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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 December 2013 (19.12.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/188522 A2

(51) International Patent Classification:
A01K 67/027 (2006.01)

(21) International Application Number:
PCT/US2013/045382

(22) International Filing Date:
12 June 2013 (12.06.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/658,670 12 June 2012 (12.06.2012) US

(71) Applicant *(for all designated States except AL, AT, BE, BG, CH, CN, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IN, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR)*: **GENENTECH, INC.** [US/US]; 1 DNA Way, South San Francisco, California 94080 (US).

(71) Applicant *(for AL, AT, BE, BG, CH, CN, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IN, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR only)*: **F. HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacherstrasse 124, CH-4070 Basel (CH).

(72) Inventors: **WARMING, Soren**; c/o Genentech, Inc., 1 DNA Way, South San Francisco, California 94080 (US). **ANDERSON, Keith R.**; c/o Genentech, Inc., 1 DNA Way, South San Francisco, California 94080 (US).

(74) Agents: **VOM WEGE, Julia** et al.; Genentech, Inc., 1 DNA Way, MS-49, South San Francisco, California 94080 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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, 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:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

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

WO 2013/188522 A2

(54) Title: METHODS AND COMPOSITIONS FOR GENERATING CONDITIONAL KNOCK-OUT ALLELES

(57) Abstract: The invention provides methods and compositions for generating conditional knock-out alleles using sequence-specific nucleases.

METHODS AND COMPOSITIONS FOR GENERATING CONDITIONAL KNOCK-OUT ALLELES

CROSS REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application No. 61/658,670, filed June 12, 2012, the disclosure of which is incorporated herein by reference in its entirety.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on June 12, 2013, is named P4905R1WO_PCTSequenceListing.txt and is 49,214 bytes in size.

FIELD OF THE INVENTION

The present invention concerns novel methods of producing genetically engineered conditional knock out alleles.

BACKGROUND

Selective inhibition or enhancement of individual gene expression has greatly assisted the study of gene function *in vitro* and *in vivo*. Gene targeting of murine embryonic stem (ES) cells using homologous recombination is a well-established method for manipulating the murine genome and has allowed creation of null or "knock-out" mice with respect to a gene under investigation. More recently, conditional or inducible knock-out technology has advanced the study of genes that, when deleted systemically, result in embryonic or perinatal lethality (*e.g.*, Lakso, M. *et al., Proc. Natl. Acad. Sci. USA* 89:6232-36 (1992); Jacks, T. *et al.*, Nature 359:295-300 (1992)). Conditional knock-out mice can also be used to study the effects of selectively deleting a gene in a particular tissue, while leaving its function intact in other tissues. However, conventional methods for creating conditional knock-out animals are laborious, inefficient and require the availability of embryonic stem cells.

Engineered sequence-specific nucleases have been used to create knock-out alleles. Examples of such sequence-specific endonucleases include zinc finger nucleases (ZFNs), which are composed of sequence-specific DNA binding domains fused to an endonuclease effector domain (Porteus, M.H. and Caroll, D., *Nat. Biotechnol.* 23, 967-973 (2005)). Another

example of sequence-specific nucleases are transcription activator-like effector nucleases (TALENs), which are composed of a nuclease domain fused to TAL effector proteins (Miller, J.C. *et al., Nat. Biotechnol.* 29, 143-148 (2011); Cermak, T. *et al., Nucleic Acid Res.* 39, e82 (2011)). Sequence-specific endonucleases are modular in nature, and DNA binding specificity is obtained by arranging one or more modules. For example, zinc finger domains in ZFNs each recognize three base pairs (Bibikova, M. *et al.*, *Mol. Cell. Biol.* 21, 289–297 (2001)), whereas individual TAL domains in TALENs each recognize one base-pair via a unique code (Boch, J. *et al., Science* 326, 1509–1512 (2009).) Another example of sequence-specific nucleases includes RNA-guided DNA nucleases, e.g., the CRISPR/Cas system.

ZFNs, TALENs and most recently CRISPR/Cas mediated gene editing have been used to efficiently and directly generate gene knock-out alleles (Geurts, A. M. *et al., Science* 325, 433 (2009); Mashimo, T. *et al.*, *PLoS ONE* 5, e8870 (2010); Carbery, I. D. *et al., Genetics* 186, 451–459 (2010); Tesson, L., *et al., Nat. Biotech.* 29, 695-696 (2011)). The knock-out alleles are thought to be produced by an error-prone non-homologous end-joining (NHEJ) of the endonuclease-mediated double-strand break (DSB).

Recently, ZFNs were successfully used for targeted insertion (knock-in) of a reporter gene by homologous recombination of the targeted chromosomal locus with a donor DNA in both mouse and rat (Meyer, M., *et al.*, *Proc. Natl. Acad. Sci. USA* 107, 15022–15026 (2010); Cui, X. *et al.*, *Nat. Biotechnol.* 29(1), 64-67 (2010)). The sequence-specific insertion of the donor sequence has been proposed to occur via a synthesis-dependent strand annealing (SDSA) model of double-strand break repair by homologous recombination between the donor and the locus at which the double-strand break occurred (Moehle, E. A. *et al., Proc Natl Acad Sci USA* 104, 3055–3060 (2007)). According to this model, after endonuclease-mediated double-strand break and strand resection, the single-stranded chromosome ends anneal to the homology regions present on the donor DNA followed by synthesis using the donor insert as template.

Despite these advances, a need in the art remains for new methods to create conditional knock-out alleles and to expand this technology to other species. The present invention fulfills this need and provides other benefits.

SUMMARY

The present invention relates to novel methods and compositions for generating conditional knock-out alleles. Specifically, the present invention relates to using specific donor constructs together with sequence-specific nucleases to generate conditional knock-out alleles.

In one aspect, a method of generating a conditional knock-out allele in a cell comprising a target gene is provided. The method comprises the steps of:

1. introducing into the cell a donor construct, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation; and
2. introducing into the cell a sequence-specific nuclease that cleaves a sequence within the target gene, thereby producing a conditional knock-out allele in the cell.

In certain embodiments, the sequence-specific nuclease is a zinc finger nuclease (ZFN), a ZFN dimer, a transcription activator-like effector nuclease (TALEN), or a RNA-guided DNA endonuclease. In certain embodiments, the sequence-specific nuclease cleaves the target gene only once. In certain embodiments, the sequence-specific nuclease is introduced into the cell as a protein, mRNA, or cDNA.

In certain embodiments, the recombinase recognition site is a *loxP* site, a *rox* site or an *frt* site. In certain embodiments, the donor sequence comprises one, two, three, four, five, six, seven, eight, nine, ten, eleven, or twelve neutral mutations. In certain embodiments, the homology between the donor sequence and the target sequence is 51-99%. In certain embodiments the homology between the donor sequence and the target sequence is 78%. In certain embodiments, the donor construct comprises the sequence shown in FIG. 4A or FIG. 4B. In certain embodiments, the 5' homology region comprises at least 1.1 kb and wherein the 3' homology region comprises at least 1 kb. In certain embodiments, the target gene is *Lrp5*.

In a further embodiment, the cell is a mammalian cell. In certain embodiments, the mammalian cell a mouse, rat, rabbit, hamster, cat, dog, sheep, horse, cow, monkey or human cell. In certain embodiments, the cell is from a non-human animal. In certain embodiments, the cell is a somatic cell, a zygote or a pluripotent stem cell.

In a further aspect, a method of generating a conditional knock-out animal is provided, the method comprising the steps of:

1. introducing a donor construct into a cell comprising a target gene, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation;
2. introducing a sequence-specific nuclease into the cell, wherein the nuclease cleaves the target gene; and

3. introducing the cell into a carrier animal to produce the conditional knock-out animal from the cell.

In some embodiments, the animal is a mouse, rat, rabbit, hamster, guinea pig, dog, sheep, pig, horse, cow or monkey. In certain embodiments, the cell is from a non-human animal. In some embodiments, the cell is a zygote or a pluripotent stem cell.

In a further aspect, a method of generating a knock-out animal is provided, the method comprising the steps of:

1. introducing a donor construct into a zygote comprising a target gene, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation;
2. introducing a sequence-specific nuclease into the zygote, wherein the nuclease cleaves the target gene;
3. introducing the zygote into a carrier animal to produce a conditional knock-out animal from the zygote; and
4. breeding the conditional knock-out animal with a transgenic animal having a transgene encoding a recombinase that catalyzes recombination at the 5' and 3' recombinase recognition sites, thereby producing the knock-out animal.

In certain embodiments, the recombinase recognition site is a *loxP* site and the recombinase is Cre recombinase. In certain embodiments, the recombinase recognition site is an *frt* site and the recombinase is flippase. In certain embodiments, the recombinase recognition site is a *rox* site and the recombinase is Dre recombinase. In certain embodiments, the transgene encoding the recombinase is under the control of a tissue-specific promoter.

In a further aspect of the invention, a composition for generating a conditional knock-out allele of a target gene is provided, comprising:

1. a donor construct comprising a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation; and
2. a sequence-specific nuclease that recognizes the target gene.

In certain embodiments, the sequence-specific nuclease is a ZFN, a ZFN dimer, a ZFNickase, a TALEN, or a RNA-guided DNA endonuclease. In certain embodiments, the recombinase recognition site is a *loxP* site, an *frt* site or a *rox* site.

In a further aspect of the invention, a donor construct comprising the sequence shown in FIG. 4A (SEQ ID NO: 30), FIG. 4B (SEQ ID NO: 31), or FIG. 14C (SEQ ID NOS: 44-46) is provided.

In a further aspect of the invention, a cell comprising the donor construct comprising the sequence shown in FIG. 4A (SEQ ID NO: 30), FIG. 4B, or FIG. 14C (SEQ ID NOS: 44-46) is provided. In certain embodiment, the cell is a mammalian cell. In certain embodiments, the mammalian cell a mouse, rat, rabbit, hamster, cat, dog, sheep, horse, cow, monkey or human cell. In certain embodiments, the cell is from a non-human animal. In certain embodiments, the cell is a somatic cell, a zygote or a pluripotent stem cell.

In a further aspect of the invention, a non-human conditional knock-out animal prepared according to the method described herein is provided.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 shows the distribution of ZFN-mediated mutant *Lrp5* alleles in live-born mice. The size of deletions and insertions are indicated in base pairs on the x-axis. Compound KO: animals with two independent mutant alleles of the same gene and no detectable wildtype allele of the gene; Multiple allele: chimeric animals carrying more than two alleles; SKG->WTD: deletion of TCCAAGGGT (ZFN cut site is underlined).

FIGS. 2A-E show vascular phenotypes of 2-month-old mice with compound in frame and out-of-frame deletions in *Lrp5*. 542: chimeric functional heterozygous mouse (control) that carried an allele with a 3bp in-frame deletion that appeared to be silent and an allele with a 1bp out-of-frame deletion; 495: mouse that carried a 4bp out-of-frame deletion allele and a 1bp out-of-frame deletion allele; 519: mouse that carried a 29bp out-of-frame deletion allele and a 17bp out-of-frame deletion allele; 555: functional heterozygous mouse that carried a 3bp in-frame deletion allele and a 1bp out-of-frame deletion allele and is a functional heterozygote; FA: fluorescent angiography; IB4: isolectin B4; NFL: nerve fiber layer; IPL: inner plexiform layer; OPL: outer plexiform layer.

FIGS. 3A-B show conditional knock-out alleles obtained from co-microinjection or co-electroporation of *Lrp5* exon2 ZFN and donor plasmid. FIG. 3A depicts a schematic of double-strand break repair by synthesis-dependent strand annealing. Arrow heads represent recombinase recognition sites; large arrow in Step 1 represents the target sequence; large arrow with asterisks represents the donor sequence; asterisks represent neutral mutations; half arrows indicate primer positions. FIG. 3B depicts the results of a polymerase chain reaction (PCR)

analysis of DNA isolated from tail samples of pups (left panel) or ES cells (right panel). The respective primer pairs used for the analysis are indicated to the left (primer positions are as depicted in FIG. 3A).

FIGS. 4A-C show the donor sequences (SEQ ID NOS: 30-32, respectively, in order of appearance) that were used in plasmids in the correct orientation and with the sequences flanking the inserts.

FIGS. 5A-B show a sequence alignment of the three *Lrp5* CKO DNA donors from 5' loxP to 3' loxP sites (SEQ ID NOS 33-35, respectively, in order of appearance). Uppercase bold letters indicate loxP sites; lowercase letters indicate intron sequences; uppercase letters indicate exon 2 (wild type or modified) sequences; dashed line boxes indicate ZFN binding sites; solid line boxes indicate silent mutations; underlined letters indicate the sequence at which the wild type exon 2 is cleaved by the ZFN.

FIGS. 6A-E show normal retinal phenotypes of mice carrying a codon-modified *Lrp5* conditional knock-out allele. FIGS. 6A-D depict confocal projections of retinal whole mounts stained with isolectin B4 (scale bars: 50 μm). FIG. 6E depicts retinal cross sections of the opposite eyes to those depicted in FIGS. 6A-D, stained with IB4, MECA32, and DAPI. Arrows point to example staining as indicated. +/+: wild type control; KO/KO: *Lrp5* homozygous knock out; KO/+: *Lrp5* heterozygous knock out; CKO/KO: *Lrp5* conditional knock out/*Lrp5* knock-out compound heterozygous; IB4: isolectin B4; NFL: nerve fiber layer; IPL: inner plexiform layer; OPL: outer plexiform layer.

FIGS. 7A-D show a graphic representation of possible mechanism that produced each of the observed donor-derived *Lrp5* alleles. Primers that bind to the resulting alleles are indicated. Neutral mutations are indicated by asterisks.

FIG. 8 depict the results of a SURVEYOR Assay following introduction of either zinc finger pairs (pZFN1+pZFN2) or Cas9 (+pRK5-hCas9) together with a guide RNA targeting *Lrp5* exon 2 (p_gRNA T2, p_gRNA T5 or p_gRNA T7) or a control plasmid (PMAXGFP) into NIH/3T3 cells or Hepa1-6 cells.

FIGS. 9A-B illustrate a summary of gRNA/Cas9 mutation rates (FIG. 9A) and deletion sizes (FIG. 9B) at the *Lrp5* exon 2 genomic locus in Hepa1-6 murine hepatoma cells. The cells received a gRNA targeting *Lrp5* together with either mRNA (Cas9 mRNA + gRNA T2, solid bars) or a plasmid (Cas9 plasmid + gRNA T2, clear bars), or two plasmids encoding zink finger pairs targeting exon2 of *Lrp5* (ZFN plasmid, grey bars).

FIG. 10 depicts the result of PCR analysis using a forward primer specific for the COexon2 sequence and a reverse primer outside of the homology arm in the genomic locus to identify integration of the donor exon in the *Lrp5* locus. Murine Hepa1-6 cells received plasmid (pRK5-hCas9) or mRNA (hCas9 mRNA) encoding Cas9 together with either the guide RNA alone (p_gRNA T2), the guide RNA and the donor plasmid (p_gRNA T2+ p_donor1) or a control plasmid (PMAXGFP). Some cells received the donor together with the *Lrp5* zink finger pair (pZFN1+pZFN2+p_donor1).

FIG. 11 depicts the result of PCR analysis using primers that detect 5' (top, primers P9 and P10) and 3' (bottom, primers P11 and P12) loxP site integration in the *Lrp5* genomic locus. The treatment groups are as described in FIG. 10. DNA from a heterozygous *Lrp5* conditional knock out (mouse CKO/wt) was used as positive control.

FIG. 12 depicts the results of a SURVEYOR Assay following introduction of Cas9 (p_hCas9) together with a guide RNA and respective donor construct targeting *Lrp5* (*Lrp5* exon 2; p_gRNA T7 + p_Lrp5_donor1), *Usp10* (*Usp10* exon3; p_gRNA T1 + p_Usp10_donor1) or *Notch3* (*Notch3* exon3; p_gRNA T1 + p_Notch3_donor1) into Hepa1-6 cells.

FIG. 13 depicts the result of PCR analysis using primers that detect 5' loxP site integration in the *Nnmt* exon2 genomic locus (left panel, primers P26 and P27) or 3' loxP site integration in the *Notch3* exon3 genomic locus (right panel, primers P25 and P28) following Cas9/gRNA and donor administration.

FIG. 14A-D show the sequences (SEQ ID NOS: 36-46, respectively, in order of appearance) for Cas9/CRISPR targeting of mouse *Lrp5*, *Usp10*, *Nnmt*, and *Notch3* genomic loci. Sequences for guide RNA (gRNA) sequences specific for *Lrp5*, *Usp10*, *Nnmt*, and *Notch3* and donor plasmid sequences for *Usp10*, *Nnmt*, and *Notch3* are depicted. In addition, Cas9 cDNA sequence for mammalian expression and *in vitro* transcription (mRNA) are shown.

DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

I. DEFINITIONS

For purposes of interpreting this specification, the following definitions will apply and whenever appropriate, a term used in the singular will also include the plural and vice versa. In the event that any definition set forth below conflicts with any document incorporated herein by reference, the definition set forth below shall control.

The term "donor construct," as used herein, refers, unless specifically indicated otherwise, to a polynucleotide that comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region. The donor construct can further include additional sequences, such as sequences that support propagation of the donor construct or selection of cells harboring the construct.

The term "donor sequence," as used herein, refers, unless specifically indicated otherwise, to a nucleic acid having a sequence that comprises a target sequence having at least one neutral mutation compared to a portion of the sequence of the target gene. As such, the donor sequence comprises a nucleic acid that encodes a polypeptide that is functionally substantially similar to or indistinguishable from that encoded by the portion of the target gene. Consequently, the donor sequence can replace the cognate portion of the target gene at its position in the target gene without substantially changing the functional properties of the protein encoded by the target gene. The donor sequence can comprise certain non-coding sequences, such as intronic or regulatory sequences.

The term "homology region," as used herein, refers, unless specifically indicated otherwise, to a nucleic acid in the donor construct that is homologous to a nucleic acid flanking a target sequence.

The term "recombinase recognition site," as used herein, refers, unless specifically indicated otherwise, to a nucleic acid in a donor construct having a sequence that is recognized by a recombinase.

The term "recombinase," as used herein, refers, unless specifically indicated otherwise, to an enzyme that recognizes specific polynucleotide sequences (recombinase recognition sites) that flank an intervening polynucleotide and catalyzes a reciprocal strand exchange, resulting in inversion or excision of the intervening polynucleotide.

The term "target gene," as used herein, refers, unless specifically indicated otherwise, to a nucleic acid encoding a polypeptide within a cell.

The term "target sequence," as used herein, refers, unless specifically indicated otherwise, to a portion of the target gene, *e.g.*, one or more of the exon sequences of the target gene, intronic sequences, or regulatory sequences of the target gene, or a combination of exon and intron sequences, intron and regulatory sequences, exon and regulatory sequences, or exon, intron, and regulatory sequences of the target gene.

The term "sequence-specific endonuclease" or "sequence-specific nuclease," as used herein, refers, unless specifically indicated otherwise, to a protein that recognizes and binds to

a polynucleotide, *e.g.,* a target gene, at a specific nucleotide sequence and catalyzes a single- or double-strand break in the polynucleotide.

The term "RNA-guided DNA nuclease" or "RNA-guided DNA nuclease" or "RNA-guided endonuclease," as used herein, refers, unless specifically indicated otherwise, to a protein that recognizes and binds to a guide RNA and a polynucleotide, *e.g.,* a target gene, at a specific nucleotide sequence and catalyzes a single- or double-strand break in the polynucleotide.

The term "conditional knock-out allele," as used herein, refers, unless specifically indicated otherwise, to an allele comprising a polynucleotide sequence that is flanked by recombinase recognition sites but produces a phenotype that is indistinguishable from that produced by the cognate wild type allele.

The term "neutral mutation," as used herein, refers, unless specifically indicated otherwise, to a mutation in a donor sequence that reduces overall homology between the donor sequence and the target sequence but leaves the donor sequence capable of encoding a functional polypeptide. Examples of neutral mutations include silent mutations, *i.e.*, mutations that alter the nucleotide sequence but not the encoded polypeptide sequence. Examples of neutral mutations also include conservative mutations, such as point mutations (*e.g.,* substitutions), insertions and deletions, *i.e.*, mutations that alter the nucleotide sequence and the encoded polypeptide sequence but that do not substantially alter the function of the resulting polypeptide. Examples of conservative substitution mutations are shown in Table 8. Neutral mutations can also include combinations of silent mutations, combinations of conservative mutations, or combinations of silent and conservative mutations.

The term "animal," as used herein, refers, unless specifically indicated otherwise, to any non-human animal, including, but not limited to, domesticated animals (*e.g.,* cows, sheep, cats, dogs, and horses), primates (*e.g.*, non-human primates such as monkeys), rabbits, fish, rodents (*e.g.*, mice, rats, hamsters, guinea pigs), and non-vertebrates (*e.g.*, *Drosophila melanogaster* and *Caenorhabditis elegans*).

An "isolated" nucleic acid refers, unless specifically indicated otherwise, to a nucleic acid molecule that has been separated from a component of its natural environment. An isolated nucleic acid includes a nucleic acid molecule contained in cells that ordinarily contain the nucleic acid molecule, but the nucleic acid molecule is present extrachromosomally or at a chromosomal location that is different from its natural chromosomal location.

"Isolated nucleic acid encoding a protein" refers, unless specifically indicated otherwise, to one or more nucleic acid molecules encoding a protein (or fragments thereof), including such nucleic acid molecule(s) in a single vector or separate vectors, and such nucleic acid molecule(s) present at one or more locations in a host cell.

The term "sequence homology," as used herein with respect to the donor and target gene polynucleotide sequences, is defined as the percentage of nucleotide residues in a donor sequence that are identical to the nucleotide residues in the target gene sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleotide sequence homology can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ClustalW2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for aligning sequences, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

II. EMBODIMENTS OF THE INVENTION

The invention relates, in part, to the recognition and solution of technical challenges associated with creating conditional knock-out alleles using sequence-specific endonucleases in combination with a recombinase recognition sequence-flanked donor sequence. This process relies on targeting specific sequences of nucleic acid molecules, such as chromosomes, with endonucleases that recognize and bind to such sequences and induce a double-strand break in the nucleic acid molecule. The double strand break is repaired either by an error-prone non-homologous end-joining or by homologous recombination. If a template for homologous recombination is provided *in trans*, the double-strand break can be repaired using the provided template. The initial double strand break increases the frequency of targeting by several orders of magnitude, compared to conventional homologous recombination-based gene targeting. In principle, this method could be used to insert any sequence at the site of repair so long as it is flanked by appropriate regions homologous to the sequences near the double-strand break. However, this approach is associated with certain challenges when applied to creating conditional knock-out alleles. Conditional knock-out alleles typically include certain recombinase recognition sequences, such as *loxP* sites, that flank the gene or portions of the gene but leaves its function intact, such that the conditional knock-out allele produces functional polypeptides substantially similar to the unmodified allele but that can be rendered

non-functional at a certain time or within certain tissues by the presence of the recombinase recognizing the recognition sequences.

A first challenge associated with the approach described above to create conditional knock-out alleles resides in the fact that, following the double-strand break catalyzed by the sequence-specific endonuclease, undesirable recombination can occur between the donor exon and the chromosomal (target) exon, instead of the homology regions outside of the recombinase recognition sequence-flanked donor, because of their sequence identity with respect to each other. This will result in alleles that lack one or both recombinase recognition sequences. A second challenge resides in the fact that the sequence-specific endonuclease can recognize and cleave not only the target gene but also the donor exon before it can serve as a template for repair. The methods and compositions described herein provide a solution to these challenges.

A. Exemplary Methods

In various aspects of the invention, methods of generating a conditional knock-out allele in a cell comprising a target gene are provided. The method comprises the steps of introducing into the cell having a target gene a donor construct and a sequence-specific nuclease that cleaves a sequence within the target gene but does not inhibit function of the donor construct, thereby producing a conditional knock-out allele in the cell. These and further aspects of the invention are described below.

In a particular aspect of the invention, a conditional knock-out allele is produced in a cell comprising a target gene by introducing into the cell a donor construct that comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region. The donor sequence comprises the sequence of a target sequence having at least one neutral mutation. In certain embodiments, the donor sequence and the target sequence are identical except for the at least one neutral mutation. A neutral mutation means any mutation in the nucleotide sequence of the donor sequence that reduces homology between the donor sequence and the target sequence but leaves the coding potential of the donor for a functional polypeptide intact. The neutral mutation decreases the number of undesired homologous recombination events, compared to a wild type sequence, between the donor sequence and the target sequence that do not result in a conditional knock-out allele (FIG. 7B, C, D). In some embodiments, the neutral mutation also abrogates binding of the sequence-specific nuclease to the donor sequence.

Examples of neutral mutations include silent mutations, *i.e.*, mutations that alter the nucleotide sequence but not the encoded polypeptide sequence. Neutral mutations also include conservative mutations, *i.e.*, mutations that alter the nucleotide sequence and the encoded polypeptide sequence but that do not substantially alter the function of the resulting polypeptide. This is the case, for example, when one amino acid is substituted with another amino acid that has similar properties (size, charge, etc.). For example, Amino acids may be grouped according to common side-chain properties:

(1) hydrophobic: Norleucine, Met, Ala, Val, Leu, Ile;

(2) neutral hydrophilic: Cys, Ser, Thr, Asn, Gln;

(3) acidic: Asp, Glu;

(4) basic: His, Lys, Arg;

(5) residues that influence chain orientation: Gly, Pro;

(6) aromatic: Trp, Tyr, Phe.

Examples of conservative mutations are shown in Table 8. In certain embodiments, the donor sequence comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or 50 silent mutations. In certain embodiments, the homology between the donor sequence and the target sequence is 99%, 98%, 95%, 90%, 85%, 80%, 78%, 75%, 70%, 65%, 60%, 55%, or 50%. In certain embodiments, the sequence homology between donor and target sequence is less than 50%. Any number of neutral mutations can be introduced that reduce or inhibit the number of homologous recombination events between the donor sequence and the target sequence (FIG. 7B-D), rather than between the homologous regions and their cognate sequence on the targeted molecule, but maintain the ability of the donor sequence to encode a functional polypeptide. In certain embodiments, the donor comprises the sequence shown in FIG. 4A (SEQ ID NO: 30), FIG. 4B (SEQ ID NO: 31), or FIG. 14C (SEQ ID NOS: 44-46). In certain embodiments, at least one neutral mutation abrogates binding of the sequence-specific nuclease to the donor sequence. In certain embodiments, several neutral mutations are spaced along the length of the donor sequence to reduce the number of consecutive unmodified base pairs to less than 20-100 base pairs at any position in the donor sequence.

Because the mutations within the donor sequence are neutral, the donor sequence encodes a polypeptide that is functionally substantially similar to or indistinguishable from that encoded by the target sequence. The functionality of a peptide or protein can be assessed by methods well-known in the art, such as functional assays, enzymatic assays, and biochemical assays. The donor sequence can replace the target sequence at its position in the target gene

without substantially altering the functional properties of the polypeptide encoded by the target gene. However, once integrated in the target gene, subsequent removal of the donor sequence from the target gene can result in altered, reduced or loss of function of the polypeptide encoded by the target gene.

Within the donor construct, the donor sequence is flanked 5' and 3' by recombinase recognition sites. These recombinase recognition sites are nucleic acid sequences within the donor construct that are recognized by a recombinase that subsequently catalyzes recombination at the recombination recognition sites. Sequence-specific recombination is well-known in the art and includes recombinase-mediated sequence-specific cleavage and ligation of a polynucleotide flanked by the recombinase recognition sites. Examples of recombinase recognition sites include *loxP* (locus of X-over P1) sites (Hoess *et al.*, *Proc. Natl. Acad. Sci. USA* 79:3398-3401 (1982)), *frt* sites (McLeod, M., Craft, S. & Broach, J. R., *Molecular and Cellular Biology* 6, 3357–3367 (1986)) and *rox* sites (Sauer, B. and McDermott, J., *Nucleic Acids Res* 32, 6086–6095 (2004).).

The 5' homology region is located 5' or "upstream" of the 5' recombinase recognition site and is homologous to a nucleic acid flanking the target sequence in its nucleotide context. Similarly, the 3' homology region is located 3' or "downstream" of the 3' recombinase recognition site and is homologous to a nucleic acid flanking the target sequence. In one embodiment, the homology regions are more than 30bp, preferably several kb in length. For example, the homology regions can be 50bp, 100bp, 200bp, 300bp, 500bp, 800bp, 1kb, 1.1kb, 1.5kb, 2kb and 5kb in length. In certain embodiments, the 5' homology region comprises 1.1 kb and the 3' homology region comprises 1 kb. The homology regions can be homologous to regions of the target gene and also, or instead, be homologous to regions upstream or downstream of the target gene. In one embodiment, the homology regions are homologous to chromosomal regions immediately adjacent to the target sequence. For example, in the case of the 5' homology region, the homology region is homologous to a sequence having its most 3' nucleotide immediately adjacent to the first (most 5') nucleotide of the target sequence. In one embodiment, homology regions are homologous to chromosomal regions that are not immediately adjacent to the target sequence on the chromosome. In some embodiments, the 5' and 3' homologous regions are each 95-100% homologous to the cognate nucleic acid sequences flanking the target sequence.

To summarize the above-described component arrangement, the donor construct comprises, in order from 5' to 3,' a 5' homology region, a 5' recombinase recognition site, a

donor sequence, a 3' recombinase recognition site, and a 3' homology region. The donor construct can further include certain sequences that provide structural or functional support, such as sequences of a plasmid or other vector that supports propagation of the donor construct (*e.g.*, pUC19 vector). The donor construct can, optionally, also include certain selectable markers or reporters, some of which may be flanked by recombinase recognition sites for subsequent activation, inactivation, or deletion. The recombinase recognition sites flanking the optional marker or reporter can be the same or different from the recombinase recognition sites flanking the donor sequence. In certain embodiments, a single type of donor construct is used to produce the conditional knock-out allele.

Concomitant with, or sequential to, introduction of the donor construct, a sequence-specific nuclease is introduced into the cell. The sequence-specific nuclease recognizes and binds to a specific sequence within the target gene and introduces a double-strand break in the target gene. As described above, the donor sequence is modified by at least one neutral mutation to reduce homologous recombination events that do not result in conditional knock-out alleles. In certain embodiments, the sequence-specific nuclease cleaves the target gene only once, *i.e.*, a single double-strand break is introduced in the target gene during the methods described herein.

Examples of sequence-specific nucleases include zinc finger nucleases (ZFNs). ZFNs are recombinant proteins composed of DNA-binding zinc finger protein domains and effector nuclease domains. Zinc finger protein domains are ubiquitous protein domains, *e.g.*, associated with transcription factors, that recognize and bind to specific DNA sequences. One of the "finger" domains can be composed of about thirty amino acids that include invariant histidine residues in complex with zinc. While over 10,000 zinc finger sequences have been identified thus far, the repertoire of zinc finger proteins has been further expanded by targeted amino acid substitutions in the zinc finger domains to create new zinc finger proteins designed to recognize a specific nucleotide sequence of interest. For example, phage display libraries have been used to screen zinc finger combinatorial libraries for desired sequence specificity (Rebar *et al.*, *Science* 263:671-673 (1994); Jameson *et al.*, *Biochemistry* 33:5689-5695 (1994); Choo et al., *PNAS* 91:11163-11167 (1994), each of which is incorporated herein as if set forth in its entirety). Zinc finger proteins with the desired sequence specificity can then be linked to an effector nuclease domain, *e.g.*, as described in 6,824,978, such as FokI, described in PCT Application Publication Nos. WO1995/09233 and WO1994018313, each of which is incorporated herein by reference as if set forth in its entirety.

Another example of sequence-specific nucleases includes transcription activator-like effector endonucleases (TALEN), which comprise a TAL effector domain that binds to a specific nucleotide sequence and an endonuclease domain that catalyzes a double strand break at the target site. Examples of TALENs and methods of making and using are described by PCT Patent Application Publication No. WO2011072246, incorporated herein by reference as if set forth in its entirety.

Another example of a sequence-specific nuclease system that can be used with the methods and compositions described herein includes the Cas9/CRISPR system (Wiedenheft, B. et al. *Nature* 482, 331–338 (2012); Jinek, M. *et al. Science* 337, 816–821 (2012); Mali, P. *et al. Science* 339, 823–826 (2013); Cong, L. *et al. Science* 339, 819–823 (2013)). The Cas9/CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) system exploits RNA-guided DNA-binding and sequence-specific cleavage of target DNA. A guide RNA (gRNA) contains 20 nucleotides that are complementary to a target genomic DNA sequence upstream of a genomic PAM (protospacer adjacent motifs) site (NNG) and a constant RNA scaffold region. The Cas (CRISPR-associated)9 protein binds to the gRNA and the target DNA to which the gRNA binds and introduces a double-strand break in a defined location upstream of the PAM site. Cas9 harbors two independent nuclease domains homologous to HNH and RuvC endonucleases, and by mutating either of the two domains, the Cas9 protein can be converted to a nickase that introduces single-strand breaks (Cong, L. *et al. Science* 339, 819–823 (2013)). It is specifically contemplated that the inventive methods and compositions can be used with the single- or double-strand-inducing version of Cas9, as well as with other RNA-guided DNA nucleases, such as other bacterial Cas9-like systems. In some embodiments, the guide RNAs used in the methods described herein are those of SEQ ID NOS: 36-42, respectively, in order of appearance. The sequence-specific nuclease of the methods and compositions described herein can be engineered, chimeric, or isolated from an organism.

The sequence-specific nuclease can be introduced into the cell in form of a protein or in form of a nucleic acid encoding the sequence-specific nuclease, such as an mRNA or a cDNA. Nucleic acids can be delivered as part of a larger construct, such as a plasmid or viral vector, or directly, *e.g.,* by electroporation, lipid vesicles, viral transporters, microinjection, and biolistics. Similarly, the donor construct can be delivered by any method appropriate for introducing nucleic acids into a cell.

Without being limited by any particular mechanism or theory, following sequence-specific nuclease-introduced double-strand break in the target sequence (e.g., ZFN-induced

DSB; FIG. 3A, Step 1), strand resection generates 3'single-stranded chromosome ends (FIG. 3A, Step 2). To initiate repair, the single-stranded chromosome ends anneal to complementary base pairs within the homology regions present on the donor construct by strand invasion (FIG. 3A, Step 3). The donor sequence can then be used as a template to extend the 3' single-stranded ends by DNA polymerase-mediated strand extension. Following strand extension, the extended strand anneals to the single-stranded chromosome end on the other side of the original double-strand break and repair is completed by DNA synthesis, using the extended strand as template, and ligation. The resulting double-stranded DNA contains the donor sequence flanked by recombinase recognition sites (FIG. 3A, Step 4).

This synthesis-dependent strand annealing model of double-strand break repair is consistent with the observation that very large stretches of foreign DNA with little or no homology to endogenous sequence, such as a reporter gene, can be inserted precisely into the point of the double-strand break. Consequently, donor sequences flanked by recombinase recognition sites can be integrated at the double strand break by resection of the free chromosome ends to expose regions around the target sequence that are substantially homologous to the homology regions on the donor construct (FIG. 3A). The homology regions can be of any length suitable for placement in a donor construct and effective in mediating strand annealing as described above, *e.g.*, a combined length of 10-5000bp, 100-1000bp, 500-600bp, or 537 bp. These steps, thus, create a conditional knock-out allele at the site of the target gene, *i.e.,* an allele comprising the donor sequence flanked by the recombinase recognition sites that produces a phenotype that is substantially similar to, or indistinguishable from, that produced by the cognate target gene allele. Two phenotypes are substantially similar or indistinguishable if upon standard inspection by a skilled artisan the nature of the underlying allele of the target gene cannot be detected. In some embodiments, the methods described herein produce cells carrying heterozygous conditional knock-out alleles or homozygous conditional knock-out alleles, *i.e.*, less than all or all of the endogenous alleles are replaced by the conditional knock-out allele.

The target gene can be any nucleic acid molecule encoding a protein (or fragments thereof) within the genetic material of the cell that is being targeted by the donor construct to produce a conditional knock-out version of the gene. For example, a target gene can be a gene located on the chromosome of a eukaryotic cell that encodes a protein of unknown function or that is involved in a cellular process. Such gene can be composed of a series of exons and introns. A target sequence can include exon, intron (including artificial intron), or regulatory

sequences of the target gene, or various combinations thereof. A target sequence can include the entire target gene.

The cell can be any eukaryotic cell, *e.g.*, an isolated cell of an animal, such as a totipotent, pluripotent, or adult stem cell, a zygote, or a somatic cell. In certain embodiments, cells for use in the methods described herein are cells of non-human animals, such as domesticated animals (*e.g.,* cows, sheep, cats, dogs, and horses), primates (*e.g.*, non-human primates such as monkeys), rabbits, fish, rodents (*e.g.*, mice, rats, hamsters, guinea pigs), flies, and worms. In certain embodiments, cells for use in the methods are human cells. The methods and compositions described herein can be used to target any genomic locus. Several specific examples of targeting different loci are described herein. In certain embodiments, the methods and compositions described herein can be used to target more than one genomic locus within a cell, *i.e.,* for multiplex gene targeting.

In a further particular aspect of the invention, a conditional knock-out animal is produced using the methods described herein. To produce a conditional knock-out animal, a donor construct and a sequence-specific nuclease are introduced into a cell, such as a zygote or a pluripotent stem cell, such as an embryonic stem cell or an induced pluripotent stem cell, or an adult stem cell, to create at least one conditional knock-out allele in the cell. Methods for screening for the desired genotype are well known in the art and include PCR analysis, *e.g.,* as described herein in the specific examples. The cell is then introduced into a female carrier animal to produce the conditional knock-out animal from the cell, for example as disclosed by US Patent No. 7,13,608, incorporated herein by reference as if set forth in its entirety. In certain embodiments, the cell is expanded to a two-cell stage, introduced into a blastocyst, or otherwise cultured or associated with additional cells prior to introduction into the carrier animal. In certain embodiments, the resulting conditional knock-out animal carries the conditional knock-out allele in its germline such that the conditional knock-out allele can be passed on to future generations.

In a further particular aspect of the invention, the methods and compositions described herein can be used to produce a knock-out allele. This method includes excising, inverting, or otherwise inhibiting normal expression of the recombinase recognition site-flanked donor sequence, once incorporated into the genome as conditional knock-out allele. The conditional knock-out allele is converted to a knock-out allele by introducing a recombinase into the cell that specifically recognizes the recombinase recognition sites. *E.g.,* Araki *et al.*, *Proc. Natl. Acad. Sci. USA* 92:160-164 (1995). The recombinase is an enzyme that recognizes specific

polynucleotide sequences (recombinase recognition sites) that flank an intervening polynucleotide and catalyzes a reciprocal strand exchange, resulting in inversion or excision of the intervening polynucleotide. One of skill in the art recognizes the advantageous efficiency of selecting for use in the methods described herein a recombinase that specifically recognizes the recombinase recognition sites within the donor construct.

The recombinase can be introduced into the cell containing the donor construct by any method in form of a protein or nucleotide sequence encoding the recombinase protein. To produce a knock-out animal, the conditional knock-out animal, produced as described above, is crossed to a transgenic animal having a transgene encoding a recombinase protein that catalyzes recombination at the 5' and 3' recombinase recognition site. Examples of animals carrying a recombinase transgene are known in the art and disclosed, for example, by US Patent No. 7,135,608, incorporated herein by reference as if set forth in its entirety. In some embodiments, the transgene encoding the recombinase is under the control of a tissue-specific promoter, such that the recombinase is expressed and, consequently, the knock-out allele is produced, only in such tissue. In some embodiments, the transgene encoding the recombinase is under the control of an inducible promoter, such that recombinase expression can be induced at a specific time. For example, the activation of Tet-On or Tet-Off promoters can be controlled by tetracycline or one of its derivatives. In some embodiments, the recombinase-encoding transgene is expressed only at a certain stage of development or in response to a compound administered to the animal. Examples of recombinases suitable for use in the methods disclosed herein include any version of P1 Cre recombinase, any version of FLP recombinase (flippase), and any version of Dre recombinase, including any inducible version of these recombinases (*e.g.*, fusions to a hormone-responsive domain such as CreERT2 and Cre-PR, or tetracycline-regulated recombinase).

B. Exemplary Compositions

In a further specific aspect of the invention, a composition for generating a conditional knock-out allele of a target gene is provided. Such composition includes a donor construct comprising a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, as described herein. The donor sequence comprises a target sequence having at least one neutral mutation, as described herein. The composition further comprises a sequence-specific nuclease that recognizes the target gene.

In certain embodiments, the sequence-specific nuclease is a zinc finger nuclease or a transcription activator-like effector nuclease. In certain embodiments, the recombinase recognition site is a *loxP* site or an *frt* site. Optionally, the composition can also include a recombinase, as described herein.

In a further aspect of the invention, a donor construct comprising the sequence shown in FIG. 4A (SEQ ID NO: 30), FIG. 4B (SEQ ID NO: 31), or FIG. 14C (SEQ ID NOS: 44-46).

In a further aspect of the invention, a guide RNA comprising the sequence shown in FIG.14A (SEQ ID NOS: 36-42) is provided.

In a further aspect of the invention, a cell comprising the donor construct comprising the sequence shown in FIG. 4A (SEQ ID NO: 30), FIG. 4B (SEQ ID NO: 31), or FIG. 14C (SEQ ID NOS: 44-46) is provided. This cell may be isolated from an animal produced by the methods described herein.

The invention can be further understood by reference to the following non-limiting examples of certain embodiments of the invention.

III. EXAMPLES

The following are examples of methods and compositions of the invention. It is understood that various other embodiments may be practiced, given the general description provided above.

Example 1: Pronuclear microinjection of Lrp5 ZFN mRNA into C57BL/6N fertilized eggs.

A custom eHi-Fi CompoZr® ZFN pair targeting exon 2 of mouse Low-density lipoprotein receptor-related protein 5 *(Lrp5)* was obtained from Sigma-Aldrich. The ZFNs harbor an optimized (eHi-Fi) FokI endonuclease interface that significantly increases its efficiency in introducing double-strand breaks (Doyon, Y. *et al. Nat Meth* **8**, 74–79 (2011)) at 5'-gacttccagttctccaagggtgctgtgtactggacagat-3' (SEQ ID NO: 29) (ZFN cleavage site is underlined). No significant potential off-site target activity was observed. Messenger RNA (mRNA) encoding the ZFN pair was stored at -80°C prior to use. mRNA (Sigma-Aldrich) was used for pronuclear microinjection and the two plasmids encoding the ZFN pair were used for ES cell electroporation.

To determine endonuclease activity, various concentrations of mRNAs encoding the *Lrp5* ZFNs were microinjected into the pronucleus of C57BL/6N zygotes (Table 1). *Lrp5* ZFN mRNA (2 μg of each ZFN in 5 μl) was thawed and diluted to 50 ng/μl in RNase- and DNase-free microinjection buffer (10 mM Tris and 1mM of EDTA, PH 8.0). ZFN microinjections,

Lrp5 ZFN mRNA was diluted to working concentrations of 2, 3, 4, or 5 ng/μl. Mouse zygotes were obtained from superovulated C57BL/6N females mated to C57BL/6N males (Charles River) the day before microinjection. Zygotes were harvested with M2 medium and microinjected in M2 following standard procedures (Nagy, A., *et al.*, *Manipulating the Mouse Embryo: A Laboratory Manual, Third Edition* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA, (2002)) and transferred into oviducts of E0.5 pseudopregnant ICR females (Taconic), 30 embryos per pseudopregnant female. ICR females were fed a 9% high fat diet (Harlan, catalog #2019) after embryo transfer surgery until the pups were weaned.

Microinjection Experiment	mRNA conc. (ng/μl)	Zygotes transferred after injection	Pups born	birth rate %	KOs	KO rate % (KO/born)
1	2	168	48	29	20	42
2	3	114	15	13	2	13
3	3	150	39	26	15	38
4	4	108	21	19	8	38
5	5	174	57	33	36	63

Table 1. Pronuclear microinjection of *Lrp5* ZFN mRNA into C57BL/6N fertilized eggs. KO mutants include mice with one or more mutant alleles. KO = knock-out.

DNA from the resulting pups was isolated from tail tissue and analyzed by PCR amplification and subsequent sequencing to identify large and small mutations. Genomic tail DNA was purified using Extract-N-Amp Tissue PCR kit (Sigma, Cat# XNAT2) or using Qiagen DNeasy 96 Blood and Tissue kit (Qiagen Cat# 69582). To determine ZFN-mediated mutation efficiency and to characterize the types of mutations caused by NHEJ repair, a 3-step PCR approach was performed. In the first step, an outer PCR using primers P1 and P2 was performed to detect large deletions or insertions. In the second step, an inner PCR using primers P3 and P4 was performed to detect small to medium size deletions or insertions. In the third step, direct sequencing of the inner PCR reaction product using primers P3 and P4 was performed to identify 1 to 20 base pair changes. Individual chromatograms were analyzed using Sequencher 4.10.1 (Gene Codes Corp.). If two distinct traces were detected, base pair calls for each individual allele were determined manually. Alleles from a subset of mutants were further analyzed by PCR TOPO subcloning (Invitrogen, Cat# K4575-J10). Twelve to twenty-four TOPO clones per mouse were sequenced using M13F and M13R primers.

Mutation rates of up to 63% of live-born pups were observed (5 ng/μl ZFN mRNA). The mutations ranged widely from insertions of one to three bp and deletions ranging from a single bp up to ~100 bp as well as one large ~800 bp deletion (summarized in FIG. 1). Multiple chimeric animals carrying more than two alleles were identified, likely resulting from

continuing ZFN activity after the first cell division. Furthermore, five animals were compound mutants, *i.e.*, these animals carried two independent mutant alleles of the same gene and no detectable wildtype allele of the gene, indicating ZFN activity on both chromosomes at the one cell stage.

Example 2: Direct generation of functional homozygous mutant alleles by microinjection of sequence-specific endonucleases.

LRP5 plays an obligatory role in retinal vascular development by serving as a co-receptor for NORRIN. Disrupted NORRIN signaling leads to vascular defects characterized by a failure to form capillary beds in the deeper layers of the retina, as well as vascular leakage (Xia, C.-H. *et al.*, *Human Molecular Genetics* **17**, 1605–1612 (2008); Xia, C.-H., *PLoS ONE* **5**, e11676 (2010); Junge, H. J. *et al.*, *Cell* **139**, 299–311 (2009)). Thus, 2-month-old mice with compound in-frame and out-of-frame deletions in *Lrp5* were generating as described in Example 1 and were examined for retinal vascular development. Animal #542 is a chimeric functional heterozygous that served as control. This animal carries one wild-type allele (a small 3bp in-frame deletion appeared to be silent) and an allele with a 1bp out-of-frame deletion. Animal #495 contains a 4bp out-of-frame deletion allele and a 1bp out-of-frame deletion allele. Animal #519 contains a 29bp out-of-frame deletion allele and a 17bp out-of-frame deletion allele. Animal #555 has a 3bp in-frame deletion allele and a 1bp out-of-frame deletion allele and is a functional heterozygote.

For phenotypic analysis, animals carrying *Lrp5* mutations were analyzed by fluorescein angiography. Mice were anesthetized with a mixture of ketamine/xylazine (80 mg/kg; 7.5 mg/kg) and dilating the eyes with 1% Tropicamide (Akorn, Inc.). Fluorescein angiography was performed after intraperitoneal injection of sterile 10% fluorescein solution (100 μl, AK-Fluor; Akorn, Inc.). Images were captured 1 minute after fluorescein injection using imaging setting of 0 focus and 50 sensitivity.

For histologic analysis, mice were sacrificed two days after angiography, enucleated, and processed for histology. Eyes were fixed in 4% paraformaldehyde (PFA) prior to dissection of retinas for whole mount histology, or cryoprotected in 30% sucrose overnight and embedded in Tissue-Tek® OCT Compound (Sakura) for frozen sections. Isolectin-B4 staining of whole mounts and sections was performed as previously described (Gerhardt, H. *et al., J. Cell. Biol.* 161, 1163–1177 (2003)). For frozen sectioning, cornea and lens were removed and eyes were washed extensively in PBS to remove residual PFA. Frozen 12 μm sections were

prepared and stained for MECA32, an antigen of the fenestrated endothelial cell marker PLVAP, essentially as described by Junge *et al*, Cell 139, 299–311 (2009).

The retinal phenotype of three compound mutant mice (495, 519 and 555) and a control heterozygous mutant mouse carrying one wildtype allele (542) is shown in FIG. 2. Mice carrying the compound mutations displayed an *Lrp5* null phenotype. Fluorescent angiography revealed that mice 542 and 555 display no apparent neovascular defects or vessel leakage (FIG. 2A). In contrast, mice 495 and 519, which contain compound out-of-frame deletions in both alleles of *Lrp5*, displayed numerous precapillary arteriole occlusions (FIG. 2A, arrows pointing to examples of precapillary arteriole occlusions) and significant vascular leakage, as indicated by the diffuse fluorescein signal throughout the retina. Scale bar in the bottom right panel of FIG. 2 represents 200μm for all panels in FIG. 2A.

Confocal projections of isolectin-stained wholemount retinas confirmed the *Lrp5* null phenotype of compound mutant mice 495 and 519. For each mouse, a projection of the maximum depth of the retina containing all three vascular layers (FIG. 2B), and images derived from projection of a single vascular layer residing in the nerve fiber layer (NFL, FIG. 2C), inner plexiform layer (IPL, FIG. 2D), and outer plexiform layer (OPL, FIG. 2E) were analyzed. While functional heterozygous retinas (542 and 555) contain a dense, well-organized three-tiered network of vessels, compound knock-out retinas (495 and 519) have irregular vasculature with reduced density (FIG. 2B, C). In addition, 542 and 555 contain normal capillary networks in the IPL (FIG. 2D) and OPL (FIG. 2E), whereas compound KO mice (495 and 555) have abnormal neovascular clusters in the IPL (FIG. 2D) and a small number of endothelial cell clusters in the OPL (FIG. 2E). Scale bar in the bottom right panel of FIG. 2 represents 100 μm for all panels in FIG. 2B-E.

In summary, mutant 555, carrying a loss of function allele with a 1 bp deletion and a functional allele with 3 bp in-frame deletion, displayed a normal retinal phenotype, while mutant 495, carrying a 4 bp and a 1 bp deletion, and mutant 519, carrying two larger deletions (17 and 29 bp), were phenotypically homozygous null, with a phenotype recapitulating what has been reported previously (Xia, C.-H. *et al.*, *Human Molecular Genetics* 17, 1605–1612 (2008)). These results demonstrate that microinjection of sequence-specific endonucleases can produce functional homozygotes (compound mutants) directly, although it is not known if these animals are compound mutants in all cells.

Example 3: Generation of conditional knock-out alleles by co-microinjection of *Lrp5* exon2 ZFN mRNA and donor construct.

FIG. 3A depicts a schematic outline of the strategy employed to generate a conditional knock-out allele (Gu, H., *Science* 265, 103–106 (1994)) of *Lrp5*, targeting exon 2. The ZFN pair introduces a double-strand break in *Lrp5* exon 2 (indicated by interrupted block arrow). The break is repaired by invasion of the donor plasmid through strand invasion and homologous recombination between the 5' and 3' *Lrp5* homology regions of the donor plasmid and the respective homologous sequences 5' and 3' of exon 2. The resulting locus contains the codon-optimized *Lrp5* exon 2 flanked by two loxP sites (FIG. 1A, bottom).

The 5' and 3' *Lrp5* homology regions in the donor plasmid were 1.1 and 1 kb, respectively, in length. Codon-modified (donor 1, FIG. 4A) and wildtype (donor 3, FIG. 4C) donor sequences were synthesized by Blue Heron/Origene (Bothell, WA) into a modified pUC19 vector. Donor 2 (FIG. 4B) was generated from donor 3 by replacing a 300 bp MscI-BamHI fragment with a synthesized fragment containing seven silent mutations to abrogate ZFN recognition. The insert in donor 1 is in opposite orientation compared to the insert in donors 2 and 3. Therefore, PCR amplification using primers that bind the plasmid backbone in combination with *Lrp5* locus-specific primers was conducted using primer combinations of the opposite orientation. The donor sequence, with the exception of the loxP sites, corresponds to mouse genome assembly NCBI37/mm9 chr.19:3658179-3660815. Circular donor plasmids were used in all experiments.

Silent mutations were introduced into the wildtype *Lrp5* exon 2 sequence to produce a codon-optimized version maintaining the protein-coding potential of the exon, but reducing the overall homology between wildtype C57BL/6 and donor *Lrp5* exon 2 to only 78% (donor 1, FIG. 4A; FIG. 5). To preserve normal RNA splicing, the first 13 bp or the last 11 bp of exon 2 were excluded from modification. FIG. 5 depicts a sequence alignment of the three *Lrp5* conditional knock-out DNA donors, excluding the 1.1 kb 5' homology and 1 kb 3' homology regions. Alignment was done using the alignment program ClustalW2, available at http://www.ebi.ac.uk/Tools/msa/clustalw2/. The overall homology between donor 1 (codon modified) and donor 3 (wildtype) exon 2 is 311/397 = 78%. The overall homology between donor 2 (ZFN binding site-modified only) and donor 3 exon 2 is 390/397 = 98%. LoxP sites are indicated by uppercase bold letters, intron sequences are indicated by lowercase letters, and the exon 2 (wild type or modified) sequences are indicated by uppercase letters. The ZFN

binding sites are boxed with dashed lines and the sequence at which the wild type exon 2 is cleaved is underlined. Silent mutations are boxed with solid lines.

