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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 May 2017 (11.05.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/079428 A1

(51) International Patent Classification:
C12N 15/90 (2006.01) *C12N 9/22* (2006.01)
C12N 5/075 (2010.01) *C12N 15/113* (2010.01)
C12N 5/076 (2010.01)

(21) International Application Number:
PCT/US2016/060315

(22) International Filing Date:
3 November 2016 (03.11.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/250,893 4 November 2015 (04.11.2015) US

(71) Applicant: **PRESIDENT AND FELLOWS OF HARVARD COLLEGE** [US/US]; 17 Quincy Street, Cambridge, MA 02138 (US).

(72) Inventors; and
(71) Applicants : **SINCLAIR, David A.** [AU/US]; 43 Newbrook Circle, Chestnut Hill, MA 02467 (US). **SHULTZ, Michael** [US/US]; 150 Brookline Avenue, Apt. 404, Boston, MA 02215 (US). **BONKOWSKI, Michael S.** [US/US]; 53 Tremont Street, Cambridge, MA 02139 (US).

(74) Agents: **SMITH, DeAnn F.** et al.; Patent Group, Foley Hoag LLP, 155 Seaport Boulevard, Boston, MA 02210-2600 (US).

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

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

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

WO 2017/079428 A1

(54) Title: SITE SPECIFIC GERMLINE MODIFICATION

(57) Abstract: Provided herein are methods for the safe and reliable generation of inheritable genetic modifications. In some embodiments, provided herein are methods of modifying a target nucleic acid sequence in the genome of a germ stem cell using a germline targeted nuclease, such as a zinc-finger nuclease (ZFN), a transcription activator-like effector nuclease (TALEN), an RNA-guided endonuclease (RGEN) or an engineered homing meganuclease (EHM).

SITE SPECIFIC GERMLINE MODIFICATION

RELATED APPLICATIONS

This application claims the benefit of priority to Provisional Application No. 62/250,893, filed November 4, 2015, which is hereby incorporated by reference in its entirety.

BACKGROUND

The reliable generation of inheritable genetic modifications has been a tantalizing goal for many years. Such technology would, for example, allow the permanent correction of genetic defects in the germline, thereby eliminating the need to treat many genetic disorders and alleviating much pain and suffering. However, for a germline genetic modification technology to become widely accepted, it is necessary to ensure that the technology only modifies the targeted site with the intended genetic change. Unfortunately, technology for the generation of safe germline genetic modifications is not currently available.

SUMMARY

Provided herein are methods for the safe and reliable generation of inheritable genetic modifications. In certain embodiments, the genetic modifications are made to the germline of an animal. In some embodiments, the animal is a mammal. In some embodiments, the animal is a human. In some embodiments, the animal is a non-human mammal.

In certain aspects, provided herein are methods of modifying a target nucleic acid sequence in the genome of a germ stem cell (GSC, *e.g.,* a mammalian GSC, such as a human GSC). In some embodiments, the target nucleic acid sequence in the germline of the GSC is replaced with a modified target nucleic acid sequence. In some embodiments, the method includes the step of contacting the GSC with a nucleic acid molecule encoding a germline targeted nuclease. In some embodiments, the GSC is an oogonial stem cell (OSC) or a spermatogonial stem cell (SSC).

In some aspects, provided herein is a method of generating an animal carrying a modified target nucleic acid sequence in its genome. In some embodiments, the method includes the step of contacting a sample comprising a population of GSCs with a germline targeted nuclease, wherein the germline targeted nuclease is targeted to the target nucleic acid sequence. In some embodiments, the method includes the step of isolating single GSCs from the population of GSCs. In some embodiments, the method includes the step of

culturing the single GSCs to form clonal GSC colonies. In some embodiments, the method includes the step of identifying the clonal GSC colonies that contain a modified target nucleic acid sequence in its genome. In some embodiments, the method includes the step of generating an animal from a cell from an identified clonal GSC colony. In some embodiments, the GSC is an oogonial stem cell (OSC) or a spermatogonial stem cell (SSC).

In some embodiments of the methods provided herein, the target nucleic acid sequence is a gene encoding an allele associated with a disease or disorder (*e.g.*, a Brca1 or Brca2 allele associated with breast cancer, a Htt allele associated with Huntington's disease, a Cftr allele associated with cystic fibrosis, a Hbb allele associated with thalassemia or sickle cell disease). In some embodiments, the modified target nucleic acid sequence is an allele of the gene that is not associated with the disease or disorder. In some embodiments, the target nucleic acid sequence is a gene encoding an allele associated with a particular trait in an animal (*e.g.*, disease resistance, toxin resistance, cancer resistance, radiation resistance, cell stress resistance, growth rate, adult size, lifespan, milk production, meat production, leather production, endurance, wool production, egg production, domestication, allergies, food sensitivities, hair properties (*e.g.*, growth, length, thickness, color, waviness), fertility, speed, endurance). In some embodiments, the modified target nucleic acid sequence is an allele of the gene that is associated with a different trait in the animal. In some embodiments, the modified target nucleic acid sequence is a DNA sequence from another species with a particular trait. In some embodiments, the target nucleic acid sequence is located in chromosomal DNA. In some embodiments, the target nucleic acid sequence is located on mitochondrial DNA.

In some embodiments, the methods provided herein include the step of contacting the GSC with a nucleic acid molecule encoding a germline targeted nuclease, wherein the germline targeted nuclease is targeted to the target nucleic acid. In some embodiments, the GSC is transfected with a nucleic acid molecule encoding the germline targeted nuclease. In some embodiments, the GSC is contacted with a vector (*e.g.*, an expression vector) comprising a nucleic acid molecule encoding the germline targeted nuclease. In some embodiments the nucleic acid sequence encoding the germline targeted nuclease is operably linked to a regulatable or non-regulatable transcription control element (*e.g.*, an inducible promoter) such that the germline targeted nuclease is expressed in the GSC. In some embodiments, the GSC is contacted with a nucleic acid molecule encoding a marker (*e.g.*, a fluorescent protein or an antibiotic resistance protein).

In some embodiments of the methods provided herein, any germline targeted nuclease can be employed. In some embodiments, the germline targeted nuclease is a zinc-finger nuclease (ZFN), a transcription activator-like effector nuclease (TALEN), an RNA-guided endonuclease (RGEN) or an engineered homing meganuclease (EHM). In some embodiments, the germline targeted nuclease is a component of a CRISPR/Cas system, such as Cas9. In some embodiments the germline targeted nuclease is a fusion protein comprising the germline targeted nuclease and a detectable moiety (*e.g.,* a fluorescent moiety such as GFP).

In some embodiments of the methods provided herein, the GSC is transfected with a homology repair template. In some embodiments, the homology repair template comprises the modified target nucleic acid sequence flanked by homology arms. In some embodiments, each of the homology arms comprise a nucleic acid sequence homologous to the nucleic acid sequence flanking the target nucleic acid sequence in the genome of the GSC. In certain embodiments, the germline targeted nuclease induces a double-stranded DNA break or a single-stranded DNA nick in the target nucleic acid sequence, which induces homology-directed repair of the break or nick and the replacement of the target nucleic acid sequence in the genome of the GSC with the modified target nucleic acid sequence of the homology repair template.

In some embodiments of the methods provided herein, the target nucleic acid sequence is modified using a CRISPR/Cas system. In some embodiments, the Cas protein in the CRISPR/Cas system is Cas9. In some embodiments, the GSC is transfected with a nucleic acid encoding Cas9 and a nucleic acid encoding a guide-RNA that is specific to the target nucleic acid sequence. In some embodiments, Cas9 and the guide-RNA are encoded on the same nucleic acid. In some embodiments, the guide-RNA is directly transfected into the GSC. In some embodiments, the guide-RNA comprises a target-specific guide sequence (*e.g.,* a sequence that is complementary to a sequence of the target RNA sequence) and a guide-RNA scaffold sequence. In some embodiments, the target-specific guide sequence is 5' of the guide-RNA scaffold sequence. In some embodiments, the nucleic acid encoding Cas9 is codon optimized. In some embodiments, the Cas9 carries a mutation that causes it to be a DNA nickase (*e.g.,* a D10A mutation). In some embodiments, the GSC is also transfected with a homology repair template. In some embodiments, the homology repair template comprises the modified target nucleic acid sequence flanked by homology arms. In some embodiments, the homology arms comprise nucleic acid sequences homologous to the

nucleic acid sequences flanking the target nucleic acid sequence in the genome of the GSC. In certain embodiments, Cas9 is targeted to the target nucleic acid sequence by the guide-RNA, where it induces a double-stranded break or a single-stranded nick. In some embodiments, the double-stranded break or single-stranded nick induces homology-directed repair of the break or nick and the replacement of the target nucleic acid sequence in the genome with the modified target nucleic acid sequence. In some embodiments the modified target nucleic acid sequence is on a homology repair template.

In some embodiments, the methods provided herein include the step of isolating the GSC prior to contacting it with the germline targeted nuclease. In some embodiments the method includes the isolation of OSCs from ovarian tissue. In some embodiments, the method includes the step of isolating SSCs from testes tissue. In some embodiments, the GSC is isolated by fluorescent activated cell sorting (FACS). In some embodiments, the GSC is isolated based on Ddx4 expression.

