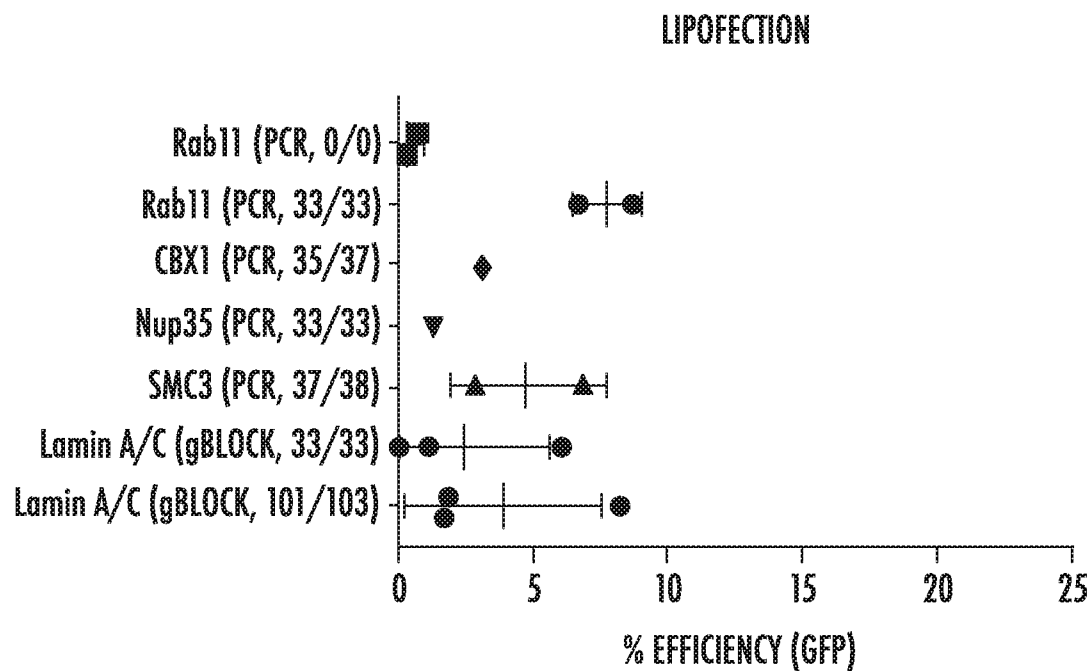


FIG. 18B

EXPRESSION PATTERNS