Different combinations of ZFN mRNA and donor constructs were co-microinjected into C57BL/6N pronuclei (Table 2), essentially as described in Experiment 1, except that ZFN mRNA and donor construct were diluted together to working concentration (2.5 - 5 ng/μl for ZFN mRNA and 2.5 or 3 ng/μl for donor construct).

Co-microinjection Experiment	mRNA + DNA conc. (ng/μl)	Zygotes transferred after injection	Pups born	birth rate %	KOs	KO rate % (KO/born)	CKOs	CKO rate % (CKO/born)
1	2.5+2.5	120	28	23	11	39	0	-
2	2.5+2.5	128	35	27	10	29	0	-
3	2.5+3	114	6	5	4	67	0	-
4	3+2.5	121	10	8	6	60	0	-
5	3+3	126	23	18	7	30	0[a]	-
6	4.5+2.5	118	18	15	5	28	0	-
7	4.5+3	146	30	21	13	43	2[b]	6.7
8	5+2.5	102	18	18	8	44	0	-

Table 2. Co-microinjection of *Lrp5* ZFN mRNA (mRNA) and CKO donor 1 plasmid. KO mutants include mice with one or more mutant alleles. KO = knock-out; CKO = conditional knock-out. [a]One mouse (#95) was a false positive (donor 1 plasmid integrated into *Lrp5* locus). [b]Mice #140 and #155.

DNA isolated from tail samples from the 168 resulting pups were analyzed to identify mice that carry a conditional knock-out allele (FIG. 3B). The respective primer pairs used for analysis of mutants in the absence (P1-P4) or presence (P5-P12) of donor plasmid are indicated in FIG. 3B. First, the overall ZFN mutation frequency was determined as described in Experiment 1 and 2. Initial screening to identify mice carrying a potential conditional knock-out allele was performed by assaying for presence of the 5' LoxP site using a 5' nuclease assay (TaqMan®, Livak, K. J., *Genet. Anal.* **14**, 143–149 (1999)). In brief, 20μl reactions were constructed with a 2X Qiagen Type-it Fast SNP Probe PCR master mix, 50 – 120 ng template DNA, 400 nM primers and 200 nM fluorogenic Locked Nucleic Acid (LNA)-based probe specific for LoxP site recognition (Weis, B., *BMC Biotechnol* 10, 75 (2010)). Reactions were thermally cycled in an Applied Biosystems 7900HT (Life Technologies). Presence of the 5' LoxP was determined by analysis with Applied Biosystems Sequence Detection Software, version 2.3 (Life Technologies), by visualization of fluorescence evolution in the multi-component and amplification plots. *Lrp5* Locus-specific PCR analysis using primers P5/P6 was then performed to detect a 5' product specific for the codon-modified *Lrp5* exon 2 sequence present on both donor 1 and 2 (but not donor 3 used for the ES cell experiment of Example 4). Similarly, PCR using primers P7/P8 was performed to analyze the 3' end. To validate the presence of both 5' and 3' loxP, PCR analysis using primers P9/P10 and P11/P12,

respectively, was performed which will result in products only if the appropriate loxP sequence is present in the *Lrp5* locus. As the DNA was isolated from a mixture of chimeric subclones, false positive results were observed, *i.e.*, PCR products appear to be positive for 5'-3'-floxed *Lrp5* alleles even in the absence of such true conditional knock-out alleles. False positive results could be produced, for example, if one allele carries only the 5' loxP site and another allele carries only the 3' loxP site. To confirm the presence of a conditional knock-out allele, as opposed to false positive, a ~2.8kb *Lrp5* exon 2 PCR product was amplified using primers P5/P8 (both primers anneal outside the donor homology arms), cloned using and TOPO cloning (Life Technologies), and fully sequenced. This analysis identified conditional knock-out alleles, alleles with only a single loxP site, and alleles with donor-derived exon 2 sequence only (*i.e.*, no loxP sites). Alleles identified as false positives by sequencing analysis were analyzed for the presence of an integrated copy of the entire donor vector in the *Lrp5* allele by additional PCR using flanking primers P5 and P8 in combination with donor plasmid backbone-specific primers. Presence of random genomic insertions was determined with primers P6 and P7 (donor 1 and donor 2) in combination with donor plasmid backbone-specific primers (P13-P14). For random insertions of donor 3, donor plasmid backbone-specific primers (P13-P14) were used in combination with primers P15 and P16 that bind to the wildtype *Lrp5* sequence of donor 3. All primer sequences and reaction conditions are set forth in Table 3. The conditions for all PCR studies are set forth in Table 7.

Two mice (#140 and #155) were confirmed as carrying conditional knock-out alleles by full sequencing of a cloned PCR product obtained using primers located outside of the homology regions. For both mice, the conditional knock-out allele was transmitted to their progeny. In addition to the conditional knock-out allele, animal #155 also had one low frequency allele (not transmitted to progeny) with the 5' loxP site only. Animal #95 was a false positive as initial PCR analysis suggested a conditional knock-out allele, but detailed analysis revealed that a full-length donor plasmid was instead integrated into *Lrp5* exon 2. The knock-out mutation rates for each combination of ZFN mRNA and donor DNA ranged from 28 to 67% (Table 2).

Primer Number	Primer Sequence (5' --> 3')	Purpose
1	CATGTGCCTTTGAAGAGCACACC (SEQ ID NO: 1)	To detect large deletions/insertions
2	ACTCCACGGTCCTGGGATTATAGA (SEQ ID NO: 2)	To detect large deletions/insertions
3	GGCCTATCACTAAGGGAGCC (SEQ ID NO: 3)	To detect small to medium deletions/insertions
4	GCCCGAGATGACAATGTTCT (SEQ ID NO: 4)	To detect small to medium deletions/insertions
5	CGAGCTTTTCTTAGTGATCTTTTAAG (SEQ ID NO: 5)	5' flanking primer (outside of homology arm)
6	CTCACGTCGGTCCAATAAACG (SEQ ID NO: 6)	To detect donor 1 and donor 2 exon2 sequence
7	CGTTTATTGGACCGACGTGAG (SEQ ID NO: 7)	To detect donor 1 and donor 2 exon2 sequence
8	CCTAGACTGCAGTGAAGGACAT (SEQ ID NO: 8)	3' flanking primer (outside of homology arm)
9	GCTCACGAGCTTTTCTTAGTGATCTTTTAAGG (SEQ ID NO: 9)	5' flanking primer (outside of homology arm)
10	GAGAATCATGCACGGATAACTTCGTATAGC (SEQ ID NO: 10)	To detect 5' loxP integration
11	CAGGATTTCTTCTGTAGAGTATAACTTCGTATAATG (SEQ ID NO: 11)	To detect 3' loxP integration
12	CCTAGACTGCAGTGAAGGACATTCAC (SEQ ID NO: 12)	3' flanking primer (outside of homology arm)
13	GGATAACAATTTCACACAGGAAACAGCTA (SEQ ID NO: 13)	To detect random insertions or plasmid integration
14	GTAAAACGACGGCCAGTGAATTGG (SEQ ID NO: 14)	To detect random insertions or plasmid integration
15	CAGGGAAAGAGAATCATGCAC (SEQ ID NO: 15)	To detect donor 3 random insertions
16	CTGCACATGGGTAAACCTCTG (SEQ ID NO: 16)	To detect donor 3 random insertions
17	CACCTGAACTACTGAAAG (SEQ ID NO: 17)	To detect 5' loxP
18	CAGGGAAAGAGAATCATG (SEQ ID NO: 18)	To detect 5' loxP
19	F-ATAACTTCG-IQ-TATAGCATACATTATAC-Q (SEQ ID NO: 19)	To detect 5' loxP

Table 3. Primer nucleotide sequences. Primer P19: F=fluorophore (fluorescein); Q=quencher (Iowa Black FQ, Integrated DNA Technologies); IQ=internal quencher (ZEN, Integrated DNA Technologies). LNA bp are underlined.

The co-injection experiment in mouse zygotes (4.5 ng/μl ZFN mRNA and 3 ng/μl donor DNA) were repeated, by co-injecting *Lrp5* ZFN mRNA and either donor 1 or a floxed codon-optimized exon 2 donor that carries seven silent mutations with respect to the wildtype sequence, which abrogates ZFN-binding and cleaving of the donor (donor 2, FIG. 4B; FIG. 5). The results of these experiments are summarized in Table 4. Co-injection of donor 1 with

Lrp5 ZFN mRNA resulted in one out of twelve pups that carried a conditional knock-out allele (#243, 8.3% conditional knock-out rate). Co-injection of donor 2 with *Lrp5* ZFN mRNA resulted in three out of thirty-five pups (8.6%) that carried donor 2 exon sequence in the *Lrp5* locus. However, only one of these was subsequently confirmed as carrying a low frequency conditional knock-out allele (#250). The second of the three animals carried an allele with the 3' loxP site only (#274); the last animal (#280) harbored one allele with donor 2 exon sequence only (no loxP sites) and another allele with fully integrated donor 2 plasmid (false positive). These results suggest that the donor plasmid with the lowest sequence homology to the endogenous *Lrp5* exon 2 sequence (donor 1, FIG. 4A) was more efficient at generating conditional knock-out alleles.

Co-microinjection Experiment	Plasmid	Zygotes transferred after injection	Pups born	birth rate %	KOs	KO rate % (KO/born)	CKOs	CKO rate % (CKO/born)
1	Donor 1	50	10	20	10	100	1[a]	10
2	Donor 1	67	2	3	1	50	0	-
3	Donor 1	70	0	-	-	-	-	-
1	Donor 2	42	11	26	4	36	1[b]	9
2	Donor 2	73	15	21	10	67	0[c]	-
3	Donor 2	78	9	11.5	7	78	0[d]	-

Table 4. Co-microinjection of *Lrp5* ZFN mRNA and CKO donor 1 or donor 2 plasmids. All experiments were performed using 4.5 ng/μl ZFN mRNA and 3 ng/μl donor plasmid DNA. Overall CKO rate was 1/12 (8.3%) for donor 1 and 1/35 (2.9%) for donor 2. [a]Mouse #243; [b]Mouse #250; [c]One mouse (#274) carried a 3'loxP site only allele; [d]One mouse (#280) carried one allele with donor 2 exon only (no loxP sites) and one false positive allele (donor 2 plasmid integrated into *Lrp5* locus).

Example 4: Generation of conditional knock-out alleles by co-electroporation of *Lrp5* exon2 ZFN and donor plasmid.

C57BL/6N ES cells were co-transfected by electroporation with plasmids encoding the two *Lrp5* ZFN pair components alone, or along with either donor plasmid used for the microinjection experiments, or with an unmodified floxed wildtype *Lrp5* exon 2 plasmid (donor 3). C2 ES cells (Gertsenstein, M. *et al.*, *PLoS ONE* **5**, e11260 (2010)) were cultured, expanded, and electroporated using established methods (Nagy, A., Gertsenstein, M., Vintersten, K. and Behringer, R. *Manipulating the Mouse Embryo: A Laboratory Manual, Third Edition.* 800 (Cold Spring Harbor Laboratory Press: 2002)). In brief, fifteen million cells were electroporated with 15 μg of each ZFN plasmid with or without 15 μg donor plasmid. Electroporated cells were recovered in media and serial dilutions were plated on 10 cm plates on a feeder layer. Cells were grown for 7-8 days after which 144 clones (1.5 96 well plate) from each experiment were picked and placed into 96-well plates with feeder cells for

expansion. Two days after plating, the cells were split 1:2 into new 96-well plates with feeder cells. One plate was then stored at -80°C and the other plate was split into a new 96-well plate with 1% gelatin only without feeders cells, for DNA analysis. DNA was isolated as described in Example 1 except that ES cells were lysed over-night and DNA was precipitated, washed, and resuspended in TE buffer the following day, essentially as described by Ramírez-Solis, R. *et al.*, *Anal Biochem* **201**, 331–335 (1992).

ES cell experiment	Plasmid	Colonies screened	KOs	KO rate % (KO/analyzed)	CKOs	CKO rate % (CKO/screened)
1	None	144	24	17	NA	NA
2	Donor 1	144	ND	ND	1[a]	0.7%
3	Donor 2	144	ND	ND	0[b]	-
2	Donor 3	144	ND	ND	1[c]	0.7%

Table 5. Electroporation of plasmids encoding the *Lrp5* ZFN pair alone or in combination with CKO donors 1, 2, or 3 into C57BL/6N ES cells. All experiments were performed using 15 μg donor DNA and/or 15 μg each of ZFN1 and ZFN2. [a]Donor 1 ES clone #C8; [b] one donor 2 clone (F5) carried a 5' loxP only allele and a donor 2 exon only allele (no loxP sites), clone H10 carried a 3' loxP only allele; [c]two donor 3 clones (E3 and E4) carried alleles with 5' loxP only. Clone E3 also carried a false positive allele (donor 3 plasmid integration). Clone E4 also carried a true CKO minor allele (one positive out of 240 TOPO clones sequenced). ND: not investigated; NA: not applicable.

Results of the DNA analysis are shown in FIG. 3B, right, and the results are summarized in Table 5. The overall frequency of knock-out alleles observed in ES cells using electroporation (17%) was lower than obtained *in vivo* via pronuclear injection. The genetic alteration patterns from the ES cell electroporation experiment were similar to those observed after microinjection. Co-electroporation of donor 1 with *Lrp5* ZFN plasmid resulted in one conditional knock-out clone (clone C8) out of 144 analyzed. Co-electroporation of donor 2 with *Lrp5* ZFN plasmid resulted in two ES cell clones out of 144 analyzed that carry alleles derived from the donor. One of these clones (H10) carried the 3' loxP site allele only; the other (F5) carried one allele with donor 2 sequence only (no loxP sites) and one allele with the 5' loxP site only. Co-electroporation of donor 3 (wildtype) with *Lrp5* ZFN mRNA resulted in two targeted ES cell clones (E3 and E4). Both contained one allele with the 5' loxP site only. In addition, E3 carried another allele resulting from integration of donor 3 plasmid (false positive). Interestingly, clone E4 also had a very rare subclone positive for both loxP sites (conditional knock-out allele), possibly resulting from subsequent re-targeting of the previously targeted allele. These results confirm that using a donor with low homology to the endogenous exon is most efficient at generating conditional knock-out alleles. Table 6 provides an overall summary of the data from the microinjection and ES cell experiments.

	Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Random
Donor 1						
Mouse #95	2bp del.	plasmid int.	-	-	-	no
Mouse #140	CKO	wt	-	-	-	yes
Mouse #155	CKO	wt	5'loxP	-	-	no
Mouse #243	3bp del.	CKO	1bp del	-	-	no
ES cell C8	CKO	9bp del.	wt	-	-	yes
Donor 2						
Mouse #250	27 bp del.	1 bp del.	CKO	wt	3' loxP w/27 bp del.	no
Mouse #274	3' loxP	1 bp del.	-	-	-	no
Mouse #280	wt	donor only	plasmid int.	-	-	no
ES cell F5	wt	5' loxP	donor only	-	-	no
ES cell H10	wt	3' loxP	14 bp deletion	-	-	yes
Donor 3						
ES cell E3	95 bp del.	5' loxP	wt	plasmid int.	-	no
ES cell E4	wt	5'loxP	4 bp deletion	CKO	-	no

Table 6. Overview of *Lrp5* alleles derived from CKO donor plasmid.

Example 5: Normal gene function of conditional knock-out allele.

To determine if the silent mutations in the conditional knock-out allele obtained from donor 1 (FIG. 4A; FIG. 5) affected normal function of the *Lrp5* gene, mice carrying one knock-out allele (#140) and one conditional knock-out allele (#155) were bred with *Lrp5* knock-out homozygous mice generated using the ZFN pair of Example 3. Age matched postnatal day 16 (P16) control mice (FIG. 6A, +/+) were derived from an *Lrp5* heterozygous cross. The other mice used for the experiments were derived from a cross between an *Lrp5* KO/KO female and an *Lrp5* CKO/+ male. The *Lrp5* KO/KO female (FIG. 6B) is the adult mother of KO/+ (FIG. 6C, P16) and CKO/KO (FIG. 6D, P16). FIGS. 6A-D show representative confocal projections of retinal whole mounts stained with isolectin B4 (IB4) (scale bars: 50 μm). For each projection shown in FIGS. 6A-D, the left image depicts the maximum XY projection and the right image depicts the Z projection displaying vasculatures in the nerve fiber layer (NFL), the inner plexiform layer (IPL), and the outer plexiform layer (OPL) (labels on the bottom right panel of FIG. 6D). The *Lrp5* null animal showed a reduced vascular complexity in the XY projection and the absence of deep vascular layers (FIG. 6B). Mice carrying the conditional knock-out allele on the null background display a normal vascular phenotype (FIG. 6D), suggesting that the conditional knock-out allele is functional. FIG. 6E shows retinal cross sections of the opposite eyes to those depicted in FIGS. 6A-D stained with IB4, MECA32, and DAPI. Homozygous knock-out mice ectopically expressed the fenestrated endothelial cell marker MECA32, whereas the CKO/KO, KO/+, and +/+ mice are MECA32 negative.

In summary, homozygous knock-out animals display the retinal phenotypes described above (FIG. 6), whereas the retinal phenotypes of mice carrying one knock-out allele and one conditional knock-out allele were indistinguishable from those of wild type mice or mice having either one knock-out allele and one wild type allele (Fig. 6), indicating that the

conditional knock-out allele is a functional allele. Together these results demonstrate that a recombinase-recognition site-flanked donor sequence having neutral mutations can be used together with a sequence-specific nuclease to generate fully functional conditional knock out alleles *in vitro* and *in vivo*.

FIG. 7 illustrates possible mechanism that gave rise to the *Lrp5* alleles observed in these studies. The overall homology between *Lrp5* genomic sequence and donor 1 is reduced by multiple silent mutations (FIG.7A, asterisks). After resection of chromosome ends, strand invasion takes place in the large regions of 100% homology outside of the loxP sites, leading to a conditional knock-out allele having both loxP sites. Due to the limited homology in the region between the loxP sites, cross-over events inside the loxP sites is rare. Donor 2 contains larger regions of 100% homology between the loxP sites, allowing for strand invasion to take place inside of the loxP sites, resulting in alleles having a 3' loxP site only (FIG. 7B), a 5' loxP site only (FIG. 7C), or no loxP site (FIG. 7D). Primer combinations P9+P10 and P11+P12 both gave rise to PCR products for events according to FIG. 7A. Use of primer pair P9+P10 resulted in a product for events depicted in FIG. 7C but not for events depicted in FIG. 7B or D. Similarly, primer pairs P11+P12 gave rise to a product for events depicted in FIG. 7B but not for events depicted in FIG. 7C or D. Primer combinations P5+P6 and P7+P8 resulted in PCR products regardless of loxP status.

Primer Pair	Reagents Used	Cycling Reaction Conditions			Cycles	Product Size (bp)
		95 °C	Annealing	72 °C		
P1/P2	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 sec	a) 62C (-0.3C/cycle) - 30 sec b) 57C - 30 sec	1 min	a) 25 b) 12	1059
P3/P4	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 sec	a) 56C (-0.3C/cycle) - 30 sec b) 53C - 30 sec	30 sec	a) 10 b) 30	370
P5/P6	Advantage GC 2 PCR kit (Clontech, Cat# 639119)	45 sec	56C - 45 sec	1 min 30 sec	40	1394
P7/P8	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 sec	63C - 45 sec	1 min 30 sec	40	1410
P9/P10	Advantage GC 2 PCR kit (Clontech, Cat# 639119)	45 sec	56.5C - 45 sec	1min 30 sec	40	1195
P11/P12	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 sec	62C - 45 sec	1 min 30 sec	40	1105
P13/P6	Advantage GC 2 PCR kit (Clontech, Cat# 639119)	45 sec	56C - 45 sec	1 min 30 sec	40	1482
P7/P14	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 sec	62C - 45 sec	1 min 30 sec	40	1462
P5/P14	LA Taq (TaKaRa, Cat# RR002M)	45 sec	55C - 45 sec	10 min	47	2836
P13/P8	LA Taq (TaKaRa, Cat# RR002M)	45 sec	55C - 45 sec	10 min	47	2871
P5/P8	Advantage GC 2 PCR kit (Clontech, Cat# 639119)	45 sec	57C - 45 sec	3 min 30 sec	40	wt - 2729 CKO - 2797
P13/P15	Advantage GC 2 PCR kit (Clontech, Cat# 639119)	45 sec	56C - 45 sec	1 min 30 sec	40	1273
P16/P14	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 sec	63C - 45 sec	1 min 30 sec	40	1211
P17/P18 P19	Type-it Fast SNP Probe PCR mix (Qiagen, Cat# 206042)	15 sec	60C - 60 sec	20 sec	40	fluorescence

Table 7. Conditions for PCR reactions used in the Examples described above.

Original Residue	Exemplary Substitutions	Conservative Substitutions
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Asp, Lys; Arg	Gln
Asp (D)	Glu; Asn	Glu
Cys (C)	Ser; Ala	Ser
Gln (Q)	Asn; Glu	Asn
Glu (E)	Asp; Gln	Asp
Gly (G)	Ala	Ala
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe; Norleucine	Leu
Leu (L)	Norleucine; Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Trp; Leu; Val; Ile; Ala; Tyr	Tyr
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Val; Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala; Norleucine	Leu

Table 8. Conservative substitutions.

Example 6: Cas9/CRISPR-mediated mutagenesis of *Lrp5* exon 2 using different guide RNAs.

To confirm that other sequence-specific endonucleases can be used with the methods and compositions described herein, modified alleles of *Lrp5* were produced using the Cas9/CRISPR system. Hepa1-6 murine hepatoma cells were cultured in RPMI supplemented with 10% FBS, L-glutamine, and antibiotics. After trypsinization and pelleting, 10^6 cells were electroporated with 2 μg per plasmid containing hCas9-encoding cDNA or 15 μg of mRNA encoding Cas9 (FIG. 14, SEQ ID NO: 43) using AMAXA Nucleofector Kit V with AMAXA Nucleofector program T-028 (Lonza) according to the manufacturer's instruction and plated into a 6 well plate. Nucleofection efficiencies reached 80-95% as assessed by GFP expression

(PMAXGFP). Fresh media was exchanged at 24hrs post-nucleofection and purified genomic DNA was harvested at 72 hr post-nucleofection using DNeasy Blood and Tissue kit (Qiagen). HCas9 mRNA was transcribed *in vitro* with MMESSAGE MMachine T7 Ultra kit (Life Technologies) following the manufacturer's protocol, including a polyA tailing reaction. mRNA was purified and concentrated using standard phenol:chloroform extraction and precipitation of RNA.

Three unique guide RNAs (gRNAs) targeting mouse *Lrp5* exon 2 were generated (FIG. 14A; Lrp5 gRNA T2, Lrp5 gRNA T5 and Lrp5 gRNA T7; SEQ ID NOS: 36-38).

NIH/3T3 cells or Hepa1-6 cells were co-transfected with either DNA encoding zinc finger pairs (pZFN1+pZFN2) or with Cas9 (+pRK5-hCas9) together with a guide RNA targeting *Lrp5* exon 2 (p_gRNA T2, p_gRNA T5 or p_gRNA T7) or a control plasmid (PMAXGFP). gRNA T7 sequence overlaps with the 3' end of the right ZFN protein binding site sequence.

To detect mutations in the *Lrp5* locus following co-transfection, indicative of Cas9-mediated cleavage and subsequent repair, SURVEYOR Assays (Transgenomic) were performed essentially according to the manufacturer's instruction. In this assay, PCR products are hybridized. In the event of mutations, the hybridization complex contains a mismatch which is cleaved by the SURVEYOR nuclease. In this example, a ~2.7kb PCR product specific for the Lrp5 exon2 genomic locus was amplified using primers P9 and P12 (SEQ ID NOS: 9 and 12) using the following parameters and LA Taq (Takara): 95°C for 3min, 35 cycles of 95°C for 45sec; 57°C for 45sec; 70°C for 2min30sec, followed by 72°C for 7min. One-third of the PCR product was used in the SURVEYOR Assay. Resulting digested products were resolved by electrophoresis on a 1.5% agarose gel. Nuclease cutting was identified by the presence of shorter fragments, which indicated the presence of mutant alleles that annealed with wildtype.

All three guide RNAs (gRNAs) targeting mouse *Lrp5* exon 2 efficiently mediated Cas9-induced mutations (FIG. 8). The activity of each gRNA/Cas9 pairing appears to be folds greater than ZFN mediated mutagenesis in these experiments. Mutation rates were calculated from sequencing TOPO cloned alleles from a 2.7kb PCR product of the *Lrp5* exon 2 genomic locus. Alignments of individual sequences to wildtype determined exact deletion (quantified above) or insertion sizes (data not shown). A 2.7kb genomic region was amplified by PCR with primers P9 and P12 as described above. The PCR products were cloned directly using TOPO-TA cloning (Invitrogen) to capture all possible deletion sizes. After transformation and

plating for single colonies, clones were selected, plasmid DNA isolated, and sequenced according to the Sanger method using primers P20 and P21. FIG. 9A-B illustrate a summary of gRNA/Cas9 mutation rates (FIG. 9A) and deletion sizes (FIG. 9B) in Hepa1-6 murine hepatoma cells.

Example 7: Cas9/CRISPR mediated gene targeting using a codon-optimized conditional knock-out donor vector.

Hepa1-6 cells were co-transfected with Cas9 plasmid or mRNA, a gRNA and the Lrp5 CKO donor 1 comprising the codon-optimized exon sequence. For comparison, some cells were co-transfected with Lrp5 ZFN plasmids and the donor plasmid (FIG. 10). After 72 hours, genomic DNA from the transfected cells was analyzed by PCR with a primer specific to the codon-optimized *Lrp5* donor exon (P7; SEQ ID NO: 7), and a primer specific to a region outside of the 3' homology arm (P12; SEQ ID NO: 12). Primers P7 and P12 were used for the PCR reaction with REDExtract-N-Amp PCR ReadyMix (Sigma) with the following conditions: 95°C for 3 min, 38 cycles of 95°C for 45sec; 63°C for 45sec; 72°C for 1min30sec, followed by 72°C for 7min. PCR products were resolved by electrophoresis on a 1% agarose gel. As described above, the *Lrp5* exon 2 donor1 vector contains a codon optimized exon (COexon2) harboring many neutral mutations, excluding from mutation the first 13bp and the last 11bp, as well as exogenous flanking loxP sites. The PCR above uses a forward primer specific for COexon2 sequence and a reverse primer outside of the homology arm in the genomic locus, therefore producing a PCR product only if the donor exon sequence was incorporated in the correct *Lrp5* locus. The use of gRNA/Cas9 resulted in donor sequence integration at the *Lrp5* locus with great efficiency, exceeding that observed when using the ZFN system and the same donor vector strategy (FIG. 10).

Example 8: Cas9/CRISPR mediated targeted introduction of loxP sites.

To determine if the donor design strategy and the Cas9/CRISPR system can be used to introduce loxP sites at a genomic locus, genomic DNA from cells transfected as described in Example 7 was analyzed by PCR analysis using one primer located outside the homology arms, and one primer anchored at either the 5' or 3' loxP sites from the donor. For the 5' genomic to 5' loxP reaction, primers P9 and P10 (SEQ ID NOS: 9 and 10) were used with the standard Expand High Fidelity PCR System (Roche) protocol except for an addition of DMSO to a final concentration of 2%. PCR parameters were as follows: 95°C for 3min, 45 cycles of 95°C for 45sec; 63°C for 45sec; 72°C for 1min30sec, followed by 72°C for 7min. For the 3' loxP to 3' genomic reaction, primers P11 and P12 were used following the standard REDExtract-N-Amp

PCR ReadyMix (Sigma) protocol. PCR parameters were as follows: 95°C for 3min, 40 cycles of 95°C for 45sec; 62.5°C for 45sec; 72°C for 1min30sec, followed by 72°C for 7min. PCR products were resolved by electrophoresis on 1% agarose gels. PCR products were obtained for the 3' loxP site from samples isolated from cells that were transfected with either of the two different Lrp5 gRNAs and the CKO donor (FIG. 11; p_gRNA T2). Similarly, PCR products were obtained from samples isolated from cells that were transfected with gRNA T7 for the 5' loxP site (FIG. 11). Thus, FIG. 11 shows that in Hepa1-6 cells, Lrp5 gRNA T2/Cas9 and Lrp5 gRNA T7/Cas9 mediated double-strand breaks resulted in introduction of loxP sites at the Lrp5 locus using the codon optimized exon donor vector strategy. Only cells electroporated with Cas9, gRNA, and donor exhibit evidence of 5' (FIG. 11, top) and 3' (FIG. 11, bottom) loxP sites in the *Lrp5* genomic locus. gRNA T7 resulted in more prominent 5'loxP presence whereas integrated loxP sites were not detectable with ZFNs. The absence of detectable loxP sites in the ZFN samples and low levels in the gRNA samples in these experiments using Hepa1-6 cells might be explained by both low homologous recombination rates in cell lines and the fact that the full cell pool transfected, not clonal subsets, were analyzed. A single mouse genomic DNA sample with an *Lrp5* CKO/wt genotype was used as a positive control. These results show that the CKO design strategy can be used in somatic cells and that it effectively reduces the frequency of undesirable cross-over events between the double strand break and the location of both the 5' and 3' loxP sites. In summary, targeting of specific genomic loci by introducing RNA-guided nuclease-mediated DNA breaks that are subsequently repaired using an engineered codon-optimized CKO donor sequence can be used to insert loxP sites and thereby produce conditional knock-out alleles.

Example 9: Targeting of *Usp10, Nnmt, and Notch3* genomic loci.

To confirm that other genes can be targeted with the inventive methods, donor and gRNAs for the *Usp10, Nnmt, and Notch3* genomic loci were generated. These Cas9/gRNAs and donors were introduced into Hepa1-6 cells as described in Example 6 and as depicted in FIG. 12 to introduce DNA double-strand breaks at the respective loci and subsequent repair using the codon-optimized donor as a template. A SURVEYOR Assays were performed essentially as described above. PCR products 2.2 - 2.7kb in size, specific for *Lrp5*, *Usp10,* and *Notch3* genomic loci were amplified using primers P9, P12, P22, P23, P24, P25 (SEQ ID NOS: 9, 12, 22, 23, 24 and 25, respectively) and the following parameters with LA Taq (Takara): 95°C for 3min, 35 cycles of 95°C for 45sec; Ta for 45sec (Lrp5=57C, Usp10 & Notch3 = 63); 70°C for 2min30sec, followed by 72°C for 7min. One-seventh, 1/3, and all of the PCR

products were used, respectively, in the SURVEYOR Assay as following the manufacturer's instruction (Transgenomic). Resulting digested products representing nuclease cutting where strands of wildtype and mutant alleles have annealed, were resolved by electrophoresis on a 1.5% agarose gel

FIG. 12 and FIG. 13 show that as observed with the *Lrp5* locus, *Usp10, Nnmt, and Notch3* genomic loci were efficiently targeted by specific gRNA/Cas9 complexes (FIG. 12) and that loxP sites were integrated (FIG. 13).

Example 10: Generation of conditional knock-out and knock out alleles of *Lrp5* using RNA-guided sequence-specific endonucleases and codon-optimized donor.

The *Lrp5* locus can be targeted with the *Lrp5*-specific gRNAs described herein to introduce a floxed codon-optimized exon thereby creating conditional knock-out alleles. Subsequent expression of the Cre recombinase protein in cells harboring the conditional knock out allele can excise the floxed exon resulting in a knock-out allele.

Primer Number	Primer Sequence (5' --> 3')	Purpose
20	AGG AAA GCT AGC TTT CCA GGA GTA TG (SEQ ID NO: 20)	Sequencing *Lrp5* genomic PCR
21	GGA AGT CAA ATC CTC CTG GTT ACG A (SEQ ID NO: 21)	Sequencing *Lrp5* genomic PCR
22	GGC GTC CAG ATT ATG CAC AC (SEQ ID NO: 22)	Amplify *Usp10 locus*
23	GAT AAT CAT GGA ATC TAA TC (SEQ ID NO: 23)	Amplify *Usp10 locus*
24	TCT TTG CCT GAC CTG GCT ATG AG (SEQ ID NO: 24)	Amplify *Notch3* locus
25	CAA TCT TTC TAA CGC TCA ACT CAG AGT C (SEQ ID NO: 25)	Amplify *Notch3* locus/Detect 3'loxP
26	CAT TGG GCT GGT ACA CGG A (SEQ ID NO: 26)	Detect *Nnmt* 5'loxP
27	GAG CTG AAG TTA TAG ATA ACT TCG TAT AGC (SEQ ID NO: 27)	Detect *Nnmt* 5'loxP
28	GGG AAC CCT ATA ACT TCG TAT AAT G (SEQ ID NO: 28)	Detect *Notch3* 3'loxP

Table 9. Primer nucleotide sequences.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, the descriptions and examples should not be construed as limiting the scope of the invention. The disclosures of all patent and scientific literature cited herein are expressly incorporated in their entirety by reference.

WHAT IS CLAIMED IS:

1. A method of generating a conditional knock-out allele in a cell comprising a target gene, the method comprising the steps of:
    a. introducing into the cell a donor construct, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation; and
    b. introducing into the cell a sequence-specific nuclease that cleaves a sequence within the target gene, thereby producing a conditional knock-out allele in the cell.
2. The method of claim 1, wherein the sequence-specific nuclease is a zinc finger nuclease (ZFN).
3. The method of claim 1, wherein the sequence-specific nuclease is a transcription activator-like effector nuclease (TALEN).
4. The method of claim 1, wherein the sequence-specific nuclease is a ZFN dimer that cleaves the target gene only once.
5. The method of claim 1, wherein the sequence-specific nuclease is an RNA-guided nuclease.
6. The method of claim 5, wherein the RNA-guided nuclease is Cas9.
7. The method of claim 1, wherein the sequence-specific nuclease is introduced as a protein, mRNA, or cDNA.
8. The method of claim 1, wherein the recombinase recognition site is a *loxP* site, an *frt* site, or a *rox* site.
9. The method of claim 1, wherein the donor sequence comprises seven silent mutations.
10. The method of claim 1, wherein sequence homology between the donor sequence and the target sequence is 98% or less.
11. The method of claim 10, wherein sequence homology between the donor sequence and the target sequence is 78%.
12. The method of claim 1, wherein the donor construct comprises the sequence of SEQ ID NO: 30, 31, 44, 45, or 46.
13. The method of claim 1, wherein the 5' homology region comprises at least 1.1 kb and wherein the 3' homology region comprises at least 1 kb.

14. The method of claim 1, wherein the target gene is selected from the group consisting of *Lrp5, Usp10, Nnmt, and Notch3.*
15. The method of claim 1, wherein the cell was isolated from a mammal.
16. The method of claim 15, wherein the mammal is selected from the group consisting of mouse, rat, rabbit, hamster, guinea pig, cat, dog, sheep, horse, cow, monkey, and human.
17. The method of claim 1, wherein the cell is a zygote or a pluripotent stem cell.
18. A method of generating a conditional knock-out animal, the method comprising the steps of:
    a. introducing a donor construct into a cell comprising a target gene, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation;
    b. introducing a sequence-specific nuclease into the cell, wherein the nuclease cleaves the target gene; and
    c. introducing the cell into a carrier animal to produce the conditional knock-out animal from the cell.
19. The method of claim 18, wherein the cell is a zygote or a pluripotent stem cell.
20. A method of generating a knock-out animal, the method comprising the steps of:
    a. introducing a donor construct into a cell comprising a target gene, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation;
    b. introducing a sequence-specific nuclease into the cell, wherein the nuclease cleaves the target gene;
    c. introducing the cell into a carrier animal to produce a transgenic animal from the transfected cell; and
    d. breeding the conditional knock-out animal with a transgenic animal having a transgene encoding a recombinase protein that catalyzes recombination at the 5' and 3' recombinase recognition site.
21. The method of claim 20, wherein the cell is a zygote or a pluripotent stem cell.

22. The method of claim 20, wherein the recombinase recognition site is a *loxP* site and wherein the recombinase is Cre recombinase.
23. The method of claim 20, wherein the recombinase recognition site is an *frt* site and wherein the recombinase is FLP recombinase.
24. The method of claim 20, wherein the recombinase recognition site is a *rox* site and wherein the recombinase is Dre recombinase.
25. The method of claim 20, wherein the transgene encoding the recombinase is under the control of a tissue-specific promoter or an inducible promoter.
26. A composition for generating a conditional knock-out allele of a target gene comprising:
    a. a donor construct comprising a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation compared to the sequence of the target gene; and
    b. a sequence-specific nuclease that recognizes the target gene.
27. The composition of claim 26, wherein the sequence-specific nuclease is selected from the group consisting of ZFN, TALEN, and RNA-guided nuclease.
28. A donor construct comprising the sequence of SEQ ID NO: 30, 31, 44, 45, or 46.
29. A cell comprising the donor construct of claim 28.
30. A non-human conditional knock-out animal prepared according to the method of claim 18.

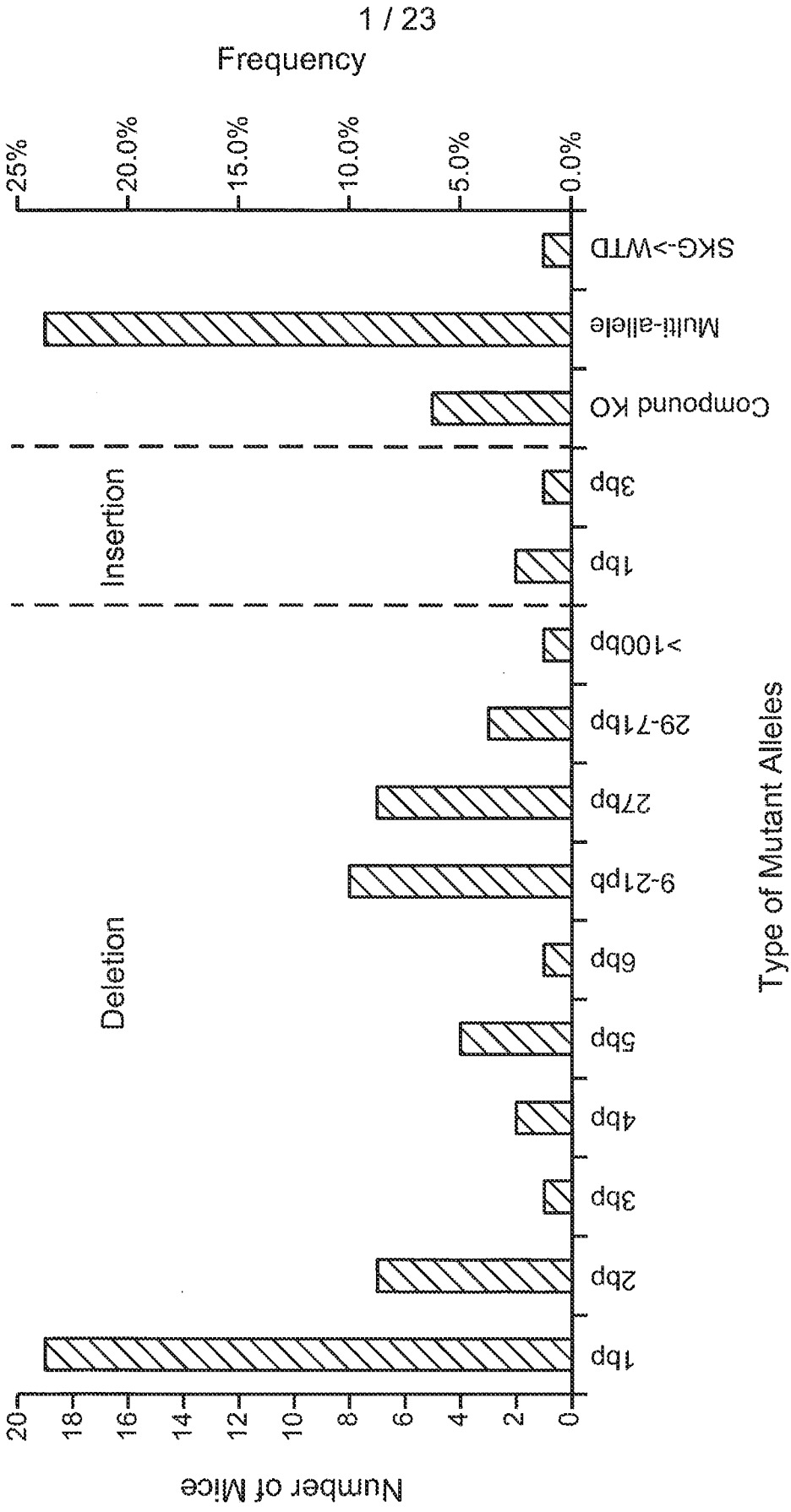

Frequency
25%
20.0%
15.0%
10.0%
5.0%
0.0%
Number of Mice
20
18
16
14
12
10
8
6
4
2
0
Deletion
Insertion
1bp
2bp
3bp
4bp
5bp
6bp
9-21bp
27bp
29-71bp
>100bp
1bp
3bp
Compound KO
Multi-allele
SKG->WTD
Type of Mutant Alleles

FIG. 1

FIG. 2A

FA

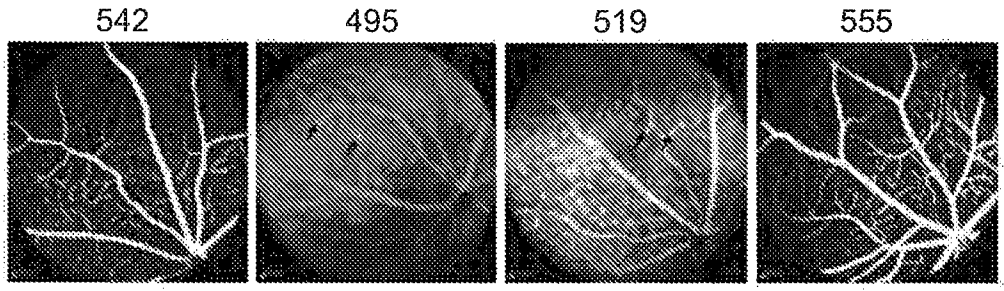


FIG. 2B

IB4, All Layers

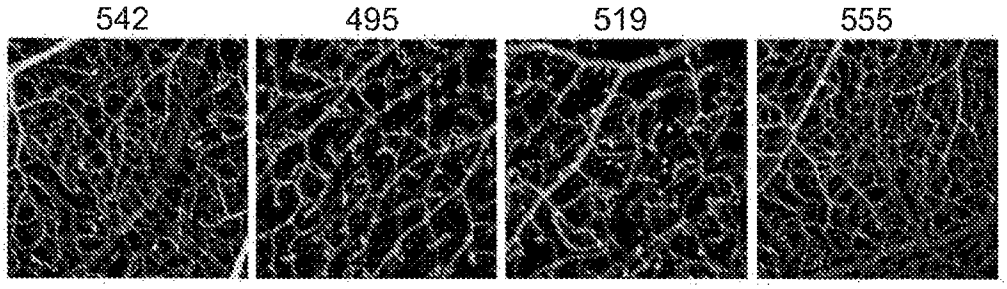


FIG. 2C

IB4, NFL

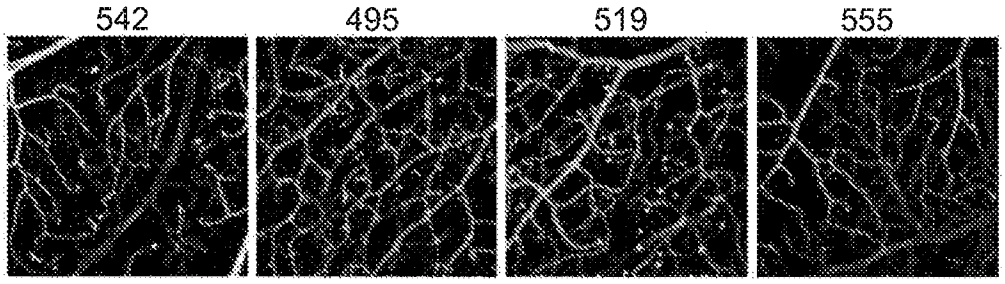


FIG. 2D

IB4, IPL

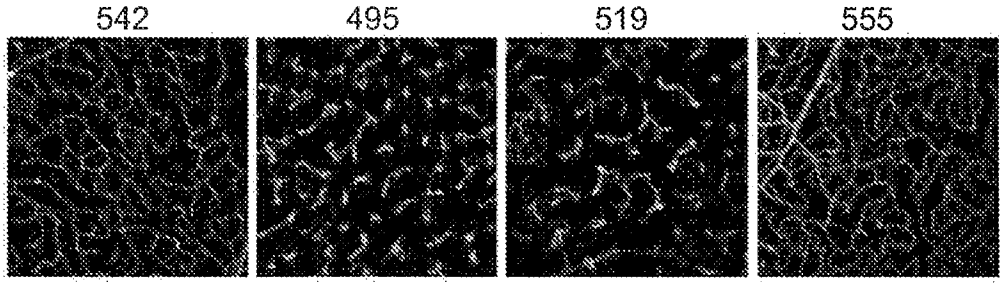
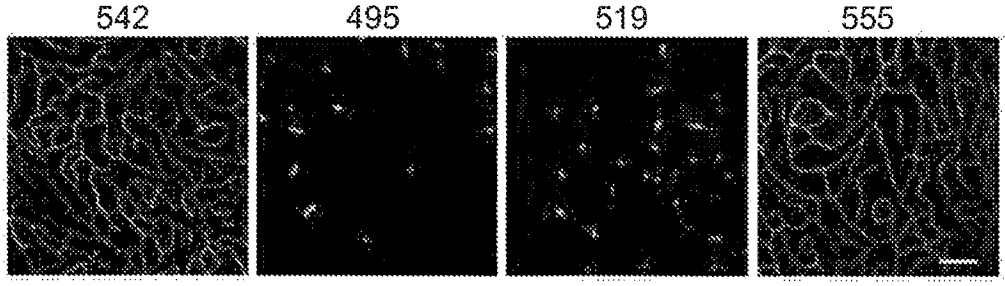


FIG. 2E

IB4, OPL



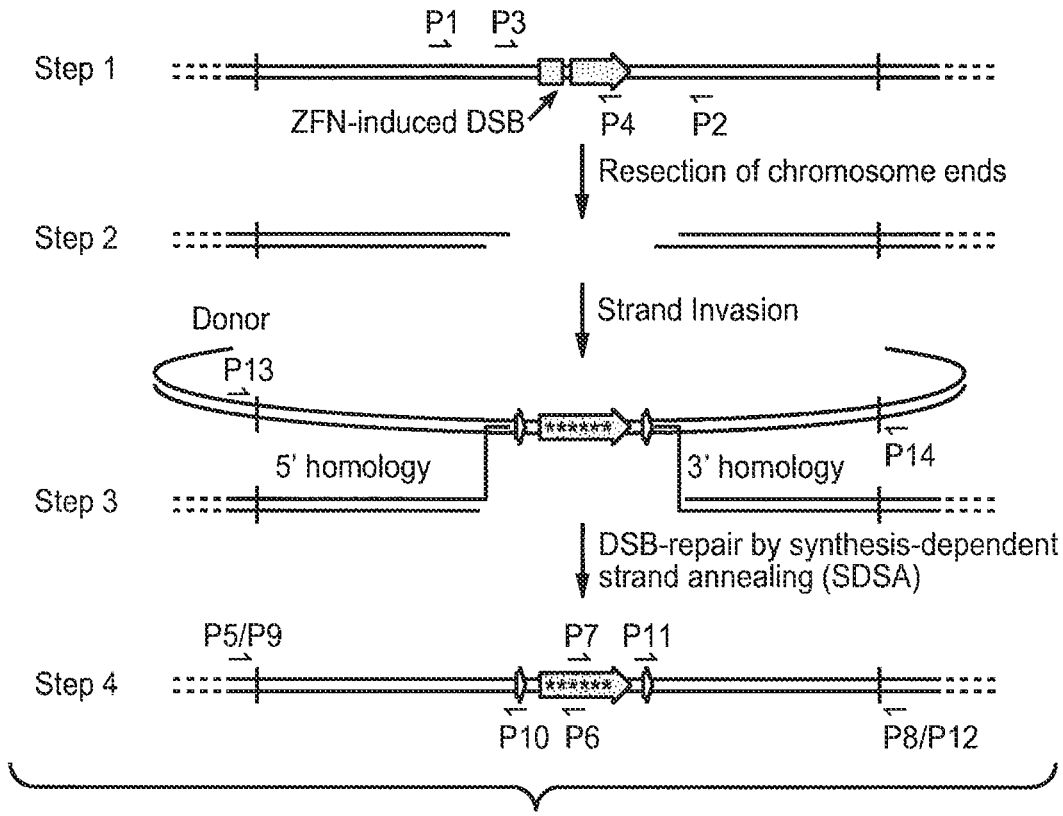

P1
P3
Step 1
ZFN-induced DSB
P4
P2
Resection of chromosome ends
Step 2
Donor
Strand Invasion
P13
5' homology
3' homology
P14
Step 3
DSB-repair by synthesis-dependent strand annealing (SDSA)
P5/P9
P7
P11
Step 4
P10
P6
P8/P12

FIG. 3A

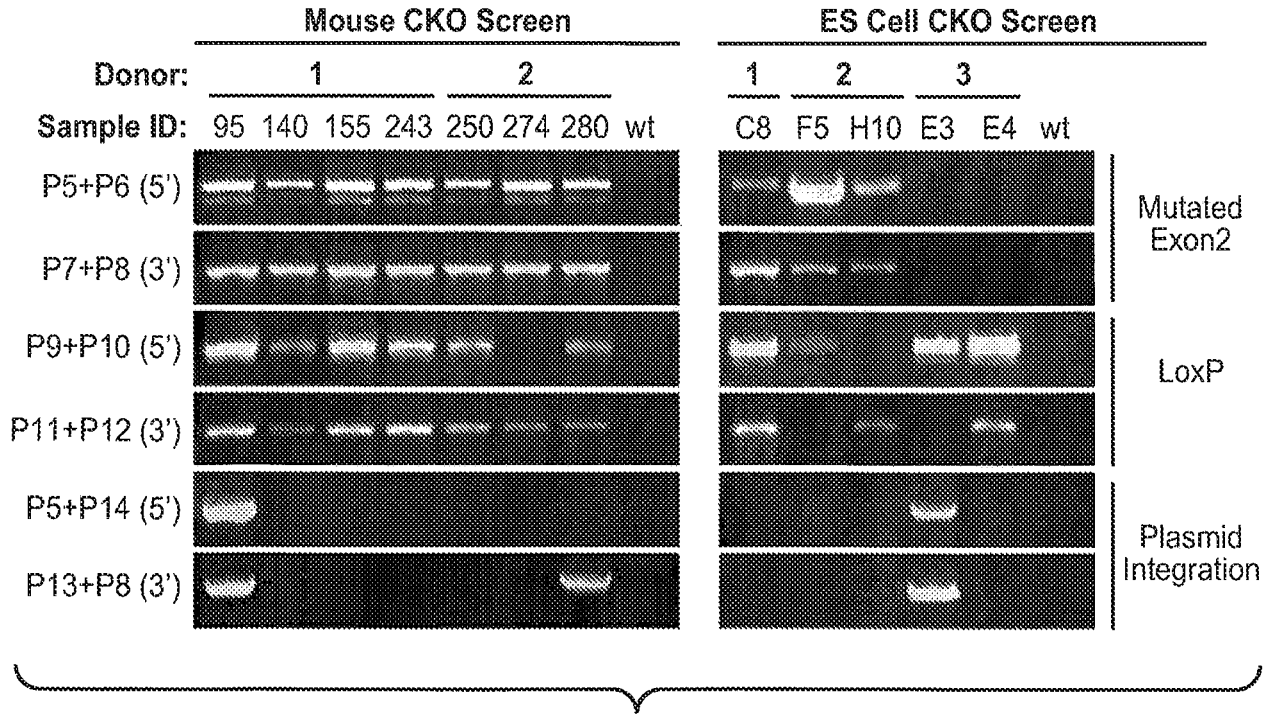

Mouse CKO Screen
ES Cell CKO Screen
Donor:
1
2
1
2
3
Sample ID:
95 140 155 243 250 274 280 wt
C8 F5 H10 E3 E4 wt
P5+P6 (5')
P7+P8 (3')
P9+P10 (5')
P11+P12 (3')
P5+P14 (5')
P13+P8 (3')
Mutated Exon2
LoxP
Plasmid Integration