In some embodiments, the methods provided herein include the step of clonally culturing the transfected GSC to generate a clonal GSC colony. In some embodiments, the GSC is isolated from other GSCs prior to being cultured (*e.g.,* by FACS). In some embodiments, GSCs that express the germline targeted nuclease are isolated from GSCs that do not express the germline targeted nuclease. In some embodiments, GSCs that express a low level of the germline targeted nuclease (*e.g.,* a low level relative to other transfected GSCs) are isolated from other GSCs prior to clonal culturing. In some embodiments, the GSCs that express a low level of the germline targeted nuclease are the about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15% 16%, 17%, 18%, 19%, 20%, 25%, 30%, 35%, 40%, 45% or 50% of GSCs that are positive for germline-targeted nuclease with the lowest level of germline targeted nuclease expression. In some embodiments, the germline-targeted nuclease is a fusion protein that comprises a detectable marker, such as a fluorescent marker. In some embodiments the fluorescent marker is GFP. In some embodiments, the level of germline targeted nuclease expression is determined by determining the amount of detectable marker in the GSC (*e.g.,* by FACS).

In some embodiments, the methods provided herein include the step of determining whether the target nucleic acid sequence has been modified in the GSC. In some embodiments, the method includes the step of determining whether the target nucleic acid sequence has been modified in cells of the clonal GSC colony. In some embodiments, whether the target nucleic acid sequence has been modified is determined by sequencing the

locus in which the target nucleic acid sequence was located in the genome of the GSC. In some embodiments, whether the target nucleic acid sequence has been modified is determined by contacting the locus in which the target nucleic acid sequence was located in the genome of the GSC with a probe that selectively hybridizes to the target nucleic acid sequence or the modified target nucleic acid sequence. In some embodiments, the locus in which the target nucleic acid sequence was located in the genome of the GSC is amplified using a nucleic acid amplification procedure and the amplification product is sequenced. In some embodiments, the locus in which the target nucleic acid sequence was located in the genome of the GSC is amplified using a nucleic acid amplification procedure and the amplification product is contacted with a probe that selectively hybridizes to the target nucleic acid sequence or the modified target nucleic acid sequence.

In some embodiments, the methods provided herein include the step of sequencing the genome of the cells from the clonal GSC colony. In some embodiments, the genome sequencing step is used to identify GSC colonies in which only the target nucleic acid sequence has been modified (*i.e.*, cells in which the germline targeted nuclease did not cause any off-target mutations).

In some embodiments of the methods provided herein, a cell from a clonal GSC colony (*e.g.*, a clonal GSC colony in which the target nucleic acid has been modified) is caused to differentiate into mature germ cells (*e.g.*, mature spermatocytes or mature oocytes). In some embodiments, the methods provided herein include the step of inducing the maturation of a cell from the clonal GSC colony into a mature germ cell. In some embodiments, the maturation of the GSC is induced by transplantation of the GSC into testes or ovarian tissue, directed differentiation *in vitro* or nuclear transfer into a mature germ cell. In some embodiments, the mature germ cell is used in an in vitro fertilization procedure. In some embodiments, the GSC is an OSC and it is transplanted into a surrogate mother. In some embodiments, the GSC is a SSC and it is transplanted into a surrogate father. In some embodiments the surrogate mother and/or the surrogate father is mated to generate offspring comprising the modified target nucleic acid.

BRIEF DESCRIPTION OF FIGURES

Figure 1 is a flowchart illustrating an exemplary method in accordance with one or more embodiments.

Figure 2 shows two FACS plots, each depicting the isolation of primary OSCs from murine ovaries.

Figure 3 shows fluorescent microscope images taken of OSCs that have been transfected using the specified amounts of a Cas9-GFP expression vector and lipofectamine 2000.

Figure 4 shows a series of FACS plots depicting the isolation of transfected OSCs expressing Cas9-GFP.

Figure 5 shows an example of a clonal colony generated from a FACS-purified single OSC transfected with components of the CRISPR/Cas9 system.

Figure 6 provides the nucleic acid sequence of human codon optimized Cas9.

Figure 7 shows an exemplary gating strategy for isolating OSCs using flow cytometry.

Figure 8 shows the expansion of sorted OSCs in culture. One-day post isolation, cells are fluorescent in the FITC channel due to Calcein uptake during sorting.

Figure 9 shows the PCR genotyping of the GFP insertion into the ROSA26 locus in the germline of OSCs. The top panel is a diagram of the edited ROSA26 locus, showing regions of homology to the donor plasmid and the primers used for genotyping. The bottom channel depicts the genotyping results for five OSC lines.

Figure 10 shows the genomic sequencing results of the genome-insert junction of an edited OSC lines.

Figure 11 shows fluorescent microscope images of two of the edited OSC cell lines depicting GFP expression by the cells.

Figure 12 shows expression of germline genes by edited OSC cell lines.

Figure 13 has four panels and shows the targeting of the AAVS1 “safe-harbor” locus in human OSCs. Panel (a) is a diagram of the human AAVS1 locus after knock-in of a GFP reporter. Panel (b) shows a representative genotyping PCR for unmodified and modified AAVS1 alleles. Panels (c) and (d) show human OSCs transfected with CRISPR plasmids of AAVS1-GFP knock-in.

Figure 14 is a flowchart illustrating an exemplary pipeline for editing OSCs.

Figure 15 has two panels. Panel (a) shows a mixed population of human OSCs in which the GFP gene has been directed to the AAV1 safe harbor locus using CRISPR/Cas9.

Panel (b) shows GFP-positive single sorted EggPCs that were selected for clonal expansion under optimized conditions.

DETAILED DESCRIPTION

General

Provided herein are methods for the reliable generation of inheritable genetic modifications. In some embodiments, provided herein are methods of modifying a target nucleic acid sequence in the genome of a germ stem cell (GSC), such as an oogonial stem cell (OSC) or a spermatogonial stem cell (SSC). In some embodiments the target nucleic acid sequence is replaced with a modified target nucleic acid sequence. In some embodiments, a germline targeted nuclease, such as a zinc-finger nuclease (ZFN), a transcription activator-like effector nuclease (TALEN), an RNA-guided endonuclease (RGEN) or an engineered homing meganuclease (EHM) is used to induce the modification. In some embodiments, the RGEN is Cas9 and the target nucleic acid sequence is modified using the CRISPR/Cas9 system.

Disclosed herein is a way to make permanent genetic alterations or inheritable genetic modifications (IGMs) in the mammalian (*e.g.,* human) germline using germline stem cells (GSCs) from males or females. Male GSCs are known as spermatogonial stem cells (SSCs). The female GSCs are known as oogonial stem cells (OSCs) or egg precursor cells (EggPCs).

In some embodiments, the methods provided herein can be used to fix mutations associated with genetic diseases, to knock-out genes, to insert transgenes, and/or to make other changes or enhancements to the germline genome, which are inherited by the next generation. In some embodiments, the mitochondrial genome of the GSC is edited prior to mitochondrial transfer into a GSC, oocyte, or zygote. In some embodiments, the nuclear genome of the GSC is edited prior to transferring the nucleus into a GSC, oocyte, or zygote. In certain embodiments, the methods provided herein can be applied to mammals for research, ecological, agricultural (*e.g.* farm animals or pets), and therapeutic purposes, including the application to humans for purposes relating to health.

In some embodiments, the methods provided herein use genome editing in GSCs as an alternative to genome editing in pluripotent cells such as embryonic (ES) cells, induced pluripotent (iPS) cells, and zygotes. Advantages over these systems include the fact that GSCs may be more easily obtained, may have less cellular defects, and are pre-meiotic cells.

A flowchart illustrating an exemplary method in accordance with one or more embodiments is provided in Figure 1. In certain embodiments, GSCs are isolated from ovarian or testes tissue using methods such as FACS (*e.g.*, as described in Woods and Tilly, *Nat Protoc.* 8:966-88 (2013), incorporated by reference in its entirety). The GSCs are transfected and/or infected with targeted nucleases to make genetic alterations (*e.g.*, using components of the CRISPR/Cas9 system). Single transfected cells that have a high chance of being error free (*e.g.* that express low levels of CAS9) are FACS purified into individual wells and grown into clonal colonies. The colonies are genotyped to identify those with desired alterations. The genomes of the colonies containing the desired alteration are then sequenced to confirm that the genetic change is correct and to exclude colonies that have additional genetic changes. Once cells with desired genetic alterations (and only the desired genetic alterations) are identified, the cells can be induced to form competent, mature spermatocytes or oocytes by any one of several methods, including transplantation to testes or ovarian tissue, directed differentiation *in vitro*, or nuclear transfer from an *in vitro*-derived immature spermatocyte or oocyte to a mature cell. Genetically-modified mature germ cells can be used to produce offspring by conventional mating or *in vitro* fertilization (IVF) techniques.

Definitions

In order that the present invention may be more readily understood, certain terms are first defined. Additional definitions are set forth throughout the detailed description.

The articles "*a*" and "*an*" are used herein to refer to one or to more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

As used herein, an "*allele*" refers to one of two or more alternative forms of a nucleotide sequence at a given position (locus) on a chromosome. An individual can be heterozygous or homozygous for any allele described herein.

As used herein, the term "*animal*" includes both human and non-human animals.

An "*expression vector*" is a vector which is capable of promoting expression of a nucleic acid incorporated therein. Typically, the nucleic acid to be expressed is "*operably linked*" to a transcriptional control element, such as a promoter and/or an enhancer, and is therefore subject to transcription regulatory control by the transcriptional control element.

The term "*gene*" is used broadly to refer to any nucleic acid associated with a biological function. The term "gene" applies to a specific genomic sequence, as well as to a cDNA or an mRNA encoded by that genomic sequence.

The term "*germline targeted nuclease*" refers to programmable nucleases capable of targeted genome modification. Examples of germline targeted nucleases include, but are not limited to, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), RNA-guided endonucleases (RGENs), such as the CRISPR/Cas9 system, and engineered homing meganucleases (EHMs).