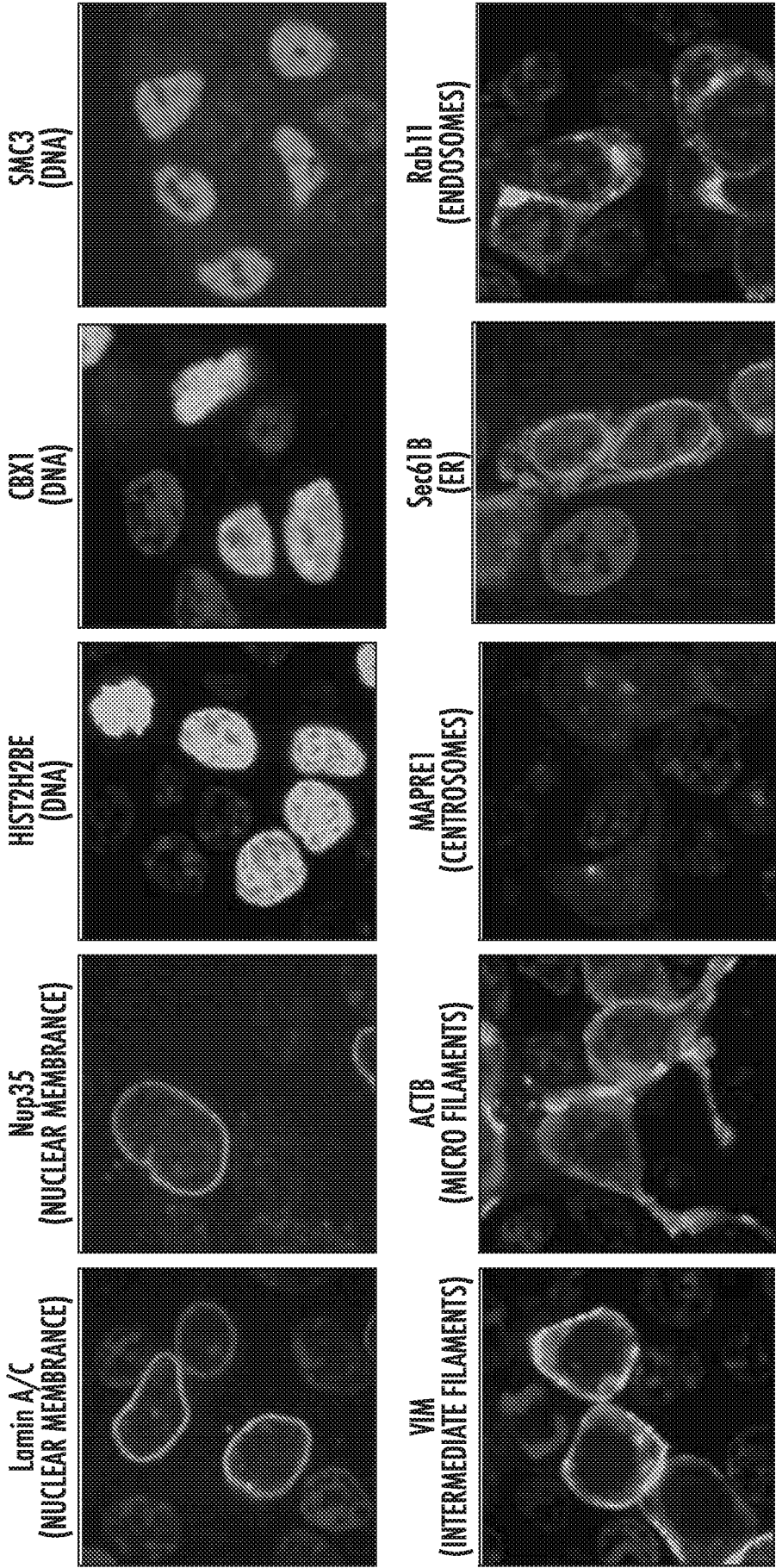


FIG. 18C