FIG. 3B

aatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgt
gagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgt
ggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagctcc
ttcctcttccagcccttcctctttcactgactgactgactggaagacacacctgcggccgcgca
agctgggccatggggcttaacccctaacatcactcttggtgtgacttggcacaagtgtatgctt
ctttatgaaggctgcagctctgggcctggagaatgtataccaggacccctagtctaccaggggc
tgtgcatttgagaggcatgggtcacctggtaggggaggaacccgtaggaccagcctcctgactg
aggctggcactgtgtgtttctaccctgggaactttgcatgggttttaatgaggcctccctaac
cactgtatgtagtgtgcagccttctcctgttctgttcccactcctgtcctgtccctgactctgg
cccattcacagcattaacctgacagatgtcactacctgcttttggtctttgcccggtgtctgtt
tctccactggaacgggaaccttggaggcccctggtacaagcagcccctgcttgccagcctatga
aggcatcttggttttgccactgtggctctttttctttcactcccagaactttctttttctcctc
ctgagctttcctttgcagcttctgcccttgaaaaggagaggctggtggagccttggggtgatac
agtgtctgtgttgtctaggattccccaggcgggggaggagaggagaagcagcctgccccaggtg
agggtggggagccacccacattctggcctgcctcactgtgggtccctggggccactgccaggcc
tcttagaaggaaagctagctttccaggagtatgctgtgggttccccttaaaacttagggattaa
gtgtgtggtttctttacttttcagagctgggagcactgccccaccatctcccagtgagggttgg
agaaaaaaaaagtcctaggttacaaatccacccccagcctccctgtccctgcccctgcccccctc
ccctcccatgtgcctttgaagagcacacctgaactactgaaagcaagttgaggtgaggaagtgg
tctttccagtttggggagcagtagagccagagcttggggatttggggagccttcctgcctctct
atactggcctatcactaagggagcctttgggcaactggaaggaaggagtcttgggagcaggggc
ccggggtagataacttcgtataatgtatgctatacgaagttatccgtgcatgattctctttccc
tgggagctggttcctgtcctgcttagccatcctgaccctgtgctctgcccacagcctcaccgct
cctcctctttgctaaccgccgagatgttcgactcgtagacgctggcggcgtgaagctggagagt
acaatcgttgctagtggactggaggatgctgccgctgtcgacttccagttctcaaagggagccg
tttattggaccgacgtgagcgaagaggcaataaaacagacataccttaaccagactggagctgc
cgcccagaatatcgtaatctctggcttggtgagccctgatggactcgcttgtgattgggtcgga
aaaaagctttattggactgatagtgaaacgaatcggatagaagtagcgaaccttaacggtacat
caagaaaggtcctgttttggcaggatcttgaccagcctagagcaatcgccctcgaccctgcaca
tgggtaaacctctgtttttctcctgctcctatggggaggcgcggggaacaggatttcttctgta
gagtataacttcgtataatgtatgctatacgaagttatacctgaggcaggctgatgaggtgcct
tgggaagcagcccttgctgggaatcacctcccttccagatcctgaactgtaaaacatggcatgt
ttagccagggatggtggttcacacctacaatcctagcattctggaggctgaggcaggagtgttg
ccacaagttccaggccaacctgagctataaagaccgtcttaaaaatatacaggaactggggaga
tgcctccattggtgatgtacttacctcgtaaccaggaggatttgacttccatccatgaagaaag
ttaggtgcggggcgtcacatgtctataatcccaggaccgtggagttagagacacagggatctgt
gaggcttgctggctgcctgcagtcaaattgcccagcccctggtcccagtgagagacaaggtgtt
ttaaggtatttatattgccgtcttaaaaactccatgaccaagaacagcttggggatgaatcagt
cagtatacttgatcagacaatactcaagtcatgccccacctctgagtaaagtcagggcaggacc
tcaaggcaggaacccagagtcaagaactaaactgcttagtggtttgctctccattggttgttca
gattgctctcttgtagcacccaggaccatcagcccaggcactgaactgcccacagtgacctgta
ccctcccccaccaatcatcagccaaggaaatgcttcccaggcttagaggccaatgcaaactgac
tgtagcttagctgagcaggacgttgtgtctcaatcaaacacacaaacaagggcagatgaggatg
atgcttgggtttgaagcttagcatccttctgcactaggacaggtgacccaggtaagttcaggga
cctcccacccccttctgtgactgttggctgtggtggctctgtgtctggatgtagcacctactctg
tgtattctggaagttacctatattggtgcagtcaactcactctgaccctggaaagcgcaggaca
gccatgaaggccacgcggccgcacgacagtcttcactgactgactgactggaaagtcctctcca
ctgactgtagcctccaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcg
ttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggc
ccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtat
tttctccttacgcatctgtgcggtatttcacaccgcatacgtcaaagcaaccatagtacgcgcc
ctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgcc
agcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcacctcga
ccccaaaaaacttgatttgggtgatggttcacgtagtgggccatcgc

FIG. 4A-1

cctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgtt
ccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagggattttgccg
atttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaa
tattaacgtttacaattttatggtgcactctcagtacaatctgctctgatgccgcatagttaag
ccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatcc
gcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcac
cgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataat
aatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgttta
tttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaat
aatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgc
ggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagat
cagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagtt
ttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtatt
atcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttg
gttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgca
gtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggacc
gaaggagctaaccgcttttttgcacaacatggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaa
caacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaataga
ctggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggttt
attgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccag
atggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacg
aaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtt
tactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaaga
tccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga
ccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttg
caaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactcttt
ttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgta
gttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtta
ccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttac
cggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaac
gacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaaggg
agaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttc
cagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctttta
cggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctg
tggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcg
cagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctctccccgcgcgt
tggccgattcatt

FIG. 4A-2

tttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattagg
caccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataaca
atttcacacaggaaacagctatgaccatgattacgccaagctccttcctcttccagcccttcct
ctttccagtcagtcagtcagtggagactgtcgtgcggccgcgtggccttcatggctgtcctgcg
ctttccagggtcagagtgagttgactgcaccaatataggtaacttccagaatacacagagtagg
tgctacatccagacacagagccaccacagccaacagtcacagaaggggtgggaggtccctgaac
ttacctgggtcacctgtcctagtgcagaaggatgctaagcttcaaacccaagcatcatcctcat
ctgcccttgtttgtgtgtttgattgagacacaacgtcctgctcagctaagctacagtcagtttg
cattggcctctaagcctgggaagcatttccttggctgatgattggtgggggagggtacaggtca
ctgtgggcagttcagtgcctgggctgatggtcctgggtgctacaagagagcaatctgaacaacc
aatggagagcaaaccactaagcagtttagttcttgactctgggttcctgccttgaggtcctgcc
ctgactttactcagaggtggggcatgacttgagtattgtctgatcaagtatactgactgattca
tccccaagctgttcttggtcatggagtttttaagacggcaatataaataccttaaaacaccttg
tctctcactgggaccaggggctgggcaatttgactgcaggcagccagcaagcctcacagatccc
tgtgtctctaactccacggtcctgggattatagacatgtgacgccccgcacctaactttcttca
tggatggaagtcaaatcctcctggttacgaggtaagtacatcaccaatggaggcatctccccag
ttcctgtatatttttaagacggtctttatagctcaggttggcctggaacttgtggcaacactcc
tgcctcagcctccagaatgctaggattgtaggtgtgaaccaccatccctggctaaacatgccat
gttttacagttcaggatctggaagggaggtgattcccagcaagggctgcttcccaaggcacctc
atcagcctgcctcaggtataacttcgtatagcatacattatacgaagttatactctacagaaga
aatcctgttccccgcgcctccccataggagcaggagaaaaacagaggtttacccatgtgcagga
tccagggcaatggcccttggctggtccaggtcctgccagaagagaaccttacgggacgtcccat
tgaggttggcaacctcaatgcggttggtctcggagtccgtccagtacagcttcttgccaaccca
gtcacaggccaggccatcaggtgacacgaggcccgagatgacaatgttctgtgcagcagctcca
gtctggttcaggtaggtctgtttgatggcctcctcgctcacgtcggtccaataaacggctccct
ttgagaactggaagtctacagcagctgcatcctccaggccactggccacaatggtggactccag
cttcactccgccggcatccactagccgcacatcccggcggttggcaaacaacaggagcggtgag
gctgtgggcagagcacagggtcaggatggctaagcaggacaggaaccagctcccagggaaagag
aatcatgcacggataacttcgtatagcatacattatacgaagttatctaccccgggcccctgct
cccaagactccttccttccagttgcccaaaggctcccttagtgataggccagtatagagaggca
ggaaggctccccaaatccccaagctctggctctactgctccccaaactggaaagaccacttcct
cacctcaacttgctttcagtagttcaggtgtgctcttcaaaggcacatgggaggggagggggca
ggggcagggacagggaggctgggggtggatttgtaacctaggactttttttttctccaaccctc
actgggagatggtggggcagtgctcccagctctgaaaagtaaagaaaccacacacttaatccct
aagttttaaggggaacccacagcatactcctggaaagctagctttccttctaagaggcctggca
gtggccccagggacccacagtgaggcaggccagaatgtgggtggctccccaccctcacctgggg
caggctgcttctcctctcctcccccgcctggggaatcctagacaacacagacactgtatcaccc
caaggctccaccagcctctccttttcaagggcagaagctgcaaaggaaagctcaggaggagaaa
aagaaagttctgggagtgaaagaaaaagagccacagtggcaaaaccaagatgccttcataggct
ggcaagcaggggctgcttgtaccaggggcctccaaggttcccgttccagtggagaaacagacac
cgggcaaagaccaaaagcaggtagtgacatctgtcaggttaatgctgtgaatgggccagagtca
gggacaggacaggagtgggaacagaacaggagaaggctgcacactacatacagtggttagggga
ggcctcattaaaacccatgcaaagttcccagggtagaaacacacagtgccagcctcagtcagga
ggctggtcctacgggttcctcccctaccaggtgacccatgcctctcaaatgcacagcccctggt
agactaggggtcctggtatacattctccaggcccagagctgcagccttcataaagaagcataca
cttgtgccaagtcacaccaagagtgatgttaggggttaagccccatggcccagcttgcgcggcc
gcaggtgtgtcttccagtcagtcagtcagtgaaagtcctctccactgactgtagcctccaattc
actggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcctt
gcagcacatcccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttccc
aacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtg
cggtatttcacaccgcatacgtcaaagcaaccatagtacgcgccctgtagcggcgcattaagcg
cggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcc
tttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgg
gggctcccttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgatttgg
gtgatggttcacgtagt

FIG. 4B-1

gggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtg
gactcttgttccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagg
gattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaat
tttaacaaaatattaacgtttacaattttatggtgcactctcagtacaatctgctctgatgccg
catagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgct
cccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttca
ccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatg
tcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccc
tatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataa
atgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattc
ccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaaga
tgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatc
cttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtg
gcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctca
gaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaaga
gaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacga
tcggaggaccgaaggagctaaccgcttttttgcacaacatggggatcatgtaactcgccttga
tcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgta
gcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggc
tggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagca
ctggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaacta
tggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtc
agaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatc
taggtgaagatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccact
gagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaat
ctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagcta
ccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctag
tgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgct
aatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaaga
cgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagct
tggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgct
tcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacg
agggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgac
ttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgc
ggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaa
cgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctct
ccccgcgcgttggccgattcattaatgcagctggcacgacagg

FIG. 4B-2

tttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattagg
caccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataaca
atttcacacaggaaacagctatgaccatgattacgccaagctccttcctcttccagcccttcct
ctttccagtcagtcagtcagtggagactgtcgtgcggccgcgtggccttcatggctgtcctgcg
ctttccagggtcagagtgagttgactgcaccaatataggtaacttccagaatacacagagtagg
tgctacatccagacacagagccaccacagccaacagtcacagaaggggtgggaggtccctgaac
ttacctgggtcacctgtcctagtgcagaaggatgctaagcttcaaacccaagcatcatcctcat
ctgcccttgtttgtgtgtttgattgagacacaacgtcctgctcagctaagctacagtcagtttg
cattggcctctaagcctgggaagcatttccttggctgatgattggtgggggagggtacaggtca
ctgtgggcagttcagtgcctgggctgatggtcctgggtgctacaagagagcaatctgaacaacc
aatggagagcaaaccactaagcagtttagttcttgactctgggttcctgccttgaggtcctgcc
ctgactttactcagaggtggggcatgacttgagtattgtctgatcaagtatactgactgattca
tccccaagctgttcttggtcatggagtttttaagacggcaatataaataccttaaaacaccttg
tctctcactgggaccaggggctgggcaatttgactgcaggcagccagcaagcctcacagatccc
tgtgtctctaactccacggtcctgggattatagacatgtgacgccccgcacctaactttcttca
tggatggaagtcaaatcctcctggttacgaggtaagtacatcaccaatggaggcatctccccag
ttcctgtatatttttaagacggtctttatagctcaggttggcctggaacttgtggcaacactcc
tgcctcagcctccagaatgctaggattgtaggtgtgaaccaccatccctggctaaacatgccat
gttttacagttcaggatctggaagggaggtgattcccagcaagggctgcttcccaaggcacctc
atcagcctgcctcaggtataacttcgtatagcatacattatacgaagttatactctacagaaga
aatcctgttccccgcgcctccccataggagcaggagaaaaacagaggtttacccatgtgcagga
tccagggcaatggcccttggctggtccaggtcctgccagaagagaaccttacgggacgtcccat
tgaggttggcaacctcaatgcggttggtctcggagtccgtccagtacagcttcttgccaaccca
gtcacaggccaggccatcaggtgacacgaggcccgagatgacaatgttctgtgcagcagctcca
gtctggttcaggtaggtctgtttgatggcctcctcgctcacatctgtccagtacacagcaccct
tggagaactggaagtctacagcagctgcatcctccaggccactggccacaatggtggactccag
cttcactccgccggcatccactagccgcacatcccggcggttggcaaacaacaggagcggtgag
gctgtgggcagagcacagggtcaggatggctaagcaggacaggaaccagctcccagggaaagag
aatcatgcacggataacttcgtatagcatacattatacgaagttatctaccccgggcccctgct
cccaagactccttccttccagttgcccaaaggctcccttagtgataggccagtatagagaggca
ggaaggctccccaaatccccaagctctggctctactgctccccaaactggaaagaccacttcct
cacctcaacttgctttcagtagttcaggtgtgctcttcaaaggcacatgggaggggagggggca
ggggcagggacagggaggctgggggtggatttgtaacctaggactttttttttctccaaccctc
actgggagatggtggggcagtgctcccagctctgaaaagtaaagaaaccacacacttaatccct
aagttttaaggggaacccacagcatactcctggaaagctagctttccttctaagaggcctggca
gtggccccagggacccacagtgaggcaggccagaatgtgggtggctccccaccctcacctgggg
caggctgcttctcctctcctcccccgcctggggaatcctagacaacacagacactgtatcaccc
caaggctccaccagcctctccttttcaagggcagaagctgcaaaggaaagctcaggaggagaaa
aagaaagttctgggagtgaaagaaaaagagccacagtggcaaaaccaagatgccttcataggct
ggcaagcaggggctgcttgtaccaggggcctccaaggttcccgttccagtggagaaacagacac
cgggcaaagaccaaaagcaggtagtgacatctgtcaggttaatgctgtgaatgggccagagtca
gggacaggacaggagtgggaacagaacaggagaaggctgcacactacatacagtggttagggga
ggcctcattaaaacccatgcaaagttcccagggtagaaacacacagtgccagcctcagtcagga
ggctggtcctacgggttcctcccctaccaggtgacccatgcctctcaaatgcacagcccctggt
agactaggggtcctggtatacattctccaggcccagagctgcagccttcataaagaagcataca
cttgtgccaagtcacaccaagagtgatgttaggggttaagccccatggcccagcttgcgcggcc
gcaggtgtgtcttccagtcagtcagtcagtgaaagtcctctccactgactgtagcctccaattc
actggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcctt
gcagcacatcccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttccc
aacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtg
cggtatttcacaccgcatacgtcaaagcaaccatagtacgcgccctgtagcggcgcattaagcg
cggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcc
tttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgg
gggctcccttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgatttgg
gtgatggttcacgtagt

FIG. 4C-1

gggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtg
gactcttgttccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagg
gattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaat
tttaacaaaatattaacgtttacaattttatggtgcactctcagtacaatctgctctgatgccg
catagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgct
cccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttca
ccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatg
tcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccc
tatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataa
atgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattc
ccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaaga
tgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatc
cttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtg
gcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctca
gaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaaga
gaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacga
tcggaggaccgaaggagctaaccgcttttttgcacaacatggggatcatgtaactcgccttga
tcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgta
gcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggc
tggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagca
ctggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaacta
tggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtc
agaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatc
taggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccact
gagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaat
ctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagcta
ccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctag
tgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgct
aatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaaga
cgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagct
tggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgct
tcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacg
agggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgac
ttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgc
ggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaa
cgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctct
ccccgcgcgttggccgattcattaatgcagctggcacgacagg

FIG. 4C-2


Donor 1 ATAACTTCGTATAATGTATGCTATACGAAGTTATccgtgcatgattctctttccctggga 60
Donor 2 ATAACTTCGTATAATGTATGCTATACGAAGTTATccgtgcatgattctctttccctggga 60
Donor 3 ATAACTTCGTATAATGTATGCTATACGAAGTTATccgtgcatgattctctttccctggga 60
Donor 1 gctggttcctgtcctgcttagccatcctgaccctgtgctctgcccacagCCTCACCGCTC 120
Donor 2 gctggttcctgtcctgcttagccatcctgaccctgtgctctgcccacagCCTCACCGCTC 120
Donor 3 gctggttcctgtcctgcttagccatcctgaccctgtgctctgcccacagCCTCACCGCTC 120
Donor 1 CTCCTCTTTGCTAACCGCCGAGATGTTCGACTCGTAGACGCTGGCGGCGTGAAGCTGGAG 180
Donor 2 CTGTTGTTTGCCAACCGCCGGGATGTGCGGCTAGTGGATGCCGGCGGAGTGAAGCTGGAG 180
Donor 3 CTGTTGTTTGCCAACCGCCGGGATGTGCGGCTAGTGGATGCCGGCGGAGTGAAGCTGGAG 180
Donor 1 AGTACAATCGTTGCTAGTGGACTGGAGGATGCTGCCGCTGTCGACTTCCAGTTCTCAAAG 240
Donor 2 TCCACCATTGTGGCCAGTGGCCTGGAGGATGCAGCTGCTGTAGACTTCCAGTTCTCAAAG 240
Donor 3 TCCACCATTGTGGCCAGTGGCCTGGAGGATGCAGCTGCTGTAGACTTCCAGTTCTCCAAG 240
Donor 1 GGAGCCGTTATTGGACCGACGTGAGCGAAGAGGCAATAAAACAGACATACCTTAACCAG 300
Donor 2 GGAGCCGTTATTGGACCGACGTGAGCGAGGAGGCCATCAAACAGACCTACCTGAACCAG 300
Donor 3 GGTGCTGTGTACTGGACAGATGTGAGCGAGGAGGCCATCAAACAGACCTACCTGAACCAG 300
Donor 1 ACTGGAGCTGCCGCCCAGAATATCGTAATCTCTGGCTTGGTGAGCCCTGATGGACTCGCT 360
Donor 2 ACTGGAGCTGCTGCACAGAACATTGTCATCTCGGGCCTCGTGTCACCTGATGGCCTGGCC 360
Donor 3 ACTGGAGCTGCTGCACAGAACATTGTCATCTCGGGCCTCGTGTCACCTGATGGCCTGGCC 360

FIG. 5A


Donor 1 TGTGATTGGGTCGGAAAAAAGCTTTATTGGACTGATAGTGAAACGAATCGGATAGAAGTA 420
Donor 2 TGTGACTGGGTTGGCAAGAAGCTGTACTGGACGGACTCCGAGACCAACCGCATTGAGGTT 420
Donor 3 TGTGACTGGGTTGGCAAGAAGCTGTACTGGACGGACTCCGAGACCAACCGCATTGAGGTT 420
Donor 1 GCGAACCTTAACGGTACATCAAGAAAGGTCCTGTTTGGCAGGATCTTGACCAGCCTAGA 480
Donor 2 GCCAACCTCAATGGGACGTCCCGTAAGGTTCTCTTCTGGCAGGACCTGGACCAGCCAAGG 480
Donor 3 GCCAACCTCAATGGGACGTCCCGTAAGGTTCTCTTCTGGCAGGACCTGGACCAGCCAAGG 480
Donor 1 GCAATCGCCCTCGACCCTGCACATGGgtaaacctctgttttctcctgctcctatgggga 540
Donor 2 GCCATTGCCCTGGATCCTGCACATGGgtaaacctctgttttctcctgctcctatgggga 540
Donor 3 GCCATTGCCCTGGATCCTGCACATGGgtaaacctctgttttctcctgctcctatgggga 540
Donor 1 ggcgcggggaacaggatttcttctgtagagtATAACTTCGTATAATGTATGCTATACGAA 600
Donor 2 ggcgcggggaacaggatttcttctgtagagtATAACTTCGTATAATGTATGCTATACGAA 600
Donor 3 ggcgcggggaacaggatttcttctgtagagtATAACTTCGTATAATGTATGCTATACGAA 600
Donor 1 GTTAT 605
Donor 2 GTTAT 605
Donor 3 GTTAT 605

FIG. 5B

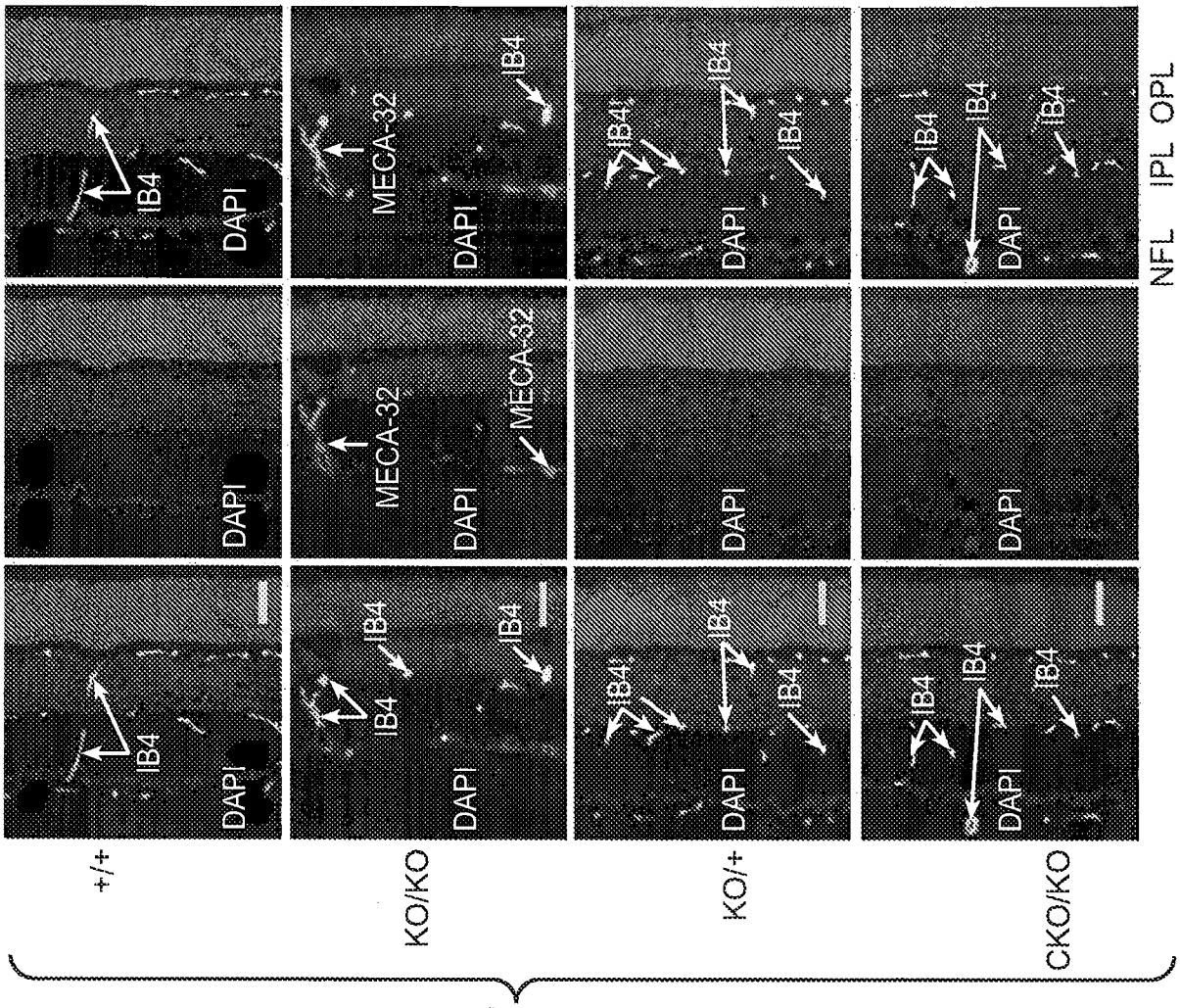


FIG. 6E

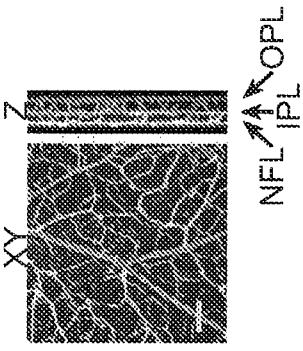


FIG. 6A +/+

FIG. 6B KO/KO

FIG. 6C KO/+

FIG. 6D CKO/KO

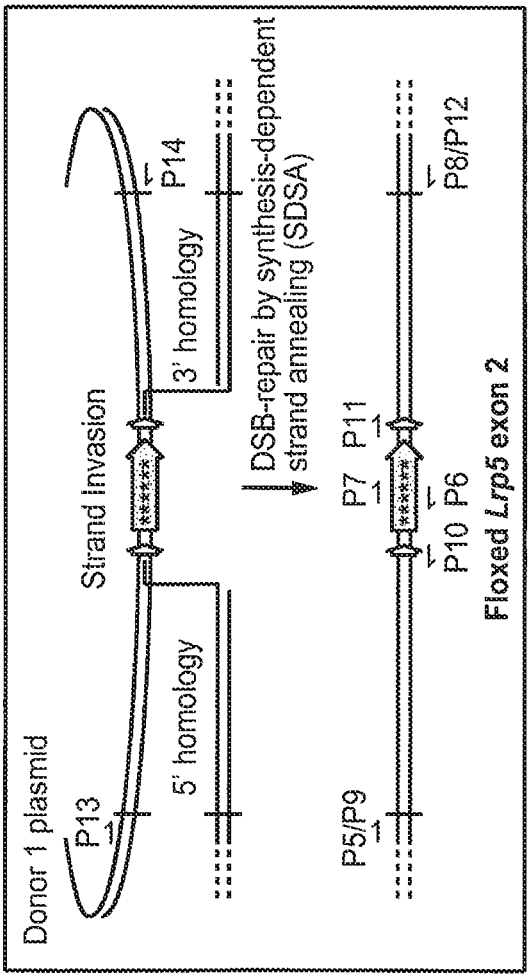


FIG. 7A

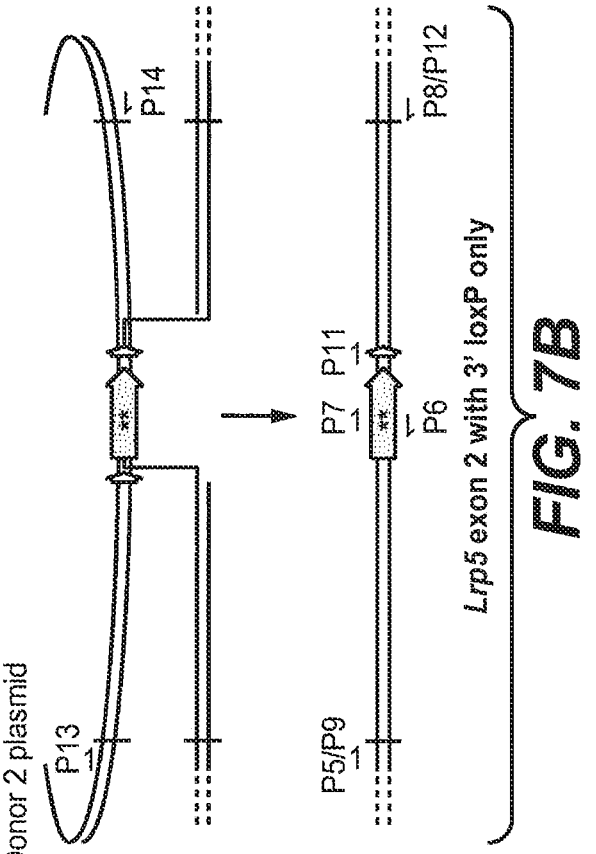


FIG. 7B

FIG. 7C

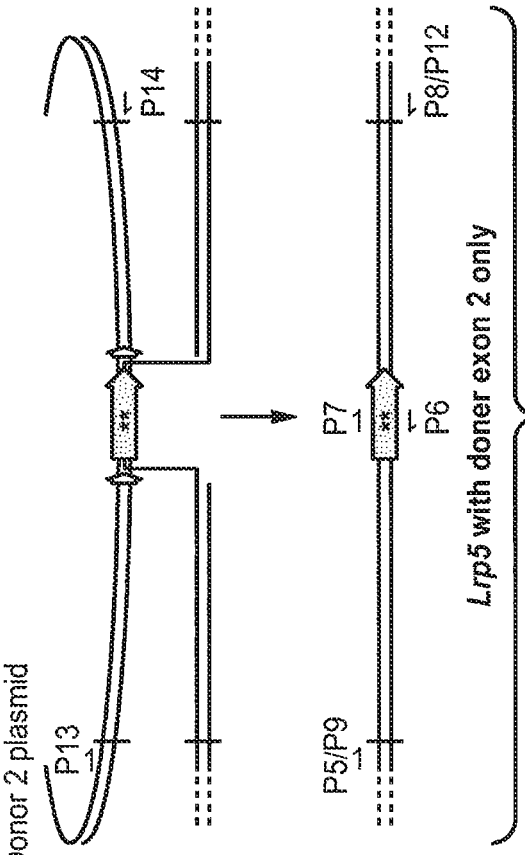


FIG. 7D

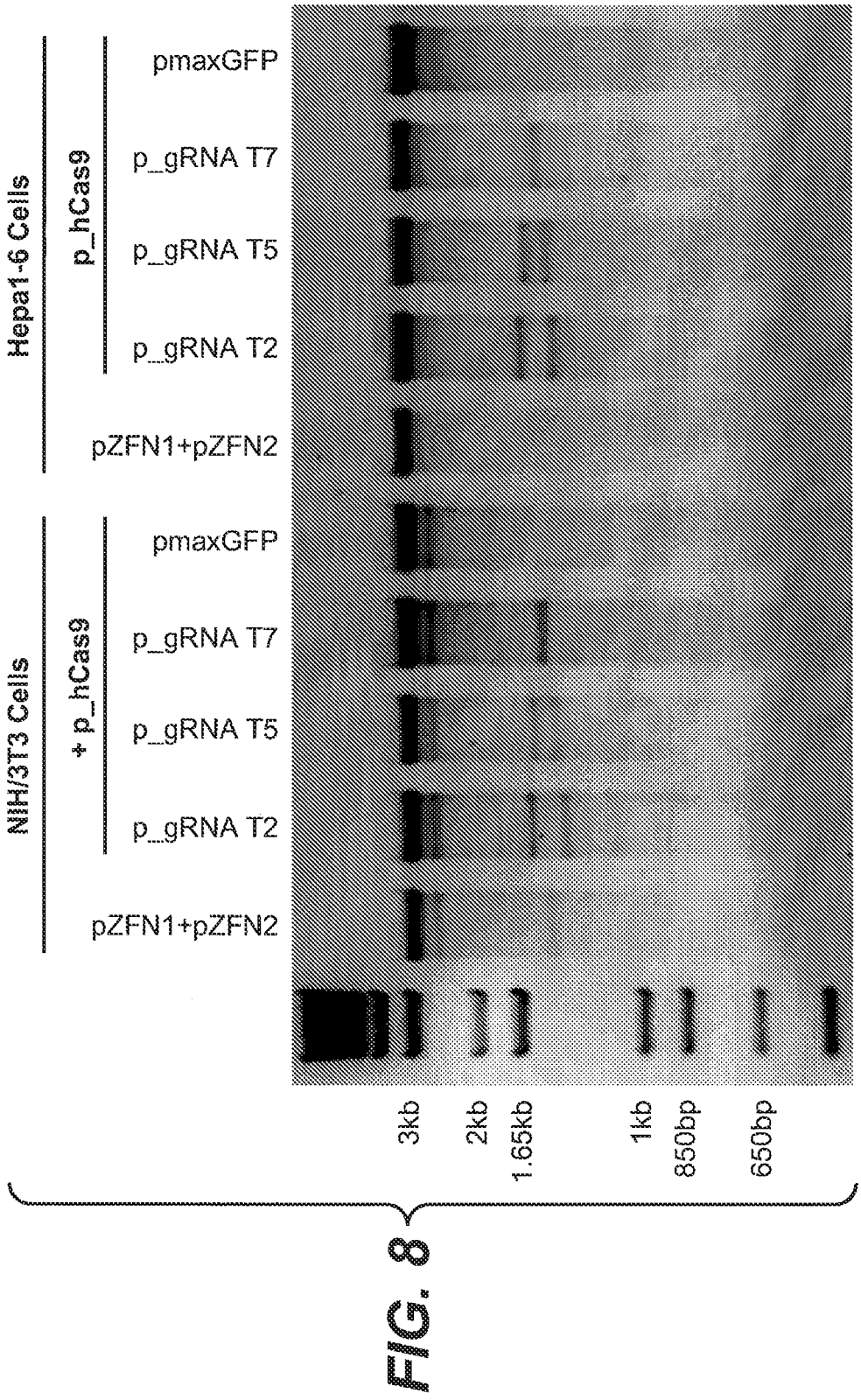

Hepa1-6 Cells
p_hCas9
pmaxGFP
p_gRNA T7
p_gRNA T5
p_gRNA T2
pZFN1+pZFN2
NIH/3T3 Cells
+ p_hCas9
pmaxGFP
p_gRNA T7
p_gRNA T5
p_gRNA T2
pZFN1+pZFN2
3kb
2kb
1.65kb
1kb
850bp
650bp

FIG. 8

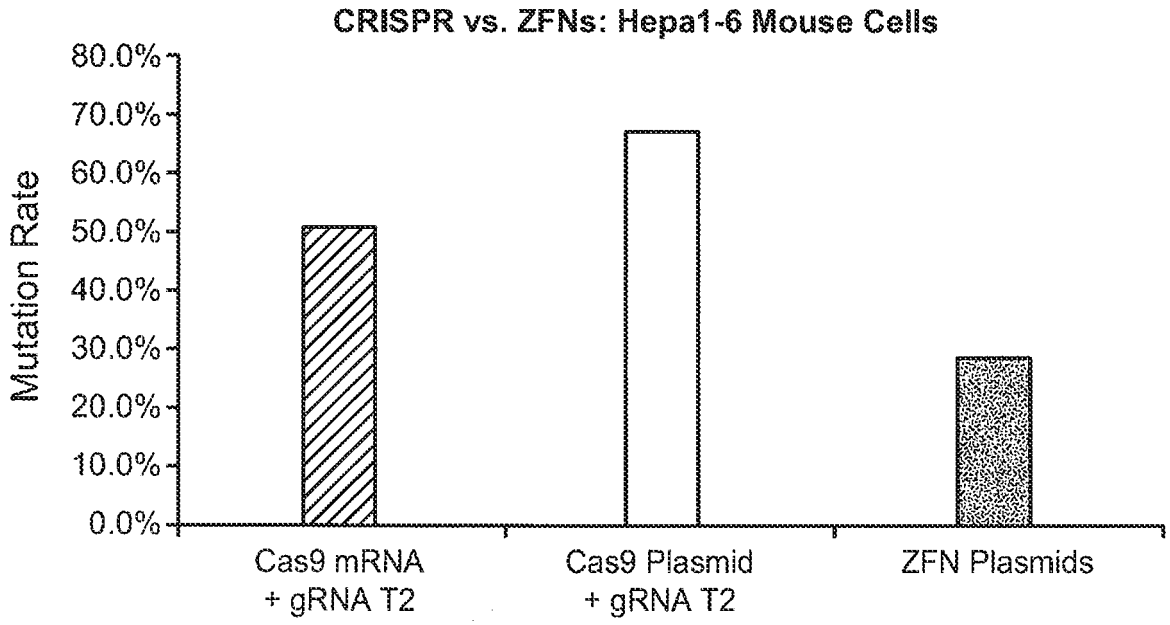

CRISPR vs. ZFNs: Hepa1-6 Mouse Cells
Mutation Rate
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
Cas9 mRNA + gRNA T2
Cas9 Plasmid + gRNA T2
ZFN Plasmids

FIG. 9A

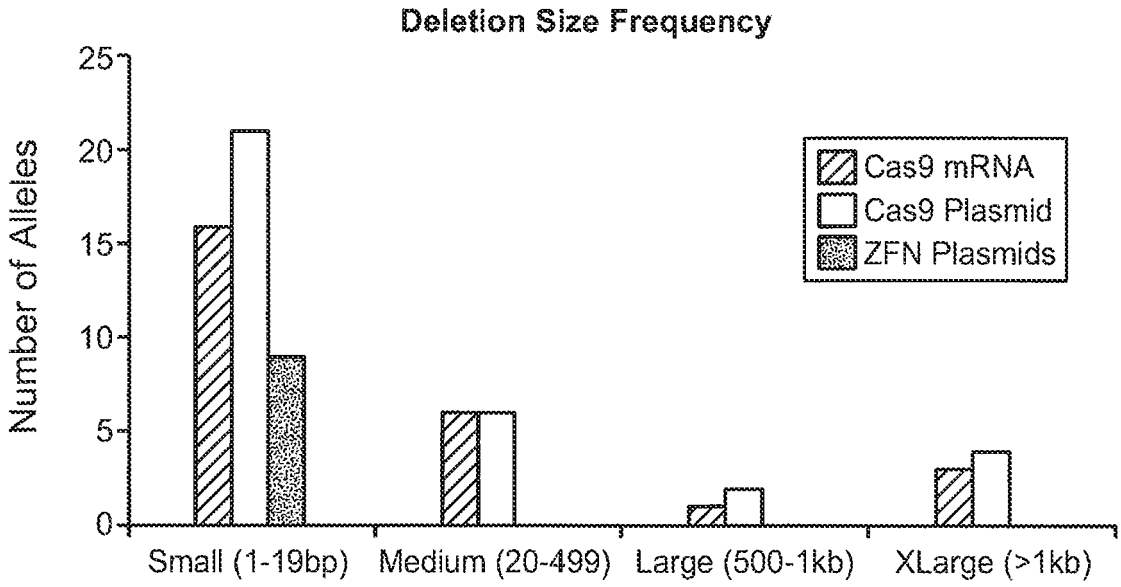

Deletion Size Frequency
Number of Alleles
25
20
15
10
5
0
Cas9 mRNA
Cas9 Plasmid
ZFN Plasmids
Small (1-19bp)
Medium (20-499)
Large (500-1kb)
XLarge (>1kb)

FIG. 9B

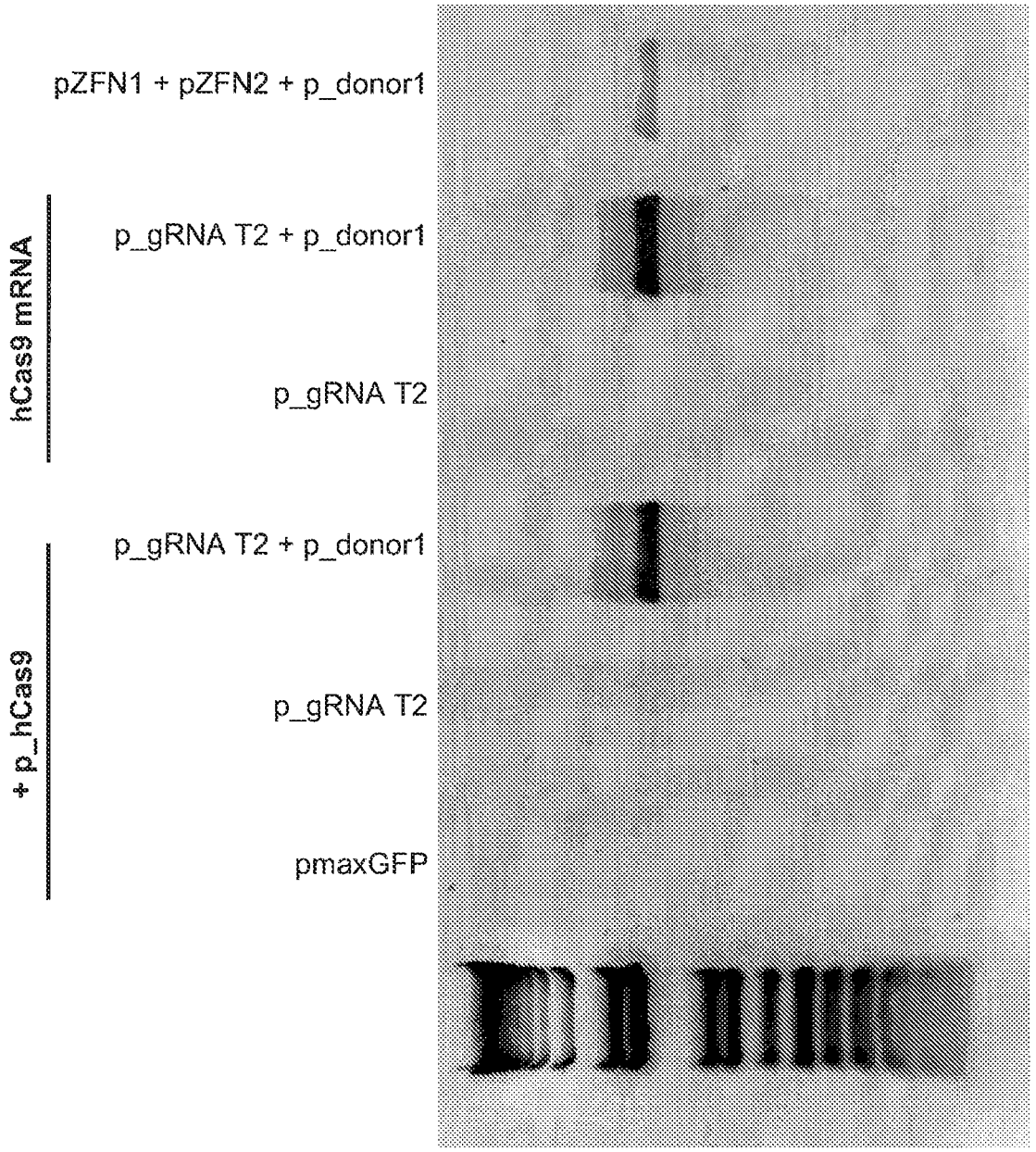

pZFN1 + pZFN2 + p_donor1
p_gRNA T2 + p_donor1
p_gRNA T2
p_gRNA T2 + p_donor1
p_gRNA T2
pmaxGFP
hCas9 mRNA
+ p_hCas9
2kb
1.65kb
1kb

FIG. 10

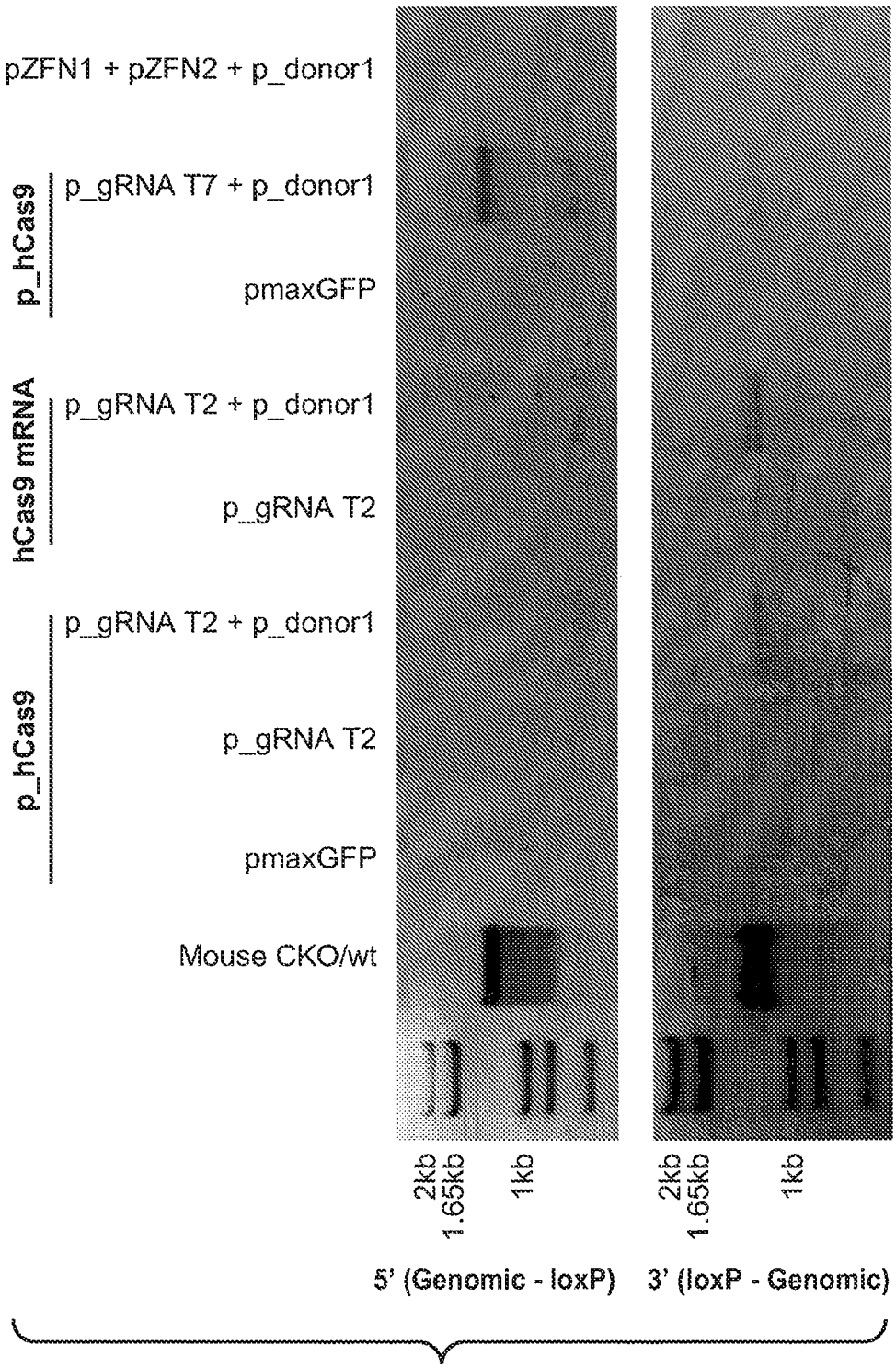

pZFN1 + pZFN2 + p_donor1
p_hCas9
p_gRNA T7 + p_donor1
pmaxGFP
hCas9 mRNA
p_gRNA T2 + p_donor1
p_gRNA T2
p_hCas9
p_gRNA T2 + p_donor1
p_gRNA T2
pmaxGFP
Mouse CKO/wt
2kb
1.65kb
1kb
2kb
1.65kb
1kb
5' (Genomic - loxP)
3' (loxP - Genomic)

FIG. 11

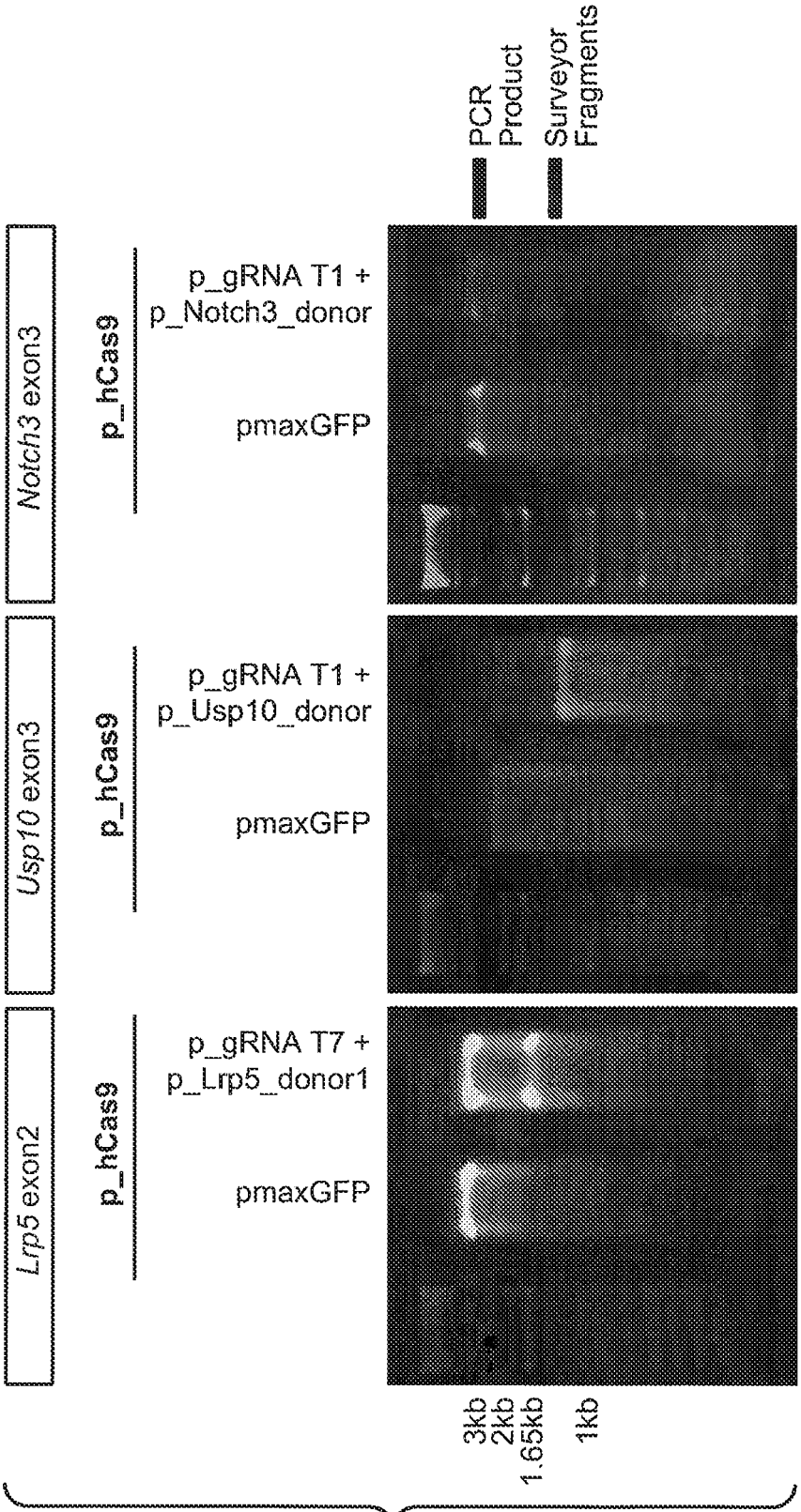
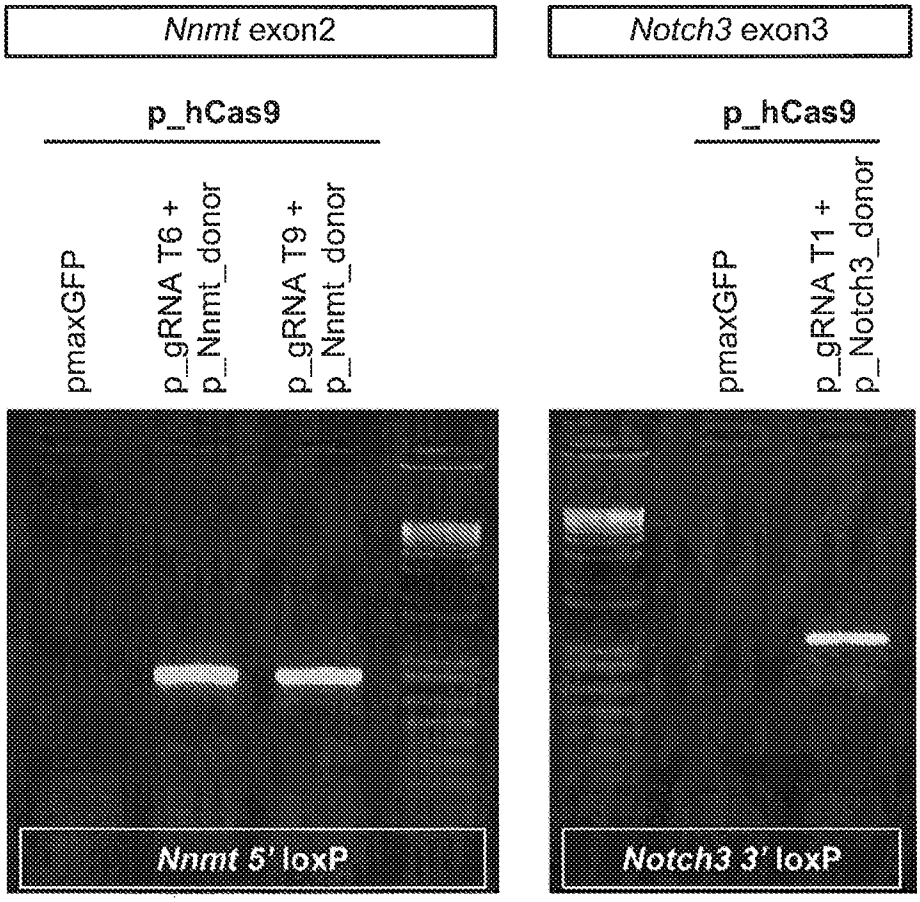