The terms "*polynucleotide*", and "*nucleic acid*" are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three-dimensional structure, and may perform any function. The following are non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. A polynucleotide may be further modified, such as by conjugation with a labeling component. In all nucleic acid sequences provided herein, U nucleotides are interchangeable with T nucleotides.

The term "*vector*" refers to the means by which a nucleic acid can be propagated and/or transferred between animals, cells, or cellular components. Vectors include plasmids, viruses, bacteriophage, pro-viruses, phagemids, transposons, and artificial chromosomes, and the like, that may or may not be able to replicate autonomously or integrate into a chromosome of a host cell.

Isolation of Germline Stem Cells

In some embodiments, the methods provided herein include the step of isolating a population of GSCs from an animal. In some embodiments, the GSCs can be isolated from any animal. In some embodiments, the animal is a non-human animal. In some embodiments, the animal is a mammal. In some embodiments, the animal is a non-human mammal. In some embodiments the mammal is a domesticated mammal (*e.g.,* a cow, a pig, a horse, a

donkey, a goat, a camel, a cat, a dog, a guinea pig, a rat, a mouse, a sheep, a zebu, a water buffalo, a yak, a llama, an alpaca, a ferret, a rabbit, a caribou, a reindeer).

In some embodiments, the GSCs are OSCs. In some embodiments, the OSCs are isolated from the ovarian tissue of the animal. Any method for the isolation of OSCs can be used. Methods for the isolation of OSCs from ovarian tissue are known in the art and described, for example, in Woods and Tilly, *Nat Protoc.* 8:966-88 (2013) and White *et al., Nat. Med.* 18:413-421 (2012), each of which is incorporated by reference in its entirety.

In some embodiments, the GSCs are SSCs. In some embodiments, SSCs are isolated from testes tissue. Any method for the isolation of SSCs can be used. Methods for the isolation of SSCs from testes tissue are known in the art and described, for example, in Guan *et al., Nature* 440:1199-1203 (2006) and Kossack *et al., Stem Cells* 27:138-149 (2009), each of which is incorporated by reference in its entirety.

In some embodiments, the GSC is isolated by fluorescent activated cell sorting (FACS). In some embodiments, the GSC are stained with an anti-Ddx4 antibody and then isolated based on Ddx4 expression.

Transfection of Germline Stem Cells

In some embodiments, the methods provided herein require the introduction of an extrinsic nucleic acid molecule into a GSC. For example, in some embodiments, the methods provided herein require the introduction of a nucleic acid molecule encoding a germline targeted nuclease (*e.g.,* a germline targeted nuclease targeted to a germline target nucleic acid sequence) into the GSC. In some embodiments, the methods require the introduction of a homology repair template into the GSC. In some embodiments, the methods require the introduction of a guide-RNA and/or a nucleic acid molecule encoding a guide-RNA into the GSC.

In certain embodiments, any method can be used to facilitate the introduction of an extrinsic nucleic acid molecule into the GSC. In some embodiments, the GSC is transfected with the nucleic acid molecule. As used herein, "transfection" encompasses any process of introducing a nucleic acid molecule into a cell, including, but not limited to, chemical-based methods (*e.g.,* using calcium phosphate, dendrimers, liposomes and/or cationic polymers), electroporation, sono-poration, optical transfection, protoplast fusion, impalefection, viral methods (*e.g.,* using adenovirus vectors, adeno-associated virus vectors, retrovirus vectors and/or lentivirus vectors), and particle-based methods (*e.g.,* using a gene gun,

magnetofection or particle bombardment). In some embodiments, the transfection is transient transfection.

In some embodiments, the GSC is transfected by contacting the GSC with a vector (*e.g.*, an expression vector) comprising a nucleic acid molecule encoding a germline targeted nuclease and/or a guide-RNA. In some embodiments the nucleic acid sequence encoding the germline targeted nuclease or guide-RNA is operably linked to a transcription control element (*e.g.*, a promoter and/or an enhancer) such that the germline targeted nuclease is expressed in the GSC. In some embodiments, the transcription control element is an inducible promoter. In some embodiments, the transcription control element is a constitutive promoter. In some embodiments, the transfected nucleic acid molecule also encodes a detectable moiety, such as a fluorescent moiety. For example, in some embodiments the germline targeted nuclease is a fusion protein that includes a fluorescent moiety. In general, any fluorescent protein can be used as the detectable moiety. Examples of fluorescent proteins useful in the methods described herein include, but are not limited to, EGFP, mPlum, mCherry, mOrange, mKO, EYFP, mCitrine, Venus, YPet, Emerald, Cerulean and CyPet.

Nucleic acid molecules can be delivered to the GSC in any desired vector. These include viral or non-viral vectors, including adenovirus vectors, adeno-associated virus vectors, retrovirus vectors, lentivirus vectors, and plasmid vectors. Exemplary types of viruses include HSV (herpes simplex virus), AAV (adeno associated virus), HIV (human immunodeficiency virus), BIV (bovine immunodeficiency virus), and MLV (murine leukemia virus). Nucleic acid molecules can be administered in any desired format that provides sufficiently efficient delivery levels, including in virus particles, in liposomes, in nanoparticles, and complexed to polymers.

In some embodiments, the nucleic acid molecule is associated with a liposome to form a gene delivery vehicle. Liposomes are small, lipid vesicles comprised of an aqueous compartment enclosed by a lipid bilayer, typically spherical or slightly elongated structures several hundred Angstroms in diameter. Under appropriate conditions, a liposome can fuse with the plasma membrane of a cell or with the membrane of an endocytic vesicle within a cell which has internalized the liposome, thereby releasing its contents into the cytoplasm. Prior to interaction with the surface of a cell, however, the liposome membrane acts as a relatively impermeable barrier that sequesters and protects its contents, for example, from degradative enzymes. Additionally, because a liposome is a synthetic structure, specially

designed liposomes can be produced which incorporate desirable features. See Stryker, Biochemistry, pp. 236-240, 1975 (W.H. Freeman, San Francisco, CA); Soak *et al.*, Biochip. Biopsy's. Acta 600:1, 1980; Bayer *et al.*, Biochip. Biopsy's. Acta. 550:464, 1979; Rivnay *et al.*, Meth. Enzymol. 149:119, 1987; Wang *et al.*, PROC. NATL. ACAD. SCI. U.S.A. 84: 7851, 1987, Plant *et al.*, Anal. Biochem. 176:420, 1989, and U.S. Patent 4,762,915. Liposomes can encapsulate a variety of nucleic acid molecules including DNA, RNA, plasmids, and expression constructs comprising growth factor polynucleotides such those disclosed in the present invention.

Liposomal preparations useful in the methods described herein include cationic (positively charged), anionic (negatively charged) and neutral preparations. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner *et al.*, Proc. Natl. Acad. Sci. USA 84:7413-7416, 1987), mRNA (Malone *et al.*, Proc. Natl. Acad. Sci. USA 86:6077-6081, 1989), and purified transcription factors (Debs *et al.*, J. Biol. Chem. 265:10189-10192, 1990), in functional form. Cationic liposomes are readily available. For example, N[1-2,3-dioleyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. See also Felgner *et al.*, Proc. Natl. Acad. Sci. USA 91: 5148-5152.87, 1994. Other commercially available liposomes include Transfectace (DDAB/DOPE) and DOTAP/DOPE (Boerhinger). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, *e.g.*, Soak *et al.*, Proc. Natl. Acad. Sci. USA 75:4194-4198, 1978; and WO 90/11092 for descriptions of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphoshatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

One or more nucleic acid sequence (*e.g.*, encoding a germline targeted nuclease and a guide-RNA) of interest may be introduced into a GSC as part of a single nucleic acid molecule. Alternatively, separate nucleic acid molecules can be used to introduce multiple nucleic acid sequences into the GSC (*e.g.,* a germline targeted nuclease, a guide-RNA and/or

a homology repair template co-transfected into the GSC on separate vectors). Transfection of the multiple vectors can be performed simultaneously or sequentially.

Germline Targeted Nucleases

In some embodiments of the methods provided herein, a target nucleic acid in the genome of a GSC is modified using a germline targeted nuclease. In some embodiments, any germline targeted nuclease can be used. In some embodiments, the germline targeted nuclease is a zinc-finger nuclease (ZFN), a transcription activator-like effector nuclease (TALEN), an RNA-guided endonuclease (RGEN) or an engineered homing meganuclease (EHM).

A germline targeted nuclease is a programmable nuclease capable of targeted genome modification. In general, germline targeted nucleases are able to induce double-stranded DNA breaks or single-stranded DNA nicks within the a specific nucleic acid sequence in the genome of a cell. The presence of the break or nick stimulates the cell's DNA repair machinery. In the case of a break, the DNA damage can be repaired either by non-homologous end joining (NHEJ) or homology directed repair (HDR). In the case of a nick, HDR is induced.

NHEJ is a pathway that repairs double-stranded DNA breaks without the need of a homologous template. The NHEJ process is error prone, and will frequently result in an insertion or deletion at the position of the break. Targeting two adjacent sites in a genome with a germline targeted nuclease can be used to delete the intervening sequence. Induction of NHEJ at a locus using a germline targeted nuclease can therefore be used to inactivate the gene present at the locus. This process can be used, for example, to inactivate or delete alleles associated with autosomal dominant phenotypes.

HDR is a DNA repair mechanism in which a cell uses a homologous DNA sequence to repair a DNA break or nick. Induction of HDR at a target nucleic acid sequence requires the presence of a DNA molecule having a homologous sequence to the target nucleic acid sequence (*i.e.*, a homology repair template). If the template used in the HDR process contains minor nucleic acid sequence differences compared to the target nucleic acid sequence, those minor differences can be incorporated into the repaired target sequence, thereby creating a modified target nucleic acid sequence.