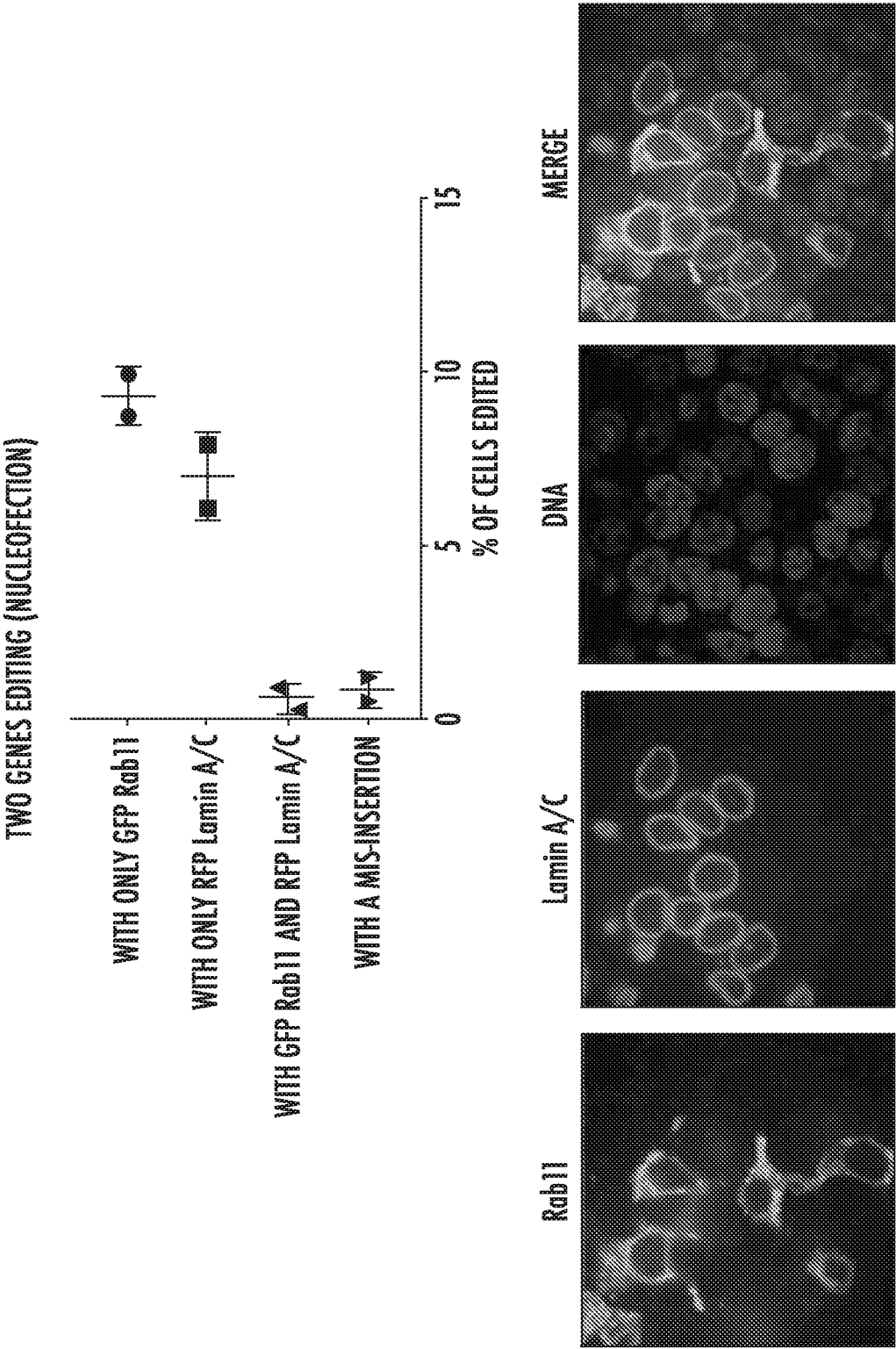
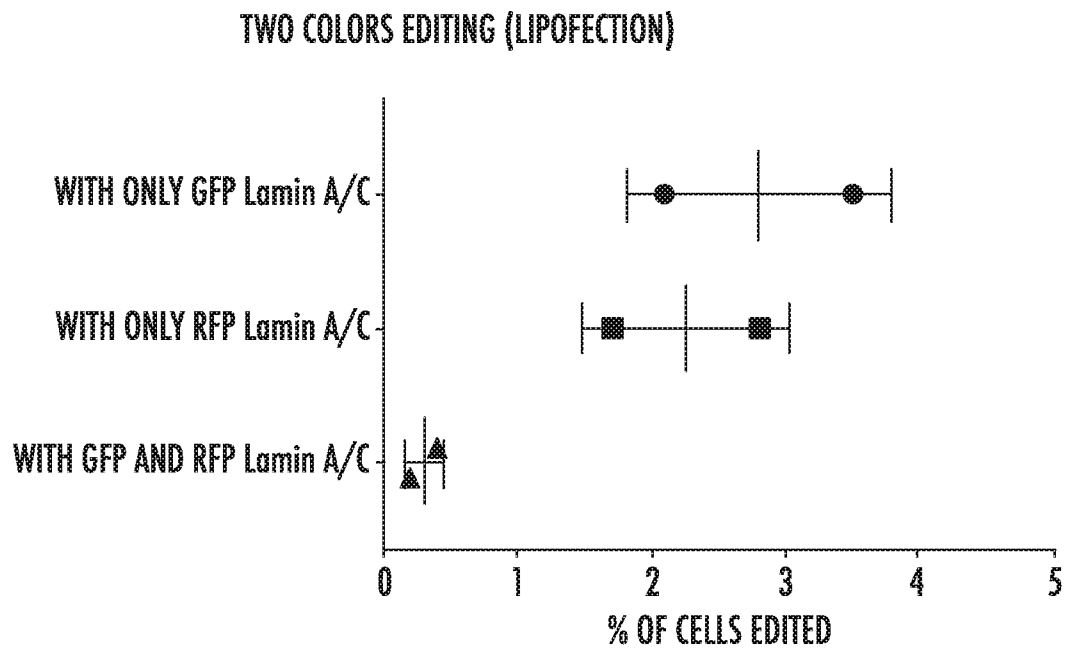