PCR Product
Surveyor Fragments
Notch3 exon3
p_hCas9
p_gRNA T1 + p_Notch3_donor
pmaxGFP
Usp10 exon3
p_hCas9
p_gRNA T1 + p_Usp10_donor
pmaxGFP
Lrp5 exon2
p_hCas9
p_gRNA T7 + p_Lrp5_donor1
pmaxGFP
3kb
2kb
1.65kb
1kb

FIG. 12

FIG. 13


Nnmt exon2
Notch3 exon3
p_hCas9
p_hCas9
pmaxGFP
p_gRNA T6 + p_Nnmt_donor
p_gRNA T9 + p_Nnmt_donor
pmaxGFP
p_gRNA T1 + p_Notch3_donor
Nnmt 5' loxP
Notch3 3' loxP

Lrp5 gRNA T2
ggtgcggctagtggatgccgggttttagagctagaaatagcaagttaaaataaggctagtccgt
tatcaacttgaaaaagtggcaccgagtcggtgctttttt

Lrp5 gRNA T5
ggagtccaccattgtggccaggttttagagctagaaatagcaagttaaaataaggctagtccgt
tatcaacttgaaaaagtggcaccgagtcggtgctttttt

Lrp5 gRNA T7
ggtactggacagatgtgagcggttttagagctagaaatagcaagttaaaataaggctagtccgt
tatcaacttgaaaaagtggcaccgagtcggtgctttttt

Usp10 gRNA T1
gctgaagatgaactgccagagttttagagctagaaatagcaagttaaaataaggctagtccgtt
atcaacttgaaaaagtggcaccgagtcggtgctttttt

Nnmt gRNA T6
gatctccgtgaaggactcacgttttagagctagaaatagcaagttaaaataaggctagtccgtt
atcaacttgaaaaagtggcaccgagtcggtgctttttt

Nnmt gRNA T9
gaccactggggaccagtcaagttttagagctagaaatagcaagttaaaataaggctagtccgtt
atcaacttgaaaaagtggcaccgagtcggtgctttttt

Notch3 gRNA T1
gaggcgtttgccagagttcaggttttagagctagaaatagcaagttaaaataaggctagtccgt
tatcaacttgaaaaagtggcaccgagtcggtgctttttt

hCas9 cDNA
atggacaagaagtacagcatcggcctggacatcggcaccaactctgtgggctgggccgtgatca
ccgacgagtacaaggtgcccagcaagaaattcaaggtgctgggcaacaccgaccggcacagcat
caagaagaacctgatcggcgccctgctgttcgacagcggagaaacagccgaggccacccggctg
aagagaaccgccagaagaagatacaccagacggaagaaccggatctgctatctgcaagagatct
tcagcaacgagatggccaaggtggacgacagcttcttccacagactggaagagtccttcctggt
ggaagaggataagaagcacgagcggcaccccatcttcggcaacatcgtggacgaggtggcctac
cacgagaagtaccccaccatctaccacctgagaaagaaactggtggacagcaccgacaaggccg
acctgcggctgatctatctggccctggcccacatgatcaagttccggggccacttcctgatcga
gggcgacctgaaccccgacaacagcgacgtggacaagctgttcatccagctggtgcagacctac
aaccagctgttcgaggaaaaccccatcaacgccagcggcgtggacgccaaggccatcctgtctg
ccagactgagcaagagcagacggctggaaaatctgatcgcccagctgccccggcgagaagaagaa
tggcctgttcggcaacctgattgccctgagcctgggcctgaccccaacttcaagagcaacttc
gacctggccgaggatgccaaactgcagctgagcaaggacacctacgacgacgacctggacaacc
tgctggcccagatcggcgaccagtacgccgacctgtttctggccgccaagaacctgtccgacgc
catcctgctgagcgacatcctgagagtgaacaccgagatcaccaaggcccccctgagcgcctct
atgatcaagagatacgacgagcaccaccaggacctgaccctgctgaaagctctcgtgcggcagc
agctgcctgagaagtacaaagagattttcttcgaccagagcaagaacggctacgccggctacat
cgatggcggagccagccaggaagagttctacaagttcatcaagcccatcctggaaaagatggac
ggcaccgaggaactgctcgtgaagctgaacagagaggacctgctgcggaagcagcggaccttcg

FIG. 14A

acaacggcagcatcccccaccagatccacctgggagagctgcacgccattctgcggcggcagga
agatttttacccattcctgaaggacaaccgggaaaagatcgagaagatcctgaccttccgcatc
ccctactacgtgggccctctggccaggggaaacagcagattcgcctggatgaccagaaagagcg
aggaaaccatcacccctggaacttcgaggaagtggtggacaagggcgccagcgcccagagctt
catcgagcggatgaccaacttcgataagaacctgcccaacgagaaggtgctgcccaagcacagc
ctgctgtacgagtacttcaccgtgtacaacgagctgaccaaagtgaaatacgtgaccgagggaa
tgagaaagcccgccttcctgagcggcgagcagaaaaaagccatcgtggacctgctgttcaagac
caaccggaaagtgaccgtgaagcagctgaaagaggactacttcaagaaaatcgagtgcttcgac
tccgtggaaatctccggcgtggaagatcggttcaacgcctccctgggcacataccacgatctgc
tgaaaattatcaaggacaaggacttcctggacaatgaggaaaacgaggacattctggaagatat
cgtgctgaccctgacactgtttgaggacagagagatgatcgaggaacggctgaaaacctatgcc
cacctgttcgacgacaaagtgatgaagcagctgaagcggcggagatacaccggctggggcaggc
tgagccggaagctgatcaacggcatccgggacaagcagtccggcaagacaatcctggatttcct
gaagtccgacggcttcgccaacagaaacttcatgcagctgatccacgacgacagcctgaccttt
aaagaggacatccagaaagcccaggtgtccggccagggcgatagcctgcacgagcacattgcca
atctggccggcagccccgccattaagaagggcatcctgcagacagtgaaggtggtggacgagct
cgtgaaagtgatgggccggcacaagcccgagaacatcgtgatcgaaatggccagagagaaccag
accacccagaagggacagaagaacagccgcgagagaatgaagcggatcgaagagggcatcaaag
agctgggcagccagatcctgaaagaacaccccgtggaaaacacccagctgcagaacgagaagct
gtacctgtactacctgcagaatgggcgggatatgtacgtggaccaggaactggacatcaaccgg
ctgtccgactacgatgtggaccatatcgtgcctcagagctttctgaaggacgactccatcgata
acaaagtgctgactcggagcgacaagaaccggggcaagagcgacaacgtgccctccgaagaggt
cgtgaagaagatgaagaactactggcgccagctgctgaatgccaagctgattacccagaggaag
ttcgacaatctgaccaaggccgagagaggcggcctgagcgaactggataaggccggcttcatca
agagacagctggtggaaacccggcagatcacaaagcacgtggcacagatcctggactcccggat
gaacactaagtacgacgagaacgacaaactgatccgggaagtgaaagtgatcaccctgaagtcc
aagctggtgtccgatttccggaaggatttccagttttacaaagtgcgcgagatcaacaactacc
accacgcccacgacgcctacctgaacgccgtcgtgggaaccgccctgatcaaaaagtaccctaa
gctggaaagcgagttcgtgtacggcgactacaaggtgtacgacgtgcggaagatgatcgccaag
agcgagcaggaaatcggcaaggctaccgccaagtacttcttctacagcaacatcatgaacttttt
tcaagaccgagattaccctggccaacggcgagatccggaagcggcctctgatcgagacaaacgg
cgaaacaggcgagatcgtgtgggataagggccgggactttgccaccgtgcggaaagtgctgtct
atgccccaagtgaatatcgtgaaaaagaccgaggtgcagacaggcggcttcagcaaagagtcta
tcctgcccaagaggaacagcgacaagctgatcgccagaaagaaggactgggaccctaagaagta
cggcggcttcgacagccccaccgtggcctattctgtgctggtggtggccaaagtggaaaagggc
aagtccaagaaactgaagagtgtgaaagagctgctggggatcaccatcatggaaagaagcagct
tcgagaagaatcccatcgactttctggaagccaagggctacaaagaagtgaaaaaggacctgat
catcaagctgcctaagtactccctgttcgagctggaaaacggccggaagagaatgctggcctct
gccggcgaactgcagaagggaaacgaactggccctgccctccaaatatgtgaacttcctgtacc
tggccagccactatgagaagctgaagggctcccccgaggataatgagcagaaacagctgtttgt
ggaacagcacaaacactacctggacgagatcatcgagcagatcagcgagttctccaagagagtg
atcctggccgacgctaatctggacaaggtgctgagcgcctacaacaagcacagagacaagccta
tcagagagcaggccgagaatatcatccacctgtttaccctgaccaatctgggagcccctgccgc
cttcaagtactttgacaccaccatcgaccggaagaggtacaccagcaccaaagaggtgctggac
gccaccctgatccaccagagcatcaccggcctgtacgagacacggatcgacctgtctcagctgg
gaggcgacgcctatccctatgacgtgcccgattatgccagcctgggcagcggctcccccaagaa
aaaacgcaaggtggaagatcctaagaaaaagcggaaagtggactga

FIG. 14B

CO_Usp10ex3 donor

cgtctgtcccctagtaccatccacagagcagacattctctgtaagggtgtcttgccctttttgtg
acctagagacgggatttggtgagcagccgtggtgatgttgctgtaccacactgcttctgtgtgc
tgtcagcgatggcatgtggaagagtttgcctgtgtcaaggtgctaagtcctaggctgccaggtc
ctgagtggttgggctcctttccagccatggttggtttttcttgctgtttgtcatcacctcctg
gggagagaacccagcactttatttttggggaaggaagtaaccttgcctcgttagaactgttgca
ttggctagccagagttcatctggagacacgagcctaggaggccacgggcacagcagacagtgat
tgccctgggcaccagctagagcctcctggactacttcacagatgctgcagcagggtggtgtttg
gggctcggttctgtcagctgactcacttctcagacacccccaaatcctaaataaataatagaat
tagcaggaagaagtatttgagaggttcataaaatttgccaggttttcacaaagtagccagacct
atgttgttttcttcctttcttcatacaaggaaggtgacgttaaaccattaactcatctgcatgg
cgcactgtaaaatgacagttaagggcccagccctttcttggggactctggctcacttctgccag
tggtgctgactcttcagttcccactgttaacagatggagtgttggggttgggatcatcctgcct
gcctgacactcaggcagtatcccagtctccgttgcagtgtatttagcagtggtaaacgattgaa
agcattcctacactgtttggttcagattggggcgaatctgtgtcttcctgggaacgtgatgtca
tggccagttgcttctgcagagcttagaggaagggttgagctggtgcagtggcttctgtcataac
ttcgtataatgtatgctatacgaagttatgggctttgataatctgttaagatttgtttctgact
tctctcgcagctccctccatacagcgggactctttgtagcatccaggctgaagacgaactgcca
gatggtaagccgggttgcatggactgggtgggcaggaaacctgcaaacatgtccataacttcgt
ataatgtatgctatacgaagttataacattgggtcatcttttctcaacacaaatttttatttat
tcatttattttggagtgtgcttttgtgtggaggtcagaggtcagctctctggaatatgtcctct
tctgcttttgcctgggtttggtgattggacagaggtcatcgggcttatgcagcaagtctcttgc
tgagctgtctccctgatctgtccctgttgactcccatcctttgtaggttagcaggctcaggctt
gtgcccagagctgcagtgtccgttcagccggaggcgcatttcattagagcttgaagcttctgct
ctgacctgcagctcacttaactcctgtcttacagtttccattgctgtggagagacaccatgacc
agagcagccagtatgaaggcaaacatttatttggggctggtttacaggttcagaggttcagtcc
gttatcatggagggaaacatggcagcatccgggcatgcatggcactgggagctgagaattctct
atgttgttctcaaggcaaacaggagaagtcttgcaagcagtgaggagggtctcatagcccaccc
gcacagtgccagttcctctaacaaggccacacctacccacagtgcctcggcccatgggccaagc
atactcaaaccaccacaactgccaagtcactcagattggcaggtgtgctttcaaggttgtgcca
cgctcactctgggcatggtaagttgtggtcaggagagaggtgctctagtctcctcagcaggccc
tggtgttgtgggtgtgtcatcccacatggctgtgctcaactctgccgggtccagggagttgtta
tagcagttggtgttcatgacccatgcctgggcatttaggcctccagcatgagccctggccgatt
tatgcataagtgtcatgtttatcaaggtgcagcagtgtttgttactgaggtgtgtttaagatat
cctggctatgcagt

CO_Nnmtex2 donor

gaaagtttaattctactcttatttgaaaggaatgtttgtttactcatgaagtaaaggacccccc
aatattaagaactgggtgatttgattattaattaatctcagtgttttaaattttaatatatata
catctaagcatatgtgtacatggtctatgcaggtacatgcatggggccagaagagggcaccaga
tgtatgcttccatcactctctccacctattcttgtggggcagcgtcccctcctaaacctggggc
ttgtgtgttgactagactggaagccagaaagcctcagagagcctctcctctctgctcctctcag
agatgggagcataggtatttgtgagatgcctgacttgttacacaggttctgggacctgaatgct
ggccctcatctctggaccatcactccaggccttaaaatcagttgttacacaacaacaacgacaa
caccaccaccaccacaacaaaaatggcacaatcccagtaaagacaataacttcgtataatgtat
gctatacgaagttatctataacttcagctctgcctgttccctttgatgcctgattccctgccc
ctcttctgtcctaggtgctgtaaaaggagaactcctcattgacatcggctctggcccaaccatc
taccagcttctgtctgcctgtgagtctttcacggagatcatcgtctctgactatacagaccaaa

FIG. 14C

acctctgggaactgcagaaatggctcaagaaggagccaggagcctttgattggtccccagtcgt
cacctatgtgtgcgatcttgaaggcaacaggtagaggaatgagtatctgctcttcaactttctg
aagggacctgcataacttcgtataatgtatgctatacgaagttatatgtttctatcagcccaaa
ctaccactaagatccagatgaaaaatgcaaaagaaactcaaagaccaaagggaactcgggatag
aagatgagggctcaggagaatagctgggttgtcagatgtaccaccagagtactatttggttacc
agtcatccaacacaaaggtggtggctaatgcctaatgcccacaccaccggaaacagcaaaacac
tggcagatgtatttttttagccaacagaaataacctgaagcaattatggggaaagaattcatta
cttaccatctataaagttagaatctttaggatctgggagacggtcagtgggtaaagtgcttgtc
acagaggtttgtcctttctccagtgtgtcctcttagcatctttattgaaaaccagctggctgcag
attcatggcttagtcctatactgtctattctgttccattattctgtgttctctggctggtgatg
ttttgtggaaagactttgctggtttgatgatatagtttttttgttgttgttgttgctgctgctg
ctgctgttgttgaagctaagtagtgtgtgatttctgatattttcttggtgctatttgggctttg
tgtgtacggttgcctaaaaattctagaattgttttctagtcctgtcattagtattttaagagga
atggcattggctacactaattgctttgggtagtatggacatattcatgatcttgattcttcaga
tcagcgaacagatctttccattttgatggtgtctgcactattcttctccattgatgttttcagt
gtaaagatcattggcttgttttagcaaaattcattttgaatttatatattatatttattaataa
atacattaataaaataatattaaatattacattttatttatttttaatagctattgcaatagga
ttgacttcttgatttctgaatcaaataattctctattggg

CO_Notch3ex3 donor
gacagaggcaggtaaattactgtgaattccagtctagccagcgatatgtagtgagaccctgcct
aaaatatattaaaaaaaaaaaaaaaaaggggaaggaagagtaaatggatttgtctgatagtctgt
ctggcacgagtgttgtttgataaacgcatcttgtgataacttcgtataatgtatgctatacgaa
gttatttatctgtctggcattgccatgcttttataccgtcccgaccacacatcttcccacaggt
gcctgccaggctgggttggtgagcgctgccagctcgaagacccctgccactccggcccttgtgc
cggccgaggcgtctgccagagctcagtggtggctggcaccgcacgattctcctgccggtgcctt
cgaggcttccaaggtgaaggggtgtgtctggacgggaaccctataacttcgtataatgtatgct
atacgaagttattggtaggcgagaatgtagtcagacccaagctcaccctctcctggttcttcca
ggcccagactgctcccagccagacccctgcgtcagcaggccctgtgttcatggtgccccctgct
cagtggggccggatggccgatttgcctgtgcctgcccacctggctaccagggtcaaagctgcca
aagtgacatagatgagtgccgatctggtacaacttgccgtcatggtggtacctgtctcaataca
cctggatccttccgctgccagtgtcctcttggttatacagggctgctgtgtgagaaccccgtag
tgccctgtgccccttccccgtgtcgtaatggtggcacctgtaggcagagcagtgatgtcacata
tgactgtgcttgccttcctggtaagtaagttgtgcccagggaaggcagctggggacaataggct
agcctcttagtgaccattgtcaccttgtcctcccctacgaggcttcgagggccagaactgtgaa
gtcaacgtggatgactgtcctggacatcggtgtctcaatgggggaacgtgtgtagacggtgtca
atacttacaactgccagtgccctccggagtggacaggtgggcatcagggctgcagagaaccagg
gtggctgacctcaggtgggcacacgggcaacttagactagcacatctttgtgccctaggccagt
tctgtacagaagatgtggatgagtgtcagctgcagcccaatgcctgccacaatgggggtacctg
cttcaacctactgggtggccacagctgtgtatgtgtcaatggctggacgggtgagagctgcagt
cagaatatcgatgactgtgctacagccgtgtgtttccatggggccacctgccatgaccgtgtgg
cctctttctactgtgcctgccctatggggaagacaggtgagtggccctttctttgtaggcaac
agaatggtttcagcatgaaaggt

FIG. 14D

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 December 2013 (19.12.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/188522 A3

(51) International Patent Classification:
A01K 67/00 (2006.01) *C12N 15/00* (2006.01)
A01K 67/027 (2006.01)

(21) International Application Number:
PCT/US2013/045382

(22) International Filing Date:
12 June 2013 (12.06.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/658,670 12 June 2012 (12.06.2012) US

(71) Applicant *(for all designated States except AL, AT, BE, BG, CH, CN, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IN, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR)*: **GENENTECH, INC.** [US/US]; 1 DNA Way, South San Francisco, California 94080 (US).

(71) Applicant *(for AL, AT, BE, BG, CH, CN, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IN, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR only)*: **F. HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacherstrasse 124, CH-4070 Basel (CH).

(72) Inventors: **WARMING, Soren**; c/o Genentech, Inc., 1 DNA Way, South San Francisco, California 94080 (US). **ANDERSON, Keith R.**; c/o Genentech, Inc., 1 DNA Way, South San Francisco, California 94080 (US).

(74) Agents: **VOM WEGE, Julia** et al.; Genentech, Inc., 1 DNA Way, MS-49, South San Francisco, California 94080 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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, 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))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

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

(88) Date of publication of the international search report:
10 April 2014

(54) Title: METHODS AND COMPOSITIONS FOR GENERATING CONDITIONAL KNOCK-OUT ALLELES

(57) Abstract: The disclosure provides methods and compositions for generating conditional knock-out alleles using donor constructs together with sequence-specific nucleases to generate conditional knock-out alleles. Specifically, the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region. Further disclosed are the donor sequences each comprises a target sequence having at least one neutral mutation. Different sequence-specific nucleases can be used with the donor constructs are further disclosed.

WO 2013/188522 A3

INTERNATIONAL SEARCH REPORT

013/045382 19.0

International application No.
PCT/US 13/45382

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A01K 67/00; A01K 67/027; C12N 15/00 (2013.01)
USPC - 800/13; 800/14, 800/18; 435/455
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - A01K 67/00; A01K 67/027; C12N 15/00 (2013.01)
USPC - 800/13; 800/14, 800/18; 435/455; 800/15, 800/16, 800/17

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
IPC(8) - A01K 67/00; A01K 67/027; C12N 15/00 (2013.01) - see keyword below
USPC - 800/13; 800/14, 800/18; 435/455; 800/15, 800/16, 800/17 - see keyword below

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST(USPT,PGPB,EPAB,JPAB); PatBase; Medline, Google: conditional knock-out, KO, homology, 5', 3', recognition site, recombinase, recombination, inducible, nuclease, cleavage, tissue-specific, promoter, zinc finger, ZFN, dimer, Cre, loxP, neutral, silent, mutation, stem cell, pluripotent, Cas9, TALEN, upstream, downstream, flank, allele, Lrp5,

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y --- A	US 2011/0023143 A1 (WEINSTEIN et al.) 27 January 2011 (27.01.2011), para [0005], [0006], [0013], [0014], [0019], [0056], [0060], [0064], [0066], [0067], [0069], [0078], [0079], [0085], [0088], [0089], [0091], [0100], [0104], and [0108]	1-2, 4, 7-11,13-22, 25-27, 30 --- 12, 28-29
Y -- A	XIE et al. Cloning-Independent and Counterselectable Markerless Mutagenesis System in Streptococcus mutans. Appl Environ Microbiol. 2011, Vol. 77(22), p. 8025-33. pg 8026, col 2, para 4; pg 8030, Fig 5, and pg 8032, col 1, top para	1-2, 4, 7-11,13-22, 25-27, 30 --- 12, 28-29
Y	US 2009/0136507 A1 (ALLEN et al.) 28 May 2009 (28.05.2009), para [0001], [0045], [0115], [0468], and [0589]	14
A	US 2009/0092578 A1 (SU et al.) 09 April 2009 (09.04.2009), para [0700], and SEQ ID NO: 86(3171 nt), the region between nucleotides 1-2974	12, 28-29
A	US 2007/0218071 A1 (MORRIS et al.) 20 September 2007 (20.09.2007), Abstract, para [0020], and SEQ ID NO: 21(108566 nt), the region between nucleotides 25612-28284	12, 28-29
A	GenBank_AC120144, Mus musculus chromosome 8, clone RP24-427K8, complete sequence, 15 October 2004 [online]. [Retrieved on 2012.11.21]. Retrieved from the Internet: <URL: http://www.ncbi.nlm.nih.gov/nuccore/AC120144> Source, and Origin: sequence the region between nucleotides 95807-93750	12, 28-29

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

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

Date of the actual completion of the international search	Date of mailing of the international search report
22 January 2014 (22.01.2014)	19 FEB 2014
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GenBank_AC113548, Mus musculus chromosome 9, clone RP23-268F15, complete sequence, 28 July 2004 [online]. [Retrieved on 2012.11.21]. Retrieved from the Internet: <URL: http://www.ncbi.nlm.nih.gov/nuccore/AC113548> Source, and Origin: sequence the region between nucleotides 59239-57476	12, 28-29
A	GenBank_AC090975, Mus musculus strain C57BL6/J chromosome 17 clone RP23-290I19, WORKING DRAFT SEQUENCE, 15 unordered pieces, 15 May 2002 [online]. [Retrieved on 2012.11.21]. Retrieved from the Internet: <URL: http://www.ncbi.nlm.nih.gov/nuccore/AC090975> Source, and Origin: sequence the region between nucleotides 93683-92321	12, 28-29
A	FRIEDEL et al. Generating conditional knockout mice. Methods Mol Biol. 2011, Vol. 693, p. 205-31. Abstract	1-2, 4, 7-22, 25-30

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

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

   a. (means)

   ☐ on paper

   ☒ in electronic form

   b. (time)

   ☒ in the international application as filed

   ☐ together with the international application in electronic form

   ☐ subsequently to this Authority for the purposes of search

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

2013/045382 19.02.2014

International application No.
PCT/US 13/45382

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

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

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

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

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

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

This International Searching Authority found multiple inventions in this international application, as follows:
This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I+, claims 1-30, drawn to methods of and compositions for generating a knock-out/conditional knock-out allele in a cell and animal. The first invention is restricted to the first named sequence-specific nuclease that is a zinc finger nuclease (ZFN)(claim 2), the first named recombinase recognition site that is a loxP site (claim 7), the first named recombinase cre (claim 22). Group I+ will be searched to the extent that it reads on ZFN, Cre/loxP, and the donor construct comprises the sequence of SEQ ID NO: 30, 31, 44, 45, or 46, without fee. It is believed that claims 1-2, 4, 7-17, 18-22, 25-30 read on this first named invention. Applicants must indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An exemplary election for (an) additional claim(s) to be searched would be: wherein the sequence-specific nuclease is a transcription activator-like effector nuclease (TALEN) (claim 3) and wherein the recombinase recognition site is an frt site and wherein the recombinase is FLP recombinase (claims 23), for paying an additional fee for the pair of election. **see extra sheet***

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

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

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

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-2, 4, 7-22, 25-30

Remark on Protest

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

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

☐ No protest accompanied the payment of additional search fees.

Continuation of:
Box No III (unity of invention is lacking)

The inventions listed as Group I+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special Technical Feature
Among Group I+, zinc finger nuclease (ZFN) in the first named invention is structurally and functionally different from all other nucleases as it targets to different sequence-specific sites than TALEN or Cas9; and loxP recoombinase regognition site in the first named invention is also structurally different from all other recombination sites including frt and rox, each requires a different specific recombinase for catalyzing the recombination.

Common Technical Features
The inventions of Group I+ share the technical feature of (claim 20) a method of generating a knock-out /conditional knock-out animal, the method comprising the steps of:
a. introducing into the cell a donor construct, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence having at least one neutral mutation;
b. introducing into the cell a sequence-specific nuclease that cleaves a sequence within the target gene,
c. introducing the cell into a carrier animal to produce a transgenic animal from the transfected cell; and
d. breeding the conditional knock-out animal with a transgenic animal having a trans gene encoding a recombinase protein that catalyzes recombination at the 5' and 3' recombinase recognition site.

However, these shared technical features do not represent a contribution over prior art as being obvious over US 2011/0023143 A1 to WEINSTEIN et al. (hereinafter 'Weinstein'), in view of an article entitled 'Cloning-Independent and Counterselectable Markerless Mutagenesis System in Streptococcus mutans' by XIE et al. (hereinafter 'Xie'; Appl Environ Microbiol. 2011, Vol. 77(22), p. 8025-33) as follows:

Weinstein teaches a method of generating a knock-out or conditional knock-out animal (para [0013]- "knock out" or a "conditional knock out." [0014]- [0019]), comprising generating a conditional knock-out allele in a cell comprising a target gene (para [0013] - 'a genetically modified ... animal cell ...edited chromosomal sequence encoding a neurodevelopmental protein- 'conditional knock out', wherein 'edited chromosomal sequence' is 'a target gene'; para [0014] - 'the chromosomal sequence may be modified ... mutation'; para [0006] - 'The genetic modifications ... include ... temporal-specific knockouts using loxP-flanked ("floxed") alleles ... Cre-recombinase.....or over-expression of alleles'), the method comprising the steps of:

a. introducing into the cell a donor construct, wherein the donor construct comprises a 5' homology region, a donor sequence, and a 3' homology region, wherein the donor sequence comprises a target sequence (para [0013]; para [0085] - 'editing chromosomal sequences ...introducing ...donor polynucleotide ... encoding a neurodevelopmental protein into... cell. A donor ... three components ...sequence encoding the protein is flanked by the upstream and downstream sequence, wherein the upstream and downstream sequences share sequence similarity with either side of the site of integration in the chromosome', wherein 'sequence encoding the protein is flanked by the upstream and downstream sequence...share sequence similarity with either side' are ' the donor construct comprises a 5' homology region, a donor sequence, and a 3' homology region', and wherein 'upstream and downstream sequences' are ' 5' homology region' and 'a 3' homology region', respectively; para [0091] - 'to construct a donor polynucleotide ...well-known standard recombinant techniques'); and

b. introducing into the cell a sequence-specific nuclease that cleaves a sequence within the target gene, thereby producing a conditional knock-out allele in the cell (para [0067] - 'the genetically modified ...cell ... using a zinc finger nuclease-mediated genome editing process. ... comprises:(a) introducing into ... cell at least one nucleic acid encoding a zinc finger nuclease that recognizes a target sequence ... able to cleave a site in the chromosomal sequence, and ... (i) at least one donor polynucleotide comprising a sequence for integration flanked by an upstream sequence and a downstream sequence that share substantial sequence identity with either side of the cleavage site'; para [0013] - 'a genetically modified ... cell ...edited chromosomal sequence...conditional knock out').

Weinstein further discloses the system used for generating a conditional knock-out allele in a cell is a zinc-finger system (para [0067] - 'the genetically modified ...cell ... using a zinc finger nuclease-mediated genome editing process.... cleavage site' and [0013]). Weinstein further discloses methods of generating a conditional knock-out allele in a cell include Cre-lox recombination system comprising a donor construct comprises a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site (para [0019] - 'A Cre-lox recombination system comprises a Cre recombinase enzyme ... to produce temporal and tissue specific expression ... generated with lox sites flanking a chromosomal sequence, such as a chromosomal sequence encoding a neurodevelopmental protein'; 'lox sites flanking a chromosomal sequence' is a sequence comprising 'a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site'; para [0006] - 'The genetic modifications ... include ... temporal-specific knockouts using loxP-flanked ("floxed") alleles ... Cre-recombinase.....or over-expression of alleles').

One of ordinary skill in the art at the time the invention was made would have been motivated to combine both Zinc-finger system and Cre-lox recmination system to generate a donor construct, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, because both systems are featured by flanking a donor sequence with sequences that specific for the respective system, that is to have 5' and 3' homology regions for Zinc-finger system, and 5' and 3' recombinase recognition sites for Cre-lox recombination system, in order to use a donor construct that can be regulated by two different systems for facilitating generating a conditional knock-out allele in a cell comprising target gene with the possibility to compare both systems in the same set of cells for achieving a desired result with expected success without undue experimentation.

Weinstein further discloses 5' homology region and 3' homology region sequences may share different percentage of homology with the targeted chromosomal sequence (para [0088] - 'The upstream and downstream sequences in the donor polynucleotide may share about 75%...100% sequence identity with the targeted chromosomal sequence').

**********Continued in the next extra sheet**********

Continuation of:
The previous extra sheet - Box No III (unity of invention is lacking)

Weinstein does not specifically teach wherein the donor sequence comprises a target sequence having at least one neutral mutation. Xie discloses a method of solving a problem caused by high rate of recombination of donor sequence and targeted chromosomal sequence by introducing at least one neutral mutation into the donor sequence comprising a target sequence to reduce the homology of a donor sequence with a chromosomal sequence (Pg 8029, col 2, para 1: By engineering a series of silent mutations in the remaining codons after the pheS314AG mutation site, we created a new pheS* cassette (mpheS) that still retained the amino acid sequence of the original cassette but exhibited much lower homology at the nucleotide level downstream of pheS314AG (Fig. 5A and B); pg 8030, Fig 5: a series of silent mutations were engineered in the region downstream of codon 314 to create the new mpheS negative-selection cassette. Three overlapping oligonucleotides containing the desired pheS silent mutations were synthesized. The oligonucleotides were mixed and subjected to overlapping PCR to produce an amplicon carrying the 3' portion of pheS; pg 8032, col 1, top para - 'high rate of recombination between the pheS* cassette and the chromosomal copy of pheS....The problem was finally solved by introducing a series of silent mutations downstream of the pheS* point mutation site to reduce the overall homology of the cassette with the wild-type pheS', wherein 'silent mutations' are 'neutral mutations'; Specification: pg 9, ln 15 - ' Examples of neutral mutations include silent mutations'). It would have been obvious to one of ordinary skill in the art at the time the invention was made to combine the teachings of Weinstein and Xie, to obtain a method of generating a conditional knock-out allele in a cell comprising a target gene, comprising the steps of: a. introducing into the cell a donor construct, wherein the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region, wherein the donor sequence comprises a target sequence; and b. introducing into the cell a sequence-specific nuclease that cleaves a sequence within the target gene, thereby producing a conditional knock-out allele in the cell, based on the teaching of Weinstein, and further wherein the donor sequence comprises a target sequence having at least one neutral mutation, based on the combination of Xie and Weinstein, in order to use the method taught by Xie to produce a series of donor constructs with different degree of reduced homology for selecting a desired homology of the donor sequence with the target sequence to obtain a donor construct that will produce a desired result with expected success without undue experimentation.

Weinstein further discloses c. introducing the cell into a carrier animal to produce a transgenic animal from the transfected cell (para [0066] - 'the cell may be a stem cell... embryonic stem cells'; para [0104] - 'an embryo may be cultured in vivo by transferring the embryo into the uterus of a female host? result in a live birth of an animal derived from the embryo'; para [0100] - 'methods of introducing the nucleic acids to the embryo or cell include ... calcium phosphate-mediated transfection, cationic transfection'; [0019]- 'a genetically modified animal is generated with lox sites flanking a chromosomal sequence, such as a chromosomal sequence encoding a neurodevelopmental protein'; Specification: pg 17: ln 19-20 - 'The cell is then introduced into a female carrier animal to produce the conditional knock-out animal from the cell').

d. breeding the conditional knock-out animal with a transgenic animal having a trans gene encoding a recombinase protein that catalyzes recombination at the 5' and 3' recombinase recognition site (para [0108]; para [0019] - 'The genetically modified animal comprising the lox-flanked chromosomal sequence encoding a neurodevelopmental protein may then be crossed with another genetically modified animal expressing Cre recombinase. Progeny animals comprising the lox-flanked chromosomal sequence and the Cre recombinase are then produced...leading to deletion or inversion of the chromosomal sequence encoding a neurodevelopmental protein', wherein 'be crossed with another genetically modified animal expressing Cre recombinase' is 'breeding the conditional knock-out animal with a transgenic animal having a trans gene encoding a recombinase protein that catalyzes recombination at the 5' and 3' recombinase recognition site'; para [0006] - 'The genetic modifications ... include ... temporal-specific knockouts using loxP-flanked ("floxed") alleles ... Cre-recombinase'; Please also see: Friedel et al.; Methods Mol Biol. 2011, Vol. 693, p. 205-31: Abstract - for how Cre works). Without a shared special technical feature, the inventions lack unity with one another.

Group I+ therefore lack unity under PCT Rule 13 because they do not share a same or corresponding special technical feature providing a contribution over prior art.

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104540382 A
(43)申请公布日 2015.04.22

(21)申请号 201380042657.8
(22)申请日 2013.06.12
(30)优先权数据
61/658,670 2012.06.12 US
(85)PCT国际申请进入国家阶段日
2015.02.11
(86)PCT国际申请的申请数据
PCT/US2013/045382 2013.06.12
(87)PCT国际申请的公布数据
WO2013/188522 EN 2013.12.19
(71)申请人 弗·哈夫曼-拉罗切有限公司
地址 瑞士巴塞尔
(72)发明人 S·瓦明 K·R·安德森
(74)专利代理机构 北京市中咨律师事务所 11247
代理人 凌立 黄革生
(51)Int.Cl.
A01K 67/00(2006.01)
A01K 67/027(2006.01)
C12N 15/00(2006.01)

权利要求书2页 说明书24页
序列表29页 附图25页

(54)发明名称

用于产生条件性敲除等位基因的方法和组合物

(57)摘要

本发明提供用序列特异性核酸酶产生条件性敲除等位基因的方法和组合物。

CN 104540382 A

1. 在包含靶基因的细胞中产生条件性敲除等位基因的方法，所述方法包括步骤：

a. 向所述细胞中引入供体构建体，其中供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区，其中供体序列包含具有至少一个中性突变的靶序列；和

b. 向所述细胞中引入切割所述靶基因内的序列的序列特异性核酸酶，从而在所述细胞中产生条件性敲除等位基因。

2. 权利要求 1 的方法，其中序列特异性核酸酶是锌指核酸酶 (ZFN)。

3. 权利要求 1 的方法，其中序列特异性核酸酶是转录激活剂样效应核酸酶 (TALEN)。

4. 权利要求 1 的方法，其中序列特异性核酸酶是仅切割所述靶基因一次的 ZFN 二聚体。

5. 权利要求 1 的方法，其中序列特异性核酸酶是 RNA 指导的核酸酶。

6. 权利要求 5 的方法，其中 RNA 指导的核酸酶是 Cas9。

7. 权利要求 1 的方法，其中序列特异性核酸酶作为蛋白质、mRNA 或 cDNA 引入。

8. 权利要求 1 的方法，其中重组酶识别位点是 loxP 位点、frt 位点或 rox 位点。

9. 权利要求 1 的方法，其中供体序列包含七个沉默突变。

10. 权利要求 1 的方法，其中供体序列和靶序列之间的序列同源性为 98% 或更小。

11. 权利要求 10 的方法，其中供体序列和靶序列之间的序列同源性为 78%。

12. 权利要求 1 的方法，其中供体构建体包含 SEQ ID NO:30、31、44、45 或 46 的序列。

13. 权利要求 1 的方法，其中 5' 同源区包含至少 1.1kb，且其中 3' 同源区包含至少 1kb。

14. 权利要求 1 的方法，其中靶基因选自 Lrp5、Usp10、Nnmt 和 Notch3。

15. 权利要求 1 的方法，其中细胞分离自哺乳动物。

16. 权利要求 15 的方法，其中哺乳动物选自小鼠、大鼠、兔、仓鼠、豚鼠、猫、狗、绵羊、马、牛、猴和人。

17. 权利要求 1 的方法，其中细胞是合子或多能干细胞。

18. 产生条件性敲除动物的方法，所述方法包括步骤：

a. 向包含靶基因的细胞中引入供体构建体，其中供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区，其中供体序列包含具有至少一个中性突变的靶序列；

b. 向所述细胞中引入序列特异性核酸酶，其中核酸酶切割所述靶基因；和

c. 将所述细胞引入载体动物中，以从所述细胞产生所述条件性敲除动物。

19. 权利要求 18 的方法，其中细胞是合子或多能干细胞。

20. 产生敲除动物的方法，所述方法包括步骤：

a. 向包含靶基因的细胞中引入供体构建体，其中供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区，其中供体序列包含具有至少一个中性突变的靶序列；

b. 向所述细胞中引入序列特异性核酸酶，其中核酸酶切割所述靶基因；

c. 将所述细胞引入载体动物中，以从转染的细胞产生转基因动物；和

d. 使所述条件性敲除动物与下述转基因动物交配，所述转基因动物具有编码重组酶蛋白质的转基因，所述重组酶蛋白质催化 5' 重组酶识别位点和 3' 重组酶识别位点处的重组。

21. 权利要求 20 的方法，其中细胞是合子或多能干细胞。

22. 权利要求 20 的方法，其中重组酶识别位点是 loxP 位点，且其中重组酶是 Cre 重组酶。

23. 权利要求 20 的方法，其中重组酶识别位点是 frt 位点，且其中重组酶是 FLP 重组酶。

24. 权利要求 20 的方法，其中重组酶识别位点是 rox 位点，且其中重组酶是 Dre 重组酶。

25. 权利要求 20 的方法，其中编码所述重组酶的转基因处于组织特异性启动子或诱导型启动子的控制下。

26. 用于产生靶基因的条件性敲除等位基因的组合物，其包含：

a. 包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区的供体构建体，其中供体序列包含与所述靶基因的序列相比具有至少一个中性突变的靶序列；和

b. 识别所述靶基因的序列特异性核酸酶。

27. 权利要求 26 的组合物，其中序列特异性核酸酶选自 ZFN、TALEN 和 RNA 介导的核酸酶。

28. 供体构建体，其包含 SEQ ID NO:30、31、44、45 或 46 的序列。

29. 细胞，其包含权利要求 28 的供体构建体。

30. 非人条件性敲除动物，其按照权利要求 18 的方法制备。

用于产生条件性敲除等位基因的方法和组合物

[0001] 相关申请的交叉参考

[0002] 本申请要求2012年6月12日提交的美国临时申请号61/658,670的权益，该美国临时申请的公开内容在此完整引入作为参考。

[0003] 序列表

[0004] 本申请包含序列表，该序列表已通过EFS-Web以ASCII格式提交，在此完整引入作为参考。于2013年6月12日创建的该ASCII拷贝命名为P4905R1WO_PCTSequenceListing.txt，大小为49,214字节。

技术领域

[0005] 本发明涉及产生遗传改造的条件性敲除等位基因的新方法。

背景技术

[0006] 单个基因表达的选择性抑制或增强已极大地辅助了体外和体内基因功能的研究。使用同源重组的鼠胚胎干(ES)细胞的基因靶向是用于操作鼠基因组的完善方法，且已允许创造就所研究的基因而言的无效或“敲除”小鼠。最近，条件性或诱导性敲除技术推进了在全身缺失时导致胚胎或围产期致死率的基因的研究(例如Lakso,M.等,Proc.Natl.Acad.Sci.USA89:6232-36(1992);Jacks,T.等,Nature 359:295-300(1992))。条件性敲除小鼠还可以用来研究在特定组织中选择性缺失基因的作用，同时保持其在其他组织中的功能完整。但是，用于产生条件性敲除动物的常规方法费力、低效，且需要得到胚胎干细胞。

[0007] 改造的序列特异性核酸酶已用于产生敲除等位基因。这类序列特异性内切核酸酶的实例包括锌指核酸酶(ZFN)，其由与内切核酸酶效应结构域融合的序列特异性DNA结合结构域组成(Porteus,M.H.和Carroll,D.,Nat.Biotechnol.23,967-973(2005))。序列特异性核酸酶的另一实例是转录激活剂样效应核酸酶(TALEN)，其由与TAL效应蛋白融合的核酸酶结构域组成(Miller,J.C.等,Nat.Biotechnol.29,143-148(2011);Cermak,T.等,Nucleic Acid Res.39,e82(2011))。序列特异性内切核酸酶是自然界中的模块，通过排列一个或多个模块来获得DNA结合特异性。例如，ZFN中的锌指结构域各识别三个碱基对(Bibikova,M.等,Mol.Cell.Biol.21,289-297(2001))，而TALEN中的单个TAL结构域各通过唯一密码识别一个碱基对(Boch,J.等,Science 326,1509-1512(2009))。序列特异性核酸酶的另一实例包括RNA指导的DNA核酸酶，例如CRISPR/Cas系统。

[0008] 已用ZFN、TALEN和最近的CRISPR/Cas介导的基因编辑来有效和直接地产生基因敲除等位基因(Geurts,A.M.等,Science 325,433(2009);Mashimo,T.等,PLoS ONE 5,e8870(2010);Carbery,I.D.等,Genetics 186,451-459(2010);Tesson,L.等,Nat.Biotech.29,695-696(2011))。认为敲除等位基因由内切核酸酶介导的双链断裂(DSB)的易错非同源末端连接(NHEJ)产生。

[0009] 最近，ZFN在小鼠和大鼠二者中成功地用于通过所靶向的染色体基因座与供体DNA的同源重组来靶向插入(敲入)报道基因(Meyer,M.等,Proc.Natl.Acad.Sci.USA

107,15022 - 15026(2010) ;Cui,X. 等 ,Nat.Biotechnol.29(1),64-67(2010))。已提出供体序列的序列特异性插入经由通过供体和发生双链断裂的基因座之间的同源重组进行的双链断裂修复的合成依赖性链退火 (SDSA) 模型而发生 (Moehle,E.A. 等 ,Proc Natl Acad Sci USA 104,3055 - 3060(2007))。根据此模型,内切核酸酶介导的双链断裂和链切除之后,单链染色体末端退火至存在于供体 DNA 上的同源区,然后用供体插入片段作为模板进行合成。

[0010] 尽管有这些进展,但本领域仍存在对产生条件性敲除等位基因和将此技术扩展至其他物种的新方法的需要。本发明满足此需要,并提供其他益处。

[0011] 发明概述

[0012] 本发明涉及用于产生条件性敲除等位基因的新方法和组合物。具体而言,本发明涉及用特异性供体构建体与序列特异性核酸酶一起来产生条件性敲除等位基因。

[0013] 在一方面,提供在包含靶基因的细胞中产生条件性敲除等位基因的方法。该方法包括步骤 :

[0014] 1. 向该细胞中引入供体构建体,其中该供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区,其中该供体序列包含具有至少一个中性突变的靶序列 ;和

[0015] 2. 向该细胞中引入切割靶基因内的序列的序列特异性核酸酶,从而在该细胞中产生条件性敲除等位基因。

[0016] 在某些实施方案中,该序列特异性核酸酶是锌指核酸酶 (ZFN)、ZFN 二聚体、转录激活剂样效应核酸酶 (TALEN) 或 RNA 指导的 DNA 内切核酸酶。在某些实施方案中,该序列特异性核酸酶仅切割靶基因一次。在某些实施方案中,将该序列特异性核酸酶作为蛋白质、mRNA 或 cDNA 引入细胞。

[0017] 在某些实施方案中,该重组酶识别位点是 loxP 位点、rox 位点或 frt 位点。在某些实施方案中,该供体序列包含 1、2、3、4、5、6、7、8、9、10、11 或 12 个中性突变。在某些实施方案中,该供体序列和该靶序列之间的同源性为 51-99%。在某些实施方案中,该供体序列和该靶序列之间的同源性为 78%。在某些实施方案中,该供体构建体包含图 4A 或图 4B 中所示的序列。在某些实施方案中,该 5' 同源区包含至少 1.1kb,且其中该 3' 同源区包含至少 1kb。在某些实施方案中,该靶基因是 Lrp5。

[0018] 在另一实施方案中,该细胞是哺乳动物细胞。在某些实施方案中,该哺乳动物细胞是小鼠、大鼠、兔、仓鼠、猫、狗、绵羊、马、牛、猴或人细胞。在某些实施方案中,该细胞来自非人动物。在某些实施方案中,该细胞是体细胞、合子或多能干细胞。

[0019] 在另一方面,提供产生条件性敲除动物的方法,该方法包括步骤 :

[0020] 1. 向包含靶基因的细胞中引入供体构建体,其中该供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区,其中该供体序列包含具有至少一个中性突变的靶序列 ;

[0021] 2. 向该细胞中引入序列特异性核酸酶,其中该核酸酶切割靶基因 ;和

[0022] 3. 将该细胞引入载体动物中,以从该细胞产生条件性敲除动物。

[0023] 在一些实施方案中,该动物是小鼠、大鼠、兔、仓鼠、豚鼠、狗、绵羊、猪、马、牛或猴。在某些实施方案中,该细胞来自非人动物。在一些实施方案中,该细胞是合子或多能干细

胞。

[0024] 在另一方面，提供产生敲除动物的方法，该方法包括步骤：

[0025] 1. 向包含靶基因的合子中引入供体构建体，其中该供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区，其中该供体序列包含具有至少一个中性突变的靶序列；

[0026] 2. 向该合子中引入序列特异性核酸酶，其中该核酸酶切割靶基因；

[0027] 3. 将该合子引入载体动物中，以从该合子产生条件性敲除动物；和

[0028] 4. 使该条件性敲除动物与具有编码催化 5' 和 3' 重组酶识别位点处的重组的重组酶的转基因的转基因动物交配，从而产生敲除动物。

[0029] 在某些实施方案中，该重组酶识别位点是 loxP 位点，该重组酶是 Cre 重组酶。在某些实施方案中，该重组酶识别位点是 frt 位点，该重组酶是 flippase。在某些实施方案中，该重组酶识别位点是 rox 位点，该重组酶是 Dre 重组酶。在某些实施方案中，该编码重组酶的转基因处于组织特异性启动子的控制下。

[0030] 在本发明的另一方面，提供用于产生靶基因的条件性敲除等位基因的组合物，其包含：

[0031] 1. 包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区的供体构建体，其中该供体序列包含具有至少一个中性突变的靶序列；和

[0032] 2. 识别该靶基因的序列特异性核酸酶。

[0033] 在某些实施方案中，该序列特异性核酸酶是 ZFN、ZFN 二聚体、ZFNickase、TALEN 或 RNA 指导的 DNA 内切核酸酶。在某些实施方案中，该重组酶识别位点是 loxP 位点、frt 位点或 rox 位点。

[0034] 在本发明的另一方面，提供包含图 4A(SEQ ID NO:30)、图 4B(SEQ ID NO:31) 或图 14C(SEQ ID NOS:44-46) 中所示的序列的供体构建体。

[0035] 在本发明的另一方面，提供包含含有图 4A(SEQ ID NO:30)、图 4B 或图 14C(SEQ ID NOS:44-46) 中所示的序列的供体构建体的细胞。在某些实施方案中，该细胞是哺乳动物细胞。在某些实施方案中，该哺乳动物细胞是小鼠、大鼠、兔、仓鼠、猫、狗、绵羊、马、牛、猴或人细胞。在某些实施方案中，该细胞来自非人动物。在某些实施方案中，该细胞是体细胞、合子或多能干细胞。

[0036] 在本发明的另一方面，提供按照本文所述的方法制备的非人条件性敲除动物。

[0037] 附图简述

[0038] 图 1 显示 ZFN 介导的突变体 Lrp5 等位基因在活产小鼠中的分布。缺失和插入的大小在 x 轴上以碱基对表示。复合 KO：具有同一基因的两个独立的突变体等位基因而检测不到该基因的野生型等位基因的动物；多等位基因：携带超过两个等位基因的嵌合动物；SKG->WTD：TCCAAGGGT(ZFN 切割位点下划线）的缺失。

[0039] 图 2A-E 显示在 Lrp5 中具有复合的符合读框和不符合读框的缺失的 2 月龄小鼠的血管表型。542：携带具有表现为沉默的 3bp 符合读框的缺失的等位基因和具有 1bp 不符合读框的缺失的等位基因的嵌合功能性杂合小鼠（对照）；495：携带 4bp 不符合读框的缺失等位基因和 1bp 不符合读框的缺失等位基因的小鼠；519：携带 29bp 不符合读框的缺失等位基因和 17bp 不符合读框的缺失等位基因的小鼠；555：携带 3bp 符合读框的缺失等位基

因和 1bp 不符合读框的缺失等位基因且是功能性杂合子的功能性杂合小鼠；FA：荧光血管造影术；IB4：同工凝集素 B4；NFL：神经纤维层；IPL：内网层；OPL：外网层。

[0040] 图 3A-B 显示从 Lrp5 外显子 2ZFN 和供体质粒的共显微注射或共电穿孔获得的条件性敲除等位基因。图 3A 显示通过合成依赖性链退火进行的双链断裂修复的示意图。箭头代表重组酶识别位点；步骤 1 中的大箭头代表靶序列；带星号的大箭头代表供体序列；星号代表中性突变；半箭头表示引物位置。图 3B 显示从幼犬的尾样品（左图）或 ES 细胞（右图）分离的 DNA 的聚合酶链反应 (PCR) 分析的结果。左侧显示用于分析的各引物对（引物位置如图 3A 中所示）。

[0041] 图 4A-C 显示以正确的方向用于质粒且具有掺入片段侧翼序列的供体序列（分别为 SEQ ID NOS:30-32，按出现的顺序）。

[0042] 图 5A-B 显示从 5' loxP 至 3' loxP 位点的三个 Lrp5CKO DNA 供体（分别为 SEQ ID NOS 33-35，按出现的顺序）的序列比对。大写黑体字母表示 loxP 位点；小写字母表示内含子序列；大写字母表示外显子 2（野生型或修饰）序列；虚线框表示 ZFN 结合位点；实线框表示沉默突变；下划线字母表示 ZFN 切割野生型外显子 2 处的序列。

[0043] 图 6A-E 显示携带密码子修饰的 Lrp5 条件性敲除等位基因的小鼠的正常视网膜表型。图 6A-D 显示用同工凝集素 B4 染色的视网膜全样载片的共焦投射（比例尺：50μm）。图 6E 显示用 IB4、MECA32 和 DAPI 染色的图 6A-D 中所示的那些的对侧眼的视网膜横切片。箭头指向所示的实例染色。+/+：野生型对照；KO/KO：Lrp5 纯合敲除；KO/+：Lrp5 杂合敲除；CKO/KO：Lrp5 条件性敲除 /Lrp5 敲除复合杂合；IB4：同工凝集素 B4；NFL：神经纤维层；IPL：内网层；OPL：外网层。

[0044] 图 7A-D 显示产生所观察到的供体衍生的 Lrp5 等位基因中的每一种的可能的机制的图示。显示与所得到的等位基因结合的引物。星号表示中性突变。

[0045] 图 8 显示将锌指对 (pZFN1+pZFN2) 或 Cas9(+pRK5-hCas9) 与靶向 Lrp5 外显子 2 的指导 RNA(p_gRNA T2、p_gRNA T5 或 p_gRNA T7) 或对照质粒 (PMAXGFP) 一起引入 NIH/3T3 细胞或 Hepa1-6 细胞后的 SURVEYOR 测定的结果。