In some embodiments of the methods provided herein, the GSC is transfected with a homology repair template along with the germline targeted nuclease. In some embodiments, the homology repair template comprises a modified target nucleic acid sequence flanked by

homology arms. In some embodiments, each of the homology arms comprise a nucleic acid sequence that is homologous to the nucleic acid sequences flanking the target nucleic acid sequence in the genome of the GSC. As a result, in some embodiments when the germline targeted nuclease induces a double-stranded DNA break or a single-stranded DNA nick in the target nucleic acid sequence, the break or nick is repaired via HDR using the homology repair template, causing the replacement of the target nucleic acid sequence in the genome of the GSC with the modified target nucleic acid sequence. Thus, induction of HDR can be used to replace an allele of a gene associated with an undesired trait with an allele of a gene associated with a desired trait.

In some embodiments, the methods include a step wherein the expression level of the germline targeted nuclease is determined in a GSC. For example, in some embodiments, GSCs that express the germline targeted nuclease are isolated from GSCs that do not express the germline targeted nuclease prior to clonal expansion. Moreover, in some embodiments it is desirable to identify GSCs that express a certain level of the germline targeted nuclease. For example, in some embodiments GSCs expressing a low level of germline targeted nuclease are selected to reduce the likelihood of off-target genome modification. In some embodiments, GSCs that express a low level of the germline targeted nuclease (*e.g.*, a low level relative to other transfected GSCs) are isolated from other GSCs prior to clonal culturing. In some embodiments, the GSCs that express a low level of the germline targeted nuclease are the about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15% 16%, 17%, 18%, 19%, 20%, 25%, 30%, 35%, 40%, 45% or 50% of GSCs that are positive for germline-targeted nuclease with the lowest level of germline targeted nuclease expression. In some embodiments, GSCs expressing a high level of the germline targeted nuclease are selected to increase the likelihood of modification of the target nucleic acid sequence. In some embodiments, the GSCs that express a high level of the germline targeted nuclease are the about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15% 16%, 17%, 18%, 19%, 20%, 25%, 30%, 35%, 40%, 45% or 50% of GSCs with the highest level of germline targeted nuclease expression.

To facilitate monitoring of germline targeted nuclease expression, in some embodiments the germline targeted nuclease is a fusion protein that comprises a detectable moiety, such as a fluorescent protein. In general, any fluorescent protein can be used as the detectable moiety. Examples of fluorescent proteins useful in the methods described herein include, but are not limited to, EGFP, mPlum, mCherry, mOrange, mKO, EYFP, mCitrine,

Venus, YPet, Emerald, Cerulean and CyPet. In some embodiments, the level of germline targeted nuclease expression is determined by determining the amount of detectable moiety in the GSC (*e.g.*, by FACS).

RNA-Guided Endonucleases

RGENs are endonucleases capable of forming double-stranded DNA breaks (or, in some cases, single stranded DNA nicks) at a nucleic acid sequence to which the RGEN is targeted by a "guide-RNA" molecules. RGENs are generally derived from a prokaryotic immune system known as the CRISPR-Cas system that is found in many bacteria and archaea. In prokaryotes, CRISPR (clustered regularly interspaced short palindromic repeats) and CRISPR-associated proteins (Cas) provide a defense against viruses and plasmids by taking up and storing short fragments (about 30 bp) of extraneous genetic sequence in arrays separated by direct repeats. The arrays are transcribed and processed into small CRISPR RNAs ("crRNAs" or "guide-RNAs") that direct the cleavage of complementary viral or bacterial DNA by the Cas-encoded endonucleases.

RGENs use the CRISPR-Cas system to facilitate targeted genome editing in eukaryotic cells, including mammalian cells, such as human cells. To facilitate genome editing, the cell to be modified is co-transfected with an expression vector encoding the Cas machinery along with a guide-RNA molecule or an expression vector encoding a guide-RNA molecule.

While many different CRISPR-Cas systems could be modified to facilitate targeted genome modification, the most commonly used CRISPR-Cas system in targeted genome modification is the CRISPR-Cas9 system. The CRISPR-Cas9 system requires only a single protein, Cas9, to catalyze blunt double-stranded DNA breaks at sites targeted by a guide-RNA molecule. The Cas9 guide-RNA includes a 20-nucleotide target-specific guide sequence that is complementary to the targeted nucleic acid sequence and that is positioned 5′ of a guide RNA scaffold sequence. In some embodiments, the guide RNA scaffold sequence is, from 5′ to 3′:

GUUUUAGAGCUAGAAAUAGCUUAAAAUAAGGCUAGUCCUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCUUUU

Multiple guide RNA sequences can be encoded in a single CRISPR array to facilitate the simultaneous editing of multiple sites within a cell's genome. For example, a pair of guide RNAs can target proximally located sequences to facilitate the deletion of the intervening sequence. In addition, Cas9 can be converted into a nicking enzyme by

engineering an aspartate-to-alanine substitution (D10A) in the Cas9 RuvCI domain. Use of the engineered Cas9 nickase results in HDR repair of the targeted sequence, without the risk of NHEJ and the resulting mutations. In some embodiments, Cas9 is encoded by a codon-optimized sequence. Plasmids encoding Cas9, including codon-optimized plasmids and plasmids encoding engineered Cas9 nickase are publicly available from Addgene (http://www.addgene.org/CRISPR/). An exemplary human codon optimized nucleic acid sequence encoding Cas9 is provided in Figure 6.

Additional information on the application of CRISPR-Cas systems to targeted genome engineering can be found in Jinek *et al., Science* 337:816-821 (2012); Cho *et al., Nature Biotechnology* 31:230-232 (2013); Cong *et al., Science* 339:819-823 (2013); Jinek *et al., eLife* 2:e00471 (2013); Mali *et al., Science* 339:823-826 (2013); Qi *et al., Cell* 152:1173-1183 (2013); Fu *et al., Nature Biotechnology* 31:822-826 (2013); Fu *et al., Nature Biotechnology* 31:822-826 (2013); Hsu *et al., Nature Biotechnology* 31:827-832 (2013); Mali *et al., Nature Biotechnology* 31:833-838 (2013); Pattanayak *et al., Nature Biotechnology* 31:839-843 (2013) and WO/2013/142578, each of which is hereby incorporated by reference in its entirety.

In some embodiments of the methods provided herein, the target nucleic acid sequence is modified using a CRISPR/Cas system. In some embodiments, the CRISPR/Cas system is a CRISPR/Cas9 system. In some embodiments, the GSC is transfected with a nucleic acid encoding Cas9 and a nucleic acid encoding a guide-RNA that is specific to the target nucleic acid sequence. In some embodiments, the GSC is transfected with a nucleic acid encoding Cas9 and a guide-RNA that is specific to the target nucleic acid sequence. In some embodiments, the nucleic acid encoding the guide-RNA encodes multiple guide RNAs in a tandem array. In some embodiments, Cas9 and the guide-RNA are encoded on the same nucleic acid.

In some embodiments, the guide-RNA comprises a target-specific guide sequence (*e.g.*, a sequence that is complementary to a sequence of the target RNA sequence) and a guide-RNA scaffold sequence. In some embodiments, the target-specific guide sequence is 20 nt in length. In some embodiments, at least 14, 15, 16, 17, 18, 19 or 20 nt of the guider RNA sequence are complementary to the target sequence. In some embodiments, the target-specific guide sequence is 5' of the guide-RNA scaffold sequence.

In some embodiments, the nucleic acid encoding Cas9 is codon optimized for the organism from which the GSC was obtained. In some embodiments, the nucleic acid

sequence encoding Cas9 is human codon optimized. In some embodiments, the Cas9 carries a mutation that causes it to be a DNA nickase (*e.g.*, a D10A mutation).

Transcription Activator-Like Effector Nucleases

TALENs are artificial restriction enzymes generated by fusing a Transcription activator-like effectors (TALE) DNA binding domain with a DNA cleavage domain.

TALEs are nucleic acid binding proteins secreted by numerous species of proteobacteria. TALEs contain a DNA binding domain composed of tandem repeats of highly conserved monomer polypeptides that are predominantly 33, 34 or 35 amino acids in length and that differ from each other mainly in amino acid positions 12 and 13 (the "repeat variable di-residues" or "RVD"). Polypeptide monomers have a nucleotide binding affinity that is determined by the identity of the amino acids in its RVD. For example, polypeptide monomers with an RVD of NI preferentially bind to adenine, monomers with an RVD of NG preferentially bind to thymine, monomers with an RVD of HD preferentially bind to cytosine, monomers with an RVD of NN preferentially bind to both adenine and guanine, and monomers with an RVD of HN or NH preferentially bind to guanine. Thus, the number and order of the polypeptide monomer repeats in the nucleic acid binding domain of a TALE determines its nucleic acid target specificity. The structure and function of TALEs is further described in, for example, Moscou *et al., Science* 326:1501 (2009); Boch *et al., Science* 326:1509-1512 (2009); and Zhang *et al., Nature Biotechnology* 29:149-153 (2011), each of which is incorporated by reference in its entirety.

In TALENs, the TAL DNA binding domain is fused to a non-specific DNA cleavage domain, generally the cleavage domain of the FokI endonuclease. In this way, the design of the TALE DNA binding domain determines the nucleic acid sequence at which FokI cleavage domain to generates double-stranded DNA breaks.

Additional information of the construction and use of TALENs is provided in Miller *et al., Nucleic Acids Research* 39:e82 (2011); Miller *et al., Nature Biotechnology* 29:143-148 (2011); Mussolino *et al., Nucleic Acids Research* 39:9283-9293 (2011); Moirbitzer *et al., Nucleic Acids Research* 39:5790-5799 (2011); U.S. Pat. Pubs. US/2013/0117869, US/2011/0145940, US/2011/0301073, US/2013/0217131; and U.S. Pat. Nos. 8,450,471, 8,440,432, 8,440,431, each of which is hereby incorporated by reference in its entirety.