FIG. 19A

32/37

**FIG. 19B**

33/37

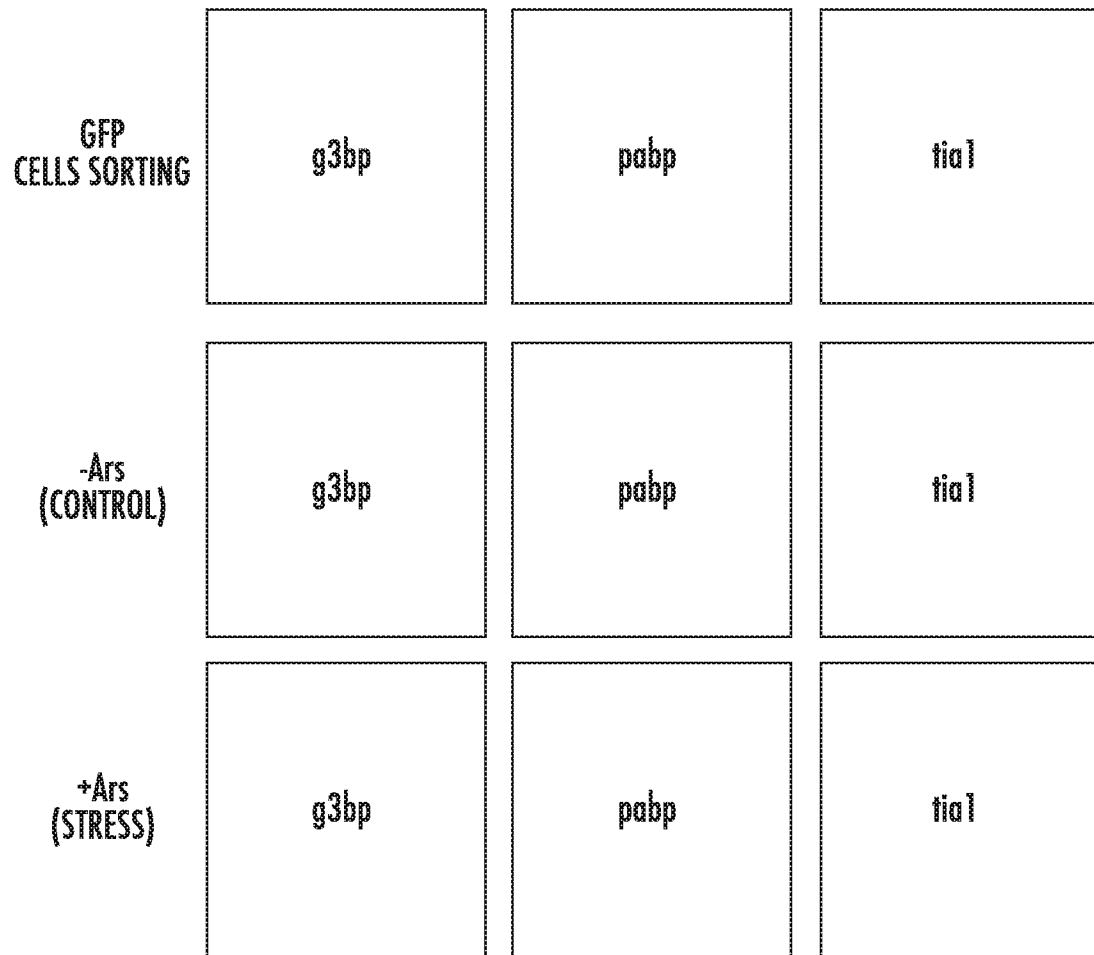


FIG. 20

34/37

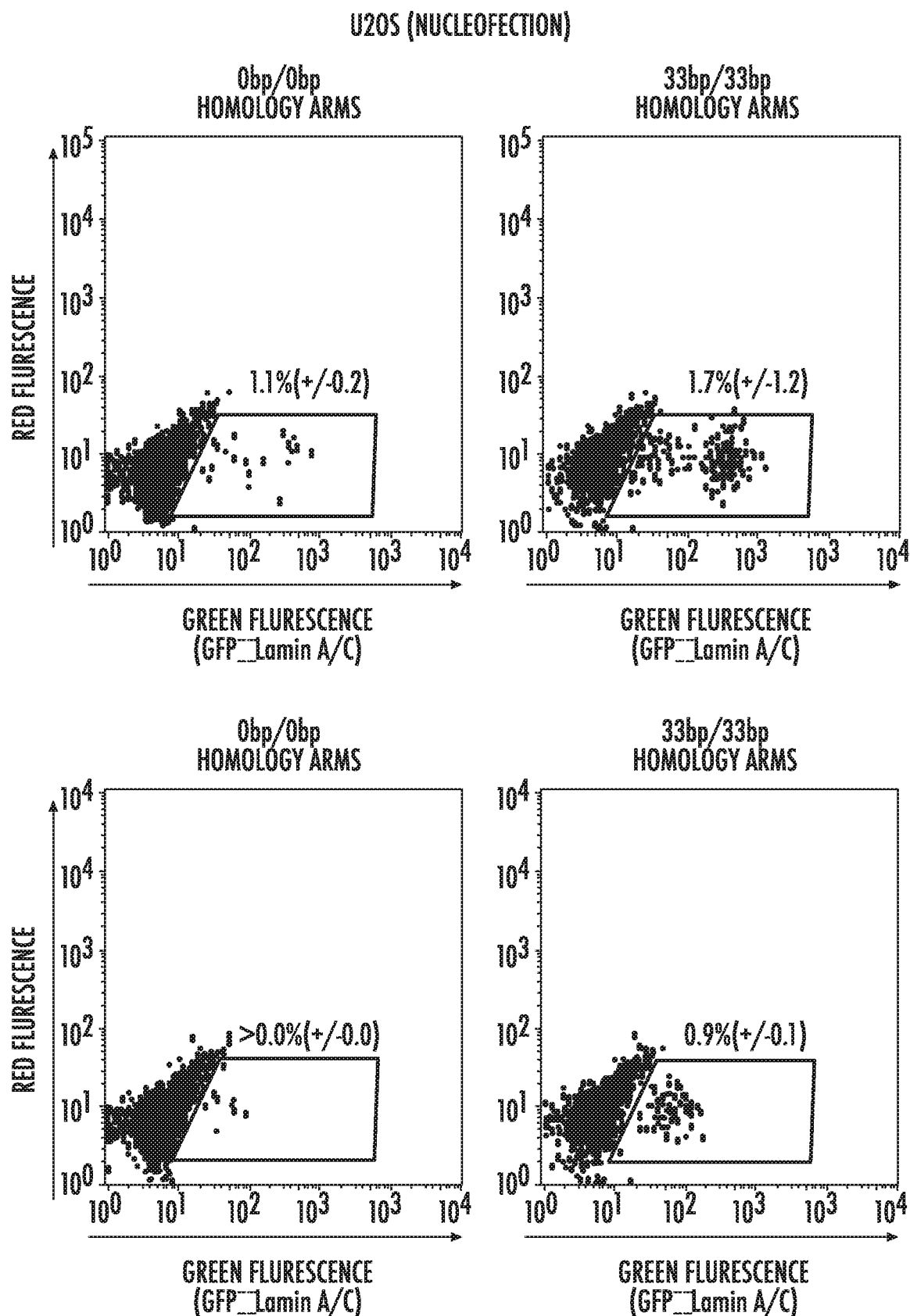
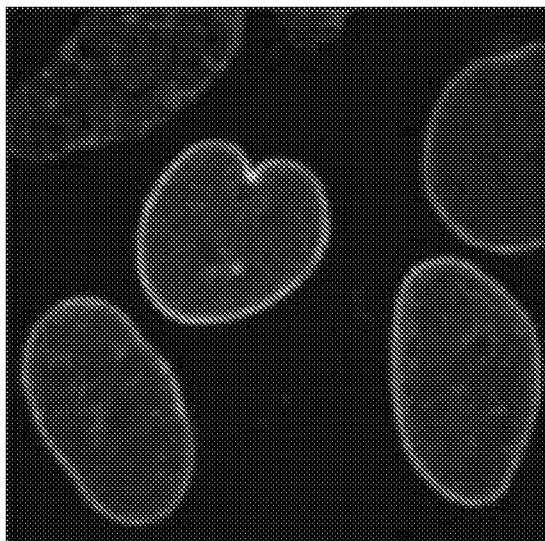


FIG. 21A

35/37

Lamin A/C
(NUCLEAR MEMBRANCE)



Rab11
(ENDOSOMES)

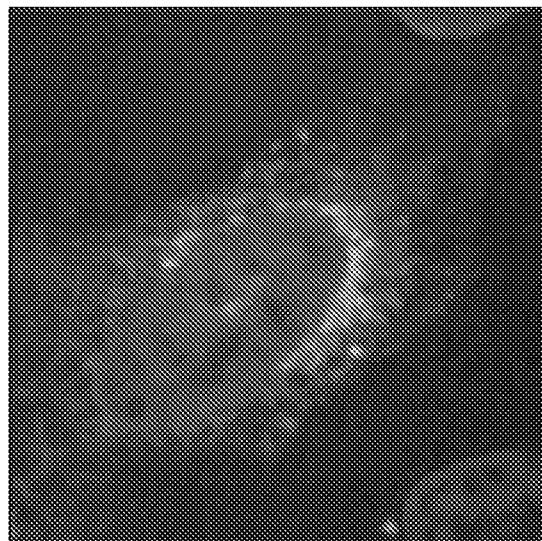


FIG. 21A CONTINUED

36/37

DLD1 (NUCLEOFECTION)