[0046] 图 9A-B 显示 Hepa1-6 鼠肝癌细胞中 Lrp5 外显子 2 基因组基因座处的 gRNA/Cas9 突变率（图 9A）和缺失大小（图 9B）的总结。细胞接受靶向 Lrp5 的 gRNA 连同 mRNA(Cas9mRNA+gRNA T2，实心条）或质粒 (Cas9 质粒 +gRNA T2，空心条），或编码靶向 Lrp5 的外显子 2 的锌指对的两种质粒 (ZFN 质粒，灰色条）。

[0047] 图 10 显示使用对 CO 外显子 2 序列特异的正向引物和同源臂外侧的反向引物，以鉴定供体外显子在 Lrp5 基因座中的整合的 PCR 分析的结果。鼠 Hepa1-6 细胞接受编码 Cas9 的质粒 (pRK5-hCas9) 或 mRNA(hCas9mRNA) 连同指导 RNA 单独 (p_gRNA T2)、指导 RNA 和供体质粒 (p_gRNAT2+p_ 供体 1) 或对照质粒 (PMAXGFP)。一些细胞接受供体连同 Lrp5 锌指对 (pZFN1+pZFN2+p_ 供体 1)。

[0048] 图 11 显示使用检测 Lrp5 基因组基因座中的 5'（上，引物 P9 和 P10）和 3'（下，引物 P11 和 P12）loxP 位点整合的引物的 PCR 分析的结果。处理组如图 10 中所述。用来自杂合 Lrp5 条件性敲除（小鼠 CKO/wt）的 DNA 作为阳性对照。

[0049] 图 12 显示将 Cas9(p_hCas9) 与指导 RNA 和靶向 Lrp5(Lrp5 外显子 2；p_gRNA T7+p_Lrp5_ 供体 1)、Usp10(Usp10 外显子 3；p_gRNAT1+p_Usp10_ 供体 1) 或 Notch3(Notch3 外显

子 3 ;p_gRNA T1+p_Notch3_ 供体 1) 的各供体构建体一起引入 Hepa1-6 细胞后的 SURVEYOR 测定的结果。

[0050] 图 13 显示 Cas9/gRNA 和供体施用后，使用检测 Nnmt 外显子 2 基因组基因座中的 5' loxP 位点整合（左图，引物 P26 和 P27）或 Notch3 外显子 3 基因组基因座中的 3' loxP 位点整合（右图，引物 P25 和 P28）的引物的 PCR 分析的结果。

[0051] 图 14A-D 显示用于小鼠 Lrp5、Usp10、Nnmt 和 Notch3 基因组基因座的 Cas9/CRISPR 靶向的序列（分别为 SEQ ID NOS:36-46，按出现的顺序）。显示指导 RNA(gRNA) 的序列，对 Lrp5、Usp10、Nnmt 和 Notch3 特异的序列，及 Usp10、Nnmt 和 Notch3 的供体质粒序列。此外，显示用于哺乳动物表达的 Cas9cDNA 序列和体外转录 (mRNA)。

[0052] 发明详述

[0053] I. 定义

[0054] 为了解释本说明书的目的，以下定义将适用，无论何时适当，以单数形式使用的术语还将包括复数形式，反之亦然。在下文所示的任意定义与本文引入作为参考的任意文件冲突时，将以下文所示的定义为准。

[0055] 除非明确地另有说明，本文所用的术语“供体构建体”指包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区的多核苷酸。供体构建体可以进一步包含附加序列，如支持供体构建体的繁殖或含有该构建体的细胞的选择的序列。

[0056] 除非明确地另有说明，本文所用的术语“供体序列”指具有这样的序列的核酸，该序列与靶基因的序列的部分相比包含具有至少一个中性突变的靶序列。因此，供体序列包含这样的核酸，该核酸编码与靶基因的部分编码的多肽在功能上基本相似或不可区分的多肽。因此，供体序列可以在其在靶基因中的位置处取代靶基因的同源 (cognate) 部分而基本上不改变靶基因编码的蛋白质的功能特性。供体序列可以包含某些非编码序列，如内含子序列或调节序列。

[0057] 除非明确地另有说明，本文所用的术语“同源区”指供体构建体中与靶序列的侧翼核酸同源的核酸。

[0058] 除非明确地另有说明，本文所用的术语“重组酶识别位点”指供体构建体中具有重组酶识别的序列的核酸。

[0059] 除非明确地另有说明，本文所用的术语“重组酶”指这样的酶，该酶识别间插多核苷酸侧翼的特异性多核苷酸序列（重组酶识别位点），并催化交互链交换，导致间插多核苷酸的倒位或切除。

[0060] 除非明确地另有说明，本文所用的术语“靶基因”指编码细胞内的多肽的核酸。

[0061] 除非明确地另有说明，本文所用的术语“靶序列”指靶基因的部分，例如靶基因一个或多个外显子序列，内含子序列，或靶基因的调节序列，或靶基因的外显子和内含子序列、内含子和调节序列、外显子和调节序列、或外显子、内含子和调节序列的组合。

[0062] 除非明确地另有说明，本文所用的术语“序列特异性内切核酸酶”或“序列特异性核酸酶”指在特异性核苷酸序列处识别和结合多核苷酸（例如靶基因）并催化该多核苷酸中的单链或双链断裂的蛋白质。

[0063] 除非明确地另有说明，本文所用的术语“RNA 指导的 DNA 核酸酶”或“RNA 指导的 DNA 核酸酶”或“RNA 指导的内切核酸酶”指在特异性核苷酸序列处识别和结合指导 RNA 和

多核苷酸（例如靶基因）并催化该多核苷酸中的单链或双链断裂的蛋白质。

[0064] 除非明确地另有说明，本文所用的术语“条件性敲除等位基因”指包含侧翼为重组酶识别位点的多核苷酸序列但产生与同源野生型等位基因所产生的表型不可区分的表型的等位基因。

[0065] 除非明确地另有说明，本文所用的术语“中性突变”指供体序列中的突变，其降低供体序列和靶序列之间的总体同源性，但保持供体序列能够编码功能性多肽。中性突变的实例包括沉默突变，即改变核苷酸序列但不改变所编码的多肽序列的突变。中性突变的实例还包括保守突变，如点突变（例如取代）、插入和缺失，即改变核苷酸序列和所编码多肽序列但基本上不改变所产生的多肽的功能的突变。保守取代突变的实例在表 8 中显示。中性突变还可以包括沉默突变的组合、保守突变的组合或沉默突变和保守突变的组合。

[0066] 除非明确地另有说明，本文所用的术语“动物”指任意非人动物，包括但不限于饲养的动物（例如牛、绵羊、猫、狗和马）、灵长类（例如非人灵长类，如猴）、兔、鱼、啮齿类（例如小鼠、大鼠、仓鼠、豚鼠）和非脊椎动物（例如黑腹果蝇 (Drosophila melanogaster) 和秀丽隐杆线虫 (Caenorhabditis elegans)）。

[0067] 除非明确地另有说明，“分离的”核酸指已从其天然环境的成分分开的核酸分子。分离的核酸包括这样的核酸分子，该核酸分子包含在通常含有该核酸分子的细胞中，但该核酸分子存在于染色体外或存在于不同于其天然染色体位置的染色体位置处。

[0068] 除非明确地另有说明，“分离的编码蛋白质的核酸”指一个或多个编码蛋白质（或其片段）的核酸分子，包括单个载体或分开的载体中的这种（类）核酸分子，及存在于宿主细胞中的一个或多个位置处的这种（类）核酸分子。

[0069] 本文就供体和靶基因多核苷酸序列所用的术语“序列同源性”定义为在比对序列并根据需要引入缺口来达到最大百分比序列同一性之后，供体序列中与靶基因序列中的核苷酸残基相同的核苷酸残基的百分比。可以以本领域技术之内的多种方式达到以测定百分比核苷酸序列同源性为目的的比对，例如，使用公开可得的计算机软件，如 BLAST、BLAST-2、ALIGN、ClustalW2 或 Megalign(DNASTAR) 软件。本领域技术人员可以确定适当的参数用于比对序列，包括在所比较的序列全长内达到最大比对所需的任意算法。

[0070] II. 本发明的实施方案

[0071] 本发明部分涉及与用序列特异性内切核酸酶与侧翼有重组酶识别序列的供体序列的组合产生条件性敲除等位基因相关的技术挑战的识别和解决。此方法依赖于用识别和结合这类序列并在核酸分子中诱导双链断裂的内切核酸酶靶向核酸分子（如染色体）的特异性序列。双链断裂通过易错非同源末端连接或通过同源重组修复。如果反式提供用于同源重组的模板，则双链断裂可以用所提供的模板修复。与基于常规同源重组的基因靶向相比，最初的双链断裂使靶向的频率提高了几个数量级。基本上，此方法可以用来在修复位点处插入任意序列，只要它的侧翼是与双链断裂附近的序列同源的适当区域。但是，在应用于产生条件性敲除等位基因时，此方法与某些挑战相关。条件性敲除等位基因通常包括某些重组酶识别序列，如 loxP 位点，其位于基因或基因的部分的侧翼，但保持其功能完整，使得条件性敲除等位基因产生基本上类似于未修饰的等位基因的功能性多肽，但可以通过识别该识别序列的重组酶的存在来在某时间或在某些组织内使其无功能。

[0072] 与上述产生条件性敲除等位基因残基的方法相关的第一挑战在于这样的事实，在

序列特异性内切核酸酶催化的双链断裂之后，由于它们就彼此而言的序列同一性，不希望的重组可以发生在供体外显子和染色体（靶）外显子之间，而不是侧翼为重组酶识别序列的供体外侧的同源区。这将产生缺乏一个或两个重组酶识别序列的等位基因。第二挑战在于这样的事实，序列特异性内切核酸酶不仅可以识别和切割靶基因，还可以在供体外显子可以作为修复的模板发挥作用之前识别和切割它。本文所述的方法和组合物提供了这些挑战的解决方案。

[0073] A. 示例性方法

[0074] 在本发明的多种方面，提供在包含靶基因的细胞中产生条件性敲除等位基因的方法。该方法包括向具有靶基因的细胞中引入供体构建体和切割该靶基因内的序列但不抑制该供体构建体的功能的序列特异性核酸酶的步骤，从而在该细胞中产生条件性敲除等位基因。下文描述本发明的这些及其他方面。

[0075] 在本发明的具体方面中，通过向细胞中引入包含 5'同源区、5'重组酶识别位点、供体序列、3'重组酶识别位点和 3'同源区的供体构建体来在包含靶基因的细胞中产生条件性敲除等位基因。该供体序列包含具有至少一个中性突变的靶序列的序列。在某些实施方案中，除该至少一个中性突变外，该供体序列和该靶序列相同。中性突变意指供体序列的核苷酸序列中的任意突变，该突变降低供体序列和靶序列之间的同源性，但保持供体编码完整的功能性多肽的潜能。与野生型序列相比，中性突变减少供体序列和靶序列之间不产生条件性敲除等位基因的不希望的同源重组事件的数目（图 7B、C、D）。在一些实施方案中，中性突变还废除了序列特异性核酸酶与供体序列的结合。

[0076] 中性突变的实例包括沉默突变，即改变核苷酸序列但不改变所编码的多肽序列的突变。中性突变还包括保守突变，即改变核苷酸序列和所编码的多肽序列但基本上不改变所产生的多肽的功能的突变。这是例如用另一具有相似特性（大小、电荷等）的氨基酸取代一个氨基酸时的情况。例如，氨基酸可以按照共有的侧链特性分组：

[0077] (1) 疏水：正亮氨酸、Met、Ala、Val、Leu、Ile；

[0078] (2) 中性亲水：Cys、Ser、Thr、Asn、Gln；

[0079] (3) 酸性：Asp、Glu；

[0080] (4) 碱性：His、Lys、Arg；

[0081] (5) 影响链方向的残基：Gly、Pro；

[0082] (6) 芳族：Trp、Tyr、Phe。

[0083] 保守取代的实例在表 8 中显示。在某些实施方案中，供体序列包含 1、2、3、4、5、6、7、8、9、10、15、20 或 50 个沉默突变。在某些实施方案中，供体序列和靶序列之间的同源性是 99%、98%、95%、90%、85%、80%、78%、75%、70%、65%、60%、55%或 50%。在某些实施方案中，供体序列和靶序列之间的同源性小于 50%。可以引入任意数目的中性突变，该中性突变减少或抑制供体序列和靶序列之间（图 7B-D）而不是同源区和它们在所靶向的分子上的同源序列之间同源重组事件的数目，但保持供体序列编码功能性多肽的能力。在某些实施方案中，供体包含图 4A(SEQ ID NO:30)、图 4B(SEQ ID NO:31) 或图 14C(SEQ ID NOS:44-46) 中所示的序列。在某些实施方案中，至少一个中性突变废除序列特异性核酸酶与供体序列的结合。在某些实施方案中，几个中性突变沿供体序列的长度间隔开，以在供体序列中的任意位置处将连续的未修饰碱基对的数目减少至少于 20-100 碱基对。

[0084] 由于供体序列内的突变为中性，供体序列编码在功能上与靶序列编码的多肽基本相似或不可区分的多肽。肽或蛋白质的功能性可以通过本领域公知的方法来评估，如功能测定、酶测定和生物化学测定。供体序列可以在其在靶基因内的位置处取代靶序列而基本上不改变靶基因编码的多肽的功能特性。但是，一旦整合在靶基因中，则随后从靶基因去除供体序列可以导致靶基因编码的多肽的功能的改变、减少或丧失。

[0085] 在供体构建体内，重组酶识别位点处于供体序列的5’和3’侧翼。这些重组酶识别位点是供体构建体内由随后催化重组酶识别位点处的重组的重组酶识别的核酸序列。序列特异性重组为本领域公知，且包括重组酶介导的侧翼为重组酶识别位点的多核苷酸的序列特异性切割和连接。重组酶识别位点的实例包括loxP(locus of X-over P1)位点(Hoess等，Proc.Natl.Acad.Sci.USA 79:3398-3401(1982))、frt位点(McLeod,M.,Craft,S.&Broach,J.R.,Molecular and Cellular Biology 6,3357-3367(1986))和rox位点(Sauer,B.和McDermott,J.,Nucleic Acids Res 32,6086-6095(2004))。

[0086] 5’同源区定位在5’重组酶识别位点的5’或“上游”，且与在其核苷酸背景中处于靶序列侧翼的核酸同源。类似地，3’同源区定位在3’重组酶识别位点的3’或“下游”，且与靶序列侧翼的核酸同源。在一个实施方案中，同源区的长度超过30bp，优选几个kb。例如，同源区的长度可以是50bp、100bp、200bp、300bp、500bp、800bp、1kb、1.1kb、1.5kb、2kb和5kb。在某些实施方案中，5’同源区包含1.1kb，3’同源区包含1kb。同源区可以与靶基因的区域同源，以及（或相反）与靶基因上游或下游的区域同源。在一个实施方案中，同源区与紧邻靶序列的染色体区域同源。例如，在5’同源区的情况下，同源区与其最靠近3’的核苷酸紧邻靶序列的第一（最靠近5’）核苷酸的序列同源。在一个实施方案中，同源区与染色体上并非紧邻靶序列的染色体区域同源。在一些实施方案中，5’和3’同源区各与靶序列侧翼的同源核酸序列95-100%同源。

[0087] 总结上述成分排列，供体构建体按从5’至3’的顺序包含5’同源区、5’重组酶识别位点、供体序列、3’重组酶识别位点和3’同源区。供体构建体可以进一步包含某些提供结构或功能支持的序列，如质粒或其他支持供体构建体的繁殖的载体（例如pUC19载体）的序列。供体构建体还可以可选地包含某些选择标记或报道基因，其中一些可以由重组酶识别位点位于侧翼，用于随后的激活、失活或删除。位于可选的标记或报道基因侧翼的重组酶识别位点可以与位于供体序列侧翼的重组酶识别位点相同或不同。在某些实施方案中，用单类供体构建体来产生条件性敲除等位基因。

[0088] 与引入供体构建体同时或在引入供体构建体之后，向细胞中引入序列特异性核酸酶。序列特异性核酸酶识别和结合靶序列内的特异性序列，并在靶基因中引入双链断裂。如上文所述，通过至少一个中性突变来修饰供体序列，以减少不产生条件性敲除等位基因的同源重组事件。在某些实施方案中，序列特异性核酸酶仅切割靶基因一次，即在本文所述方法过程中在靶基因中引入单个双链断裂。

[0089] 序列特异性核酸酶的实例包括锌指核酸酶(ZFN)。ZFN是由DNA结合锌指蛋白结构域和效应核酸酶结构域组成的重组蛋白质。锌指蛋白结构域是例如与识别和结合特异性DNA序列的转录因子相关遍在蛋白质结构域。“指”结构域之一可以由包括与锌复合的不变的组氨酸残基的约三十个氨基酸组成。虽然迄今已鉴定出超过10,000个锌指序列，但已通

过锌指结构域中的靶向氨基酸取代产生设计用于识别目的特异性核苷酸序列的新锌指蛋白质来进一步扩展了锌指蛋白质的库。例如，已用噬菌体展示文库来针对希望的序列特异性筛选锌指组合文库（Rebar 等，Science 263:671-673(1994)；Jameson 等，Biochemistry 33:5689-5695(1994)；Choo 等，PNAS91:11163-11167(1994)，每篇文献在此以其整体引入）。然后可以将具有希望的序列特异性的锌指蛋白质连接至例如 6,824,978 中所述的效应核酸酶结构域，如 PCT 申请公开号 WO1995/09233 和 WO1994018313 中所述的 FokI，每篇文献在此完整引入作为参考。

[0090] 序列特异性核酸酶的另一实例包括转录激活剂样效应内切核酸酶（TALEN），其包含与特异性核苷酸序列结合的 TAL 效应结构域和催化靶位点处的双链断裂的内切核酸酶结构域。PCT 专利申请公开号 WO2011072246 描述了 TALEN 的实例及制备和使用的方法，该专利申请在此完整引入作为参考。

[0091] 可以用于本文所述的方法和组合物的序列特异性核酸酶系统的另一实例包括 Cas9/CRISPR 系统（Wiedenheft, B. 等 Nature 482,331 - 338(2012)；Jinek, M. 等 Science 337,816 - 821(2012)；Mali, P. 等 Science 339,823 - 826(2013)；Cong, L. 等 Science 339,819 - 823(2013)）。Cas9/CRISPR（成簇的规律间隔的短回文重复序列）系统利用 RNA 指导的 DNA 结合和靶序列的序列特异性切割。指导 RNA(gRNA) 包含 20 个与基因组 PAM（原间隔区邻近基序）位点（NNG）上游的的靶基因组 DNA 序列互补的核苷酸和恒定的 RNA 支架区。Cas(CRISPR 相关)9 蛋白质结合 gRNA，靶向该 gRNA 结合的 DNA，并在 PAM 位点上游确定的位置中引入双链断裂。Cas9 包含与 HNH 和 RuvC 内切核酸酶同源的两个独立的核酸酶结构域，可以通过突变两个结构域中的任一个来将 Cas9 蛋白质转化为引入单链断裂的切口酶（Cong, L. 等 Science 339,819-823(2013)）。尤其考虑的是，本发明的方法和组合物可以用于 Cas9 的单链或双链诱导形式，以及其他 RNA 指导的 DNA 核酸酶，如其他细菌 Cas9 样系统。在一些实施方案中，用于本文所述方法的指导 RNA 按出现的顺序分别是 SEQ ID NOS:36-42 的那些。本文所述方法和组合物的序列特异性核酸酶可以是改造的、嵌合的，或分离自生物。

[0092] 序列特异性核酸酶可以以蛋白质的形式或以编码序列特异性核酸酶的核酸（如 mRNA 或 cDNA）的形式引入序列特异性核酸酶。核酸可以例如通过电穿孔、脂质载体、病毒运载体、显微注射和生物射弹作为更大的构建体（如质粒或病毒载体）的部分递送或直接递送。类似地，供体构建体可以通过任意适合用于将核酸引入细胞的方法递送。

[0093] 不受限于任意具体机制或理论，序列特异性核酸酶在靶序列中引入双链断裂（例如 ZFN 诱导的 DSB；图 3A，步骤 1）后，链切除产生 3' 单链染色体末端（图 3A，步骤 2）。为了起始修复，单链染色体末端通过链侵入退火至存在于供体构建体上的同源区内的互补碱基对（图 3A，步骤 3）。然后用供体序列作为模板，通过 DNA 聚合酶介导的链延长来延长 3' 单链末端。链延长后，延长的链退火至最初的双链断裂的另一侧的单链染色体末端，并通过用延长的链作为模板的 DNA 合成和连接来完成修复。产生的双链 DNA 包含侧翼为重组酶识别位点的供体序列（图 3A，步骤 4）。

[0094] 双链断裂修复的此合成依赖性链退火模型与这样的观察一致，与内源序列具有很小的同源性或无同源性的外来 DNA 的非常大的序列（如报道基因）可以精确地插入双链断裂的点。因此，侧翼为重组酶识别位点的供体序列可以通过切除游离染色体末端暴露靶序

列周围与供体构建体上的同源区基本同源的区域而整合在双链断裂处（图 3A）。同源区可以具有适合用于供体构建体中的取代且在介导上文所述的链退火中有效的任意长度，例如 10-5000bp、100-1000bp、500-600bp 或 537bp 的组合长度。因此，这些步骤在靶基因的位点处产生条件性敲除等位基因，即产生与由同源靶基因等位基因产生的表型基本相似或不可区分的表型的包含侧翼为重组酶识别位点的供体序列的等位基因。如果由熟练的技术人员进行的标准检查不能检测到靶基因的潜在等位基因的性质，则两种表型基本相似或不可区分。在一些实施方案中，本文所述的方法产生携带杂合条件性敲除等位基因或纯合条件性敲除等位基因，即条件性敲除等位基因取代了并非全部或全部内源等位基因。

[0095] 靶基因可以是通过供体构建体靶向来产生该基因的条件性敲除形式的细胞的遗传物质内的任意编码蛋白质（或其片段）的核酸分子。例如，靶基因可以是定位在真核细胞的染色体上的编码未知功能的蛋白质或涉及细胞过程的蛋白质的基因。这类基因可以由一系列外显子和内含子组成。靶序列可以包括靶基因的外显子、内含子（包括人工内含子）或调节序列，或其多种组合。靶序列可以包括整个靶基因。

[0096] 细胞可以是任意真核细胞，例如动物的分离的细胞，如全能、多能或成体干细胞、合子、或体细胞。在某些实施方案中，用于本文所述方法的细胞是非人动物，如饲养的动物（例如牛、绵羊、猫、狗和马）、灵长类（例如非人灵长类，如猴）、兔、鱼、啮齿类（例如小鼠、大鼠、仓鼠、豚鼠）、两翼昆虫和蠕虫的细胞。在某些实施方案中，用于该方法的细胞是人细胞。本文所述的方法和组合物可以用来靶向任意基因组基因座。本文描述了靶向不同基因座的几个具体实例。在某些实施方案中，本文所述的方法和组合物可以用来靶向细胞内的一个以上基因组基因座，即用于多重基因靶向。

[0097] 在本发明的另一具体方面，用本文所述的方法产生条件性敲除动物。为了产生条件性敲除动物，向细胞（如合子或多能干细胞，如胚胎干细胞或诱导型多能干细胞、或成体干细胞）中引入供体构建体和序列特异性核酸酶，以在该细胞中产生至少一个条件性敲除等位基因。用于筛选希望的基因型的方法为本领域公知，且包括例如本文在具体实施例中所述的 PCR 分析。然后将该细胞引入雌性载体动物，以从该细胞产生例如美国专利号 7,13,608 中所公开的条件性敲除动物，该专利在此完整引入作为参考。在某些实施方案中，将该细胞扩大至两细胞阶段，引入胚泡，或以其他方式培养或与其他细胞结合，然后引入载体动物。在某些实施方案中，产生的条件性敲除动物在其种系中携带条件性敲除等位基因，使得该条件性敲除等位基因可以传递至以后的世代。

[0098] 在本发明的另一具体方面，本文所述的方法和组合物可以用来产生敲除等位基因。此方法包括，一旦作为条件性敲除等位基因掺入基因组，即切除、反转或以其他方式抑制侧翼为重组酶识别位点的供体序列的正常表达。通过将特异性识别重组酶识别位点的重组酶引入细胞来将条件性敲除等位基因转化为敲除等位基因。例如，Araki 等，Proc. Natl. Acad. Sci. USA92:160-164(1995)。重组酶是这样的酶，该酶识别间插多核苷酸侧翼的特异性多核苷酸序列（重组酶识别位点），并催化交互链交换，导致间插多核苷酸的倒位或缺失。本领域技术人员认可选择特异性识别供体构建体内的重组酶识别位点的重组酶用于本文所述的方法的有利效率。

[0099] 重组酶可以通过任意方法以蛋白质或编码重组酶蛋白质的核苷酸序列的形式引入包含供体构建体的细胞。为了产生敲除动物，使按照上文所述产生的条件性敲除动物与

转基因动物杂交，该转基因动物具有编码催化 5’ 和 3’ 重组酶识别位点处的重组的重组酶蛋白质的转基因。携带重组酶转基因的动物的实例为本领域已知，并例如由美国专利号 7,135,608 公开，该专利在此完整引入作为参考。在一些实施方案中，编码重组酶的转基因处于组织特异性启动子的控制下，使得重组酶仅在这种组织中表达，因此仅在这种组织中产生敲除等位基因。在一些实施方案中，编码重组酶的转基因处于诱导型启动子的控制下，使得可以在特定时间诱导重组酶表达。例如，可以通过四环素或其衍生物之一来控制 Tet-On 或 Tet-Off 启动子的活化作用。在一些实施方案中，编码重组酶的转基因仅在发育的某阶段或响应对该动物施用的化合物而表达。适合用于本文公开的方法的重组酶的实例包括任意形式的 P1 Cre 重组酶、任意形式的 FLP 重组酶 (flippase) 和任意形式的 Dre 重组酶，包括这些重组酶的任意可诱导形式（例如，与激素应答性结构域的融合，如 CreERT2 和 Cre-PR，或四环素调节的重组酶）。

[0100] B. 示例性组合物

[0101] 在本发明的另一具体方面，提供用于产生靶基因的条件性敲除等位基因的组合物。这种组合物包括本文所述的含有 5’ 同源区、5’ 重组酶识别位点、供体序列、3’ 重组酶识别位点和 3’ 同源区的供体构建体。该供体序列包含本文所述的具有至少一个中性突变的靶序列。该组合物进一步包含识别靶基因的序列特异性核酸酶。

[0102] 在某些实施方案中，该序列特异性核酸酶是锌指核酸酶或转录激活剂样效应核酸酶。在某些实施方案中，该重组酶识别位点是 loxP 位点或 frt 位点。可选地，该组合物还可以包括本文所述的重组酶。

[0103] 在本发明的另一方面，包含图 4A(SEQ ID NO:30)、图 4B(SEQ ID NO:31) 或图 14C(SEQ ID NOS:44-46) 中所示的序列的供体构建体。

[0104] 在本发明的另一方面，提供包含图 14A(SEQ ID NOS:36-42) 中所示的序列的指导 RNA。

[0105] 在本发明的另一方面，提供包含含有图 4A(SEQ ID NO:30)、图 4B(SEQ ID NO:31) 或图 14C(SEQ ID NOS:44-46) 中所示的序列的供体构建体的细胞。此细胞可以分离自通过本文所述的方法产生的动物。

[0106] 可以通过参考本发明的某些实施方案的以下非限制性实例来进一步理解本发明。

[0107] III. 实施例

[0108] 以下是本发明的方法和组合物的实例。应理解，鉴于上文提供的一般描述，可以实施多种其他实施方案。

[0109] 实施例 1 :Lrp5 ZFN mRNA 前核显微注射入 C57BL/6N 受精卵

[0110] 靶向小鼠低密度脂蛋白受体相关蛋白 5(Lrp5) 的外显子 2 的定制 eHi-Fi **CompoZr®** ZFN 对获自 Sigma-Aldrich。ZFN 包含显著提高其在 5’-gacttccag ttctccaagggtgctgtgtactggacagat-3’ (SEQ ID NO:29)(ZFN 切割位点下划线）处引入双链断裂的效率 (Doyon, Y. 等 Nat Meth 8, 74 - 79(2011)) 的优化的 (eHi-Fi)FokI 内切核酸酶界面。未观察到显著的潜在脱靶 (off-site target) 活性。编码 ZFN 对的信使 RNA(mRNA) 在使用前保存在 -80℃。用 mRNA(Sigma-Aldrich) 进行前核显微注射，用编码 ZFN 对的两个质粒进行 ES 细胞电穿孔。

[0111] 为了测定内切核酸酶活性，将多种浓度的编码 Lrp5 ZFN 的 mRNA 显微注射入 C57BL/6N 合子的前核（表 1）。融化 Lrp5 ZFN mRNA(2μg 的每种 ZFN 于 5μl 中），并在无 RNA 酶和 DNA 酶的显微注射缓冲液（10mMTris 和 1mM EDTA，PH 8.0）中稀释至 50ng/μl。对于 ZFN 显微注射，将 Lrp5 ZFN mRNA 稀释至 2、3、4 或 5ng/μl 的工作浓度。从在显微注射前一天与 C57BL/6N 雄性（Charles River）交配的超排卵 C57BL/6N 雌性获得小鼠合子。用 M2 培养基收集合子，按标准方法在 M2 中显微注射（Nagy, A. 等，Manipulating the Mouse Embryo:A Laboratory Manual, Third Edition(Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA, (2002))，并转入 E0.5 假孕 ICR 雌性（Taconic）的输卵管，每只假孕雌性 30 个胚胎。ICR 雌性在胚胎转移手术后饲喂 9％高脂肪饮食（Harlan, 目录号 2019)，直至幼崽断奶。

[0112]

显微注射实验	mRNA 浓度 (ng/μl)	注射后转移的合子	出生的幼崽	出生率%	KO	KO 率% (KO/出生)
1	2	168	48	29	20	42
2	3	114	15	13	2	13
3	3	150	39	26	15	38
4	4	108	21	19	8	38
5	5	174	57	33	36	63

[0113] 表 1. Lrp5 ZFN mRNA 前核显微注射入 C57BL/6N 受精卵。KO 突变体包括具有一个或多个突变体等位基因的小鼠。KO ＝敲除。

[0114] 从尾组织分离来自所得到的幼崽的 DNA，并通过 PCR 扩增和随后的测序来分析，以鉴定大突变和小突变。用 Extract-N-Amp 组织 PCR 试剂盒（Sigma, Cat#XNAT2）或用 Qiagen DNeasy 96 血液和组织试剂盒（Qiagen Cat#69582）纯化基因座尾 DNA。为了测定 ZFN 介导的突变效率和表征由 NHEJ 修复引起的突变的类型，进行了 3 步 PCR 法。在第一步中，进行使用引物 P1 和 P2 的外侧 PCR 来检测大的缺失或插入。在第二步中，进行使用引物 P3 和 P4 的内侧 PCR 来检测小至中等大小的缺失或插入。在第三步中，进行使用引物 P3 和 P4 的内侧 PCR 反应产物的直接测序来鉴定 1 至 20 个碱基对改变。用 Sequencher 4.10.1(Gene Codes Corp.）分析单个层析图。如果检测到两条不同径迹，则手动测定各单个等位基因的碱基对要求。通过 PCR TOPO 亚克隆（Invitrogen, Cat#K4575-J10）进一步分析了来自突变体亚组的等位基因。用 M13F 和 M13R 引物测序了每只小鼠的 20 至 24 个 TOPO 克隆。

[0115] 观察到了活产幼崽的至多 63％的突变率（5ng/μl ZFN mRNA)。突变从 1 至 3bp 的插入和从单个 bp 至～ 100bp 的缺失以及一个大的～ 800bp 缺失广泛变化（总结在图 1 中）。鉴定出携带超过两个等位基因的多重嵌合动物，其可能源自首次细胞分裂后持续的 ZFN 活性。此外，五只动物为复合突变体，即这些动物携带同一基因的两个独立的突变体等位基因，且检测不到该基因的野生型等位基因，表明单细胞阶段对两条染色体的 ZFN 活性。

[0116] 实施例 2：通过显微注射序列特异性内切核酸酶直接产生功能性纯合突变体等位基因

[0117] LRP5 通过作为 NORRIN 的共同受体而在视网膜血管发育中发挥重要作用。受破坏的 NORRIN 信号发放导致血管缺陷，其表征为未能在更深层的视网膜中形成毛细血管床，以及血管渗漏（Xia, C.-H. 等，Human Molecular Genetics 17, 1605 - 1612(2008)；

Xia, C.-H., PLoS ONE 5, e11676(2010) ;Junge, H. J. 等, Cell 139, 299-311(2009))。因此,按实施例 1 中所述产生在 Lrp5 中具有复合的符合读框和不符合读框的缺失的 2 月龄小鼠,并检查视网膜血管发育。动物 #542 是作为对照的嵌合功能性杂合。此动物携带一个野生型等位基因(表现为沉默的小的 3bp 符合读框的缺失)和具有 1bp 不符合读框的缺失的等位基因。动物 #495 包含 4bp 不符合读框的缺失等位基因和 1bp 不符合读框的缺失等位基因。动物 #519 包含 29bp 不符合读框的缺失等位基因和 17bp 不符合读框的缺失等位基因。动物 #555 具有 3bp 符合读框的缺失等位基因和 1bp 不符合读框的缺失等位基因且是功能性杂合子。

[0118] 对于表型分析,通过荧光血管造影术分析了携带 Lrp5 突变的动物。用氯胺酮 / 甲苯噻嗪 (80mg/kg ;7.5mg/kg) 的混合物麻醉小鼠,并用 1%托品酰胺 (Akorn, Inc.) 扩张眼。腹膜内注射无菌 10%荧光素溶液 (100μl, AK-Fluor ;Akorn, Inc.) 后进行荧光血管造影术。荧光素注射 1 分钟后用 0 聚焦和 50 敏感性的成像设置捕获图像。

[0119] 对于组织学分析,血管造影后两天处死小鼠,并处理用于组织学。在 4%多聚甲醛 (PFA) 中固定眼,然后解剖视网膜用于全样载片组织学,或在 30%蔗糖中冷冻保护过夜,并包埋在 **Tissue-Tek®** OCT Compound(Sakura) 中进行冷冻切片。按之前所述 (Gerhardt, H. 等, J. Cell. Biol. 161, 1163-1177(2003)) 进行全样载片和切片的同工凝集素 B4 染色。对于冷冻切片,去除角膜和晶状体,并在 PBS 中充分洗涤眼来去除残留的 PFA。制备冷冻的 12μm 切片,基本上按 Junge 等, Cell 139, 299-311(2009) 所述染色 MECA32(有孔内皮细胞标记 PLVAP 的抗原)。

[0120] 三只复合突变体小鼠 (495、519 和 555) 和携带一个野生型等位基因的对照杂合突变体小鼠 (542) 的视网膜表型显示在图 2 中。携带复合突变的小鼠显示 Lrp5 无效表型。荧光血管造影术揭示,小鼠 542 和 555 未显示明显的新血管缺陷或血管渗漏 (图 2A)。相反,在 Lrp5 的两个等位基因中都包含复合的不符合读框的缺失的小鼠 495 和 519 显示许多前毛细血管小动脉阻塞 (图 2A,箭头指向前毛细血管小动脉阻塞的实例) 及由整个视网膜内弥散的荧光素信号显示的明显的血管渗漏。对于图 2A 中的所有图,图 2 的右下图中的比例尺代表 200μm。

[0121] 同工凝集素染色的全样载片视网膜的共焦投射确认了复合突变体小鼠 495 和 519 的 Lrp5 无效表型。对于每只小鼠,分析了包含全部三个血管层的视网膜的最大深度的投射 (图 2B),及源自处于神经纤维层 (NFL,图 2C)、内网层 (IPL,图 2D) 和外网层 (OPL,图 2E) 中的单个血管层的投射。虽然功能性杂合视网膜 (542 和 555) 包含致密、组织良好的三层血管网络,但复合敲除视网膜 (495 和 519) 具有密度降低的无规则脉管系统 (图 2B、C)。此外,542 和 555 在 IPL(图 2D) 和 OPL(图 2E) 中包含正常的毛细血管网络,而复合 KO 小鼠 (495 和 555) 具有 IPL 中的异常新血管簇 (图 2D) 和 OPL 中的少数内皮细胞簇 (图 2E)。对于图 2B-E 中的所有图,图 2 的右下图中的比例尺代表 100μm。

[0122] 总的来说,携带具有 1bp 缺失的功能丧失等位基因和具有 3bp 符合读框的缺失的功能性等位基因的突变体 555 显示正常的视网膜表型,而携带 4bp 和 1bp 缺失的突变体 495 及携带两个更大的缺失 (17 和 29bp) 的突变体 519 在表型上纯合无效,具有总结之前已报道的表型 (Xia, C.-H. 等, Human Molecular Genetics 17, 1605-1612(2008)) 的表型。这些结果证明,序列特异性内切核酸酶的显微注射可以直接产生功能性纯合子 (复合突变

体），但不知道这些动物是否在所有细胞中都是复合突变体。

[0123] 实施例 3：通过共显微注射 Lrp5 外显子 2ZFN mRNA 和供体构建体来产生条件性敲除等位基因

[0124] 图 3A 显示靶向外显子 2 的用来产生 Lrp5 的条件性敲除等位基因 (Gu, H., Science 265, 103 - 106(1994)) 的策略的示意图。ZFN 对在 Lrp5 外显子 2 中产生双链断裂（中断块箭头所示）。通过供体质粒经链侵入的侵入及供体质粒的 5' 和 3' Lrp5 同源区与外显子 2 的各同源序列 5' 和 3' 之间的同源重组来修复断裂。产生的基因座包含侧翼为两个 loxP 位点的密码子优化的 Lrp5 外显子 2（图 1A，底部）。

[0125] 供体质粒中的 5' 和 3' Lrp5 同源区的长度分别为 1.1 和 1kb。密码子修饰（供体 1，图 4A）和野生型（供体 3，图 4C）的供体序列由 Blue Heron/Origene(Bothell, WA) 合成入修饰的 pUC19 载体中。通过用包含七个沉默突变来废除 ZFN 识别的合成片段取代 300bp MscI-BamHI 片段来从供体 3 产生供体 2（图 4B）。与供体 2 和 3 中的掺入片段相比，供体 1 中的插入片段方向相反。因此，用相反方向的引物组合进行使用与 Lrp5 基因座特异性引物组合的结合质粒主链的引物的 PCR 扩增。除 loxP 位点外，供体序列对应于小鼠基因组装配 NCBI37/mm9 chr. 19:3658179-3660815。将环状供体质粒用于所有实验。

[0126] 向野生型 Lrp5 外显子 2 序列中引入沉默突变来产生保持外显子的蛋白质编码潜能但使野生型 C57BL/6 和供体 Lrp5 外显子 2 之间的总体同源性降低至仅 78% 的密码子优化形式（供体 1，图 4A；图 5）。为了保留正常的 RNA 剪接，从修饰排除外显子 2 的前 13bp 或最后 11bp。图 5 显示三个 Lrp5 条件性敲除 DNA 供体的序列比对，排除 1.1kb 5' 同源区和 1kb 3' 同源区。用可在 http://www.ebi.ac.uk/Tools/msa/clustalw2/ 获得的比对程序 ClustalW2 进行比对。供体 1（密码子修饰）和供体 3（野生型）外显子 2 之间的总体同源性为 311/397 = 78%。供体 2（仅修饰 ZFN 结合位点）和供体 3 外显子 2 之间的总体同源性为 390/397 = 98%。LoxP 位点表示为大写黑体字母，内含子序列表示为小写字母，外显子 2（野生型或修饰）序列表示为大写字母。ZFN 结合位点在虚线框内，切割野生型外显子 2 处的序列下划线。沉默突变在实线框内。

[0127] 除将 ZFN mRNA 和供体构建体一起稀释至工作浓度 (ZFN mRNA 为 2.5-5ng/μl，供体构建体为 2.5 或 3ng/μl) 外，基本上按实施例 1 中所述将 ZFN mRNA 和供体构建体的不同组合共显微注射入 C57BL/6N 前核（表 2）。

[0128]

共显微注射实验	mRNA + DNA 浓度(ng/μl)	注射后转移的合子	出生的幼崽	出生率%	KO	KO 率% (KO/出生)	CKO	CKO 率 % (CKO/出生)
1	2.5+2.5	120	28	23	11	39	0	-
2	2.5+2.5	128	35	27	10	29	0	-
3	2.5+3	114	6	5	4	67	0	-
4	3+2.5	121	10	8	6	60	0	-
5	3+3	126	23	18	7	30	0[a]	-
6	4.5+2.5	118	18	15	5	28	0	-
7	4.5+3	146	30	21	13	43	2[b]	6.7
8	5+2.5	102	18	18	8	44	0	-

[0129] 表 2. Lrp5 ZFN mRNA(mRNA) 和 CKO 供体 1 质粒的共显微注射。KO 突变体包括具有一个或多个突变体等位基因的小鼠。KO =敲除 ;CKO =条件性敲除。[a]一只小鼠 (#95) 为假阳性 (供体 1 质粒整合入 Lrp5 基因座)。[b]小鼠 #140 和 #155。

[0130] 分析了从来自所得到的 168 只幼崽的尾样品分离的 DNA 来鉴定携带条件性敲除等位基因的小鼠 (图 3B)。用于分析缺乏 (P1-P4) 或存在 (P5-P12) 供体质粒的情况下的突变体的各引物对在图 3B 中显示。首先,按实验 1 和 2 中所述测定了总体 ZFN 突变频率。通过用 5' 核酸酶测定 (**TaqMan®**, Livak, K. J., Genet. Anal. 14, 143 - 149(1999)) 测定 5' LoxP 位点的存在来进行鉴定携带潜在的条件性敲除等位基因的小鼠的初始筛查。简言之,用 2XQiagen Type-it Fast SNP Probe PCR 主混合物、50-120ng 模板 DNA、400nM 引物和 200nM 对 LoxP 位点识别特异的基于荧光锁核酸 (LNA) 的探针构建 20μl 反应 (Weis, B., BMC Biotechnol 10, 75(2010))。使反应在 Applied Biosystems 7900HT(Life Technologies) 中热循环。通过用 Applied Biosystems 序列检测软件 2.3 版 (Life Technologies) 分析,通过显示多成分和扩增曲线中的荧光演变来确定 5' LoxP 的存在。然后进行使用引物 P5/P6 的 Lrp5 基因座特异性 PCR 分析来检测对存在于供体 1 和 2 二者 (而不是用于实施例 4 的 ES 细胞实验的供体 3) 上的密码子修饰的 Lrp5 外显子 2 序列特异的 5' 产物。类似地,进行使用引物 P7/P8 的 PCR 来分析 3' 端。为了验证 5' 和 3' loxP 二者的存在,进行了分别使用引物 P9/P10 和 P11/P12 的 PCR 分析,其将仅在 Lrp5 基因座中存在适当的 LoxP 序列时产生产物。由于 DNA 分离自嵌合亚克隆的混合物,观察到了假阳性结果,即甚至在缺乏这类真正的条件性敲除等位基因的情况下,PCR 产物也表现为 5' -3' 侧翼为 LoxP 位点的 Lrp5 等位基因阳性。例如,在一个等位基因仅携带 5' loxP 位点而另一个等位基因仅携带 3' loxP 位点时,可以产生假阳性结果。为了确认条件性敲除等位基因的存在,与假阳性相反,用引物 P5/P8(两条引物都在供体同源臂外侧退火) 扩增～ 2.8kb 的 Lrp5 外显子 2PCR 产物,用 TOPO cloning(Life Technologies) 克隆,并充分测序。此分析鉴定出条件性敲除等位基因、仅具有单个 loxP 位点的等位基因和仅具有供体衍生的外显子 2 序列 (即无 loxP 位点) 的等位基因。通过使用与供体质粒主链特异性引物组合的侧翼引物 P5 和 P8 的附加 PCR 来针对整合的整个供体载体在 Lrp5 等位基因中的存在分析了通过测序分析鉴定为假阳性的等位基因。用引物 P6 和 P7(供体 1 和供体 2) 与供体质粒主链特异性引物 (P13-P14) 组合测定了随机基因组插入的存在。对于供体 3 的随机插入,将供体质粒主链特异性引物 (P13-P14) 与结合供体 3 的野生型 Lrp5 序列的引物 P15 和 P16 组合使用。所有引物序列和反应条件在表 3 中显示。所有 PCR 研究的条件在表 7 中显示。

[0131] 通过用定位在同源区外侧的引物获得的克隆 PCR 产物的完全测序将两只小鼠 (#140 和 #155) 确认为携带条件性敲除等位基因。对于两只小鼠,条件性敲除等位基因传递至其后代。除条件性敲除等位基因外,动物 #155 还具有一个仅具有 5' loxP 位点的低频率等位基因 (不传递至后代)。动物 #95 为假阳性,因为最初的 PCR 分析提示条件性敲除等位基因,但详细的分析揭示全长供体质粒整合入 Lrp5 外显子 2。ZFN mRNA 和供体 DNA 的各组合的敲除突变率在从 28% 至 67% 的范围内 (表 2)。

[0132]

引物编号	引物序列(5'→3')	目的
1	CATGTGCCTTTGAAGAGCACACC (SEQ ID NO: 1)	检测大的缺失/插入
2	ACTCCACGGTCCTGGGATTATAGA (SEQ ID NO: 2)	检测大的缺失/插入
3	GCCCTATCACTAAGGGAGCC (SEQ ID NO: 3)	检测小至中等的缺失/插入
4	GCCCGAGATGACAATGTTCT (SEQ ID NO: 4)	检测小至中等的缺失/插入
5	CGAGCTTTTCTTAGTGATCTTTTAAG (SEQ ID NO: 5)	5'侧翼引物(同源臂外侧)
6	CTCACGTCGGTCCAATAAACG (SEQ ID NO: 6)	检测供体 1 和供体 2 外显子 2 序列
7	CGTTTATTGGACCGACGTGAG (SEQ ID NO: 7)	检测供体 1 和供体 2 外显子 2 序列
8	CCTAGACTGCAGTGAAGGACAT (SEQ ID NO: 8)	3'侧翼引物(同源臂外侧)
9	GCTCACGAGCTTTTCTTAGTGATCTTTTAAGG (SEQ ID NO: 9)	5'侧翼引物(同源臂外侧)
10	GAGAATCATGCACGGATAACTTCGTATAGC (SEQ ID NO: 10)	检测 5' loxP 整合
11	CAGGATTTCTTCTGTAGAGTATAACTTCGTATAATG (SEQ ID NO: 11)	检测 3' loxP 整合
12	CCTAGACTGCAGTGAAGGACATTCAC (SEQ ID NO: 12)	3'侧翼引物(同源臂外侧)
13	GGATAACAATTTCACACAGGAAACAGCTA (SEQ ID NO: 13)	检测随机插入或质粒整合
14	GTAAAACGACGGCCAGTGAATTGG (SEQ ID NO: 14)	检测随机插入或质粒整合
15	CAGGGAAAGAGAATCATGCAC (SEQ ID NO: 15)	检测 3'随机插入
16	CTGCACATGGGTAAACCTCTG (SEQ ID NO: 16)	检测 3'随机插入
17	CACCTGAACTACTGAAAG (SEQ ID NO: 17)	检测 5' loxP
18	CAGGGAAAGAGAATCATG (SEQ ID NO: 18)	检测 5' loxP
19	F-ATAACTTCG-IQ-TATAGCATACATTATAC-Q (SEQ ID NO: 19)	检测 5' loxP

[0133] 表 3. 引物核苷酸序列。引物 P19 :F = 荧光团(荧光素);Q = 猝灭剂(Iowa Black FQ, Integrated DNA Technologies);IQ = 内部猝灭剂(ZEN, Integrated DNA Technologies)。LNA bp 下划线。

[0134] 通过共注射 Lrp5 ZFN mRNA 和供体 1 或就野生型序列而言携带七个沉默突变(其废除了 ZFN 结合和供体的切割)的侧翼为 LoxP 的密码子优化外显子 2 供体(供体 2,图 4B ;图 5)来重复了小鼠合子中的共注射实验(4.5ng/μl ZFN mRNA 和 3ng/μl 供体 DNA)。这些实验的结果总结在表 4 中。供体 1 与 Lrp5 ZFN mRNA 的共注射产生 1/12 携带条件性敲除等位基因的幼崽(#248,8.3%条件性敲除率)。供体 2 与 Lrp5 ZFN mRNA 的共注射产生 3/35 在 Lrp5 基因座中携带供体 2 外显子序列的幼崽(8.6%)。但是,这些幼崽中只有一只随后确认为携带低频率条件性敲除等位基因(#250)。三只动物中的第二只携带仅具有 3' loxP 位点的等位基因(#274) ;最后一只动物(#280)包含一个仅具有供体 2 外显子序列(无 loxP 位点)的等位基因和另一个具有完全整合的供体 2 质粒的等位基因(假阳性)。这些结果表明,与内源 Lrp5 外显子 2 序列具有最低序列同源性的供体质粒(供体 1,图 4A)在产生条件性敲除等位基因上更有效。

[0135]

共显微注射实验	质粒	注射后转移的合子	出生的幼崽	出生率%	KO	KO 率 % (KO/出生)	CKO	CKO 率 % (CKO/出生)
1	供体 1	50	10	20	10	100	1[a]	10
2	供体 1	67	2	3	1	50	0	-
3	供体 1	70	0	-	-	-	-	-
1	供体 2	42	11	26	4	36	1[b]	9
2	供体 2	73	15	21	10	67	0[c]	-
3	供体 2	78	9	11.5	7	78	0[d]	-

[0136] 表 4. Lrp5 ZFN mRNA 和 CKO 供体 1 或供体 2 的共显微注射。所有实验都用 4.5ng/μl ZFN mRNA 和 3ng/μl 供体质粒 DNA 进行。总体 CKO 率对供体 1 为 1/12(8.3%),对供体 2 为 1/35(2.9%)。[a]小鼠 #243 ;[b]小鼠 #250 ;[c]一只小鼠 (#274) 携带仅有 3' loxP 位点的等位基因 ;[d]一只小鼠 (#280) 携带一个仅具有供体 2 外显子 (无 loxP 位点) 的等位基因和一个假阳性等位基因 (供体 2 质粒整合入 Lrp5 基因座)。

[0137] 实施例 4 :通过共电穿孔 Lrp5 外显子 2ZFN 和供体质粒来产生条件性敲除等位基因

[0138] 通过用编码两个 Lrp5 ZFN 对成分的质粒单独或连同用于显微注射实验的供体质粒、或用未修饰的侧翼为 loxP 的野生型 Lrp5 外显子 2 质粒 (供体 3) 电穿孔来共转染 C57BL/6N ES 细胞。用已建立的方法 (Nagy, A., Gertsenstein, M., Vintersten, K. 和 Behringer, R. Manipulating the Mouse Embryo:A Laboratory Manual, 第 3 版 800(Cold Spring Harbor Laboratory Press:2002)) 培养、扩大和电穿孔 C2ES 细胞 (Gertsenstein, M. 等 , PLoS ONE 5, e11260(2010))。简言之,用含或不含 15μg 供体质粒的 15μg 各 ZFN 质粒电穿孔 $15x10^6$ 个细胞。将电穿孔的细胞回收在培养基中,并将系列稀释液接种在 10cm 平板上的饲养层上。培养细胞 7-8 天,然后挑取来自每个实验的 144 个克隆 (1.596 孔板),并放入具有饲养细胞的 96 孔板中进行扩大。接种两天后,将细胞 1:2 分入新的具有饲养细胞的 96 孔板中。然后将一个平板保存在 -80℃,将另一个平板分入新的仅具有 1% 明胶而无饲养细胞的 96 孔板中,用于 DNA 分析。除过夜裂解 ES 细胞外,按实施例 1 中所述分离 DNA,基本按 Ramírez-Solis, R. 等 , Anal Biochem201, 331 - 335(1992) 所述沉淀 DNA,洗涤,并重悬在 TE 缓冲液中。

[0139]

ES 细胞实验	质粒	筛查的克隆	KO	KO 率% (KO/分析的)	CKO	CKO 率% (CKO/筛查的)
1	无	144	24	17	NA	NA
2	供体 1	144	ND	ND	1[a]	0.7%
3	供体 2	144	ND	ND	0[b]	-
2	供体 3	144	ND	ND	1[c]	0.7%

[0140] 表 5. 编码 Lrp5 ZFN 对的质粒单独或与 CKO 供体 1、2 或 3 组合电穿孔入 C57BL/6N ES 细胞。所有实验都用 15μg 供体 DNA 和 / 或 15μg 各 ZFN1 和 ZFN2 进行。[a]供体 1ES 克隆 #C8 ;[b]一个供体 2 克隆 (F5) 携带仅具有 5' loxP 的等位基因和仅具有供体 2 外显子的等位基因 (无 loxP 位点),克隆 H10 携带仅具有 3' loxP 的等位基因 ;[c]两个供体 3 克隆 (E3 和 E4) 携带仅具有 5' loxP 的等位基因。克隆 E3 还携带假阳性等位基因 (供体 3 质粒整合)。