Zinc-Finger Nucleases

ZFNs are similar to TALENS, except that a zinc finger DNA binding domain (instead of a TALE DNA binding domain) is fused to a DNA cleavage domain to target DNA cleavage to a particular sequence.

A zinc finger is a small protein structural motif that is characterized by the coordination of one or more zinc ions to stabilize the protein structure. There are a large number of unique zinc finger structures, many of which have their own DNA binding characteristics. The DNA-binding domain of a ZFN typically contains between three and six zinc finger repeats, enabling it to recognize a specific nucleic acid sequence of between 9 and 18 bp in length. This DNA binding domain is fused to a non-specific DNA cleavage domain, generally the cleavage domain of the FokI endonuclease. In this way, the design of the zinc finger DNA binding domain determines the nucleic acid sequence at which FokI cleavage domain to generates double-stranded DNA breaks.

Additional information regarding the construction and use of ZFNs is provided in Urnov *et al., Nature* 435:646-651 (2005); Bibikova *et al., Science* 300:764 (2003); Maeder *et al., Molecular Cell* 31:294-301 (2008); Porteus *et al., Nature Biotechnology* 23:967-973 (2005); U.S. Pat. Pub. Nos. US/2012/0329067, US/2012/0214241, US/2012/0219959, US/2012/0309091; and U.S. Pat. Nos. 8,524,221, 8,106.255, each of which is hereby incorporated by reference in its entirety.

Engineered Homing Meganucleases

EHMs are derived from meganuclease endodeoxyribonucleases that are characterized by a large recognition site that can recognize double-stranded DNA sequences of 12 to 40 bp in length.

Among the homing endonucleases, those of the LAGLIDADG family are the most widely used for genome engineering. EHMs are engineered to recognize a particular DNA sequence either by mutating the amino acid sequence of a naturally occurring homing meganuclease DNA binding domain or by fusing DNA binding domains from multiple different homing meganucleases. Precision Biosciences offers a EHM design program called Directed Nuclease Editor that is capable of creating EHMs that target and modify a user-defined location in the genome.

Additional information regarding the construction and use of EHMs is provided in Grizot *et al., Nucleic Acids Research* 38:2006-2018 (2009); Seligman *et al., Nucleic Acids Research* 30:3870-3879 (2002); Rosen *et al., Nucleic Acids Research* 34:4791-4800 (2006); Arnould *et al., Journal of Molecular Microbiology* 355:443-458 (2006); Smith *et al., Nucleic*

Acids Research 34:e149 (2006); Chevalier *et al.*, *Molecular Cell* 10:895-905 (2002); U.S. Pat. Pub. Nos. US/2013/0203840, US/2010/0325745, US/2011/0207199, US/2011/0041194; and U.S. Pat. Nos. 8,530,214, 7,897,372, 8,143,015, 8,163,514, 8,119,381, each of which is hereby incorporated by reference in its entirety.

Screening Transfected GSCs

In some embodiments, GSCs transfected according to a method described herein are screened to determine whether the target nucleic acid sequence has been modified in the genome of the GSC. In some embodiments, GSCs transfected according to a method described herein are screened to determine whether any non-target nucleic acid sequences have been modified in the genome of the GSC.

To facilitate the screening and/or use of genome-modified GSCs, in some embodiments, the methods provided herein include the step of clonally culturing transfected GSCs to generate clonal GSC colonies. In some embodiments, the GSCs are separated from each other prior to being cultured. For example, single GSCs can be placed into separate wells of a microtiter plate via FACS. The individual cells can be grown in culture using known GSC culture techniques until clonal colonies are formed.

In some embodiments, the methods provided herein include the step of determining whether the target nucleic acid sequence has been modified in the GSC. In some embodiments, the method includes the step of determining whether the target nucleic acid sequence has been modified in cells of the clonal GSC colony. In some embodiments, whether the target nucleic acid sequence has been modified is determined by sequencing the locus in which the target nucleic acid sequence was located in the genome of the GSC. Nucleic acid sequencing processes include, but are not limited to chain termination sequencing, sequencing by ligation, sequencing by synthesis, pyrosequencing, ion semiconductor sequencing, single-molecule real-time sequencing, 454 sequencing, and/or Dilute-'N'-Go sequencing. In some embodiments, whether the target nucleic acid sequence has been modified is determined by contacting the locus in which the target nucleic acid sequence was located in the genome of the GSC a probe that selectively hybridizes to the target nucleic acid sequence or the modified target nucleic acid sequence (*e.g.*, a molecular beacon). In some embodiments, the locus in which the target nucleic acid sequence was located in the genome of the GSC is amplified using a nucleic acid amplification procedure and the amplification product is sequenced. In some embodiments, the locus in which the target nucleic acid sequence was located in the genome of the GSC is amplified using a

nucleic acid amplification procedure and the amplification product is contacted with a probe that selectively hybridizes to the target nucleic acid sequence or the modified target nucleic acid sequence. Examples of nucleic acid amplification processes include, but are not limited to, polymerase chain reaction (PCR), LATE-PCR a non-symmetric PCR method of amplification, ligase chain reaction (LCR), strand displacement amplification (SDA), transcription mediated amplification (TMA), self-sustained sequence replication (3SR), Qβ replicase based amplification, nucleic acid sequence-based amplification (NASBA), repair chain reaction (RCR), boomerang DNA amplification (BDA) and/or rolling circle amplification (RCA). Probes that can be used in conjunction with nucleic acid amplification processes include, but are not limited to, molecular beacons, molecular torches, TaqMan probes, scorpion probes and HPA probes.

In some embodiments, the methods provided herein include the step of sequencing the genome of the cells from the clonal GSC colony. In some embodiments, the genome sequencing step is used to identify GSC colonies in which only the target nucleic acid sequence has been modified (*i.e.,* cells in which the germline targeted nuclease did not cause any off-target mutations). Nucleic acid sequencing processes include, but are not limited to chain termination sequencing, sequencing by ligation, sequencing by synthesis, pyrosequencing, ion semiconductor sequencing, single-molecule real-time sequencing, 454 sequencing, and/or Dilute-'N'-Go sequencing.

Creation of Organisms from GSCs

In some embodiments of the methods provided herein, a GSC containing a modified target nucleic acid sequence in its genome is used to generate an organism having the modified target nucleic acid sequence in its genome. In some embodiments, any method of generating an organism from a GSC can be used. In some embodiments, the GSC is caused to differentiate into mature germ cells (*e.g.,* mature spermatocytes or mature oocytes). In some embodiments, maturation of the GSC is induced by transplantation of the GSC into testes or ovarian tissue, directed differentiation *in vitro* or nuclear transfer into a mature germ cell. In some embodiments where the target nucleic acid sequence is a mitochondrial DNA sequence, mitochondria from the GSC are transferred to a mature germ cell.

For example, for non-human animals, the GSC can be transplanted into a surrogate parent, which is then mated to produce offspring. In some embodiments, the GSC is an OSC and it is transplanted into a surrogate mother. In some embodiments, the GSC is a SSC and it

is transplanted into a surrogate father. In some embodiments the surrogate mother and/or the surrogate father is mated to generate offspring comprising the modified target nucleic acid.

In all animals, including humans, the GSC can be matured *in vitro,* for example by culturing the GSC under conditions that induce maturation or by nuclear transfer into a mature germ cell, which can then be used to generate offspring. For example, mature spermatocytes can be used in an artificial insemination, or either mature spermatocytes or mature oocytes can be used in an in vitro fertilization procedure.

Target Nucleic Acids

In some embodiments, the methods provided herein can be used to modify any target nucleic acid. For example, in some embodiments, the methods provided herein can be used to fix mutations associated with genetic diseases, to knock-out genes, to insert transgenes, and/or to make other changes or enhancements to the germline genome, which are inherited by the next generation. In some embodiments, the target nucleic acid sequence is located in chromosomal DNA. In some embodiments, the target nucleic acid sequence is located on mitochondrial DNA. In certain embodiments, the methods provided herein can be applied to mammals for research, ecological, agricultural (*e.g.* farm animals or pets), and therapeutic purposes, including the application to humans for purposes relating to health.

In some embodiments of the methods provided herein, the target nucleic acid sequence is a gene encoding an allele associated with a disease or disorder (*e.g.,* an Brca1 or Brca2 allele associated with breast cancer, a Htt allele associated with Huntington's disease, a Cftr allele associated with cystic fibrosis, a Hbb allele associated with thalassemia or sickle cell disease) that is modified to an allele of the gene that is not associated with the disease or disorder.