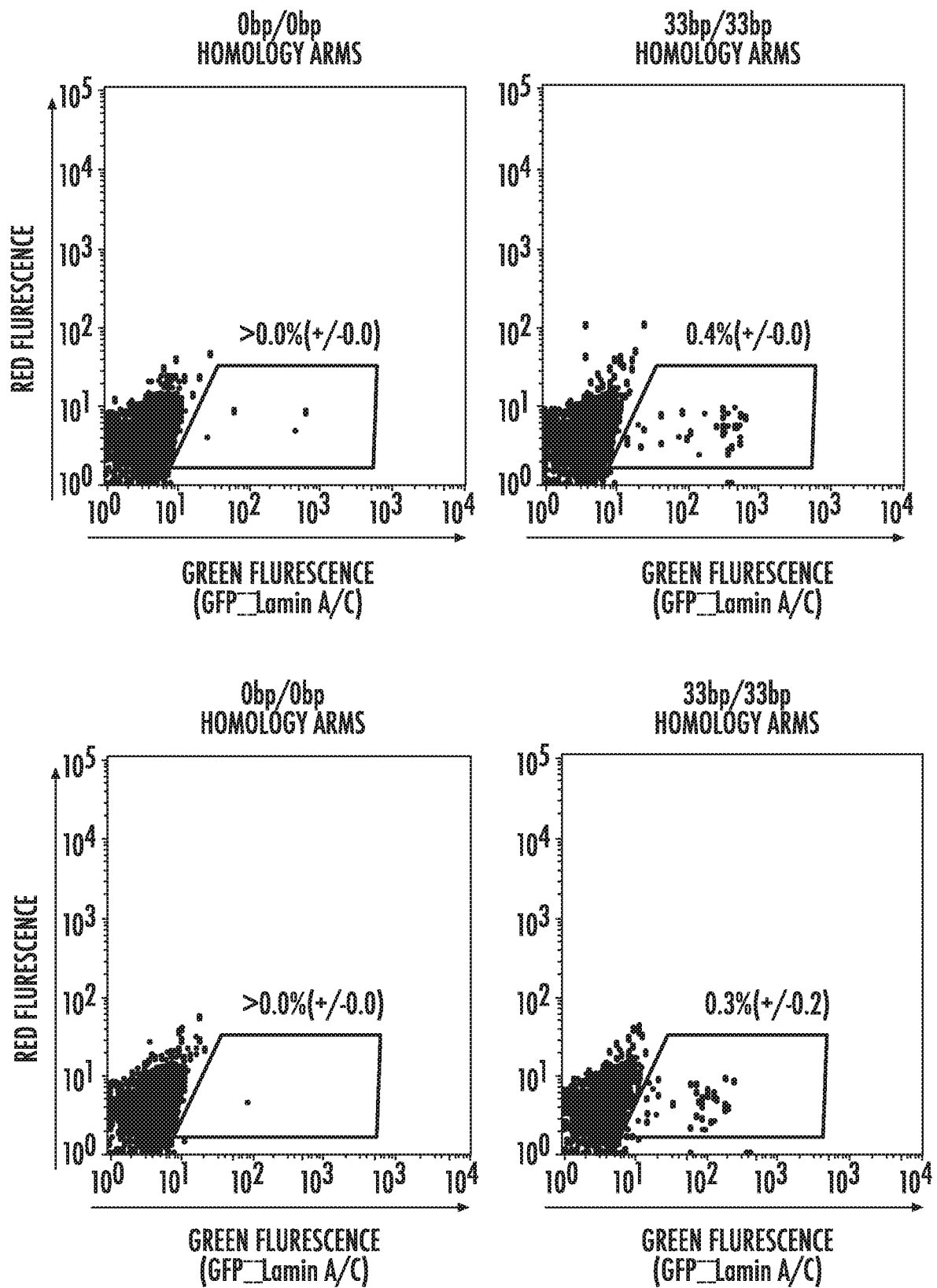


FIG. 21B

37/37

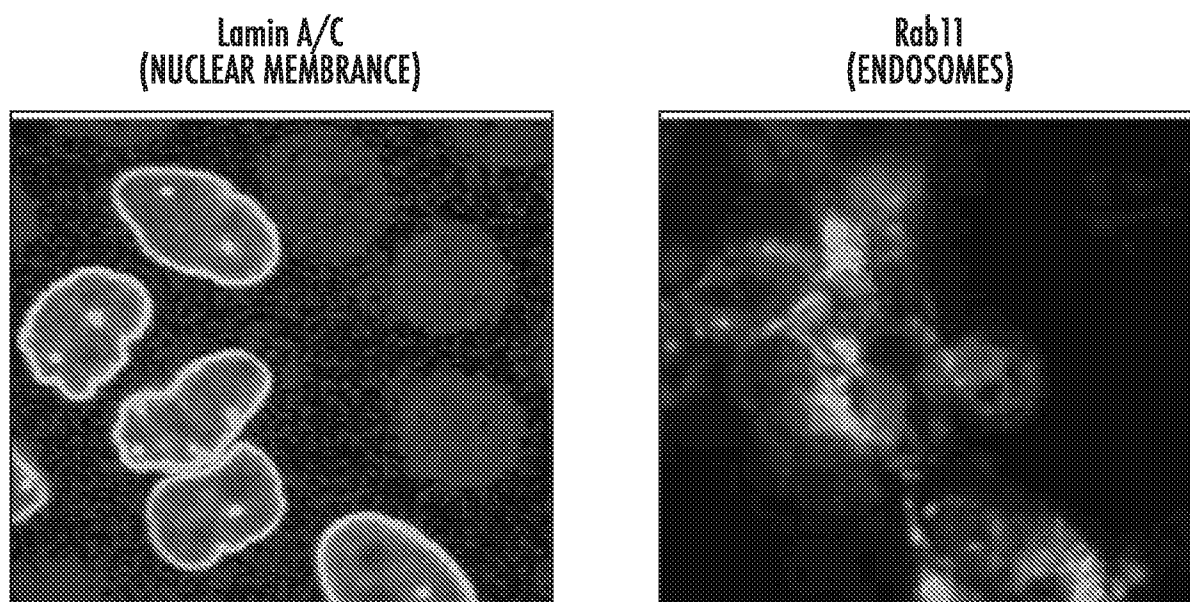


FIG. 21B CONTINUED

International application No.

PCT/US 2018/061770

A61K 48/00 (2006.01)

C12N 15/11 (2006.01)

C12N 15/90 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

Minimum documentation searched (classification system followed by classification symbols)

A61K 48/00, C12N 15/90, 15/09, 15/11

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)

PatSearch (RUPTO internal), USPTO, PAJ, Esp@cenet, Information Retrieval System of FIPS

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2016/033246 A1 (CARIBOU BIOSCIENCES, INC) 03.03.2016, paragraphs [0015], [0024], [0063], [0066]-[0070], [00106]-[00108], [00135], [00136], [00157], example 5	1-18
X	SONG FEI et al. Optimizing the DNA donor template for homology-directed repair of double-strand breaks. Molecular Therapy-Nucleic Acids, June 2017, 7: 53-60, abstract, fig.1, pp.53, 54, 57	1, 6, 10-12, 17, 18
X	WO 2017/019867 A1 (DANISCO US INC) 02.02.2017, paragraphs [0092], [0107], [0139], [0144], example, claims	1-4, 6-9, 17, 18
A	US 2017/058298 A1 (AGILENT TECHNOLOGIES, INC.) 02.03.2017	1-18

☐ Further documents are listed in the continuation of Box C.

 See patent family annex.

* Special categories of cited documents:

“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
-----	---

“A” document defining the general state of the art which is not considered to be of particular relevance

“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

“E” earlier document 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)

“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

“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

“&” document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

11 February 2019 (11.02.2019)

07 March 2019 (07.03.2019)

Name and mailing address of the ISA/RU:

Authorized officer

Federal Institute of Industrial Property,
Berezhkovskaya nab., 30-1, Moscow, G-59,
GSP-3, Russia. 125993

T. Babakova

Facsimile No: (8-495) 531-63-18, (8-499) 243-33-37

Telephone No. (495) 531-64-81