克隆 E4 还携带真正的 CKO 次要等位基因（所测序的 240 个 TOPO 克隆中一个为阳性）。ND：未考察；NA：不适用。

[0141] DNA 分析的结果显示在图 3B 右侧，结果总结在表 5 中。用电穿孔在 ES 细胞中观察到的敲除等位基因的总体频率低于经前核注射在体内获得的频率。来自 ES 细胞电穿孔实验的遗传改变模式类似于显微注射后观察到的模式。供体 1 与 Lrp5 ZFN 质粒的共电穿孔产生所分析的 144 个克隆中的一个条件性敲除克隆（克隆 C8）。供体 2 与 Lrp5 ZFN 质粒的共电穿孔产生所分析的 144 个克隆中携带源自供体的等位基因的两个 ES 细胞克隆。这些克隆中的一个 (H10) 仅携带 3' loxP 位点等位基因；另一个 (F5) 携带一个仅具有供体 2 序列的等位基因（无 loxP 位点）和一个仅具有 5' loxP 位点的等位基因。供体 3（野生型）与 Lrp5 ZFN mRNA 的共电穿孔产生两个靶向的 ES 细胞克隆 (E3 和 E4)。二者都包含一个仅具有 5' loxP 位点的等位基因。此外，E3 携带另一源自供体 3 质粒的整合的等位基因（假阳性）。有趣地，克隆 E4 还具有非常罕见的两个 loxP 位点都为阳性的亚克隆（条件性敲除等位基因），其可能源自之前靶向的等位基因随后的再靶向。这些结果确认，使用与内源外显子具有低同源性的供体在产生条件性敲除等位基因上最有效。表 6 提供来自显微注射和 ES 细胞实验的数据的总体摘要。

[0142]

	等位基因 1	等位基因 2	等位基因 3	等位基因 4	等位基因 5	随机
供体 1						
小鼠 #95	2bp 缺失	质粒整合	-	-	-	否
小鼠 #140	CKO	wt	-	-	-	是
小鼠 #155	CKO	wt	5'loxP	-	-	否
小鼠 #243	3bp 缺失	CKO	1bp 缺失	-	-	否
ES 细胞 C8	CKO	9bp 缺失	wt	-	-	是
供体 2						
小鼠 #250	27 bp 缺失	1 bp 缺失	CKO	wt	3' loxP w/27 bp 缺失	否
小鼠 #274	3' loxP	1 bp 缺失	-	-	-	否
小鼠 #280	wt	仅供体	质粒整合	-	-	否
ES 细胞 F5	wt	5' loxP	仅供体	-	-	否
ES 细胞 H10	wt	3' loxP	14 bp 缺失	-	-	是
供体 3						
ES 细胞 E3	95 bp 缺失	5' loxP	wt	质粒整合	-	否
ES 细胞 E4	wt	5'loxP	4 bp 缺失	CKO	-	否

[0143] 表 6. 源自 CKO 供体质粒的 Lrp5 等位基因的概览。

[0144] 实施例 5：条件性敲除等位基因的正常基因功能

[0145] 为了确定获自供体 1 的条件性敲除等位基因中的沉默突变（图 4A；图 5）是否影响 Lrp5 基因的正常功能，使携带一个敲除等位基因 (#140) 和一个条件性敲除等位基因 (#155) 的小鼠与用实施例 3 的 ZFN 对产生的 Lrp5 敲除纯合小鼠杂交。年龄匹配的出生后第 16 天 (P16) 的对照小鼠（图 6A，+/+）源自 Lrp5 杂合杂交。用于实验的其他小鼠源自 Lrp5 KO/KO 雌性和 Lrp5CKO/+ 雄性之间的杂交。Lrp5 KO/KO 雌性（图 6B）是 KO/+（图 6C，P16）和 CKO/KO（图 6D，P16）的成体母亲。图 6A-D 显示用同工凝集素 B4(IB4) 染色的视网膜全样载片的代表性共焦投射（比例尺：50 μ m）。对于图 6A-D 中所示的每个投射，左图显示最

大 XY 投射，右图显示 Z 投射，Z 投射显示神经纤维层 (NFL)、内膜层 (IPL) 和外膜层 (OPL)（图 6D 的右下图上的标记）中的脉管系统。Lrp5 无效动物显示 XY 投射中降低的血管复杂性和深血管层的缺乏（图 6B)。在无效背景上携带条件性敲除等位基因的小鼠显示正常的血管表型（图 6D)，表明条件性敲除等位基因有功能。图 6E 显示用 IB4、MECA32 和 DAPI 染色的图 6A-D 中所示的眼的对侧眼的视网膜横切片。纯合敲除小鼠异位表达有孔内皮细胞标记 MECA32，而 CKO/KO、KO/+ 和 +/+ 小鼠为 MECA32 阴性。

[0146] 总的来说，纯合敲除动物显示上文所述的视网膜表型（图 6)，而携带一个敲除等位基因和一个条件性敲除等位基因的小鼠的视网膜表型与野生型小鼠或具有一个敲除等位基因和一个野生性等位基因的小鼠的视网膜表型不可区分（图 6)，表明条件性敲除等位基因是功能性等位基因。这些结果一起证明，侧翼为重组酶识别位点的具有中性突变的供体序列可以与序列特异性核酸酶一起用来在体外和体内产生全功能条件性敲除等位基因。

[0147] 图 7 显示产生在这些研究中观察到的 Lrp5 等位基因的可能机制。多个沉默突变（图 7A，星号）降低了 Lrp5 基因组序列和供体 1 之间的总体同源性。切除染色体末端后，loxP 位点外侧 100％同源性的大区域中发生链侵入，产生具有两个 loxP 位点的条件性敲除等位基因。由于 loxP 位点之间的区域中的同源性有限，loxP 位点内侧的交换事件罕见。供体 2 包含 loxP 位点之间的 100％同源性的较大区域，允许在 loxP 位点内侧发生链侵入，产生仅具有 3' loxP（图 7B)、仅具有 5' loxP 位点（图 7C) 或无 loxP 位点（图 7D) 的等位基因。引物组合 P9+P10 和 P11+P12 都产生图 7A 的事件的 PCR 产物。引物对 P9+P10 的使用产生图 7C 中所示的事件而不是图 7B 或 D 中所示的事件的产物。类似地，引物对 P11+P12 产生图 7B 中所示的事件而不是图 7C 或 D 中所示的事件的产物。引物组合 P5+P6 和 P7+P8 产生与 loxP 状态无关的 PCR 产物。

[0148]

引物对	所用试剂	循环反应条件			循环数	产物大小(bp)
		95℃	退火	72 ℃		
P1/P2	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 秒	a) 62 ℃ (-0.3 ℃/循环) - 30 秒 b) 57℃ - 30 秒	1 分	a) 25 b) 12	1059
P3/P4	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 秒	a) 56 ℃ (-0.3 ℃/循环) - 30 秒 b) 53℃ - 30 秒	30 秒	a) 10 b) 30	370
P5/P6	Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)	45 秒	56℃ - 45 秒	1 分 30 秒	40	1394
P7/P8	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 秒	63℃ - 45 秒	1 分 30 秒	40	1410
P9/P10	Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)	45 秒	56.5℃ - 45 秒	1 分 30 秒	40	1195
P11/P12	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 秒	62℃ - 45 秒	1 分 30 秒	40	1105
P13/P6	Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)	45 秒	56℃ - 45 秒	1 分 30 秒	40	1482
P7/P14	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 秒	62℃ - 45 秒	1 分 30 秒	40	1462
P5/P14	LA Taq (TaKaRa, Cat# RR002M)	45 秒	55℃ - 45 秒	10 分	47	2836
P13/P8	LA Taq (TaKaRa, Cat# RR002M)	45 秒	55℃ - 45 秒	10 分	47	2871
P5/P8	Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)	45 秒	57℃ - 45 秒	3 分 30 秒	40	wt - 2729 CKO - 2797
P13/P15	Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)	45 秒	56℃ - 45 秒	1 分 30 秒	40	1273
P16/P14	REDExtract-N-Amp PCR ReadyMix (Sigma, Cat# R4775)	45 秒	63℃ - 45 秒	1 分 30 秒	40	1211
P17/P18 P19	Type-it Fast SNP Probe PCR 混合物 (Qiagen, Cat# 206042)	15 秒	60℃ - 60 秒	20 秒	40	荧光

[0149] 表 7. 用于上述实施例中的 PCR 反应的条件。

[0150]

原残基	示例性取代	保守取代
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Asp, Lys; Arg	Gln
Asp (D)	Glu; Asn	Glu
Cys (C)	Ser; Ala	Ser

[0151]

原残基	示例性取代	保守取代
Gln (Q)	Asn; Glu	Asn
Glu (E)	Asp; Gln	Asp
Gly (G)	Ala	Ala
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe; 正亮氨酸	Leu
Leu (L)	正亮氨酸; Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Trp; Leu; Val; Ile; Ala; Tyr	Tyr
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Val; Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala; 正亮氨酸	Leu

[0152] 表 8. 保守取代。

[0153] 实施例 6 :使用不同指导 RNA 的 Cas9/CRISPR 介导的 Lrp5 外显子 2 的诱变

[0154] 为了确认其他序列特异性内切核酸酶可以用于本文所述的方法和组合物，用 Cas9/CRISPR 系统产生了 Lrp5 的修饰等位基因。将 Hepa1-6 鼠肝癌细胞培养在补充了 10％ FBS、L- 谷氨酰胺和抗生素的 RPMI 中。胰蛋白酶消化和沉淀后，按照厂家的说明用 AMAXA Nucleofector Kit V 以 AMAXA Nucleofector 程序 T-028(Lonza) 用 2μg 包含 hCas9 编码 cDNA 的每种质粒或 15μg 编码 Cas9 的 mRNA(图 14，SEQ ID NO:43) 电穿孔 10^6 个细胞，并接种入 6 孔板。通过 GFP 表达 (PMAXGFP) 评估核转染 (Nucleofection) 效率达 80-95％。核转染后 24 小时更换新鲜培养基，核转染后 72 小时用 DNeasy Blood and Tissue kit(Qiagen) 收集纯化的基因组 DNA。按照厂家的流程 (包括 polyA 加尾反应) 用 MMESSAGE MMachine T7 Ultra 试剂盒 (Life Technologies) 体外转录 HCas9 mRNA。用 RNA 的标准酚 : 氯仿提取和沉淀来纯化和浓缩 mRNA。

[0155] 产生了靶向小鼠 Lrp5 外显子 2 的三个独特的指导 RNA(gRNA)(图 14A ;Lrp5 gRNA T2、Lrp5gRNA T5 和 Lrp5gRNA T7 ;SEQ ID NOS:36-38)。

[0156] 用编码锌指对的 DNA(pZFN1+pZFN2) 或用 Cas9(+pRK5-hCas9) 与靶向 Lrp5 外显子 2 的指导 RNA(p_gRNA T2、p_gRNA T5 或 p_gRNA T7) 或对照质粒 (PMAXGFP) 一起共转染 NIH/3T3 细胞或 Hepa1-6 细胞。gRNAT7 序列与右侧 ZFN 蛋白结合位点序列的 3' 端重叠。

[0157] 为了检测共转染后 Lrp5 基因座中的突变、指示 Cas9 介导的切割和修复，基本按照厂家的说明进行了 SURVEYOR 测定 (Transgenomic)。在此测定中，杂交 PCR 产物。在突变的情况下，杂交复合物包含 SURVEYOR 核酸酶切割的错配。在此实例中，使用以下参数和 LA Taq(Takara) :95℃ 3 分钟 ;95℃ 45 秒、57℃ 45 秒、70℃ 2 分 30 秒的 35 个循环 ;然后 72℃ 7

分钟，用引物 P9 和 P12(SEQ ID NOS:9 和 12) 扩增 Lrp5 外显子 2 基因座基因座特异的～2.7kb PCR 产物。将三分之一的 PCR 产物用于 SURVEYOR 测定。通过 1.5％琼脂糖凝胶上的电泳分离所得到的消化产物。通过更短的片段的存在来鉴定核酸酶切割，其指示与野生型退火的突变体等位基因的存在。

[0158] 靶向小鼠 Lrp5 外显子 2 的全部三个指导 RNA(gRNA) 都有效介导 Cas9 诱导的突变（图 8)。在这些实验中，各 gRNA/Cas9 配对的活性似乎比 ZFN 介导的诱变高数倍。从测序来自 Lrp5 外显子 2 基因座基因座的 2.7kbPCR 产物的 TOPO 克隆等位基因计算突变率。单个序列与野生型的比对确定精确的缺失（上文定量）或插入大小（数据未显示)。用上文所述的引物 P9 和 P12 通过 PCR 来扩增 2.7kb 基因组区域。用 TOPO-TA 克隆 (Invitrogen) 直接克隆 PCR 产物来捕获所有可能的缺失大小。转化并接种单克隆后，选择克隆，分离质粒 DNA，并用引物 P20 和 P21 按照 Sanger 法测序。图 9A-B 显示 Hepa1-6 肝癌细胞中 gRNA/Cas9 突变率（图 9A) 和缺失大小（图 9B) 的总结。

[0159] 实施例 7：使用密码子优化的条件性敲除供体载体的 Cas9/CRISPR 介导介导的基因靶向

[0160] 用 Cas9 质粒或 mRNA、gRNA 和包含密码子优化的外显子序列的 Lrp5 CKO 供体 1 共转染 Hepa1-6 细胞。为了比较，一些细胞用 Lrp5 ZFN 质粒和供体质粒共转染（图 10)。72 小时后，用对密码子优化的 Lrp5 供体外显子特异的引物 (P7 ;SEQ ID NO:7) 和对 3’同源臂的外侧区域特异的引物 (P12 ;SEQ ID NO:12) 通过 PCR 来分析来自转染细胞的基因组 DNA。用以下条件用引物 P7 和 P12 进行使用 REDExtract-N-Amp PCR ReadyMix(Sigma) 的 PCR 反应：95℃ 3 分钟；95℃ 45 秒、63℃ 45 秒、72℃ 1 分 30 秒的 38 个循环；然后 72℃ 7 分钟。通过 1％琼脂糖凝胶上的电泳来分离 PCR 产物。如上文所述，Lrp5 外显子 2 供体 1 载体包含含有许多中性突变（从突变排除前 13bp 和最后 11bp) 的密码子优化的外显子 (CO 外显子 2)，以及外源侧翼 loxP 位点。以上 PCR 使用对 CO 外显子 2 序列特异的正向引物和基因组基因座中的同源臂外侧的反向引物，因此仅在供体外显子序列掺入正确的 Lrp5 基因座时产生 PCR 产物。gRNA/Cas9 的使用以极高的效率在 Lrp5 基因座处产生供体序列整合，超过使用 ZFN 系统和相同的供体载体策略时观察到的效率（图 10)。

[0161] 实施例 8：Cas9/CRISPR 介导的 loxP 位点的靶向引入

[0162] 为了确定供体设计策略和 Cas9/CRISPR 系统是否可以用来在基因组基因座处引入 loxP 位点，使用一条定位在同源臂外侧的引物和一条锚定在来自供体的 5’或 3’loxP 位点处的引物，通过 PCR 分析来分析来自按实施例 7 中所述转染的细胞的基因组 DNA。对于 5’基因组至 5’loxP 反应，除加入 DMSO 至 2％的终浓度外，将引物 P9 和 P10(SEQ ID NOS:9 和 10) 用于标准 Expand High Fidelity PCR System(Roche) 流程。PCR 参数如下：95℃ 3 分钟；95℃ 45 秒、63℃ 45 秒、72℃ 1 分 30 秒的 45 个循环；然后 72℃ 7 分钟。对于 3’loxP 至 3’基因组反应，按标准 REDExtract-N-Amp PCR ReadyMix(Sigma) 流程使用引物 P11 和 P12。PCR 参数如下：95℃ 3 分钟；95℃ 45 秒、62.5℃ 45 秒、72℃ 1 分 30 秒的 40 个循环；然后 72℃ 7 分钟。通过 1％琼脂糖凝胶上的电泳来分离 PCR 产物。从分离自用两种不同的 Lrp5 gRNA 中的任一种和 CKO 供体转染的细胞的样品获得 3’loxP 位点的 PCR 产物（图 11；p_gRNA T2)。类似地，从分离自用 gRNA T7 转染的细胞的样品获得 5’ loxP 位点的 PCR 产物（图 11)。因此，图 11 显示，在 Hepa1-6 细胞中，Lrp5 gRNA T2/Cas9 和 Lrp5 gRNA T7/

Cas9 介导的双链断裂导致用密码子优化的外显子供体载体策略在 Lrp5 基因座处引入 loxP 位点。只有用 Cas9、gRNA 和供体电穿孔的细胞显示 Lrp5 基因组基因座中的 5'(图 11,上)和 3'(图 11,下)loxP 位点的迹象。gRNA T7 产生更显著的 5' loxP 存在,而对于 ZFN,检测不到整合的 loxP 位点。在这些使用 Hepa1-6 细胞的实验中,ZFN 样品中可检测到的 loxP 位点的缺乏和 gRNA 样品中可检测到的 loxP 位点的低水平可以解释为细胞系中的低同源重组率及分析所转染的全细胞库而不是克隆亚组的事实这二者。用具有 Lrp5 CKO/wt 基因型的单个小鼠基因组 DNA 样品作为阳性对照。这些结果显示,CKO 设计策略可以用于体细胞,它有效降低双链断裂与 5'和 3'loxP 位点二者的位置之间不希望的交换事件的频率。总的来说,通过引入 RNA 指导的核酸酶介导的 DNA 断裂来靶向特异性基因组基因座,随后用改造的密码子优化的 CKO 供体序列修复,可以插入 loxP 位点,从而产生条件性敲除等位基因。

[0163] 实施例 9 :Usp10、Nnmt 和 Notch3 基因组基因座的靶向

[0164] 为了确认可以用本发明的方法靶向其他基因,产生了 Usp10、Nnmt 和 Notch3 基因组基因座的供体和 gRNA。按实施例 6 中所述和图 12 中所示向 Hepa1-6 细胞中引入这些 Cas9/gRNA 和供体来在各基因座引入 DNA 双链断裂,随后用密码子优化的供体作为模板来修复。基本上按上文所述进行 SURVEYOR 测定。以 LA Taq(Takara) 用引物 P9、P12、P22、P23、P24、P25(分别为 SEQ ID NOS:9、12、22、23、24 和 25) 和以下参数扩增 Lrp5、Usp10 和 Notch3 基因组基因座特异的大小为 2.2-2.7kb 的 PCR 产物 :95℃ 3 分钟 ;95℃ 45 秒、Ta 45 秒 (Lrp5 = 57℃,Usp10&Notch3 = 63℃)、70℃ 2 分 30 秒的 35 个循环 ;然后 72℃ 7 分钟。按厂家的说明 (Transgenomic) 分别将 1/7、1/3 和全部的 PCR 产物用于 SURVEYOR 测定。通过 1.5% 琼脂糖凝胶上的电泳来分离所得到的代表其中野生型和突变体等位基因的链已退火的核酸酶切割的消化产物。

[0165] 图 12 和图 13 显示,与就 Lrp5 基因座观察到的一样,特异性 gRNA/Cas9 复合物有效靶向 Usp10、Nnmt 和 Notch3 基因组基因座 (图 12),且整合了 loxP 位点 (图 13)。

[0166] 实施例 10 :用 RNA 指导的序列特异性内切核酸酶和密码子优化的供体产生 Lrp5 的条件性敲除和敲除等位基因

[0167] 可以用本文所述的 Lrp5 特异性 gRNA 靶向 Lrp5 基因座来引入侧翼为 loxP 的密码子优化外显子,从而产生条件性敲除等位基因。随后 Cre 重组酶蛋白质在包含条件性敲除等位基因的细胞中的表达可以切除侧翼为 loxP 的外显子,产生敲除等位基因。

[0168]

引物编号	引物序列(5'—>3')	目的
20	AGG AAA GCT AGC TTT CCA GGA GTA TG (SEQ ID NO: 20)	测序 *Lrp5* 基因组 PCR
21	GGA AGT CAA ATC CTC CTG GTT ACG A (SEQ ID NO: 21)	测序 *Lrp5* 基因组 PCR
22	GGC GTC CAG ATT ATG CAC AC (SEQ ID NO: 22)	扩增 *Usp10* 基因座
23	GAT AAT CAT GGA ATC TAA TC (SEQ ID NO: 23)	扩增 *Usp10* 基因座
24	TCT TTG CCT GAC CTG GCT ATG AG (SEQ ID NO: 24)	扩增 *Notch3* 基因座
25	CAA TCT TTC TAA CGC TCA ACT CAG AGT C (SEQ ID NO: 25)	扩增 *Notch3* 基因座/检测 3'loxP
26	CAT TGG GCT GGT ACA CGG A (SEQ ID NO: 26)	检测 *Nnmt* 5'loxP
27	GAG CTG AAG TTA TAG ATA ACT TCG TAT AGC (SEQ ID NO: 27)	检测 *Nnmt* 5'loxP
28	GGG AAC CCT ATA ACT TCG TAT AAT G (SEQ ID NO: 28)	检测 *Notch3* 3'loxP

[0169] 表 9. 引物核苷酸序列。

[0170] 虽然已为了理解的清晰性的目的以说明和实例的方式较为详细地描述了前述发明，但该描述和实例不应解释为限制本发明的范围。本文引用的所有专利和科学文献的公开内容明确完整引入作为参考。

[0001]

序 列 表

<110>基因技术股份有限公司（GENENTECH, INC. ET AL.）

<120>用于产生条件性敲除等位基因的方法和组合物

<130> P4905R1-WO

<140>
<141>

<150> 61/658,670
<151> 2012-06-12

<160> 46

<170> PatentIn 版本 3.5

<210> 1
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成的引物

<400> 1
catgtgcctt tgaagagcac acc 23

<210> 2
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成的引物

<400> 2
actccacggt cctgggatta taga 24

<210> 3
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成的引物

<400> 3
ggcctatcac taagggagcc 20

[0002]

<210> 4
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 4
gcccgagatg acaatgttct 20

<210> 5
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 5
cgagcttttc ttagtgatct tttaag 26

<210> 6
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 6
ctcacgtcgg tccaataaac g 21

<210> 7
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 7
cgtttattgg accgacgtga g 21

<210> 8
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

[0003]

<400> 8
cctagactgc agtgaaggac at 22

<210> 9
<211> 32
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 9
gctcacgagc ttttcttagt gatcttttaa gg 32

<210> 10
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 10
gagaatcatg cacggataac ttcgtatagc 30

<210> 11
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 11
caggatttct tctgtagagt ataacttcgt ataatg 36

<210> 12
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 12
cctagactgc agtgaaggac attcac 26

<210> 13
<211> 29
<212> DNA

[0004]

<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 13
ggataacaat ttcacacagg aaacagcta 29

<210> 14
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 14
gtaaaacgac ggccagtgaa ttgg 24

<210> 15
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 15
cagggaaaga gaatcatgca c 21

<210> 16
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 16
ctgcacatgg gtaaacctct g 21

<210> 17
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 17
cacctgaact actgaaag 18

[0005]

<210> 18
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 18
cagggaaaga ggatcatg 18

<210> 19
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 19
tatagcatac attatac 17

<210> 20
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 20
aggaaagcta gctttccagg agtatg 26

<210> 21
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 21
ggaagtcaaa tcctcctggt tacga 25

<210> 22
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

[0006]

<400> 22
ggcgtccaga ttatgcacac 20

<210> 23
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 23
gataatcatg gaatctaatc 20

<210> 24
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 24
tctttgcctg acctggctat gag 23

<210> 25
<211> 28
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 25
caatctttct aacgctcaac tcagagtc 28

<210> 26
<211> 19
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 26
cattgggctg gtacacgga 19

<210> 27
<211> 30
<212> DNA
<213> 人工序列

[0007]

<220>
<223> 人工序列的描述:合成的引物

<400> 27
gagctgaagt tatagataac ttcgtatagc 30

<210> 28
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物

<400> 28
gggaacccta taacttcgta taatg 25

<210> 29
<211> 39
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的引物
oligonucleotide

<400> 29
gacttccagt tctccaaggg tgctgtgtac tggacagat 39

<210> 30
<211> 5918
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 30
aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgagcgc aacgcaatta 60

atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta 120

tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt 180

acgccaagct cctcctctt ccagccctc ctctttcact gactgactga ctggaagaca 240

cacctgcggc cgcgcaagct gggccatggg gcttaacccc taacatcact cttggtgtga 300

cttggcacaa gtgtatgctt ctttatgaag gctgcagctc tgggcctgga gaatgtatac 360

caggaccccc agtctaccag gggctgtgca tttgagaggc atgggtcacc tggtagggga 420

ggaacccgta ggaccagcct cctgactgag gcttggcactg tgtgtttcta ccctgggaac 480

[0008]

tttgcatggg ttttaatgag gcctccccta accactgtat gtagtgtgca gccttctcct 540

gttctgttcc cactcctgtc ctgtccctga ctctggccca ttcacagcat taacctgaca 600

gatgtcacta cctgctttig gtctttgccc ggtgttigtt tctccactgg aacgggaacc 660

ttggaggccc ctggtacaag cagccccigc ttgccagcct atgaaggcat cttggttttg 720

ccactgtggc tctttttctt tcactcccag aactttcttt ttctcctcct gagctttcct 780

ttgcagcttc tgcccttgaa aaggagaggc tggtggagcc ttggggtgat acagtgtctg 840

tgttgtctag gattccccag gcgggggagg agaggagaag cagcctgccc caggtgaggg 900

tggggagcca cccacattct ggcctgcctc actgtgggtc cctggggcca ctgccaggcc 960

tcttagaagg aaagctagct ttccaggagt atgctgtggg ttccccttaa aacttaggga 1020

ttaagtgtgt ggtttcttta cttttcagag ctgggagcac tgccccacca tctcccagtg 1080

agggttggag aaaaaaaaag tcctaggtta caaatccacc cccagcctcc ctgtccctgc 1140

ccctgccccc tcccctccca tgtgcctttg aagagcacac ctgaactact gaaagcaagt 1200

tgaggtgagg aagtggtctt tccagtttgg ggagcagtag agccagagct tggggatttg 1260

gggagccttc ctgcctctct atactggcct atcactaagg gagcctttgg gcaactggaa 1320

ggaaggagtc ttgggagcag gggcccgggg tagataactt cgtataatgt atgctatacg 1380

aagttatccg tgcatgattc tctttccctg ggagctggtt cctgtcctgc ttagccatcc 1440

tgaccctgtg ctctgcccac agcctcaccg ctcctcctct ttgctaaccg ccgagatgtt 1500

cgactcgtag acgctggcgg cgtgaagctg gagagtacaa tcgttgctag tggactggag 1560

gatgctgccg ctgtcgactt ccagttctca aagggagccg tttattggac cgacgtgagc 1620

gaagaggcaa taaaacagac atacottaac cagactggag ctgccgccca gaatatcgta 1680

atctctggct tggtgagccc tgatggactc gcttgtgatt gggtcggaaa aaagctttat 1740

tggactgata gtgaaacgaa tcggatagaa gtagcgaacc ttaacggtac atcaagaaag 1800

gtcctgtttt ggcaggatct tgaccagcct agagcaatcg ccctcgaccc tgcacatggg 1860

taaacctctg tttttctcct gctcctatgg ggaggcgcgg ggaacaggat ttcttctgta 1920

gagtataact tcgtataatg tatgctatac gaagttatac ctgaggcagg ctgatgaggt 1980

gccttgggaa gcagcccttg ctgggaatca cctcccttcc agatcctgaa ctgtaaaaca 2040

[0009]

tggcatgttt agccagggat ggtggttcac acctacaatc ctagcattct ggaggctgag 2100

gcaggagtgt tgccacaagt tccaggccaa cctgagctat aaagaccgtc ttaaaaatat 2160

acaggaactg gggagatgcc tccattggtg atgtacttac ctcgtaacca ggaggatttg 2220

acttccatcc atgaagaaag ttaggtgcgg ggcgtcacat gtctataatc ccaggaccgt 2280

ggagttagag acacagggat ctgtgaggct tgctggctgc ctgcagtcaa attgccccagc 2340

ccctggtccc agtgagagac aaggtgtttt aaggtattta tattgccgtc ttaaaaactc 2400

catgaccaag aacagcttgg ggatgaatca gtcagtatac ttgatcagac aatactcaag 2460

tcatgcccca cctctgagta aagtcagggc aggacctcaa ggcaggaacc cagagtcaag 2520

aactaaactg cttagtggtt tgctctccat tggttgttca gattgctctc ttgtagcacc 2580

caggaccatc agccccaggca ctgaactgcc cacagtgacc tgtaccctcc cccaccaatc 2640

atcagccaag gaaatgcttc ccaggcttag aggccaatgc aaactgactg tagcttagct 2700

gagcaggacg ttgtgtctca atcaaacaca caaacaaggg cagatgagga tgatgcttgg 2760

gtttgaagct tagcatcctt ctgcactagg acaggtgacc caggtaagtt cagggacctc 2820

ccacccctte tgtgactgtt ggctgtggtg gctctgtgtc tggatgtagc acctactctg 2880

tgtattctgg aagttaccta tattggtgca gtcaactcac tctgaccctg gaaagcgcag 2940

gacagccatg aaggccacgc ggccgcacga cagtcttcac tgactgactg actggaaagt 3000

cctctccact gactgtagcc tccaattcac tggccgtcgt tttacaacgt cgtgactggg 3060

aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca tccccctttc gccagctggc 3120

gtaatagcga agaggcccgc accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg 3180

aatggcgcct gatgcggtat tttctcctta cgcatctgtg cggtatttca caccgcatac 3240

gtcaaagcaa ccatagtacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt 3300

tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt 3360

cccttcctt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc 3420

tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg atttgggtga 3480

tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga cgttggagtc 3540

cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc ctatctcggg 3600

[0010]

ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct 3660

gatttaacaa aaatttaacg cgaattttaa caaaatatta acgtttacaa ttttatggtg 3720

cactctcagt acaatctgct ctgatgccgc atagttaagc cagccccgac acccgccaac 3780

acccgctgac gcgccctgac gggcttgtct gctcccggca tccgcttaca gacaagctgt 3840

gaccgtctcc gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag 3900

acgaaagggc ctcgtgatac gcctattttt ataggttaat gtcatgataa taatggtttc 3960

ttagacgtca ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttatttttt 4020

ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata 4080

atattgaaaa aggaagagta tgagtattca acatttccgt gtcgcccttа ttccctttttt 4140

tgcggcattt tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc 4200

tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat 4260

ccttgagagt tttcgccccg aagaacgttt tccaatgatg agcacttttа aagttctgct 4320

atgtggcgcg gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca 4380

ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg 4440

catgacagta agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa 4500

cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg 4560

ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga 4620

cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg 4680

cgaactactt actctagctt cccggcaaca attaatagac tggatggagg cggataaagt 4740

tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg 4800

agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc 4860

ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca 4920

gatcgctgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc 4980

atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat 5040

cctttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc 5100

agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg 5160

[0011]

ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct 5220

accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct 5280

tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct 5340

cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg 5400

gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc 5460

gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga 5520

gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg 5580

cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta 5640

tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg 5700

ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg 5760

ctggcctttt gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat 5820

taccgccttt gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc 5880

agtgagcgag gaagcggaag agcgcccaat acgcaaaccg cctctccccg cgcgttggcc 5940

gattcatt 5948

<210> 31
<211> 5948
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 31
tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gctcactcat 60

taggcacccc aggctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc 120

ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct cttcctctt 180

ccagccctc ctctttccag tcagtcagtc agtggagact gtcgtgcggc cgcgtggcct 240

tcatggctgt cctgcgcttt ccagggtcag agtgagttga ctgcaccaat ataggtaact 300

tccagaatac acagagtagg tgctacatcc agacacagag ccagccacagc caacagtcac 360

agaaggggtg ggaggtccct gaacttacct gggtcacctg tcctagtgca gaaggatgct 420

aagcttcaaa cccaagcatc atcctcatct gcccttgttt gtgtgtttga ttgagacaca 480

[0012]

acgtcctgct cagctaagct acagtcagtt tgcattggcc tctaagcctg ggaagcattt 540

ccttggctga tgattggtgg gggagggtac aggtcactgt gggcagttca gtgcctgggc 600

tgatggtcct gggtgctaca agagagcaat ctgaacaacc aatggagagc aaaccactaa 660

gcagtttagt tcttgactct gggttcctgc cttgaggtcc tgccctgact ttactcagag 720

gtggggcatg acttgagtat tgtctgatca agtatactga ctgattcatc cccaagctgt 780

tcttggtcat ggagttttta agacggcaat ataaatacct taaaacacct tgtctctcac 840

tgggaccagg ggctgggcaa tttgactgca ggcagccagc aagcctcaca gatccctgtg 900

tctctaactc cacggtcctg ggattataga catgtgacgc cccgcaccta actttcttca 960

tggatggaag tcaaatcctc ctggttacga ggtaagtaca tcaccaatgg aggcatctcc 1020

ccagttcctg tatattttta agacggtctt tatagctcag gttggcctgg aacttgtggc 1080

aacactcctg cctcagcctc cagaatgcta ggattgtagg tgtgaaccac catccctggc 1140

taaacatgcc atgttttaca gttcaggatc tggaagggag gtgattccca gcaagggctg 1200

cttcccaagg cacctcatca gcctgcctca ggtataactt cgtatagcat acattatacg 1260

aagttatact ctacagaaga aatcctgttc cccgcgcctc cccataggag caggagaaaa 1320

acagaggttt accccatgtgc aggatccagg gcaatggccc ttggctggtc caggtcctgc 1380

cagaagagaa ccttacggga cgtcccattg aggttggcaa cctcaatgcg gttggtctcg 1440

gagtccgtcc agtacagctt cttgccaacc cagtcacagg ccaggccatc aggtgacacg 1500

aggccccgaga tgacaatgtt ctgtgcagca gctccagtct ggttcaggta ggtctgtttg 1560

atggcctcct cgctcacgtc ggtccaataa acggctcccc ttgagaactg gaagtctaca 1620

gcagctgcat cctccaggcc actggccaca atggtggact ccagcttcac tccgccggca 1680

tccactagcc gcacatcccg gcggttggca aacaacagga gcggtgaggc tgtgggcaga 1740

gcacagggtc aggatggcta agcaggacag gaaccagctc ccagggaaag agaatcatgc 1800

acggataact tcgtatagca tacattatac gaagttatct accccgggcc cctgctccca 1860

agactccttc cttccagttg cccaaaggct ccottagtga taggccagta tagagaggca 1920

ggaaggctcc ccaaatcccc aagctctggc tctactgctc ccaaaactgg aaagaccact 1980

tcctcacctc aacttgcttt cagtagttca ggtgtgctct tcaaaggcac atgggagggg 2040

[0013]

aggggcagg ggcagggaca gggaggctgg gggtggattt gtaacctagg actttttttt 2100

tctccaaccc tcactgggag atggtgggge agtgctccca gctctgaaaa gtaaagaaac 2160

cacacactta atcgctaagt tttaagggga accgacagca tactcctgga aagctagctt 2220

tccttctaag aggcctggca gtggccccag ggacccacag tgaggcaggc cagaatgtgg 2280

gtggctcccc acgctcacct ggggcaggct gcttctcctc tcctccccgg cctggggaat 2340

cctagacaac acagacactg tatcacccca aggctccacc agcctctcct tttcaagggc 2400

agaagctgca aaggaaagct caggaggaga aaaagaaagt tctgggagtg aaagaaaaag 2460

agccacagtg gcaaaaccaa gatgccttca taggctggca agcaggggct gcttgtacca 2520

ggggcctcca aggttccgt tccagtggag aaacagacac cgggcaaaga ccaaaagcag 2580

gtagtgacat ctgtcaggtt aatgctgtga atgggccaga gtcagggaca ggacaggagt 2640

gggaacagaa caggagaagg ctgcacacta catacagtgg ttaggggagg cctcattaaa 2700

acccatgcaa agttcccagg gtagaaacac acagtgccag cctcagtcag gaggctggtc 2760

ctacgggttc ctcccctacc aggtgaccca tgcctctcaa atgcacagcc cctggtagac 2820

taggggtcct ggtatacatt ctccaggccc agagctgcag ccttcataaa gaagcataca 2880

cttgtgccaa gtcacaccaa gagtgatgtt aggggttaag ccccatggcc cagcttgcgc 2940

ggccgcaggt gtgtcttcca gtcagtcagt cagtgaaagt cctctccact gactgtagcc 3000

tccaattcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 3060

cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgc 3120

accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgcct gatgcggtat 3180

tttctcctta cgcatctgtg cggtatttca caccgcatac gtcaaagcaa ccatagtacg 3240

cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta 3300

cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt 3360

tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg 3420

ctttacggca cctcgacccc aaaaaacttg atttgggtga tggttcacgt agtgggccat 3480

cgccctgata gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggac 3540

tcttgttcca aactggaaca acactcaacc ctatctcggg ctattctttt gatttataag 3600

[0014]

ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg 3660

cgaattttaa caaaatatta acgtttacaa ttttatggtg cactctcagt acaatctgct 3720

ctgatgccgc atagttaagc cagccccgac acccgccaac acccgctgac gcgccctgac 3780

gggcttgtct gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca 3840

tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag acgaaagggc ctcgtgatac 3900

gcctatttt ataggttaat gtcatgataa taatggtttc ttagacgtca ggtggcactt 3960

ttcggggaaa tgtgcgcgga acccctattt gtttatttt ctaaatacat tcaaatatgt 4020

atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta 4080

tgagtattca acatttccgt gtcgcccttа ttcccttttt tgcggcattt tgccttcctg 4140

tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac 4200

gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg 4260

aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc 4320

gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg 4380

ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat 4440

gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg 4500

gaggaccgaa ggagctaacc gctttttttgc acaacatggg ggatcatgta actcgccttg 4560

atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc 4620

ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt 4680

cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct 4740

cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc 4800

gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca 4860

cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct 4920

cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt 4980

taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga 5040

ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca 5100

aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac 5160

[0015]

caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg 5220

taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag 5280

gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac 5340

cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt 5400

taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag cccagcttgg 5460

agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc 5520

ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc 5580

gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc 5640

acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa 5700

acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt 5760

tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg 5820

ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag 5880

agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc 5940

acgacagg 5948

<210> 32
<211> 5948
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 32
tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gctcactcat 60

taggcacccc aggctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc 120

ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct cattcctctt 180

ccagccattc ctctttccag tcagtcagtc agtggagact gtcgtgcggc cgcgtggcct 240

tcatggctgt cctgcgcttt ccagggtcag agtgagttga ctgcaccaat ataggtaact 300

tccagaatac acagagtagg tgctacatcc agacacagag ccaccacagc caacagtcac 360

agaaggggtg ggaggtccct gaacttacct gggtcacctg tcctagtgca gaaggatgct 420

aagcttcaaa cccaagcatc atcctcatct gcccttgttt gtgtgtttga ttgagacaca 480

[0016]

acgtcctgct cagctaagct acagtcagtt tgcattggcc tctaagcctg ggaagcattt 540

ccttggctga tgattggtgg gggagggtac aggtcactgt gggcagttca gtgcctgggc 600

tgatggtcct gggtgctaca agagagcaat ctgaaccaacc aatggagagc aaaccactaa 660

gcagtttagt tcttgactct gggttcctgc cttgaggtcc tgccctgact ttactcagag 720

gtggggcatg acttgagtat tgtctgatca agtatactga ctgattcatc cccaagctgt 780

tcttggtcat ggagttttta agacggcaat ataaatacct taaaacacct tgtctctcac 840

tgggaccagg ggctgggcaa tttgactgca ggcagccagc aagcctcaca gatccctgtg 900

tctctaactc cacggtcctg ggattataga catgtgacgc cccgcaccta actttcttca 960

tggatggaag tcaaatcctc ctggttacga ggtaagtaca tcaccaatgg aggcatctcc 1020

ccagttcctg tatattttta agacggtctt tatagctcag gttggcctgg aacttgtggc 1080

aacactcctg cctcagcctc cagaatgcta ggattgtagg tgtgaaccac catccctggc 1140

taaacatgcc atgttttaca gttcaggatc tggaagggag gtgattccca gcaagggctg 1200

cttcccaagg cacctcatca gcctgcctca ggtataactt cgtatagcat acattatacg 1260

aagttatact ctacagaaga aatcctgttc cccgcgcctc cccataggag caggagaaaa 1320

acagaggttt acccatgtgc aggatccagg gcaatggccc ttggctggtc caggtcctgc 1380

cagaagagaa ccttacggga cgtcccattg aggttggcaa cctcaatgcg gttggtctcg 1440

gagtccgtcc agtacagctt cttgccaacc cagtcacagg ccaggccatc aggtgacacg 1500

aggccgaga tgacaatgtt ctgtgcagca gctccagtct ggttcaggta ggtctgtttg 1560

atggcctcct cgctcacatc tgtccagtac acagcaccct tggagaactg gaagtctaca 1620

gcagctgcat cctccaggcc actggccaca atggtggact ccagcttcac tccgccggca 1680

tccactagcc gcacatcccg gcggttggca aacaacagga gcggtgaggc tgtgggcaga 1740

gcacagggtc aggatggcta agcaggacag gaaccagctc ccagggaaag agaatcatgc 1800

acggataact tcgtatagca tacattatac gaagttatct acccgggcc cctgctccca 1860

agactccttc cttccagttg cccaaaggct ccattagtga taggccagta tagagaggca 1920

ggaaggctcc ccaaatcccc aagctctggc tctactgctc cccaaactgg aaagaccact 1980

tcctcacctc aacttgcttt cagtagttca ggtgtgctct tcaaaggcac atgggagggg 2040

[0017]

aggggcagg ggcagggaca gggaggctgg gggtggattt gtaacctagg actttttttt 2100

tctccaaccc tcactgggag atggtggggc agtgctccca gctctgaaaa gtaaagaaac 2160

cacacactta atccctaagt tttaagggga acccacagca tactcctgga aagctagctt 2220

tccttctaag aggcctggca gtggccccag ggacccacag tgaggcaggc cagaatgtgg 2280

gtggctcccc acctcacct gggcaggct gcttctcctc tcctcccccg cctggggaat 2340

cctagacaac acagacactg tatcacccca aggctccacc agcctctcct tttcaagggc 2400

agaagctgca aaggaaagct caggaggaga aaaagaaagt tctgggagtg aaagaaaaag 2460

agccacagtg gcaaaaccaa gatgccttca taggctggca agcagggget gcttgtacca 2520

ggggcctcca aggttcccgt tccagtggag aaacagacac cgggcaaaga ccaaaagcag 2580

gtagtgacat ctgtcaggtt aatgctgtga atgggccaga gtcagggaca ggacaggagt 2640

gggaacagaa caggagaagg ctgcacacta catacagtgg ttaggggagg cctcattaaa 2700

acccatgcaa agttcccagg gtagaaacac acagtgccag cctcagtcag gaggctggtc 2760

ctacgggttc ctcccctacc aggtgaccca tgcctctcaa atgcacagcc cctggtagac 2820

taggggtcct ggtatacatt ctccaggccc agagctgcag ccttcataaa gaagcataca 2880

cttgtgccaa gtcacaccaa gagtgatgtt aggggttaag ccccatggcc cagcttgcgc 2940

ggccgcaggt gtgtcttcca gtcagtcagt cagtgaaagt cctctccact gactgtagcc 3000

tccaattcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 3060

cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgc 3120

accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgcct gatgcggtat 3180

tttctcctta cgcatctgtg cggtatttca caccgcatac gtcaaagcaa ccatagtacg 3240

cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta 3300

cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt 3360

tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg 3420

ctttacggca cctcgacccc aaaaaacttg atttgggtga tggttcacgt agtgggccat 3480

cgccctgata gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggac 3540

tcttgttcca aactggaaca acactcaacc ctatctcggg ctattcttt gatttataag 3600

[0018]

ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg 3660

cgaattttaa caaaatatta acgtttacaa ttttatggtg cactctcagt acaatctgct 3720

ctgatgccgc atagttaagc cagccccgac acccgccaac acccgctgac gcgccctgac 3780

gggcttgtct gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca 3840

tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag acgaaagggc ctcgtgatac 3900

gcctatttt ataggttaat gtcatgataa taatggtttc ttagacgtca ggtggcactt 3960

ttcggggaaa tgtgcgcgga acccctattt gtttatttt ctaaatacat tcaaatatgt 4020

atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta 4080

tgagtattca acatttccgt gtcgcccttа ttcccttttt tgcggcattt tgccttcctg 4140

ttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac 4200

gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg 4260

aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc 4320

gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg 4380

ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat 4440

gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg 4500

gaggaccgaa ggagctaacc gctttttttgc acaacatggg ggatcatgta actcgccttg 4560

atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc 4620

ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt 4680

cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct 4740

cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc 4800

gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca 4860

cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct 4920

cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt 4980

taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga 5040

ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca 5100

aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac 5160

[0019]

caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg 5220

taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag 5280

gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac 5340

cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt 5400

taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag cccagcttgg 5460

agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc 5520

ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc 5580

gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc 5640

acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa 5700

acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggccttt gctcacatgt 5760

tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg 5820

ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag 5880

agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc 5940

acgacagg 5948

<210> 33
<211> 605
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 33
ataacttcgt ataatgtatg ctatacgaag ttatccgtgc atgattctct ttcctggga 60

gctggttcct gtcctgctta gccatcctga ccctgtgctc tgcccacagc ctcaccgctc 120

ctcctctttg ctaaccgccg agatgttcga ctcgtagacg ctggcggcgt gaagctggag 180

agtacaatcg ttgctagtgg actggaggat gctgccgctg tcgacttcca gttctcaaag 240

ggagccgttt attggaccga cgtgagcgaa gaggcaataa aacagacata ccttaaccag 300

actggagctg ccgcccagaa tatcgtaatc tctggcttgg tgagccctga tggactcgct 360

tgtgattggg tcggaaaaaa gctttattgg actgatagtg aaacgaatcg gatagaagta 420

gcgaaccttа acggtacatc aagaaaggtc ctgttttggc aggatcttga ccagcctaga 480

[0020]

gcaatcgccc tcgaccctgc acatgggtaa acctctgttt ttctcctgct cctatgggga 540

ggcgcgggga acaggatttc ttctgtagag tataacttcg tataatgtat gctatacgaa 600

gttat 605

<210> 34
<211> 605
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 34
ataacttcgt ataatgtatg ctatacgaag ttatccgtgc atgattctct ttccctggga 60

gctggttcct gtcctgctta gccatcctga ccctgtgctc tgcccacagc ctcaccgctc 120

ctgttgtttg ccaaccgccg ggatgtgcgg ctagtggatg ccggcggagt gaagctggag 180

tccaccattg tggccagtgg cctggaggat gcagctgctg tagacttcca gttctcaaag 240

ggagccgttt attggaccga cgtgagcgag gaggccatca aacagaccta cctgaaccag 300

actggagctg ctgcacagaa cattgtcatc tcgggcctcg tgtcacctga tggcctggcc 360

tgtgactggg ttggcaagaa gctgtactgg acggactccg agaccaaccg cattgaggtt 420

gccaacctca atgggacgtc ccgtaaggtt ctcttctggc aggacctgga ccagccaagg 480

gccattgccc tggatcctgc acatgggtaa acctctgttt ttctcctgct cctatgggga 540

ggcgcgggga acaggatttc ttctgtagag tataacttcg tataatgtat gctatacgaa 600

gttat 605

<210> 35
<211> 605
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 35
ataacttcgt ataatgtatg ctatacgaag ttatccgtgc atgattctct ttccctggga 60

gctggttcct gtcctgctta gccatcctga ccctgtgctc tgcccacagc ctcaccgctc 120

ctgttgtttg ccaaccgccg ggatgtgcgg ctagtggatg ccggcggagt gaagctggag 180

tccaccattg tggccagtgg cctggaggat gcagctgctg tagacttcca gttctccaag 240

[0021]

ggtgctgtgt actggacaga tgtgagcgag gaggccatca aacagaccta cctgaaccag 300

actggagctg ctgcacagaa cattgtcatc tcgggcctcg tgtcacctga tggcctggcc 360

tgtgactggg ttggcaagaa gctgtactgg acggactccg agaccaaccg cattgaggtt 420

gccaacctca atgggacgtc ccgtaaggtt ctcttctggc aggacctgga ccagccaagg 480

gccattgccc tggatcctgc acatgggtaa acctctgttt ttctcctgct cctatgggga 540

ggcgcgggga acaggatttc ttctgtagag tataacttcg tataatgtat gctatacgaa 600

gttat 605

<210> 36
<211> 104
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 36
ggtgcggcta gtggatgccg ggttttagag ctagaaatag caagttaaaa taaggctagt 60

ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt tttt 104

<210> 37
<211> 104
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 37
ggagtccacc attgtggcca ggttttagag ctagaaatag caagttaaaa taaggctagt 60

ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt tttt 104

<210> 38
<211> 104
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 38
ggtactggac agatgtgagc ggttttagag ctagaaatag caagttaaaa taaggctagt 60

ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt tttt 104

<210> 39

[0022]

<211> 103
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 39
gctgaagatg aactgccaga gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt 103

<210> 40
<211> 103
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 40
gatctccgtg aaggactcac gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt 103

<210> 41
<211> 103
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 41
gaccactggg gaccagtcaa gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt 103

<210> 42
<211> 104
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 42
gaggcgtttg ccagagttca ggttttagag ctagaaatag caagttaaaa taaggctagt 60

ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt tttt 104

<210> 43
<211> 4206
<212> DNA
<213> 人工序列

[0023]