In some embodiments, the target nucleic acid sequence is a gene associated with hemochromatosis (HFE-related), alpha-1 antitrypsin deficiency, phenylketonuria, familial dysautonomia, canavan disease, familial hyperinsulinism (ABCC8-related), primary hyperoxaluria type 2 (PH2), primary hyperoxaluria type 2 (PH2), rhizomelic chondrodysplasia punctata type 1 (RCDP1), torsion dystonia, polycystic kidney disease, TTR-related cardiac amyloidosis, mucolipidosis IV, limb-girdle muscular dystrophy, leigh syndrome french canadian type (LSFC), congenital disorder of glycosylation type 1a (PMM2-CDG), DPD deficiency, dihydrolipoamide dehydrogenase deficiency, deuronal ceroid lipofuscinosis (PPT1-related), medium-chain acyl-coA dehydrogenase (MCAD) deficiency, Glycogen Storage Disease Type 1a, glycogen storage disease type 1b, gaucher

disease, ARSACS, G6PD deficiency, cystic fibrosis, factor XI deficiency, Zellweger syndrome spectrum, Nijmegen breakage syndrome, D-bifunctional protein deficiency, LAMB3-related junctional epidermolysis bullosa, familial mediterranean fever, TTR-related familial amyloid polyneuropathy, pendred syndrome, tyrosinemia type I, hereditary fructose intolerance, familial hypercholesterolemia type B, hypertrophic cardiomyopathy (MYBPC3 25bp-deletion), BRCA cancer mutations, connexin 26-related sensorineural hearing loss, beta thalassemia, sickle cell anemia and malaria resistance, Fanconi anemia (FANCC-related), Bloom's syndrome, salla disease, GRACILE syndrome, maple syrup urine disease type 1B, Tay-Sachs disease, agenesis of the corpus callosum with peripheral neuropathy (ACCPN) or neuronal ceroid lipofuscinosis (CLN5-related).

In some embodiments, the target nucleic acid sequence is a gene encoding an allele associated with a particular trait in an animal (*e.g.,* disease resistance, toxin resistance, cancer resistance, radiation resistance, cell stress resistance, growth rate, adult size, lifespan, milk production, meat production, leather production, endurance, wool production, egg production, domestication, allergies, food sensitivities, baldness, hair color, fertility) that is modified to an allele associated with a different trait.

In some embodiments, the target nucleic acid sequence is a gene associated with alcohol flush reaction, bitter taste perception, earwax type, hair curl, lactose intolerance, malaria resistance, muscle performance, non-ABO blood groups, norovirus resistance, resistance to HIV/AIDS (*e.g.,* CCR5 delta32 gene resistance), male pattern baldness, adiponectin levels, asparagus metabolite detection, biological aging, birth weight, blood glucose, breast morphology. C-reactive protein level, childhood and adolescent growth, resistance to chronic hepatitis B, finger length ratio, freckling, hair color, height, LDL cholesterol levels, HDL cholesterol levels, leprosy susceptibility, obesity, nearsightedness and farsightedness, persistent fetal hemoglobin, reading ability, response to diet, response to exercise, sex hormone regulation, tooth development, tuberculosis susceptibility, hypospadias, prostate-specific antigen, eating behavior, hair thickness, longevity, memory, odor detection, pain sensitivity or avoidance of errors.

EXAMPLES

Example 1: Transfection and Expansion of Primary Germline Stem Cells

Primary oogonial stem cells were isolated from mouse ovaries as set forth in Woods and Tilly *Nat. Protoc.* 8:966-988 (2013), which is hereby incorporated by reference in its entirety. Ovaries were dissected from 3 month old female mice, and the ovarian cells were

dissociated, washed and stained with anti-Ddx4 primary and APC-labeled secondary antibodies, as follows.

Ovaries were dissected from female mice, taking care to remove the attached fat pad, bursa and oviduct from each ovary. Using a scalpel blade or mincing scissors, the ovaries were minced into slurry in 0.5 ml of collagenase/DNase I solution in a glass tissue culture dish. Using a 5-ml glass serological pipette, the slurry was rinsed to the bottom of the dish with 2.5 ml of collagenase/DNase I solution and collected the by placing the entire 3 ml of solution into a 15-ml conical tube. The tube was incubated in a prewarmed (37 °C) orbital shaker for 15 min at 250 r.p.m. The tube was removed from the orbital shaker and the slurry manually dispersed with gentle pipetting using a 5-ml glass serological pipette. The tube was incubated in the orbital shaker at 37 °C for an additional 15 min at 250 r.p.m. The tube was removed from the orbital shaker and the slurry manually dispersed with gentle pipetting until no visible pieces of ovary were present. The cell suspension was filtered through a 100-μm nylon mesh cell strainer, collecting the filtrate into a new 15-ml conical tube. Ten ml of warm HBSS was added to the conical tube containing the strained cell suspension, and the tube was centrifuged at 300g for 5 min at room temperature, with the centrifuge brake turned off. After centrifugation, the liquid was decanted, taking care to remove as much of the supernatant as possible without disturbing the cell pellet. Using a sterile 5-ml glass serological pipette, the cell pellet was resuspended in 4 ml of warm HBSS. The volume of the solution was brought to 10 ml with warm HBSS. The cell suspension was centrifuged at 300g for 5 min at room temperature with the brake on the centrifuge turned off. After centrifugation, the liquid was decanted, taking care to remove as much of the supernatant as possible without disturbing the cell pellet.

The cell pellet was resuspended in 500 μl of cold antibody blocking/dilution solution, which was placed on ice for 20 minutes. 100 μl of cell suspension was added to negative control and secondary antibody only tubes, which were placed on ice. The remaining 300 μl of cell suspension were diluted to 10 ml in cold HBSS. The suspension was centrifuged at 300g for 5 min at 4 °C. The sample tube was removed from the centrifuge and the supernatant discarded, being careful not to disturb or dislodge the cell pellet. The cell pellet was resuspended in 100 μl of primary antibody solution. The suspension was placed on ice and incubated for 20 min. The resuspended cell sample was mixed with primary antibody to a total of 10 ml with cold HBSS and centrifuged at 300g for 5 min at 4 °C. The sample tube was removed from the centrifuge and the supernatant discarded. The sample tube cell pellet

was resuspended in 10 ml of cold HBSS. The suspension was centrifuged at 300g for 5 min at 4 °C. At this time, the cells in the secondary antibody only tube were also centrifuged and the supernatant discarded. The cell pellet in the sample tube, as well as the cell pellet in the secondary antibody only tube, were resuspended in 250 µl of prepared secondary antibody solution for FACS and placed on ice for 20 minutes. The volume in the sample tube and the secondary antibody only tube were brought to a total of 10 ml with cold HBSS and centrifuged at 300g for 5 min at 4 °C. The supernatants were discarded and the cell pellets resuspended in 10 ml of cold HBSS. At this time, the volume of the negative control tube was brought to 10 ml with cold HBSS. All three tubes were centrifuged at 300g for 5 min at 4 °C and the supernatants discarded. Each cell pellet (sample tube, secondary antibody only tube and negative control tube) were resuspended in in 0.5 ml of FACS buffer and isolated by FACS. As depicted in Figure 2, Ddx4 positive oogonial stem cells were isolated by FACS.

Oogonial stem cells that had been expanded in culture were transfected with various amounts of Cas9-GFP fusion protein expression vector using different amounts lipofectamine 2000. Briefly, 1.6 ug of total plasmid DNA was diluted into OptiMEM to a total volume of 100 uL, per sample. Four uL of Lipofectamine 2000 was diluted into OptiMEM to a total volume of 100 uL, per sample. After a 5 m incubation, 100 uL of each were combined and the resulting mixture incubated for 20 minutes – 5 hours. The OSCs were pelleted by spinning at 200g for 5 min, and the supernatant removed. The pellet was resuspended to a concentration of 100,000 cells per mL. One mL of cells were added to a well of a 12-well plate, to which the plasmid/Lipofectamine mixture was added drop-wise. After 18-24 hours, the media was changed. The resulting transfection efficiencies at 24 hours post-transfection are shown in Figure 3. As depicted in Figure 4, at 72 hours post-transfection, GFP positive cells were isolated by FACS.

Single OSCs were sorted, unstained, by FACS into individual wells of a 96-well plate and grown in culture. The OSCs were trypsinized and then centrifuged at 200g for 5 min, and the supernatant removed. The pellet was resuspended in 1 mL of FACS buffer (0.1% FBS in HBSS). To each well of a 96-well plate was added 50 uL of OSC media. A single cell was sorted into each well of the 96-well plate by FACS. In general, small colonies were visible within a few days. Within 1-2 weeks, colonies grew large enough to be passaged to larger wells. Between 21% and 33% of the isolated single OSCs produced clonal

OSC colonies, one of which is depicted in Figure 5. Whether the OSCs were transfected or not prior to isolation did not significantly affect the efficiency of colony formation.

Example 2: Genome Modification of Primary Germline Stem Cells

Primary OSCs were isolated from mouse ovaries as set forth in Woods and Tilly *Nat. Protoc.* 8:966-988 (2013), which is hereby incorporated by reference in its entirety. Ovaries were dissected from 2-4 month old female C57BL6/J mice, digested with Liberase TL (Roche 05401020001) and DNase I in HBSS, and stained with a Ddx4 primary antibody (Abcam 13840) and an APC-conjugated secondary antibody (Jackson ImmunoResearch 111-136-144). Cells were also stained with 1 μg/mL DAPI and 5 μg/mL Calcein-AM (Invitrogen C3099) for exclusion of dead cells and debris. Calcein high, $Ddx4^+$ cells were sorted on a BD FACSAria II cell sorter using the sorting strategy depicted in Figure 7.

Ddx4-positive cells were grown in MEM-α GlutaMax (Invitrogen 32561) with 10% (v/v) FBS (Invitrogen 26140), 1 mM sodium pyruvate (Invitrogen 11360), 0.1 mM NEAA (Invitrogen 11140), pen-strep-glutamine (Invitrogen 10378), N-2 Plus supplement (R&D Systems 212-GD-050) , 0.1 mM β-mercaptoethanol, 1000 units/mL LIF (Millipore ESG1106), 10 ng/mL EGF (Invitrogen PHG0314), 1 ng/mL bFGF (Invitrogen 13256), and 40 ng/mL GDNF (R&D Systems 212-GD-010). During culture of freshly-isolated cells, half the media was replaced every second day. Expansion of the sorted OSC colonies are depicted in Figure 8.