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 43

atggacaaga agtacagcat cggcctggac atcggcacca actctgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gcccagcaag aaattcaagg tgctgggcaa caccgaccgg 120

cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggaga aacagccgag 180

gccacccggc tgaagagaac cgccagaaga agatacacca gacggaagaa ccggatctgc 240

tatctgcaag agatcttcag caacgagatg gccaaggtgg acgacagctt cttccacaga 300

ctggaagagt ccttcctggt ggaagaggat aagaagcacg agcggcaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgagaaag 420

aaactggtgg acagcaccga caaggccgac ctgcggctga tctatctggc cctggcccac 480

atgatcaagt tccgggggca cttcctgatc gagggcgacc tgaaccccga caacagcgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggaaaaccct 600

atcaacgcca gcggcgtgga cgccaaggcc atcctgtctg ccagactgag caagagcaga 660

cggctggaaa atctgatcgc ccagctgccc ggcgagaaga agaatggcct gttcggcaac 720

ctgattgccc tgagcctggg cctgacccct aacttcaaga gcaacttcga cctggccgag 780

gatgccaaac tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttt ctggccgcca agaacctgtc cgacgccatc 900

ctgctgagcg acatcctgag agtgaacacc gagatcacca aggccccctt gagcgcctct 960

atgatcaaga gatacgacga gcaccaccag gacctgaccc tgctgaaagc tctcgtgcgg 1020

cagcagctgc ctgagaagta caaagagatt ttcttcgacc agagcaagaa cggctacgcc 1080

ggctacatcg atggcggagc cagccaggaa gagttctaca agttcatcaa gcccatcctg 1140

gaaaagatgg acggcaccga ggaactgctc gtgaagctga acagagagga cctgctgcgg 1200

aagcagcgga ccttcgacaa cggcagcatc ccccaccaga tccacctggg agagctgcac 1260

gccattctgc ggcggcagga agattttfac ccattcctga aggacaaccg ggaaaagatc 1320

gagaagatcc tgaccttccg catccccfac tacgtgggcc ctctggccag gggaaacagc 1380

agattcgcct ggatgaccag aaagagcgag gaaaccatca ccccctggaa cttcgaggaa 1440

gtggtggaca agggcgccag cgcccagagc ttcatcgagc ggatgaccaa cttcgataag 1500

[0024]

aacctgccca acgagaaggt gctgcccaag cacagcctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaagt gaaatacgtg accgagggaa tgagaaagcc cgccttcctg 1620

agcggcgagc agaaaaaagc catcgtggac ctgctgttca agaccaaccg gaaagtgacc 1680

gtgaagcagc tgaaagagga ctacttcaag aaaatcgagt gcttcgactc cgtggaaatc 1740

tccggcgtgg aagatcggtt caacgcctcc ctgggcacat accacgatct gctgaaaatt 1800

atcaaggaca aggacttcct ggacaatgag gaaaacgagg acattctgga agatatcgtg 1860

ctgaccctga cactgtttga ggacagagag atgatcgagg aacggctgaa aacctatgcc 1920

cacctgttcg acgacaaagt gatgaagcag ctgaagcggc ggagatacac cggctggggc 1980

aggctgagcc ggaagctgat caacggcatc cgggacaagc agtccggcaa gacaatcctg 2040

gatttcctga agtccgacgg cttcgccaac agaaacttca tgcagctgat ccacgacgac 2100

agcctgacct ttaaagagga catccagaaa gcccaggtgt ccggccaggg cgatagcctg 2160

cacgagcaca ttgccaatct ggccggcagc ccgccattа agaagggcat cctgcagaca 2220

gtgaaggtgg tggacgagct cgtgaaagtg atgggccggc acaagcccga gaacatcgtg 2280

atcgaaatgg ccagagagaa ccagaccacc cagaagggac agaagaacag ccgcgagaga 2340

atgaagcgga tcgaagaggg catcaaagag ctgggcagcc agatcctgaa agaacacccc 2400

gtggaaaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaatgggcgg 2460

gatatgtacg tggaccagga actggacatc aaccggctgt ccgactacga tgtggaccat 2520

atcgtgcctc agagctttct gaaggacgac tccatcgata acaaagtgct gactcggagc 2580

gacaagaacc ggggcaagag cgacaacgtg ccctccgaag aggtcgtgaa gaagatgaag 2640

aactactggc gccagctgct gaatgccaag ctgattaccc agaggaagtt cgacaatctg 2700

accaaggccg agagaggcgg cctgagcgaa ctggataagg ccggcttcat caagagacag 2760

ctggtggaaa cccggcagat cacaaagcac gtggcacaga tcctggactc ccggatgaac 2820

actaagtacg acgagaacga caaactgatc cgggaagtga aagtgattac cctgaagtcc 2880

aagctggtgt ccgatttccg gaaggatttc cagttttaca aagtgcgcga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtcgtgg gaaccgccct gatcaaaaag 3000

taccctaagc tggaaagcga gttcgtgtac ggcgactaca aggtgtacga cgtgcggaag 3060

[0025]

atgatcgcca agagcgagca ggaaatcggc aaggctaccg ccaagtactt cttctacagc 3120

aacatcatga acttttttcaa gaccgagatt accctggcca acggcgagat ccggaagcgg 3180

cctctgatcg agacaaacgg cgaaacaggc gagatcgtgt gggataaggg ccgggacttt 3240

gccaccgtgc ggaaagtgct gtctatgccc caagtgaata tcgtgaaaaa gaccgaggtg 3300

cagacaggcg gcttcagcaa agagtctatc ctgcccaaga ggaacagcga caagctgatc 3360

gccagaaaga aggactggga ccctaagaag tacggcggct tcgacagccc caccgtggcc 3420

tattctgtgc tggtggtggc caaagtggaa aagggcaagt ccaagaaact gaagagtgtg 3480

aaagagctgc tggggatcac catcatggaa agaagcagct tcgagaagaa tcccatcgac 3540

tttctggaag ccaagggcta caaagaagtg aaaaaggacc tgatcatcaa gctgcctaag 3600

tactccctgt tcgagctgga aaacggccgg aagagaatgc tggcctctgc cggcgaactg 3660

cagaagggaa acgaactggc cctgccctcc aaatatgtga acttcctgta cctggccagc 3720

cactatgaga agctgaaggg ctccccccgag gataatgagc agaaacagct gtttgtggaa 3780

cagcacaaac actacctgga cgagatcatc gagcagatca gcgagttctc caagagagtg 3840

atcctggccg acgctaatct ggacaaggtg ctgagcgcct acaacaagca cagagacaag 3900

cctatcagag agcaggccga gaatatcatc cacctgttta ccctgaccaa tctgggagcc 3960

cctgccgcct tcaagtactt tgacaccaac atcgaccgga agaggtacac cagcaccaaa 4020

gaggtgctgg acgccaccct gatccaccag agcatcaccg gcctgtacga gacacggatc 4080

gacctgtctc agctgggagg cgacgcctat ccctatgacg tgcccgatta tgccagcctg 4140

ggcagcggct ccccaagaa aaaacgcaag gtggaagatc ctaagaaaaa gcggaaagtg 4200

gactga 4206

<210> 44
<211> 2126
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 44
cgtctgtccc ctagtaccat ccacagagca gacattctct gtaagggtgt cttgccattt 60

tgtgacctag agacgggatt tggtgagcag ccgtggtgat gttgctgtac cacactgctt 120

[0026]

ctgtgtgctg tcagcgatgg catgtggaag agtttgcctg tgtcaaggtg ctaagtccta 180

ggctgccagg tcctgagtgg ttgggctcct ttccagccat ggttggtttt ttcttgctgt 240

ttgtcatcac ctcctgggga gagaacccag cactttattt ttggggaagg aagtaacctt 300

gcctcgttag aactgttgca ttggctagcc agagttcatc tggagacacg agcctaggag 360

gccacgggca cagcagacag tgattgccct gggcaccagc tagagcctcc tggactactt 420

cacagatgct gcagcagggt ggtgtttggg gctcggttct gtcagctgac tcacttctca 480

gacacccccа aatcctaaat aaataataga attagcagga agaagtattt gagaggttca 540

taaaatttgc caggttttca caaagtagcc agacctatgt tgttttcttc ctttcttcat 600

acaaggaagg tgacgttaaa ccattaactc atctgcatgg cgcactgtaa aatgacagtt 660

aagggcccag ccctttcttg gggactctgg ctcacttctg ccagtggtgc tgactcttca 720

gttcccactg ttaacagatg gagtgttggg gttgggatca tcctgcctgc ctgacactca 780

ggcagtatcc cagtctccgt tgcagtgtat ttagcagtgg taaacgattg aaagcattcc 840

tacactgttt ggttcagatt ggggcgaatc tgtgtcttcc tgggaacgtg atgtcatggc 900

cagttgcttc tgcagagctt agaggaaggg ttgagctggt gcagtggctt ctgtcaraac 960

ttcgtataat gtatgctata cgaagttatg ggctttgata atctgttaag atttgtttct 1020

gacttctctc gcagctccct ccatacagcg ggactctttg tagcatccag gctgaagacg 1080

aactgccaga tggtaagccg ggttgcatgg actgggtggg caggaaacct gcaaacatgt 1140

ccataacttc gtataatgta tgctatacga agttataaca ttgggtcatc tttctcaac 1200

acaaattttt atttattcat ttattttgga gtgtgctttt gtgtggaggt cagaggtcag 1260

ctctctggaa tatgtcctct tctgcttttg cctgggtttg gtgattggac agaggtcatc 1320

gggcttatgc agcaagtctc ttgctgagct gtctccctga tctgtccctg ttgactccca 1380

tcctttgtag gttagcaggc tcaggcttgt gcccagagct gcagtgtccg ttcagccgga 1440

ggcgcatttc attagagctt gaagcttctg ctctgaccctg cagctcactt aactcctgtc 1500

ttacagtttc cattgctgtg gagagacacc atgaccagag cagccagtat gaaggcaaac 1560

atttatttgg ggctggttta caggttcaga ggttcagtcc gttatcatgg agggaaacat 1620

ggcagcatcc gggcatgcat ggcactggga gctgagaatt ctctatgttg ttctcaaggc 1680

[0027]

aaacaggaga agtcttgcaa gcagtgagga gggtctcata gcccacccgc acagtgccag 1740

ttcctctaac aaggccacac ctacccacag tgcctcggcc catgggccaa gcatactcaa 1800

accaccacaa ctgccaagtc actcagattg gcaggtgtgc tttcaaggtt gtgccacgct 1860

cactctgggc atggtaagtt gtggtcagga gagaggtgct ctagtctcct cagcaggccc 1920

tggtgttgtg ggtgtgtcat cccacatggc tgtgctcaac tctgccgggt ccagggagtt 1980

gttatagcag ttggtgttca tgacccatgc ctgggcattt aggcctccag catgagcact 2040

ggccgattta tgcataagtg tcatgtttat caaggtgcag cagtgtttgt tactgaggtg 2100

tgtttaagat atcctggcta tgcagt 2126

<210> 45
<211> 1832
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 45
gaaagtttaa ttctactctt atttgaaagg aatgtttgtt tactcatgaa gtaaaggacc 60

cccaatatt aagaactggg tgatttgatt attaattaat ctcagtgttt taaattttaa 120

tatatataca tctaagcata tgtgtacatg gtctatgcag gtacatgcat ggggccagaa 180

gagggcacca gatgtatgct tccatcactc tctccaccta ttcttgtggg gcagcgtccc 240

ctcctaaacc tggggcttgt gtgttgacta gactggaagc cagaaagcct cagagagcct 300

ctcctctctg ctcctctcag agatgggagc ataggtattt gtgagatgcc tgacttgtta 360

cacaggttct gggacctgaa tgctggccct catctctgga ccatcactcc aggccttaaa 420

atcagttgtt acacaacaac aacgacaaca ccaccaccac cacaacaaaa atggcacaat 480

cccagtaaag acaataactt cgtataatgt atgctatacg aagttatcta taacttcagc 540

tctgcctgtt ccctttttgat gcctgattcc ctgccccctct tctgtcctag gtgctgtaaa 600

aggagaactc ctcattgaca tcggctctgg cccaaccatc taccagcttc tgtctgcctg 660

tgagtctttc acggagatca tcgtctctga ctatacagac caaaacctct gggaactgca 720

gaaatggctc aagaaggagc caggagcctt tgattggtcc ccagtcgtca cctatgtgtg 780

cgatcttgaa ggcaacaggt agaggaatga gtatctgctc ttcaacttte tgaagggacc 840

[0028]

tgcataactt cgtataatgt atgctatacg aagttatatg tttctatcag cccaaactac 900

cactaagatc cagatgaaaa atgcaaaaga aactcaaaga ccaaagggaa ctcgggatag 960

aagatgaggg ctcaggagaa tagctgggtt gtcagatgta ccaccagagt actatttggt 1020

taccagtcat ccaacacaaa ggtggtggct aatgcctaat gcccacacca ccggaaacag 1080

caaaacactg gcagatgtat ttttttagcc aacagaaata acctgaagca attatgggga 1140

aagaattcat tacttaccat ctataaagtt agaatcttta ggatctggga gacggtcagt 1200

gggtaaagtg cttgtcacag aggttgtcct ttctccagtg tgtcctctta gcatctttat 1260

tgaaaaccag ctggctgcag attcatggct tagtcctata ctgtctattc tgttccatta 1320

ttctgtgttc tctggctggt gatgttttgt ggaaagactt tgctggtttg atgatatagt 1380

ttttttgttg ttgttgttgc tgctgctgct gctgttgttg aagctaagta gtgtgtgatt 1440

tctgatattt tcttggtgct atttgggctt tgtgtgtacg gttgcctaaa aattctagaa 1500

ttgttttcta gtcctgtcat tagtatttta agaggaatgg cattggctac actaattgct 1560

ttgggtagta tggacatatt catgatcttg attcttcaga tcagcgaaca gatctttcca 1620

ttttgatggt gtctgcacta ttcttctcca ttgatgtttt cagtgtaaag atcattggct 1680

tgttttagca aaattcattt tgaatttata tattatattt attaataaat acattaataa 1740

aataatatta aatattacat tttatttatt tttaatagct attgcaatag gattgacttc 1800

ttgatttctg aatcaaataa ttctctattg gg 1832

<210> 46
<211> 1431
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述:合成的多核苷酸
<400> 46
gacagaggca ggtaaattac tgtgaattcc agtctagcca gcgatatgta gtgagaccct 60

gcctaaaata tattaaaaaa aaaaaaaaaa agggaaggaa gagtaaatgg atttgtctga 120

tagtctgtct ggcacgagtg ttgtttgata aacgcatctt gtgataactt cgtataatgt 180

atgctatacg aagttattta tctgtctggc attgccatgc ttttataccg tcccgaccac 240

acatcttccc acaggtgcct gccaggctgg gttggtgagc gctgccagct cgaagacccc 300

[0029]

tgccactccg gcccttgtgc cggccgaggc gtctgccaga gctcagtggt ggctggcacc 360

gcacgattct cctgccggtg ccttcgaggc ttccaaggtg aaggggtgtg tctggacggg 420

aaccctataa cttcgtataa tgtatgctat acgaagttat tggtaggcga gaatgtagtc 480

agacccaagc tcacctctc ctggttcttc caggcccaga ctgctcccag ccagacccct 540

gcgtcagcag gccctgtgtt catggtgccc cctgctcagt ggggccggat ggccgatttg 600

cctgtgcctg cccacctggc taccagggtc aaagctgcca aagtgacata gatgagtgcc 660

gatctggtac aacttgccgt catggtggta cctgtctcaa tacacctgga tccttccgct 720

gccagtgtcc tcttggttat acagggctgc tgtgtgagaa cccctagtg ccctgtgccc 780

cttccccgtg tcgtaatggt ggcacctgta ggcagagcag tgatgtcaca tatgactgtg 840

cttgccttcc tggtaagtaa gttgtgccca gggaaggcag ctggggacaa taggctagcc 900

tcttagtgac cattgtcacc ttgtcctccc ctacgaggct tcgagggcca gaactgtgaa 960

gtcaacgtgg atgactgtcc tggacatcgg tgtctcaatg ggggaacgtg tgtagacggt 1020

gtcaatactt acaactgcca gtgccctccg gagtggacag gtgggcatca gggctgcaga 1080

gaaccagggt ggctgacctc aggtgggcac acgggcaact tagactagca catctttgtg 1140

ccctaggcca gttctgtaca gaagatgtgg atgagtgtca gctgcagccc aatgcctgcc 1200

acaatggggg tacctgcttc aacctactgg gtggccacag ctgtgtatgt gtcaatggct 1260

ggacgggtga gagctgcagt cagaatatcg atgactgtgc tacagccgtg tgtttccatg 1320

gggccacctg ccatgaccgt gtggcctctt tctactgtgc ctgccctatg gggaagacag 1380

gtgagtggcc cttttctttg taggcaacag aatggtttca gcatgaaagg t 1431

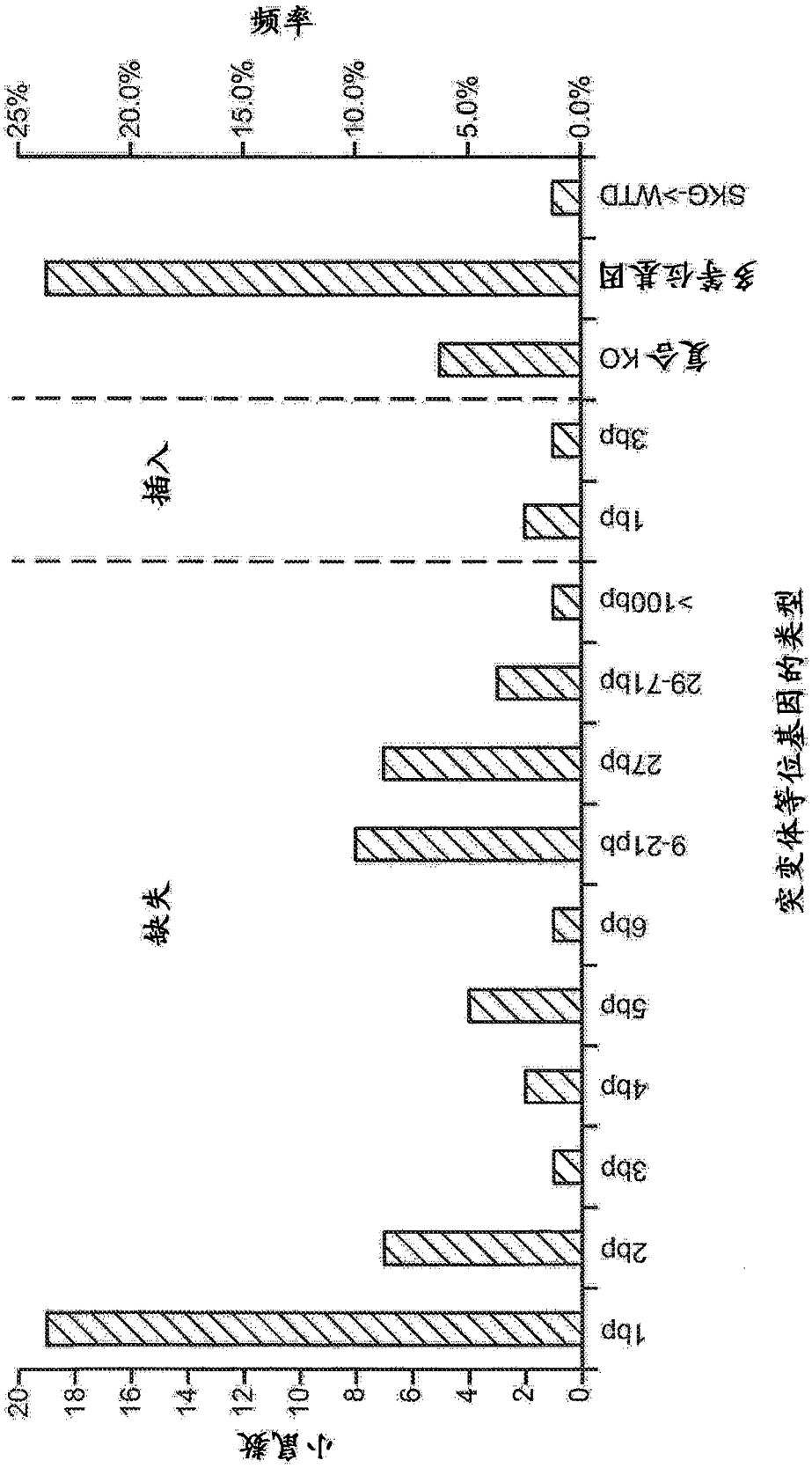

频率
25%
20.0%
15.0%
10.0%
5.0%
0.0%
SKG->WTD
多等位基因
复合KO
3bp
1bp
插入
>100bp
29-71bp
27bp
9-21pb
6bp
5bp
4bp
3bp
2bp
1bp
缺失
20
18
16
14
12
10
8
6
4
2
0
小鼠数
突变体等位基因的类型

图 1

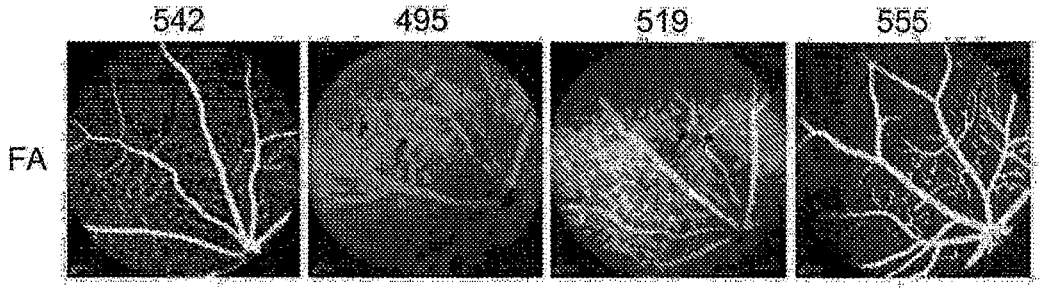


图 2A

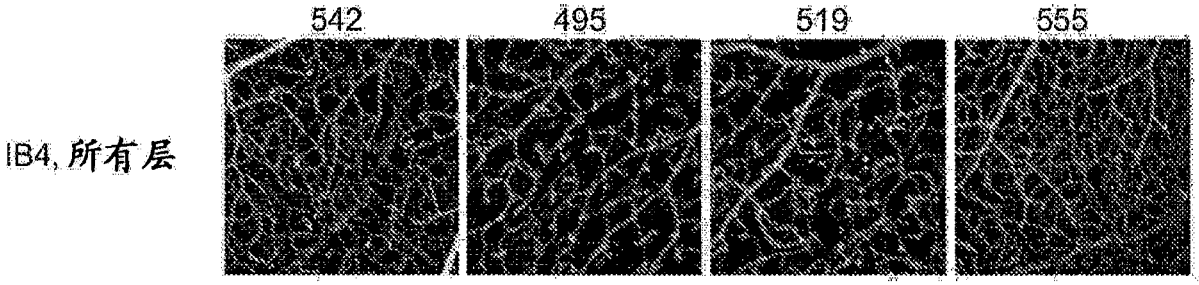


图 2B

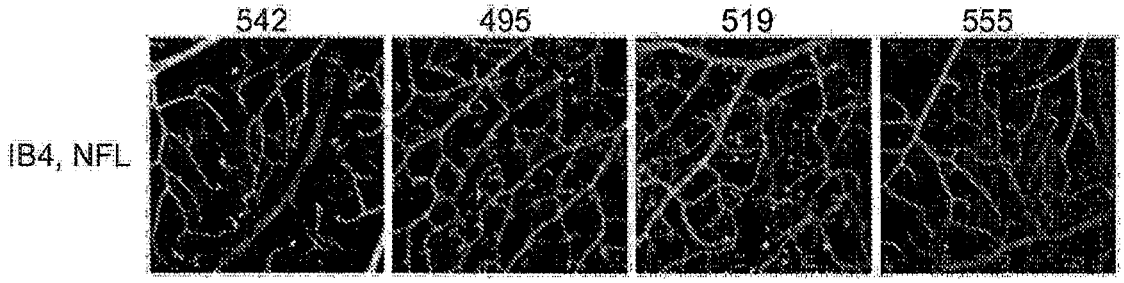


图 2C

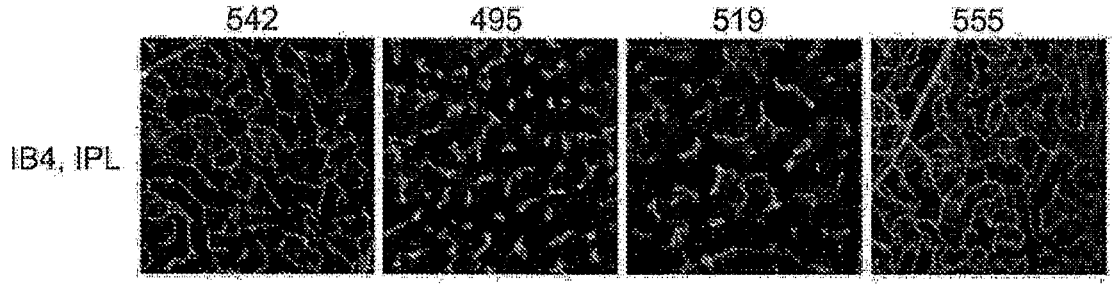


图 2D

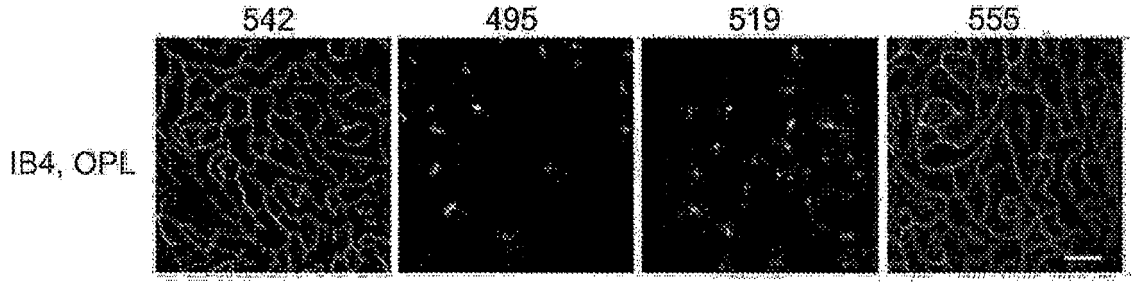

542
495
519
555
IB4, OPL

图 2E

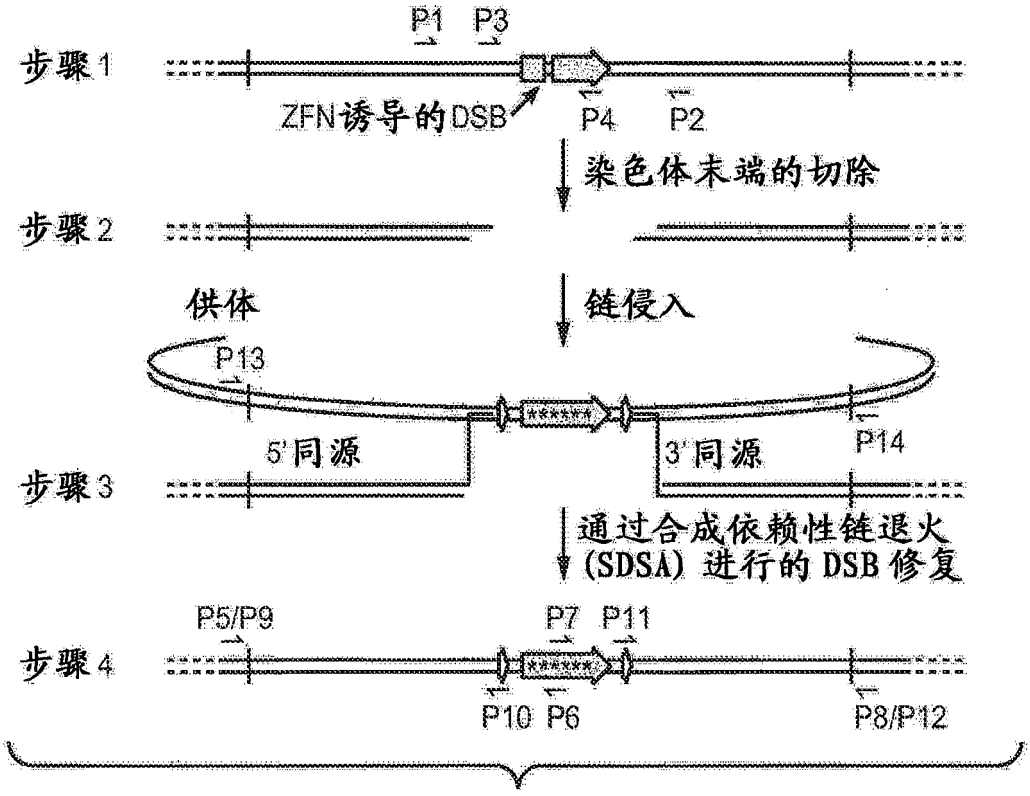

步骤 1
P1
P3
ZFN诱导的DSB
P4
P2
染色体末端的切除
步骤 2
供体
链侵入
P13
5'同源
3'同源
P14
步骤 3
通过合成依赖性链退火
(SDSA) 进行的 DSB 修复
P5/P9
P7
P11
步骤 4
P10
P6
P8/P12

图 3A

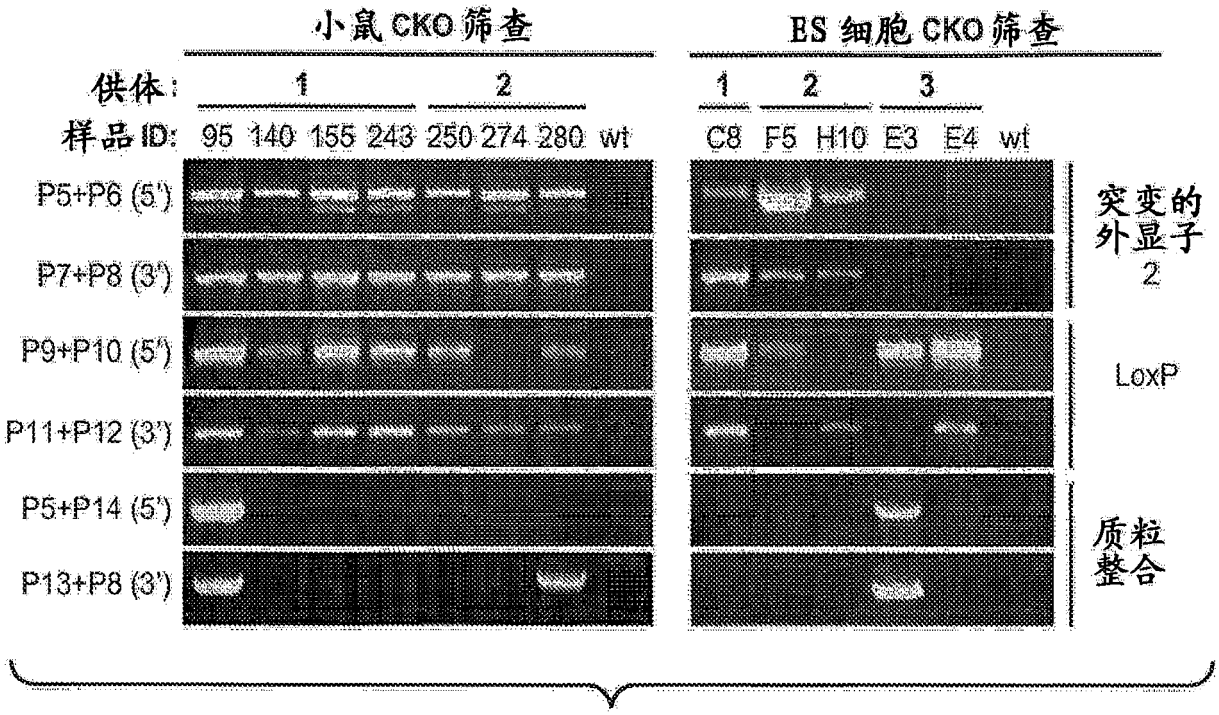

小鼠CKO筛查
ES细胞CKO筛查
供体:
1
2
1
2
3
样品ID:
95 140 155 243 250 274 280 wt
C8 F5 H10 E3 E4 wt
P5+P6 (5')
P7+P8 (3')
P9+P10 (5')
P11+P12 (3')
P5+P14 (5')
P13+P8 (3')
突变的外显子2
LoxP
质粒整合

图 3B

aatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgt
gagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgt
ggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagctcc
ttcctcttccagcccttcctctttcactgactgactgactggaagacacacctgcggccgcgca
agctgggccatggggcttaacccctaacatcactcttggtgtgacttggcacaagtgtatgctt
ctttatgaaggctgcagctctgggcctggagaatgtataccaggaccccctagtctaccaggggc
tgtgcatttgagaggcatgggtcacctggtaggggaggaacccgtaggaccagcctcctgactg
aggctggcactgtgtgtttctaccctgggaactttgcatgggttttaatgaggcctccctaac
cactgtatgtagtgtgcagccttctcctgttctgttcccactcctgtcctgtccctgactctgg
cccattcacagcattaacctgacagatgtcactacctgcttttggtctttgcccggtgtctgtt
tctccactggaacgggaaccttggaggcccctggtacaagcagcccctgcttgccagcctatga
aggcatcttggttttgccactgtggctctttttctttcactcccagaactttctttttctcctc
ctgagctttcctttgcagcttctgcccttgaaaaggagaggctggtggagccttggggtgatac
agtgtctgtgttgtctaggattccccaggcgggggaggagaggagaagcagcctgccccaggtg
agggtggggagccacccacattctggcctgcctcactgtgggtcccctggggccactgccaggcc
tcttagaaggaaagctagctttccaggagtatgctgtgggttcccccttaaaacttagggattaa
gtgtgtggtttctttacttttcagagctgggagcactgccccaccatctcccagtgagggttgg
agaaaaaaaaagtcctaggttacaaatccacccccagcctccctgtccctgcccctgccccctc
ccctcccatgtgcctttgaagagcacacctgaactactgaaagcaagttgaggtgaggaagtgg
tctttccagtttggggagcagtagagccagagcttggggatttggggagccttcctgcctctct
atactggcctatcactaagggagcctttgggcaactggaaggaaggagtcttgggagcaggggc
ccggggtagataacttcgtataatgtatgctatacgaagttatccgtgcatgattctctttccc
tgggagctggttcctgtcctgcttagccatcctgaccctgtgctctgcccacagcctcaccgct
cctcctctttgctaaccgccgagatgttcgactcgtagacgctggcggcgtgaagctggagagt
acaatcgttgctagtggactggaggatgctgccgctgtcgacttccagttctcaaagggagccg
tttattggaccgacgtgagcgaagaggcaataaaacagacataccttaaccagactggagctgc
cgcccagaatatcgtaatctctggcttggtgagccctgatggactcgcttgtgattgggtcgga
aaaaagctttattggactgatagtgaaacgaatcggatagaagtagcgaaccttaacggtacat
caagaaaggtcctgttttggcaggatcttgaccagcctagagcaatcgccctcgaccctgcaca
tgggtaaacctctgttttttctcctgctcctatggggaggcgcggggaacaggatttcttctgta
gagtataacttcgtataatgtatgctatacgaagttatacctgaggcaggctgatgaggtgcct
tgggaagcagcccttgctgggaatcacctcccttccagatcctgaactgtaaaacatggcatgt
ttagccagggatggtggttcacacctacaatcctagcattctggaggctgaggcaggagtgttg
ccacaagttccaggccaacctgagctataaagaccgtcttaaaaatatacaggaactggggaga
tgcctccattggtgatgtacttacctcgtaaccaggaggatttgacttccatccatgaagaaag
ttaggtgcggggcgtcacatgtctataatcccaggaccgtggagttagagacacagggatctgt
gaggcttgctggctgcctgcagtcaaattgcccagcccctggtcccagtgagagacaaggtgtt
ttaaggtatttatattgccgtcttaaaaactccatgaccaagaacagcttggggatgaatcagt
cagtatacttgatcagacaatactcaagtcatgccccacctctgagtaaagtcagggcaggacc
tcaaggcaggaacccagagtcaagaactaaactgcttagtggtttgctctccattggttgttca
gattgctctcttgtagcacccaggaccatcagcccaggcactgaactgcccacagtgacctgta
ccctccccccaccaatcatcagccaaggaaatgcttcccaggcttagaggccaatgcaaactgac
tgtagcttagctgagcaggacgttgtgtctcaatcaaacacacaaacaagggcagatgaggatg
atgcttgggtttgaagcttagcatccttctgcactaggacaggtgacccaggtaagttcaggga
cctccccacccccttctgtgactgttggctgtggtggctctgtgtctggatgtagcacctactctg
tgtattctggaagttacctatattggtgcagtcaactcactctgaccctggaaagcgcaggaca
gccatgaaggccacgcggccgcacgacagtcttcactgactgactgactggaaagtcctctcca
ctgactgtagcctccaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcg
ttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggc
ccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtat
tttctccttacgcatctgtgcggtatttcacaccgcatacgtcaaagcaaccatagtacgcgcc
ctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgcc
agcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttc
cccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcacctcga
ccccaaaaaacttgatttgggtgatggttcacgtagtgggccatcgc

图 4A-1

cctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgtt
ccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagggattttgccg
atttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaa
tattaacgtttacaattttatggtgcactctcagtacaatctgctctgatgccgcatagttaag
ccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatcc
gcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcac
cgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataat
aatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgttta
ttttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaat
aatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttttgc
ggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagat
cagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagtt
ttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtatt
atcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttg
gttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgca
gtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggacc
gaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa
ccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaa
caacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaataga
ctggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggttt
attgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccag
atggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacg
aaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtt
tactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaaga
tcctttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcaga
ccccgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaatctgctgcttg
caaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactcttt
ttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgta
gttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgtta
ccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttac
cggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaac
gacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaaggg
agaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttc
cagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg
atttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttta
cggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctg
tggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcg
cagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctctccccgcgcgt
tggccgattcatt

图 4A-2

tttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattagg
caccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataaca
atttcacacaggaaacagctatgaccatgattacgccaagctccttcctcttccagcccttcct
ctttccagtcagtcagtcagtggagactgtcgtgcggccgcgtggccttcatggctgtcctgcg
ctttccagggtcagagtgagttgactgcaccaatataggtaacttccagaatacacagagtagg
tgctacatccagacacagagccaccacagccaacagtcacagaaggggtgggaggtccctgaac
ttacctgggtcacctgtcctagtgcagaaggatgctaagcttcaaacccaagcatcatcctcat
ctgcccttgtttgtgtgtttgattgagacacaacgtcctgctcagctaagctacagtcagtttg
cattggcctctaagcctgggaagcatttccttggctgatgattggtgggggagggtacaggtca
ctgtgggcagttcagtgcctgggctgatggtcctgggtgctacaagagagcaatctgaacaacc
aatggagagcaaaccactaagcagtttagttcttgactctgggttcctgccttgaggtcctgcc
ctgactttactcagaggtggggcatgacttgagtattgtctgatcaagtatactgactgattca
tccccaagctgttcttggtcatggagtttttaagacggcaatataaataccttaaaacaccttg
tctctcactgggaccaggggctgggcaatttgactgcaggcagccagcaagcctcacagatccc
tgtgtctctaactccacggtcctgggattatagacatgtgacgccccgcacctaactttcttca
tggatggaagtcaaatcctcctggttacgaggtaagtacatcaccaatggaggcatctccccag
ttcctgtatatttttaagacggtctttatagctcaggttggcctggaacttgtggcaacactcc
tgcctcagcctccagaatgctaggattgtaggtgtgaaccaccatccctggctaaacatgccat
gttttacagttcaggatctggaagggaggtgattcccagcaagggctgcttcccaaggcacctc
atcagcctgcctcaggtataacttcgtatagcatacattatacgaagttatactctacagaaga
aatcctgttcccccgcgcctccccataggagcaggagaaaaacagaggtttacccatgtgcagga
tccagggcaatggcccttggctggtccaggtcctgccagaagagaaccttacgggacgtcccat
tgaggttggcaacctcaatgcggttggtctcggagtccgtccagtacagcttcttgccaaccca
gtcacaggccaggccatcaggtgacacgaggcccgagatgacaatgttctgtgcagcagctcca
gtctggttcaggtaggtctgtttgatggcctcctcgctcacgtcggtccaataaacggctccct
ttgagaactggaagtctacagcagctgcatcctccaggccactggccacaatggtggactccag
cttcactccgccggcatccactagccgcacatcccggcggttggcaaacaacaggagcggtgag
gctgtgggcagagcacagggtcaggatggctaagcaggacaggaaccagctcccagggaaagag
aatcatgcacggataacttcgtatagcatacattatacgaagttatctaccccgggcccctgct
cccaagactccttcctccagttgcccaaaggctcccttagtgataggccagtatagagaggca
ggaaggctccccaaatccccaagctctggctctactgctcccccaaactggaaagaccacttcct
cacctcaacttgctttcagtagttcaggtgtgctcttcaaaggcacatgggagggggaggggca
ggggcagggacagggaggctgggggtggatttgtaacctaggactttttttttctccaaccctc
actgggagatggtggggcagtgctcccagctctgaaaagtaaagaaaccacacacttaatccct
aagttttaaggggaacccacagcatactcctggaaagctagctttccttctaagaggcctggca
gtggccccagggacccacagtgaggcaggccagaatgtgggtggctccccaccctcacctgggg
caggctgcttctcctctcctcccccgcctggggaatcctagacaacacagacactgtatcaccc
caaggctccaccagcctctccttttcaagggcagaagctgcaaaggaaagctcaggaggagaaa
aagaaagttctgggagtgaaagaaaaagagccacagtggcaaaaccaagatgccttcataggct
ggcaagcaggggctgcttgtaccaggggcctccaaggttcccgttccagtggagaaacagacac
cgggcaaagaccaaaagcaggtagtgacatctgtcaggttaatgctgtgaatgggccagagtca
gggacaggacaggagtgggaacagaacaggagaaggctgcacactacatacagtggttagggga
ggcctcattaaaacccatgcaaagttcccagggtagaaacacacagtgccagcctcagtcagga
ggctggtcctacgggttcctccccctaccaggtgacccatgcctctcaaatgcacagcccctggt
agactaggggtcctggtatacattctccaggcccagagctgcagccttcataaagaagcataca
cttgtgccaagtcacaccaagagtgatgttaggggttaagccccatggcccagcttgcgcggcc
gcaggtgtgtcttccagtcagtcagtcagtgaaagtcctctccactgactgtagcctccaattc
actggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcctt
gcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttccc
aacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtg
cggtatttcacaccgcatacgtcaaagcaaccatagtacgcgccctgtagcggcgcattaagcg
cggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcc
tttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgg
gggctcccttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgatttgg
gtgatggttcacgtagt

图 4B-1

gggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtg
gactcttgttccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagg
gattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaat
tttaacaaaatattaacgtttacaattttatggtgcactctcagtacaatctgctctgatgccg
catagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgct
cccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttca
ccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatg
tcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccc
tatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataa
atgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattc
ccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaaga
tgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatc
cttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtg
gcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctca
gaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaaga
gaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacga
tcggaggaccgaaggagctaaccgcttttttgcacaacatggggatcatgtaactcgccttga
tcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgta
gcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggc
tggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagca
ctggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaacta
tggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtc
agaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatc
taggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccact
gagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaat
ctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagcta
ccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctag
tgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgct
aatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaaga
cgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagct
tggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgct
tcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacg
agggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgac
ttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgc
ggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaa
cgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctct
ccccgcgcgttggccgattcattaatgcagctggcacgacagg

图 4B-2

tttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattagg
caccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataaca
atttcacacaggaaacagctatgaccatgattacgccaagctccttcctcttccagcccttcct
ctttccagtcagtcagtcagtggagactgtcgtgcggccgcgtggccttcatggctgtcctgcg
ctttccagggtcagagtgagttgactgcaccaatataggtaacttccagaatacacagagtagg
tgctacatccagacacagagccaccacagccaacagtcacagaaggggtgggaggtccctgaac
ttacctgggtcacctgtcctagtgcagaaggatgctaagcttcaaacccaagcatcatcctcat
ctgcccttgtttgtgtgtttgattgagacacaacgtcctgctcagctaagctacagtcagtttg
cattggcctctaagcctgggaagcatttccttggctgatgattggtgggggagggtacaggtca
ctgtgggcagttcagtgcctgggctgatggtcctgggtgctacaagagagcaatctgaacaacc
aatggagagcaaaccactaagcagtttagttcttgactctgggttcctgccttgaggtcctgcc
ctgactttactcagaggtggggcatgacttgagtattgtctgatcaagtatactgactgattca
tccccaagctgttcttggtcatggagtttttaagacggcaatataaatacctttaaaacaccttg
tctctcactgggaccaggggctgggcaatttgactgcaggcagccagcaagcctcacagatccc
tgtgtctctaactccacggtcctgggattatagacatgtgacgccccgcacctaactttcttca
tggatggaagtcaaatcctcctggttacgaggtaagtacatcaccaatggaggcatctccccag
ttcctgtatatttttaagacggtctttatagctcaggttggcctggaacttgtggcaacactcc
tgcctcagcctccagaatgctaggattgtaggtgtgaaccaccatccctggctaaacatgccat
gttttacagttcaggatctggaagggaggtgattcccagcaagggctgcttcccaaggcacctc
atcagcctgcctcaggtataacttcgtatagcatacattatacgaagttatactctacagaaga
aatcctgttccccgcgcctccccataggagcaggagaaaaacagaggtttacccatgtgcagga
tccagggcaatggcccttggctggtccaggtcctgccagaagagaaccttacgggacgtcccat
tgaggttggcaacctcaatgcggttggtctcggagtccgtccagtacagcttcttgccaaccca
gtcacaggccaggccatcaggtgacacgaggcccgagatgacaatgttctgtgcagcagctcca
gtctggttcaggtaggtctgtttgatggcctcctcgctcacatctgtccagtacacagcaccct
tggagaactggaagtctacagcagctgcatcctccaggccactggccacaatggtggactccag
cttcactccgccggcatccactagccgcacatcccggcggttggcaaacaacaggagcggtgag
gctgtgggcagagcacagggtcaggatggctaagcaggacaggaaccagctcccagggaaagag
aatcatgcacggataacttcgtatagcatacattatacgaagttatctaccccgggccctgct
cccaagactccttccttccagttgcccaaaggctcccttagtgataggccagtatagagaggca
ggaaggctccccaaatccccaagctctggctctactgctccccaaactggaaagaccacttcct
cacctcaacttgctttcagtagttcaggtgtgctcttcaaaggcacatgggaggggagggggca
ggggcagggacagggaggctgggggtggatttgtaacctaggactttttttttctccaaccctc
actgggagatggtggggcagtgctcccagctctgaaaagtaaagaaaccacacacttaatccct
aagttttaaggggaacccacagcatactcctggaaagctagctttccttctaagaggcctggca
gtggccccagggacccacagtgaggcaggccagaatgtgggtggctccccacccctcacctgggg
caggctgcttctcctctcctccccccgcctggggaatcctagacaacacagacactgtatcaccc
caaggctccaccagcctctcctttttcaagggcagaagctgcaaaggaaagctcaggaggagaaa
aagaaagttctgggagtgaaagaaaaagagccacagtggcaaaaccaagatgccttcataggct
ggcaagcaggggctgcttgtaccaggggcctccaaggttcccgttccagtggagaaacagacac
cgggcaaagaccaaaagcaggtagtgacatctgtcaggttaatgctgtgaatgggccagagtca
gggacaggacaggagtgggaacagaacaggagaaggctgcacactacatacagtggttagggga
ggcctcattaaaacccatgcaaagttcccagggtagaaacacacagtgccagcctcagtcagga
ggctggtcctacgggttcctccctaccaggtgacccatgcctctcaaatgcacagcccctggt
agactaggggtcctggtatacattctccaggcccagagctgcagccttcataaagaagcataca
cttgtgccaagtcacaccaagagtgatgttaggggttaagccccatggcccagcttgcgcggcc
gcaggtgtgtcttccagtcagtcagtcagtgaaagtcctctccactgactgtagcctccaattc
actggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgcctt
gcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttccc
aacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtg
cggtatttcacaccgcatacgtcaaagcaaccatagtacgcgccctgtagcggcgcattaagcg
cggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcc
tttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgg
gggctcccttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgatttgg
gtgatggttcacgtagt

图 4C-1

gggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtg
gactcttgttccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagg
gattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaat
tttaacaaaatattaacgtttacaattttatggtgcactctcagtacaatctgctctgatgccg
catagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgct
cccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttca
ccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatg
tcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccc
tatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataa
atgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattc
ccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaaga
tgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatc
cttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtg
gcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctca
gaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaaga
gaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacga
tcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttga
tcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgta
gcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggc
tggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagca
ctggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaacta
tggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtc
agaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatc
taggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccact
gagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaat
ctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagcta
ccaactcttttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctag
tgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgct
aatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaaga
cgatagttaccggataaggcgcagcggtcgggctgaacggggggttcgtgcacacagcccagct
tggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgct
tcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacg
agggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgac
ttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgc
ggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcc
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaa
cgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctct
ccccgcgcgttggccgattcattaatgcagctggcacgacagg

图 4C-2

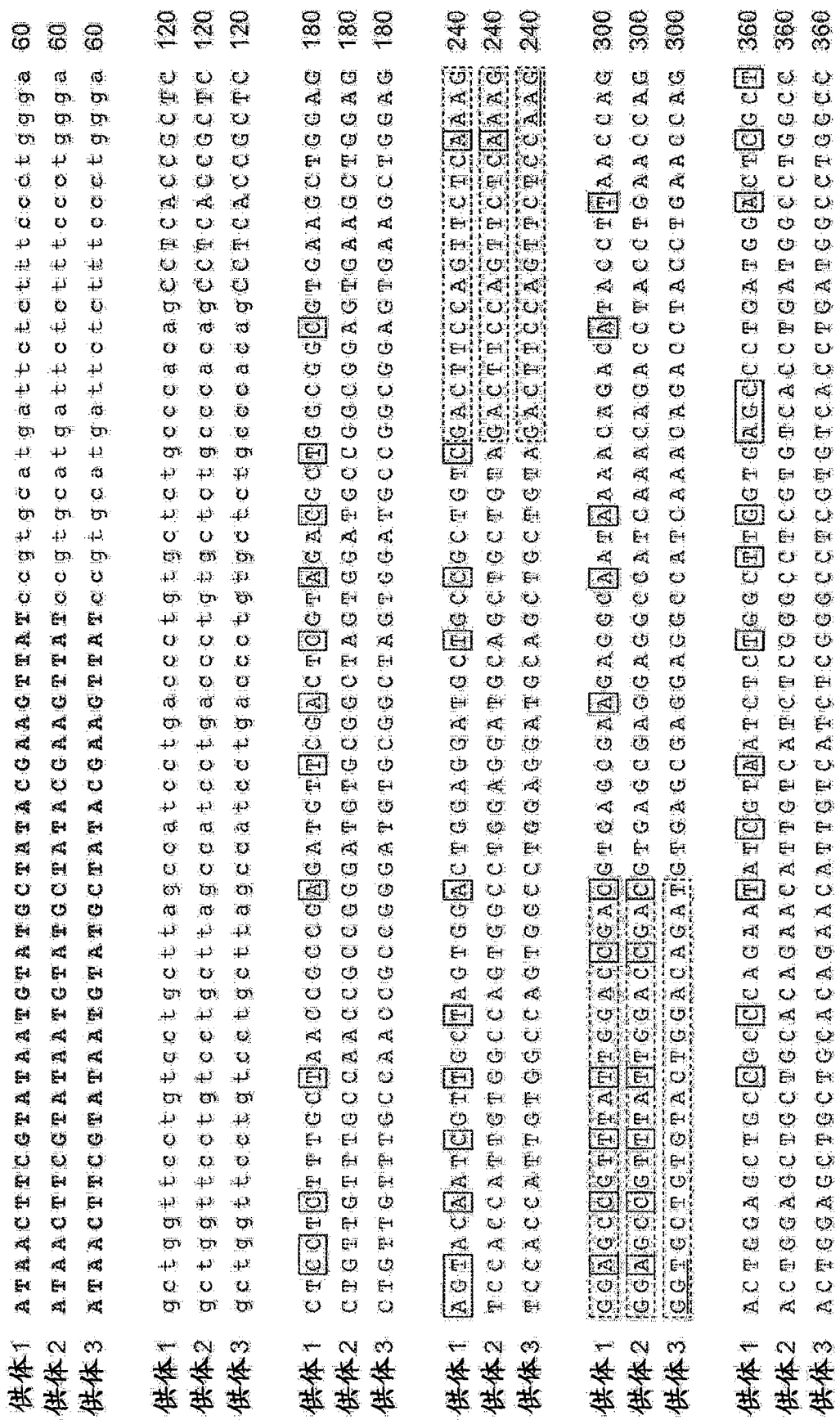


图 5A

```
供体1  TGTTGATGGGTCGCAAAAGCTTATTGGACTGATAGTGAACGGATAATCGGTAA 420
供体2  TGTTGACTGGGTTGGCCAAGAAGCCTGTACTGGACGACCCGACCAACCGTTGA 420
供体3  TGTTGACTTGGGTTGGCAAGAAGCCTGTACTGGACGACCCGACCAACCGCATTGAGGTT 420

供体1  GCCGAACCCTTACCGCTACAGCAAGGTCTTGTTGGCCAGGACCCAGCCAGCCAAGG 480
供体2  GCCAAACCCTCAATGGGACCGTCCCGTAAGGTTCCTTCTGGCAGGACCAGCCAAGG 480
供体3  GCCAACCCTCAATGGGACCGTCCCGTAAGGTTCCTTCTGGCAGGACCTGGACCCAAGG 480

供体1  GCCATCGCCCTCGACCCTGGCACATGGgtaaacctcctgctcctatgggga 540
供体2  GCCATTGCCCTGGATCCTGGCACATGGgtaaacctcctgctcctatgggga 540
供体3  GCCATTGCCCTGGATCCTGGCACATGGgtaaacctcctgctcctatgggga 540

供体1  ggcgcgggggaacagggatttcctctgtagagtATAACTTCGTATAATGCTATGCTATACGAA 600
供体2  ggcgcgggggaacagggatttcctctgtagagtATAACTTCGTATAATGCTATGCTATACGAA 600
供体3  ggcgcgggggaacagggatttcctctgtagagtATAACTTCGTATAATGCTATGCTATACGAA 600

供体1  GTTAT 605
供体2  GTTAT 605
供体3  GTTAT 605
```

图 5B

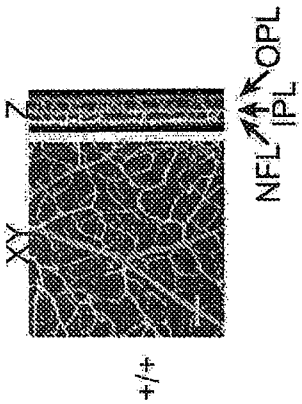


图 6A

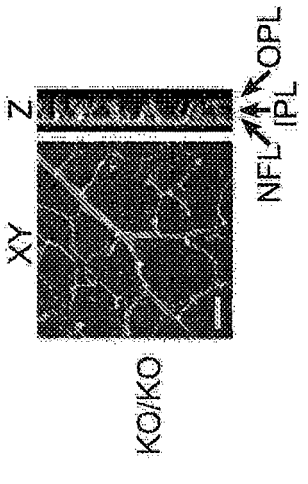


图 6B

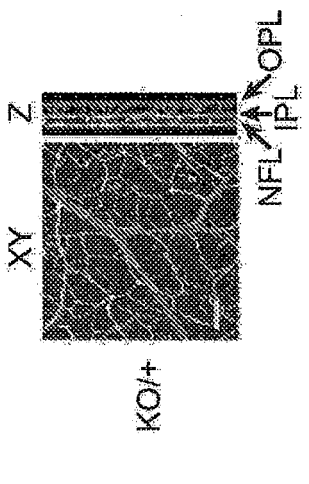


图 6C

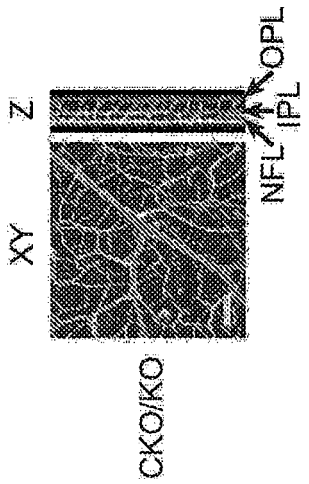


图 6D

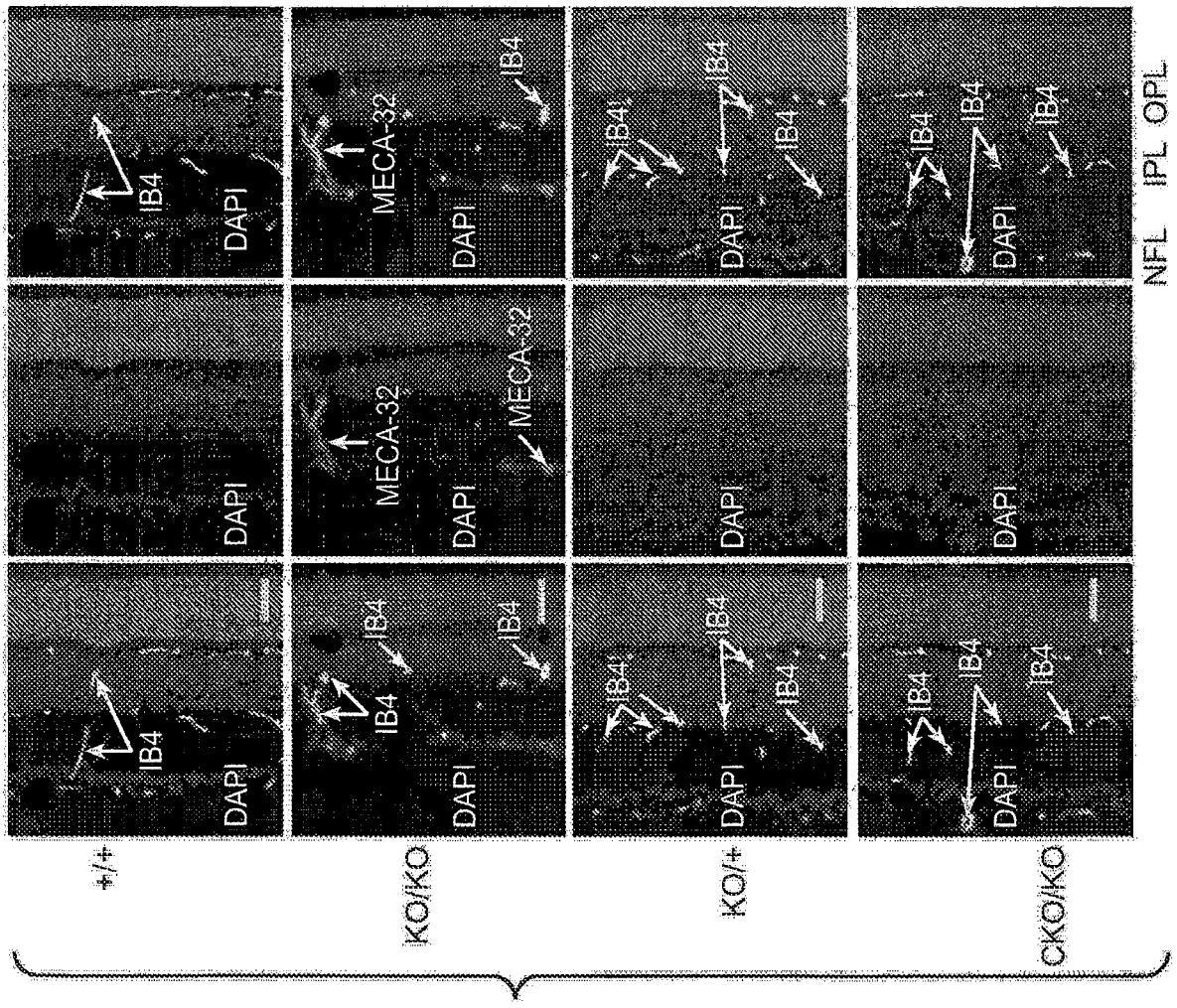


图 6E

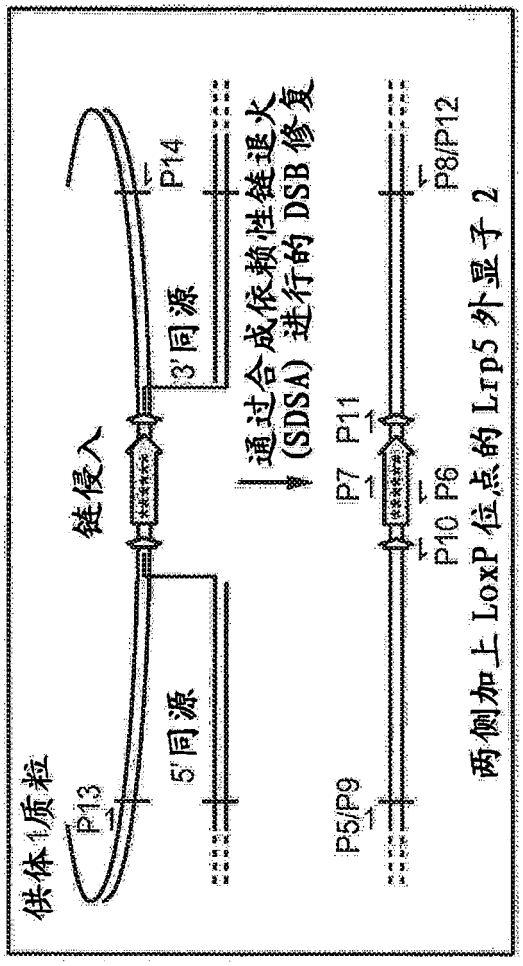

供体1/质粒
P13
链侵入
P14
3'同源
5'同源
通过合成依赖性链退火
(SDSA) 进行的 DSB 修复
P7 P11
P5/P9
P10 P6
P8/P12
两侧加上 LoxP 位点的 Lrp5 外显子 2

图 7A

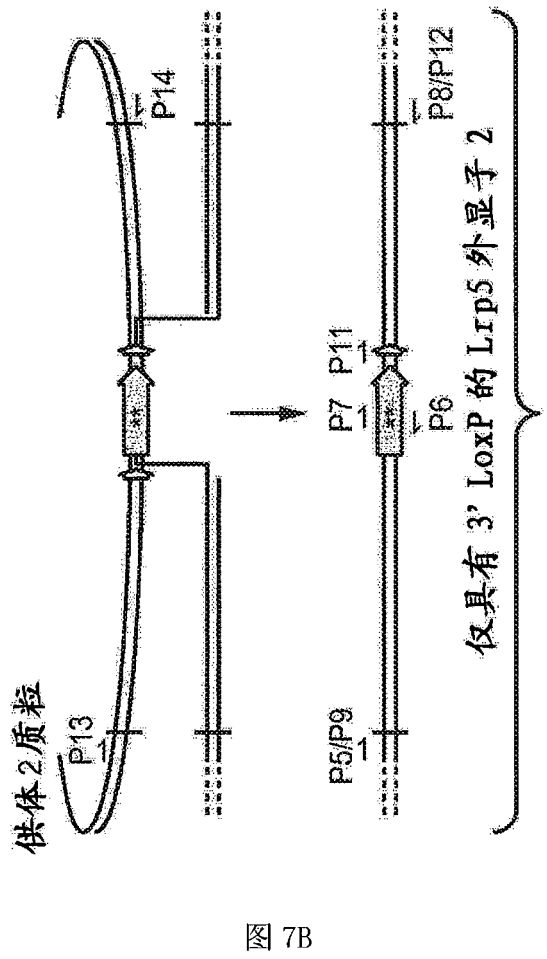

供体2/质粒
P13
P14
P5/P9
P7 P11
P6
P8/P12
仅具有 3' LoxP 的 Lrp5 外显子 2

图 7B

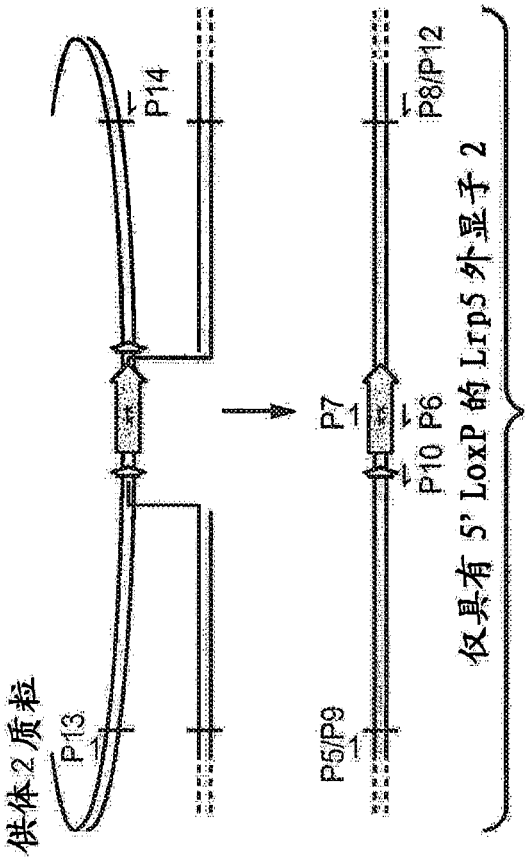

供体 2 质粒
P13
P14
P5/P9
P7
P10 P6
P8/P12
仅具有 5' LoxP 的 Lrp5 外显子 2

图 7C

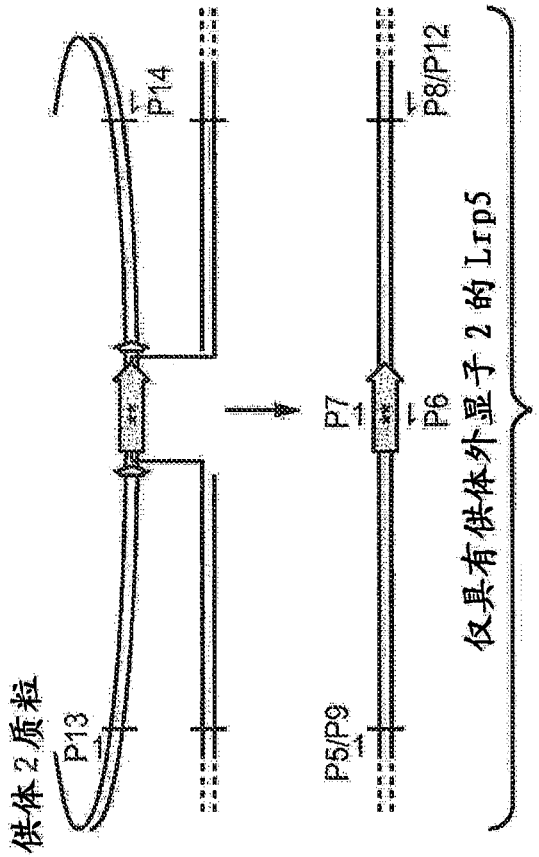

供体 2 质粒
P13
P14
P5/P9
P7
P6
P8/P12
仅具有供体外显子 2 的 Lrp5

图 7D

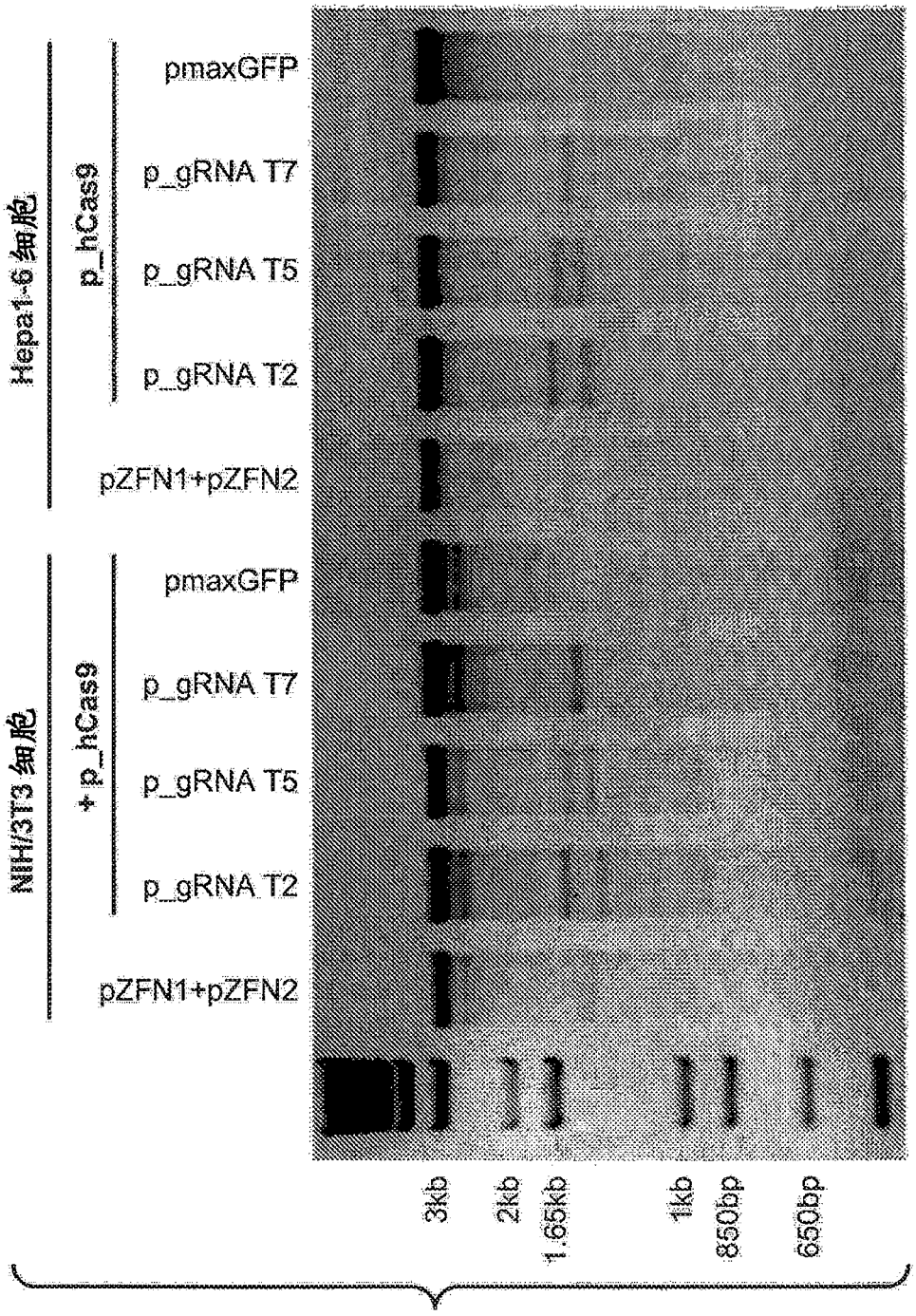

Hepa1-6 细胞
NIH/3T3 细胞
p_hCas9
+ p_hCas9
pmaxGFP
p_gRNA T7
p_gRNA T5
p_gRNA T2
pZFN1+pZFN2
pmaxGFP
p_gRNA T7
p_gRNA T5
p_gRNA T2
pZFN1+pZFN2
3kb
2kb
1.65kb
1kb
850bp
650bp

图 8

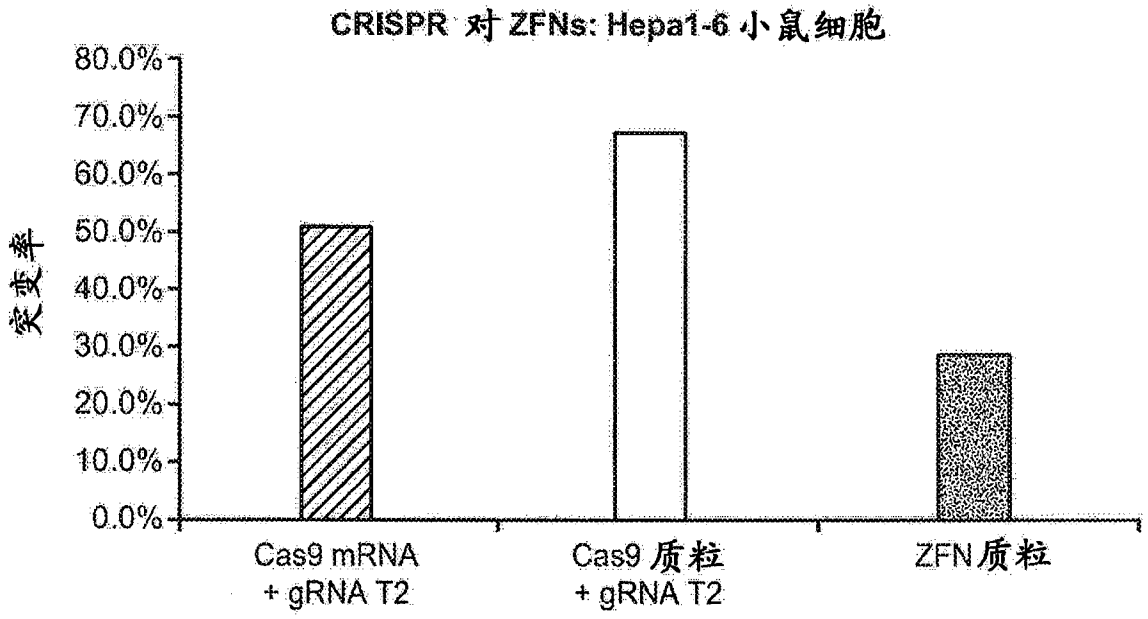

CRISPR 对 ZFNs: Hepa1-6 小鼠细胞
突变率
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
Cas9 mRNA + gRNA T2
Cas9 质粒 + gRNA T2
ZFN 质粒

图 9A

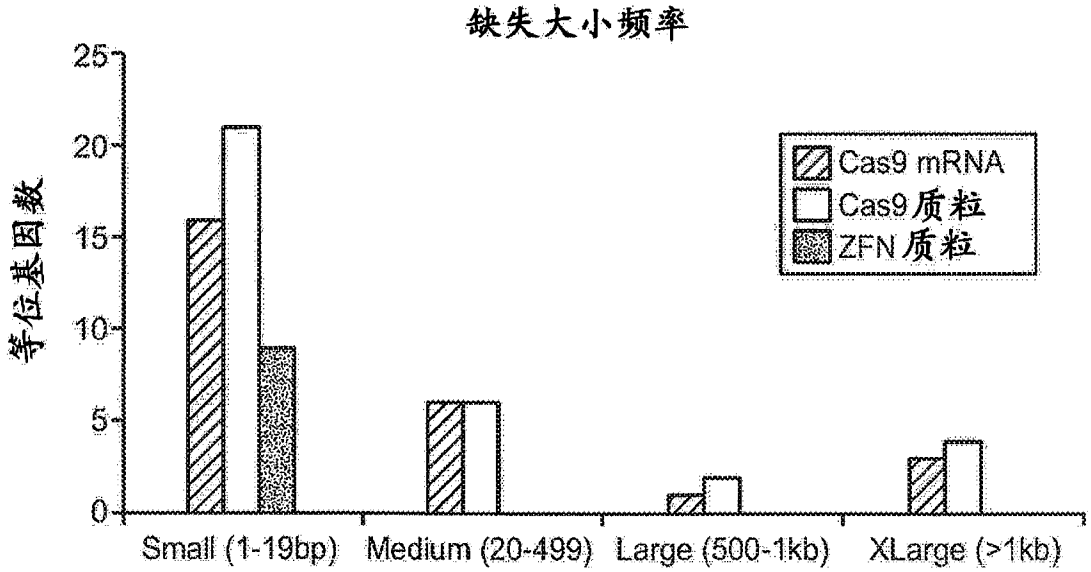

缺失大小频率
等位基因数
25
20
15
10
5
0
Cas9 mRNA
Cas9 质粒
ZFN 质粒
Small (1-19bp)
Medium (20-499)
Large (500-1kb)
XLarge (>1kb)

图 9B

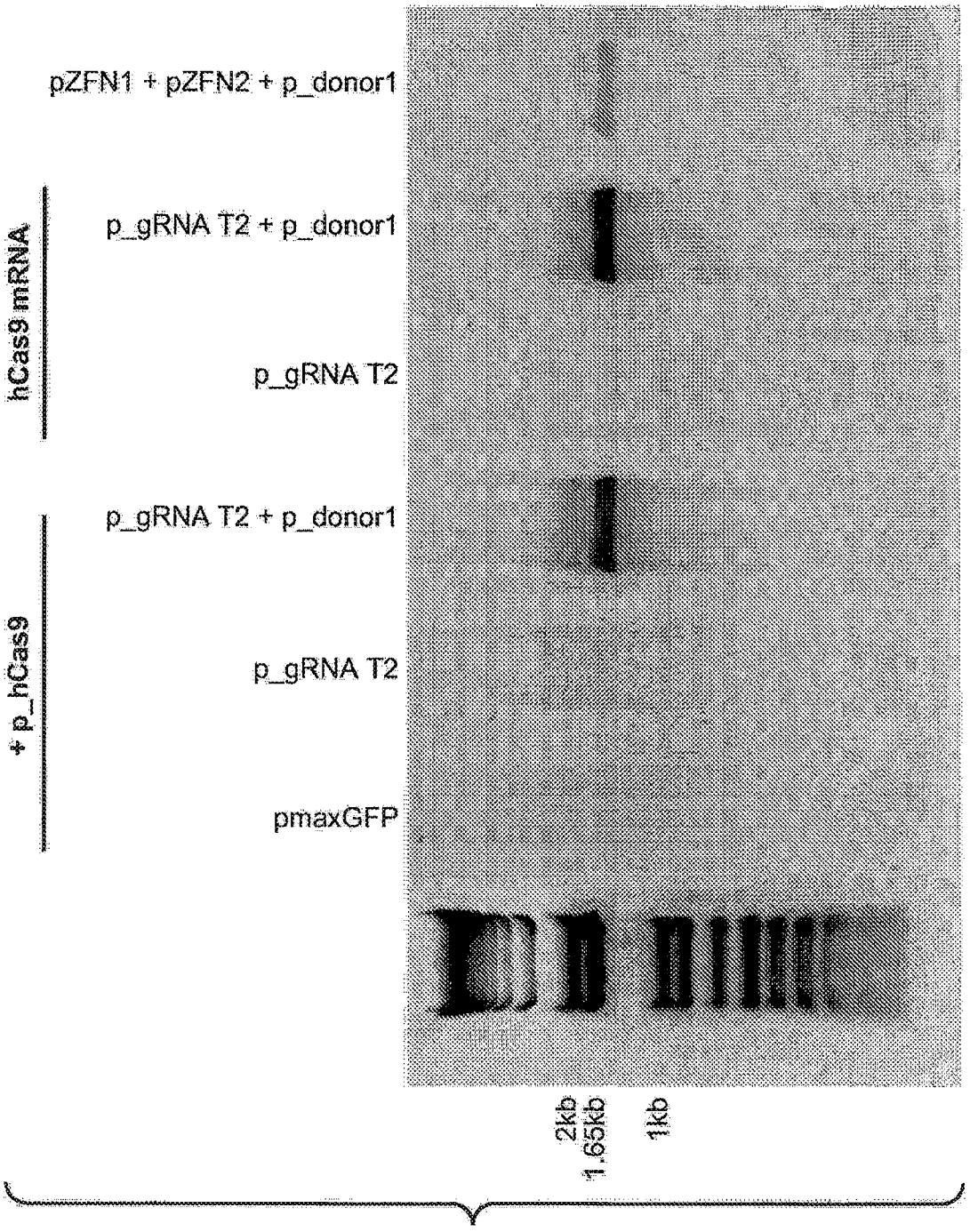

pZFN1 + pZFN2 + p_donor1
hCas9 mRNA
p_gRNA T2 + p_donor1
p_gRNA T2
+ p_hCas9
p_gRNA T2 + p_donor1
p_gRNA T2
pmaxGFP
2kb
1.65kb
1kb

图 10

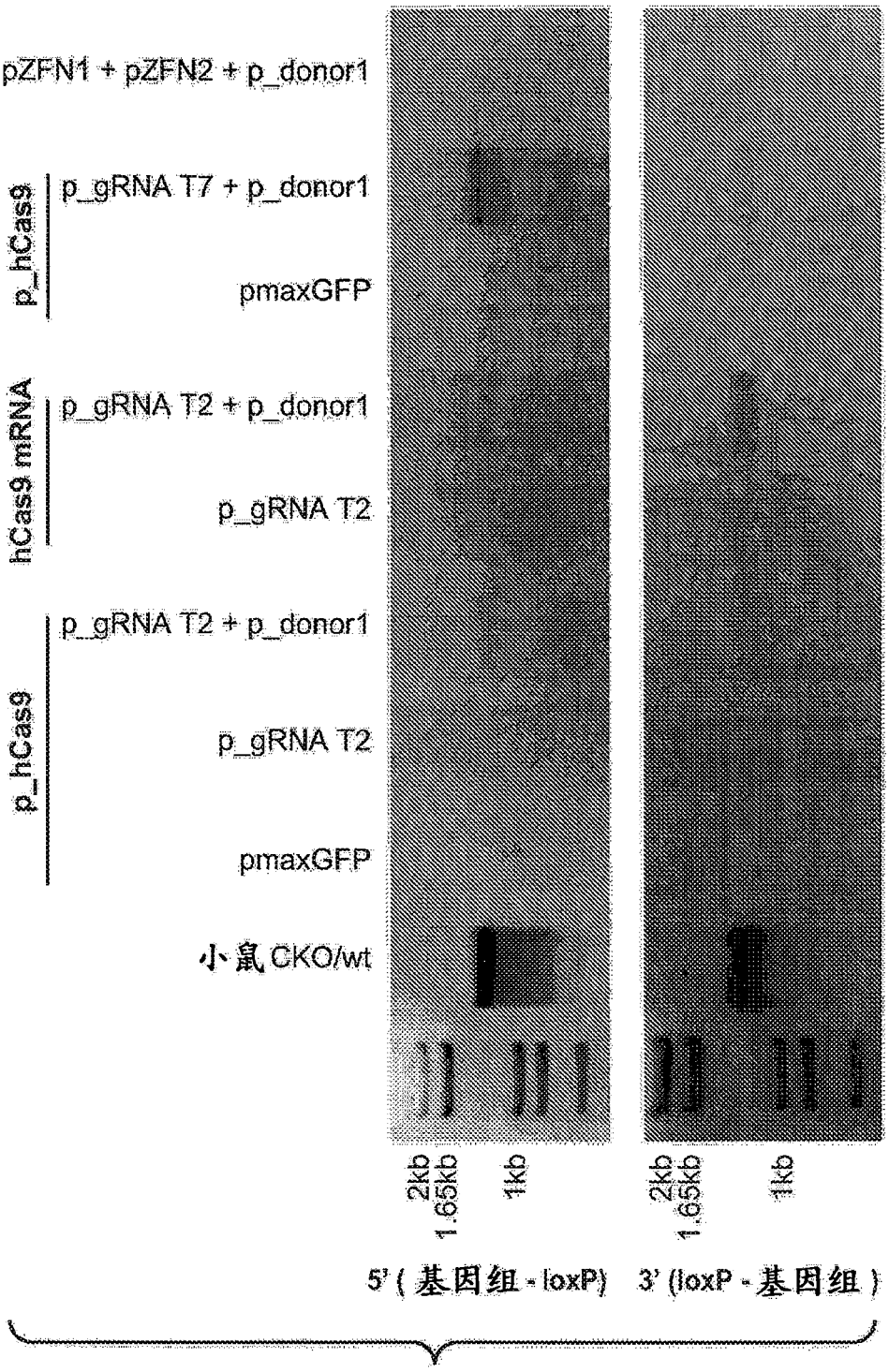

pZFN1 + pZFN2 + p_donor1
p_hCas9
p_gRNA T7 + p_donor1
pmaxGFP
hCas9 mRNA
p_gRNA T2 + p_donor1
p_gRNA T2
p_hCas9
p_gRNA T2 + p_donor1
p_gRNA T2
pmaxGFP
小鼠 CKO/wt
2kb
1.65kb
1kb
2kb
1.65kb
1kb
5' (基因组 - loxP)
3' (loxP - 基因组)

图 11

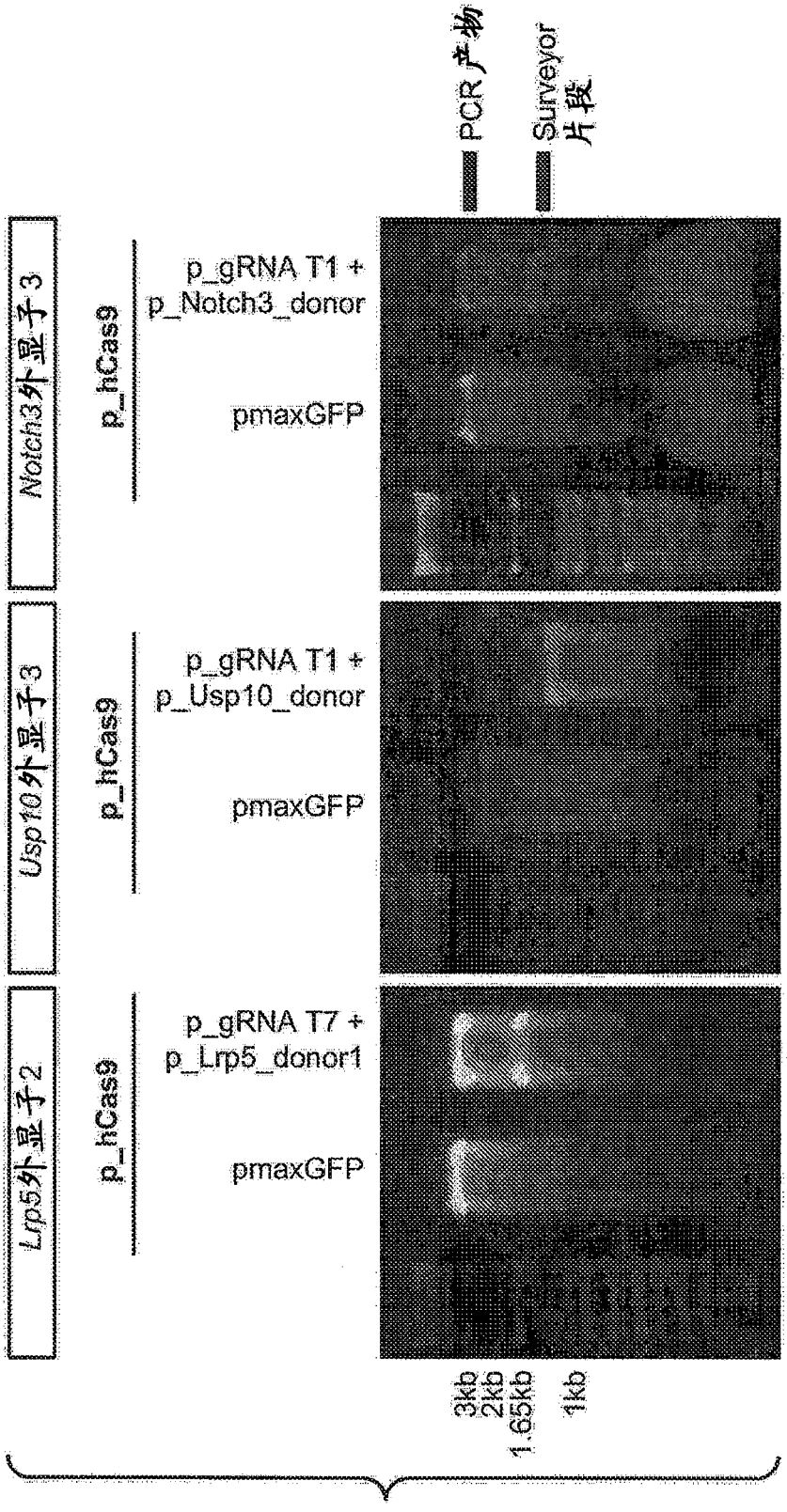

PCR产物
Surveyor片段
Notch3外显子3
p_hCas9
p_gRNA T1 +
p_Notch3_donor
pmaxGFP
Usp10外显子3
p_hCas9
p_gRNA T1 +
p_Usp10_donor
pmaxGFP
Lrp5外显子2
p_hCas9
p_gRNA T7 +
p_Lrp5_donor1
pmaxGFP
3kb
2kb
1.65kb
1kb

图 12

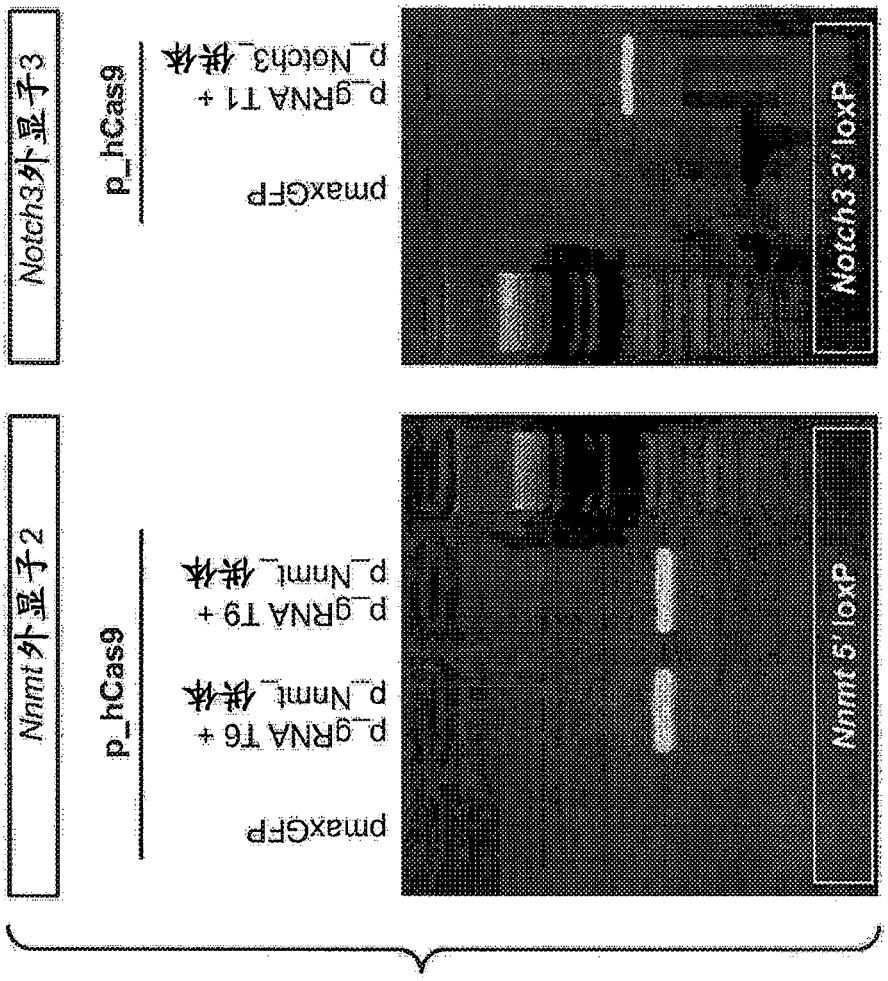

Notch3外显子3
p_hCas9
pmaxGFP
p_gRNA T1 +
p_Notch3_供体
Notch3 3' loxP
Nnmt外显子2
p_hCas9
pmaxGFP
p_gRNA T6 +
p_Nnmt_供体
p_gRNA T9 +
p_Nnmt_供体
Nnmt 5' loxP

图 13

Lrp5 gRNA T2
ggtgcggctagtggatgccgggttttagagctagaaatagcaagttaaaataaggctagtccgt
tatcaacttgaaaaagtggcaccgagtcggtgcttttttt

Lrp5 gRNA T5
ggagtccaccattgtggccaggttttagagctagaaatagcaagttaaaataaggctagtccgt
tatcaacttgaaaaagtggcaccgagtcggtgcttttttt

Lrp5 gRNA T7
ggtactggacagatgtgagcggttttagagctagaaatagcaagttaaaataaggctagtccgt
tatcaacttgaaaaagtggcaccgagtcggtgcttttttt

Usp10 gRNA T1
gctgaagatgaactgccagagttttagagctagaaatagcaagttaaaataaggctagtccgtt
atcaacttgaaaaagtggcaccgagtcggtgcttttttt

Nnmt gRNA T6
gatctccgtgaaggactcacgttttagagctagaaatagcaagttaaaataaggctagtccgtt
atcaacttgaaaaagtggcaccgagtcggtgcttttttt

Nnmt gRNA T9
gaccactggggaccagtcaagttttagagctagaaatagcaagttaaaataaggctagtccgtt
atcaacttgaaaaagtggcaccgagtcggtgcttttttt

Notch3 gRNA T1
gaggcgtttgccagagttcaggttttagagctagaaatagcaagttaaaataaggctagtccgt
tatcaacttgaaaaagtggcaccgagtcggtgcttttttt

hCas9 cDNA
atggacaagaagtacagcatcggcctggacatcggcaccaactctgtgggctgggccgtgatca
ccgacgagtacaaggtgcccagcaagaaattcaaggtgctgggcaacaccgaccggcacagcat
caagaagaacctgatcggcgccctgctgttcgacagcggagaaacagccgaggccacccggctg
aagagaaccgccagaagaagatacaccagacggaagaaccggatctgctatctgcaagagatct
tcagcaacgagatggccaaggtggacgacagcttcttccacagactggaagagtccttcctggt
ggaagaggataagaagcacgagcggcacccatcttcggcaacatcgtggacgaggtggcctac
cacgagaagtaccccaccatctaccacctgagaaagaaactggtggacagcaccgacaaggccg
acctgcggctgatctatctggccctggcccacatgatcaagttccggggccacttcctgatcga
gggcgacctgaaccccgacaacagcgacgtggacaagctgttcatccagctggtgcagacctac
aaccagctgttcgaggaaaaccccatcaacgccagcggcgtggacgccaaggccatcctgtctg
ccagactgagcaagagcagacggctggaaaatctgatcgcccagctgccccggcgagaagaagaa
tggcctgttcggcaacctgattgccctgagcctgggcctgacccccaacttcaagagcaacttc
gacctggccgaggatgccaaactgcagctgagcaaggacacctacgacgacgacctggacaacc
tgctggcccagatcggcgaccagtacgccgacctgtttctggccgccaagaacctgtccgacgc
catcctgctgagcgacatcctgagagtgaacaccgagatcaccaaggcccccctgagcgcctct
atgatcaagagatacgacgagcaccaccaggacctgaccctgctgaaagctctcgtgcggcagc
agctgcctgagaagtacaaagagattttcttcgaccagagcaagaacggctacgccggctacat
cgatggcggagccagccaggaagagttctacaagttcatcaagcccatcctggaaaagatggac
ggcaccgaggaactgctcgtgaagctgaacagagaggacctgctgcggaagcagcggaccttcg

图 14A

acaacggcagcatccccaccagatccacctgggagagctgcacgccattctgcggcggcagga
agatttttacccattcctgaaggacaaccgggaaaagatcgagaagatcctgaccttccgcatc
ccctactacgtgggccctctggccagggaaacagcagattcgcctggatgaccagaaagagcg
aggaaaccatcacccctggaacttcgaggaagtggtggacaagggcgccagcgcccagagctt
catcgagcggatgaccaacttcgataagaacctgcccaacgagaaggtgctgcccaagcacagc
ctgctgtacgagtacttcaccgtgtacaacgagctgaccaaagtgaaatacgtgaccgagggaa
tgagaaagcccgccttcctgagcggcgagcagaaaaaagccatcgtggacctgctgttcaagac
caaccggaaagtgaccgtgaagcagctgaaagaggactacttcaagaaaatcgagtgcttcgac
tccgtggaaatctccggcgtggaagatcggttcaacgcctccctgggcacataccacgatctgc
tgaaaattatcaaggacaaggacttcctggacaatgaggaaaacgaggacattctggaagatat
cgtgctgaccctgacactgtttgaggacagagagatgatcgaggaacggctgaaaacctatgcc
cacctgttcgacgacaaagtgatgaagcagctgaagcggcggagatacaccggctggggcaggc
tgagccggaagctgatcaacggcatccgggacaagcagtccggcaagacaatcctggatttcct
gaagtccgacggcttcgccaacagaaacttcatgcagctgatccacgacgacagcctgaccttt
aaagaggacatccagaaagcccaggtgtccggccagggcgatagcctgcacgagcacattgcca
atctggccggcagccccgccattaagaagggcatcctgcagacagtgaaggtggtggacgagct
cgtgaaagtgatgggccggcacaagcccgagaacatcgtgatcgaaatggccagagagaaccag
accacccagaagggacagaagaacagccgcgagagaatgaagcggatcgaagagggcatcaaag
agctgggcagccagatcctgaaagaacaccccgtggaaaacacccagctgcagaacgagaagct
gtacctgtactacctgcagaatgggcgggatatgtacgtggaccaggaactggacatcaaccgg
ctgtccgactacgatgtggaccatatcgtgcctcagagctttctgaaggacgactccatcgata
acaaagtgctgactcggagcgacaagaaccggggcaagagcgacaacgtgccctccgaagaggt
cgtgaagaagatgaagaactactggcgccagctgctgaatgccaagctgattacccagaggaag
ttcgacaatctgaccaaggccgagagaggcggcctgagcgaactggataaggccggcttcatca
agagacagctggtggaaacccggcagatcacaaagcacgtggcacagatcctggactcccggat
gaacactaagtacgacgagaacgacaaactgatccgggaagtgaaagtgatcaccctgaagtcc
aagctggtgtccgatttccggaaggatttccagttttacaaagtgcgcgagatcaacaactacc
accacgcccacgacgcctacctgaacgccgtcgtgggaaccgccctgatcaaaaagtaccctaa
gctggaaagcgagttcgtgtacggcgactacaaggtgtacgacgtgcggaagatgatcgccaag
agcgagcaggaaatcggcaaggctaccgccaagtacttcttctacagcaacatcatgaacttttt
tcaagaccgagattaccctggccaacggcgagatccggaagcggcctctgatcgagacaaacgg
cgaaacaggcgagatcgtgtgggataagggccgggactttgccaccgtgcggaaagtgctgtct
atgccccaagtgaatatcgtgaaaaagaccgaggtgcagacaggcggcttcagcaaagagtcta
tcctgcccaagaggaacagcgacaagctgatcgccagaaagaaggactgggaccctaagaagta
cggcggcttcgacagccccaccgtggcctattctgtgctggtggtggccaaagtggaaaagggc
aagtccaagaaactgaagagtgtgaaagagctgctggggatcaccatcatggaaagaagcagct
tcgagaagaatcccatcgactttctggaagccaagggctacaaagaagtgaaaaaggacctgat
catcaagctgcctaagtactccctgttcgagctggaaaacggccggaagagaatgctggcctct
gccggcgaactgcagaagggaaacgaactggccctgccctccaaatatgtgaacttcctgtacc
tggccagccactatgagaagctgaagggctcccccgaggataatgagcagaaacagctgtttgt
ggaacagcacaaacactacctggacgagatcatcgagcagatcagcgagttctccaagagagtg
atcctggccgacgctaatctggacaaggtgctgagcgcctacaacaagcacagagacaagccta
tcagagagcaggccgagaatatcatccacctgtttaccctgaccaatctgggagcccctgccgc
cttcaagtactttgacaccaccatcgaccggaagaggtacaccagcaccaaagaggtgctggac
gccaccctgatccaccagagcatcaccggcctgtacgagacacggatcgacctgtctcagctgg
gaggcgacgcctatccctatgacgtgcccgattatgccagcctgggcagcggctcccccaagaa
aaaacgcaaggtggaagatcctaagaaaaagcggaaagtggactga

图 14B

CO_Usp10ex3 donor

cgtctgtcccctagtaccatccacagagcagacattctctgtaagggtgtcttgccctttgtg
acctagagacgggatttggtgagcagccgtggtgatgttgctgtaccacactgcttctgtgtgc
tgtcagcgatggcatgtggaagagtttgcctgtgtcaaggtgctaagtcctaggctgccaggtc
ctgagtggttgggctcctttccagccatggttggtttttcttgctgtttgtcatcacctcctg
gggagagaacccagcactttatttttggggaaggaagtaaccttgcctcgttagaactgttgca
ttggctagccagagttcatctggagacacgagcctaggaggccacgggcacagcagacagtgat
tgccctgggcaccagctagagcctcctggactacttcacagatgctgcagcagggtggtgtttg
gggctcggttctgtcagctgactcacttctcagacacccccaaatcctaaataaataatagaat
tagcaggaagaagtatttgagaggttcataaaatttgccaggttttcacaaagtagccagacct
atgttgttttcttcctttcttcatacaaggaaggtgacgttaaaccattaactcatctgcatgg
cgcactgtaaaatgacagttaagggcccagccctttcttggggactctggctcacttctgccag
tggtgctgactcttcagttcccactgttaacagatggagtgttggggttgggatcatcctgcct
gcctgacactcaggcagtatcccagtctccgttgcagtgtatttagcagtggtaaacgattgaa
agcattcctacactgtttggttcagattggggcgaatctgtgtcttcctgggaacgtgatgtca
tggccagttgcttctgcagagcttagaggaagggttgagctggtgcagtggcttctgtcataac
ttcgtataatgtatgctatacgaagttatgggctttgataatctgttaagatttgtttctgact
tctctcgcagctccctccatacagcgggactctttgtagcatccaggctgaagacgaactgcca
gatggtaagccgggttgcatggactgggtgggcaggaaacctgcaaacatgtccataacttcgt
ataatgtatgctatacgaagttataacattgggtcatcttttctcaacacaaatttttatttat
tcatttatttggagtgtgcttttgtgtggaggtcagaggtcagctctctggaatatgtcctct
tctgcttttgcctgggtttggtgattggacagaggtcatcgggcttatgcagcaagtctcttgc
tgagctgtctccctgatctgtccctgttgactcccatcctttgtaggttagcaggctcaggctt
gtgcccagagctgcagtgtccgttcagccggaggcgcatttcattagagcttgaagcttctgct
ctgacctgcagctcacttaactcctgtcttacagtttccattgctgtggagagacaccatgacc
agagcagccagtatgaaggcaaacatttatttggggctggtttacaggttcagaggttcagtcc
gttatcatggagggaaacatggcagcatccgggcatgcatggcactgggagctgagaattctct
atgttgttctcaaggcaaacaggagaagtcttgcaagcagtgaggagggtctcatagcccaccc
gcacagtgccagttcctctaacaaggccacacctacccacagtgcctcggcccatgggccaagc
atactcaaaccaccacaactgccaagtcactcagattggcaggtgtgctttcaaggttgtgcca
cgctcactctgggcatggtaagttgtggtcaggagagaggtgctctagtctcctcagcaggccc
tggtgttgtgggtgtgtcatcccacatggctgtgctcaactctgccgggtccagggagttgtta
tagcagttggtgttcatgacccatgcctgggcatttaggcctccagcatgagccctggccgatt
tatgcataagtgtcatgtttatcaaggtgcagcagtgtttgttactgaggtgtgtttaagatat
cctggctatgcagt

CO_Nnmtex2 donor

gaaagtttaattctactcttatttgaaaggaatgtttgtttactcatgaagtaaaggaccccccc
aatattaagaactgggtgatttgattattaattaatctcagtgttttaaattttaatatatata
catctaagcatatgtgtacatggtctatgcaggtacatgcatgggggccagaagagggcaccaga
tgtatgcttccatcactctctccacctattcttgtgggcagcgtcccctcctaaacctgggggc
ttgtgtgttgactagactggaagccagaaagcctcagagagcctctcctctctgctcctctcag
agatgggagcataggtatttgtgagatgcctgacttgttacacaggttctgggacctgaatgct
ggccctcatctctggaccatcactccaggccttaaaatcagttgttacacaacaacaacgacaa
caccaccaccaccacaacaaaaatggcacaatcccagtaaagacaataacttcgtataatgtat
gctatacgaagttatctataacttcagctctgcctgttccctttgatgcctgattccctgccc
ctcttctgtcctaggtgctgtaaaaggagaactcctcattgacatcggctctggcccaaccatc
taccagcttctgtctgcctgtgagtctttcacggagatcatcgtctctgactatacagaccaaa

图 14C

acctctgggaactgcagaaatggctcaagaaggagccaggagcctttgattggtccccagtcgt
cacctatgtgtgcgatcttgaaggcaacaggtagaggaatgagtatctgctcttcaactttctg
aagggacctgcataacttcgtataatgtatgctatacgaagttatatgtttctatcagcccaaa
ctaccactaagatccagatgaaaaatgcaaaagaaactcaaagaccaaagggaactcgggatag
aagatgagggctcaggagaatagctgggttgtcagatgtaccaccagagtactatttggttacc
agtcatccaacacaaaggtggtggctaatgcctaatgcccacaccaccggaaacagcaaaacac
tggcagatgtatttttttagccaacagaaataacctgaagcaattatggggaaagaattcatta
cttaccatctataaagttagaatctttaggatctgggagacggtcagtgggtaaagtgcttgtc
acagaggttgtcctttctccagtgtgtcctcttagcatctttattgaaaaccagctggctgcag
attcatggcttagtcctatactgtctattctgttccattattctgtgttctctggctggtgatg
ttttgtggaaagactttgctggtttgatgatatagtttttttgttgttgttgctgctgctg
ctgctgttgttgaagctaagtagtgtgtgatttctgatattttcttggtgctatttgggctttg
tgtgtacggttgcctaaaaattctagaattgttttctagtcctgtcattagtatttaagagga
atggcattggctacactaattgctttgggtagtatggacatattcatgatcttgattcttcaga
tcagcgaacagatctttccattttgatggtgtctgcactattcttctccattgatgttttcagt
gtaaagatcattggcttgttttagcaaaattcattttgaatttatatattatatttattaataa
atacattaataaaataatattaaatattacattttatttatttttaatagctattgcaatagga
ttgacttcttgatttctgaatcaaataattctctattggg

CO_Notch3ex3 donor
gacagaggcaggtaaattactgtgaattccagtctagccagcgatatgtagtgagaccctgcct
aaaatatattaaaaaaaaaaaaaaaaaggggaaggaagagtaaatggatttgtctgatagtctgt
ctggcacgagtgttgtttgataaacgcatcttgtgataacttcgtataatgtatgctatacgaa
gttatttatctgtctggcattgccatgcttttataccgtcccgaccacacatcttcccacaggt
gcctgccaggctgggttggtgagcgctgccagctcgaagacccctgccactccggcccttgtgc
cggccgaggcgtctgccagagctcagtggtggctggcaccgcacgattctcctgccggtgcctt
cgaggcttccaaggtgaaggggtgtgtctggacgggaaccctataacttcgtataatgtatgct
atacgaagttattggtaggcgagaatgtagtcagacccaagctcaccctctcctggttcttcca
ggcccagactgctcccagccagacccctgcgtcagcaggccctgtgttcatggtgcccctgct
cagtggggccggatggccgatttgcctgtgcctgcccacctggctaccagggtcaaagctgcca
aagtgacatagatgagtgccgatctggtacaacttgccgtcatggtggtacctgtctcaataca
cctggatccttccgctgccagtgtcctcttggttatacagggctgctgtgtgagaaccccgtag
tgccctgtgccccttccccgtgtcgtaatggtggcacctgtaggcagagcagtgatgtcacata
tgactgtgcttgccttcctggtaagtaagttgtgcccagggaaggcagctgggggacaataggct
agcctcttagtgaccattgtcaccttgtcctcccctacgaggcttcgagggccagaactgtgaa
gtcaacgtggatgactgtcctggacatcggtgtctcaatgggggaacgtgtgtagacggtgtca
atacttacaactgccagtgccctccggagtggacaggtgggcatcagggctgcagagaaccagg
gtggctgacctcaggtgggcacacgggcaacttagactagcacatctttgtgccctaggccagt
tctgtacagaagatgtggatgagtgtcagctgcagcccaatgcctgccacaatgggggtacctg
cttcaacctactgggtggccacagctgtgtatgtgtcaatggctggacgggtgagagctgcagt
cagaatatcgatgactgtgctacagccgtgtgtttccatggggccacctgccatgaccgtgtgg
cctctttctactgtgcctgccctatggggaagacaggtgagtggccctttctttgtaggcaac
agaatggtttcagcatgaaaggt

图 14D

Abstract

The disclosure provides methods and compositions for generating conditional knock-out alleles using donor constructs together with sequence-specific nucleases to generate conditional knock-out alleles. Specifically, the donor construct comprises a 5' homology region, a 5' recombinase recognition site, a donor sequence, a 3' recombinase recognition site, and a 3' homology region. Further disclosed are the donor sequences each comprises a target sequence having at least one neutral mutation. Different sequence-specific nucleases can be used with the donor constructs are further disclosed.