The genomes of the cultured OSCs were modified by the targeted integration of GFP into the ROSA26 locus (Figure 9). The established OSC lines were reverse-transfected with 4 μL/mL of Lipofectamine 2000 (Life Technologies 11668027), and 3 μg/mL of DNA. hCas9 (AddGene 41815), ROSA26 gRNA (modified from AddGene 41824), and ROSA26-GFP-puro donor (AddGene 26890) plasmids were transfected in an equimolar ratio. GFP-expressing cells were sorted as single cells into wells of a 96-well plate 1-2 weeks after transfection. Two to three weeks later, colonies were screened using prepGEM Tissue (Zygem PTI0500) to isolate genomic DNA, and KAPA Hifi polymerase (KAPA Biosystems KK2602) to PCR-amplify the genome-insertion junction, with the primers TAGGTAGGGGATCGGGACTC and GAAAGACCGCGAAGAGTTTG (Figure 9). Modifications were verified using DNeasy Blood & Tissue kit (QIAGEN 69581) to re-isolate genomic DNA, and by sequencing with the PCR-primers (DF/HCC DNA Resource Core) (Figure 10). Expression of GFP by the modified cells was detected by fluorescent microscopy (Figure 11).

Expression of germline genes was determined in the edited OSCs (Figure 12). RNA was isolated using RNeasy Mini kit (QIAGEN 74140) and converted to cDNA using iScript cDNA Synthesis kit (BIO-RAD 170-8890). Real-time quantitative PCR was done using LightCycler 480 SYRB Green I Master (Roche 04707516001) using primers having the following sequences.

Gene	Forward	Reverse
Beta-actin	TGG CGC TTT TGA CTC AGG AT	GGG ATG TTT GCT CCA ACC AA
GAPDH	CTT TGT CAA GCT CAT TTC CTG G	TCT TGC TCA GTG TCC TTG C
Ddx4	TAT GAT GCG GGA TGG AAT AAC TG	ACA GCT CTT ACA CAA GTC CCA A
Blimp1	CTA CTG TAT TGC TTT GGG TTG C	ATT AAG CCT ATC CCT GCC AAC
Tert	CTG AGT CTC ACC AGT ACA AGT GT	TTG GCA CCC ATG ATT TGC CT
Fragilis	GGT CTG GTC CCT GTT CAA TAC	CTC CAG TCA CAT CAC CCA C

Example 3: Targeting the AAVS1 "Safe-Harbor" Locus in Human OSCs

Human OSCs were modified using CRISPR-Cas9 to insert a GFP reporter into the AAVS1 "safe harbor" locus as depicted in Figure 13a. Primary human OSCs were transfected with X-tremeGENE HP DNA Transfection Reagent (Roche) and four plasmids: Cas9, two gRNAs (guide sequences: GTCCCCTCCACCCCACAGTG and GGGGCCACTAGGGACAGGAT) and a donor plasmid encoding a GFP and puroR genes with two homology arms to the AAVS1 locus, each of about 800 bp. Cells are grown for 1-2 weeks, after which GFP^+ cells were sorted into individual wells of a 96-well plate to form clonal colonies. Colonies are genotyped with primers flanking the target site (CCCCTTACCTCTCTAGTCTGTGC and CTCAGGTTCTGGGAGAGGGTAG) and modified colonies were identified (Figure 13b). Expression of GFP in the positive colonies was confirmed by fluorescent microscopy (Figures 13c and 13d).

Example 4: Genome Editing of Human Oogonial Stem Cells

A pipeline was developed to edit human oogonial stem cells (OSCs). This pipeline utilizes nine quality control checkpoints to ensure successful genome editing of OSCs (Figure 14). During the maintenance phase of the pipeline, isolated OSCs are confirmed to advance to a sufficient passage number for genome editing to take place (checkpoint 1) and tested for mycoplasma (checkpoint 2). Early-passage OSCs are then genome edited using

reverse transfection delivery of CRISPR/Cas9 directed towards a gene of interest (checkpoints 3-4). Enrichment/isolation of edited cells is conducted in three arms (checkpoints 5-7) using either single cell fluorescence-activated cell sorting (FACS) (checkpoint 5), limiting dilution (checkpoint 6) or mixed population FACS enrichment (checkpoint 7). Finally, successful genome editing is confirmed through PCR (checkpoints 8, 9). Figure 15A shows a mixed population of human OSCs in which the GFP gene has been directed to the AAV1 safe harbor locus using CRISPR/Cas9. GFP-positive single sorted EggPCs were selected for clonal expansion under optimized conditions (Figure 15B).

All publications, patents, and patent applications mentioned herein are hereby incorporated by reference in their entirety as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed is:

1. A method of modifying a target nucleic acid sequence in the genome of a germ stem cell (GSC) comprising contacting the GSC with a germline targeted nuclease, wherein the germline targeted nuclease modifies the target nucleic acid sequence.

2. The method of claim 1, wherein the GSC is contacted with the germline targeted nuclease by transfecting the GSC with an expression vector or RNA encoding the germline targeted nuclease or introducing a functional targeted nuclease protein.

3. The method of claim 1, wherein the germline targeted nuclease is a zinc-finger nuclease (ZFN), a transcription activator-like effector nuclease (TALEN), an RNA-guided endonuclease (RGEN) or an engineered homing meganuclease (EMH).

4. The method of claim 1, wherein the germline targeted nuclease is Cas9.

5. The method of claim 1, wherein the germline targeted nuclease is a fusion protein comprising a detectable moiety.

6. The method of claim 1, wherein the GSC is an oogonial stem cell or a spermatogonial stem cell.

7. The method of claim 1, further comprising the step of isolating the GSC prior to contacting it with the germline targeted nuclease.

8. The method of claim 7, wherein the GSC is isolated by FACS.

9. The method of claim 1, further comprising the step of clonally culturing the transfected GSC to generate a clonal GSC colony.

10. The method of claim 9, wherein the transfected GSC is isolated from other GSCs prior to being cultured.

11. The method of claim 10, wherein the GSC is isolated by FACS.

12. The method of claim 11, wherein the GSC is isolated based on its expression of the germline targeted nuclease.

13. The method of claim 12, wherein the GSC expresses a low level of the germline targeted nuclease.

14. The method of claim 9, further comprising the step of determining whether the target nucleic acid sequence has been modified in cells in the clonal GSC colony.

15. The method of claim 14, wherein whether the target nucleic acid sequence has been modified is determined by sequencing the target nucleic acid sequence.

16. The method of claim 9, further comprising the step of sequencing the genome of cells in the clonal GSC colony.

17. The method of claim 9, further comprising the step of inducing the maturation of a cell from the clonal GSC colony into a mature germ cell.

18. The method of claim 17, wherein the maturation of the GSC is induced by transplantation of the GSC into testes or ovarian tissue, directed differentiation *in vitro* or nuclear transfer into a mature germ cell.

19. The method of claim 9, wherein the target nucleic acid sequence is a mitochondrial DNA sequence.

20. The method of claim 19, further comprising the step of transferring mitochondria from a cell in the GSC colony to a mature germ cell.

21. The method of claim 17, further comprising the step of using the mature germ cell in an in vitro fertilization procedure.

22. The method of claim 9, wherein the GSC is an oogonial stem cell and further comprising transferring a cell of the clonal GSC colony to a surrogate mother.

23. The method of claim 1, wherein the target nucleic acid sequence is a gene encoding a disease causing allele.

24. The method of claim 23, wherein the gene is modified such that it does not encode a disease causing allele.

25. A method of generating an animal carrying a modified target nucleic acid sequence in its genome, the method comprising the steps of:

a) contacting a sample comprising a population of GSCs with a germline targeted nuclease, wherein the germline targeted nuclease is targeted to the target nucleic acid sequence;

b) isolating single GSCs from the population of GSCs;

c) culturing the single GSCs to form clonal GSC colonies;

d) identifying the clonal GSC colonies that comprise a modified target nucleic acid sequence; and

e) generating an animal from a cell of a clonal GSC colony identified in step d).

26. The method of claim 23, wherein the GSC is contacted with the germline targeted nuclease by transfecting the GSC with an expression vector or RNA encoding the germline targeted nuclease or introduction of a germline targeted nuclease protein.

27. The method of claim 25, wherein the germline targeted nuclease is a zinc-finger nuclease (ZFN), a transcription activator-like effector nuclease (TALEN) an RNA-guided endonuclease (RGEN) or an engineered homing meganuclease (EMH).

28. The method of claim 25, wherein the germline targeted nuclease is Cas9.

29. The method of claim 25, wherein the germline targeted nuclease is a fusion protein comprising the detectable moiety.

30. The method of claim 29, wherein the single GSCs isolated in step b) are cells that express the germline targeted nuclease.

31. The method of claim 30, wherein the single GSCs isolated in step b) are cells that express a low level of the germline targeted nuclease.

32. The method of claim 25, wherein the GSCs are an oogonial stem cells or spermatogonial stem cells.

33. The method of claim 25, wherein the single GSCs are isolated in step b) by FACS.

34. The method of claim 25, further comprising the step of sequencing the genome of the clonal GSC colonies that comprise a modified target nucleic acid.

35. The method of claim 25, wherein step e) comprises inducing the maturation of a cell from the clonal GSC colony into a mature germ cell and the animal is generated from the mature germ cell.

36. The method of claim 35, wherein the maturation of the GSC is induced by transplantation of the GSC into testes or ovarian tissue, directed differentiation *in vitro* or nuclear transfer into a mature germ cell.

37. The method of claim 35, wherein the animal is generated in step e) from the mature germ cell using in vitro fertilization.

38. The method of claim 25, wherein the GSC is an oogonial stem cell and wherein the animal is generated in step e) by transferring a cell from a clonal GSC colony comprising a modified target nucleic acid to a surrogate mother and mating the surrogate mother.

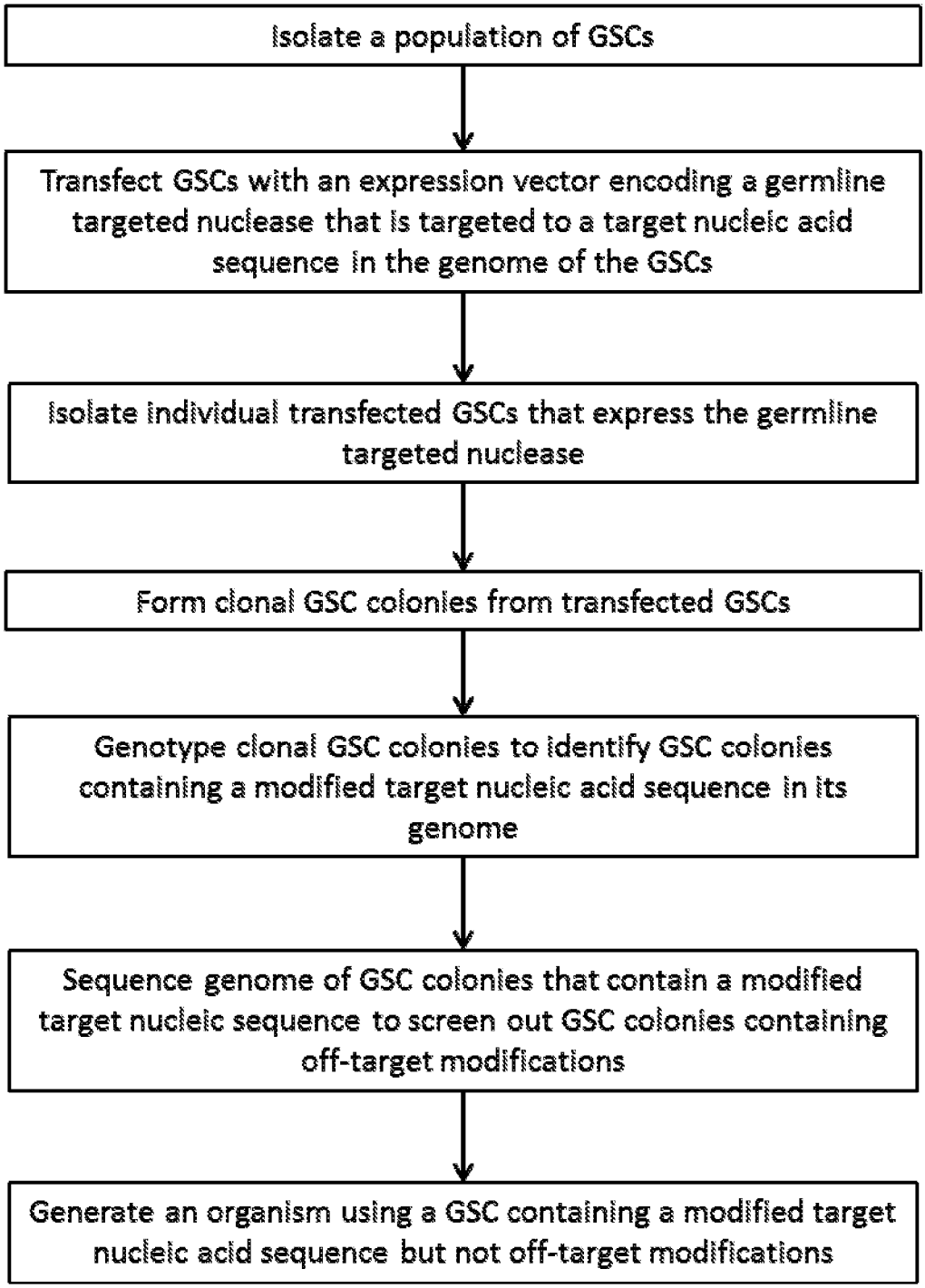
39. The method of claim 25, further comprising the step of isolating the population of GSCs prior to step a).

40. The method of claim 25, wherein the target nucleic acid sequence is a gene encoding a disease causing allele.

41. The method of claim 40, wherein the gene is modified such that it does not encode a disease causing allele.

42. The method of claim 25, wherein the animal is a mammal.

Figure 1


Isolate a population of GSCs
Transfect GSCs with an expression vector encoding a germline targeted nuclease that is targeted to a target nucleic acid sequence in the genome of the GSCs
Isolate individual transfected GSCs that express the germline targeted nuclease
Form clonal GSC colonies from transfected GSCs
Genotype clonal GSC colonies to identify GSC colonies containing a modified target nucleic acid sequence in its genome
Sequence genome of GSC colonies that contain a modified target nucleic sequence to screen out GSC colonies containing off-target modifications
Generate an organism using a GSC containing a modified target nucleic acid sequence but not off-target modifications

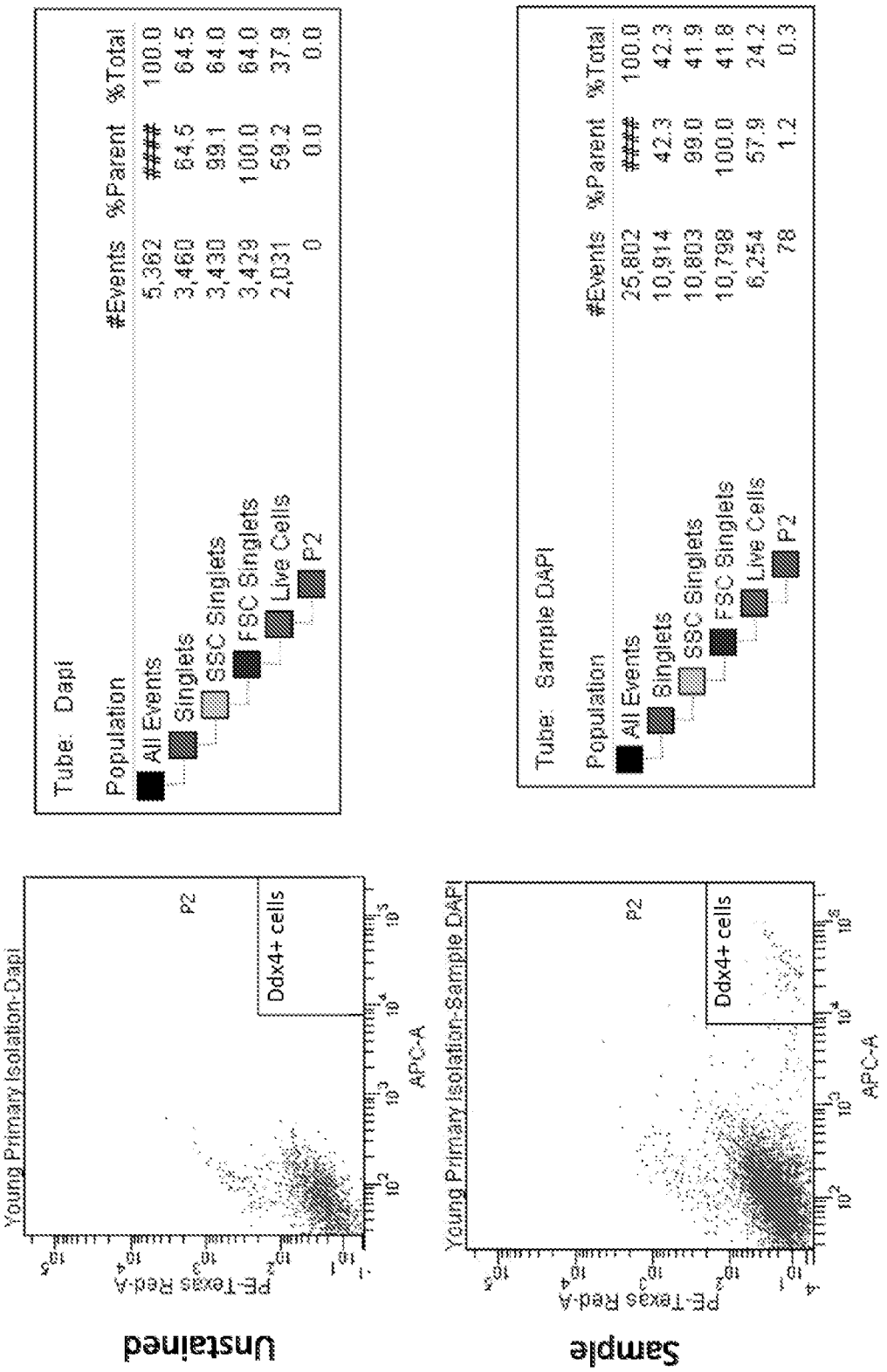

Tube: Dapi
Population
#Events
%Parent
%Total
All Events
5,362
####
100.0
Singlets
3,460
64.5
64.5
SSC Singlets
3,430
99.1
64.0
FSC Singlets
3,429
100.0
64.0
Live Cells
2,031
59.2
37.9
P2
0
0.0
0.0
Tube: Sample DAPI
Population
#Events
%Parent
%Total
All Events
25,802
####
100.0
Singlets
10,914
42.3
42.3
SSC Singlets
10,803
99.0
41.9
FSC Singlets
10,798
100.0
41.8
Live Cells
6,254
57.9
24.2
P2
78
1.2
0.3
Young Primary Isolation-Dapi
P2
Ddx4+ cells
PE-Texas Red-A
APC-A
Unstained
Young Primary Isolation-Sample DAPI
P2
Ddx4+ cells
PE-Texas Red-A
APC-A
Sample

Figure 2

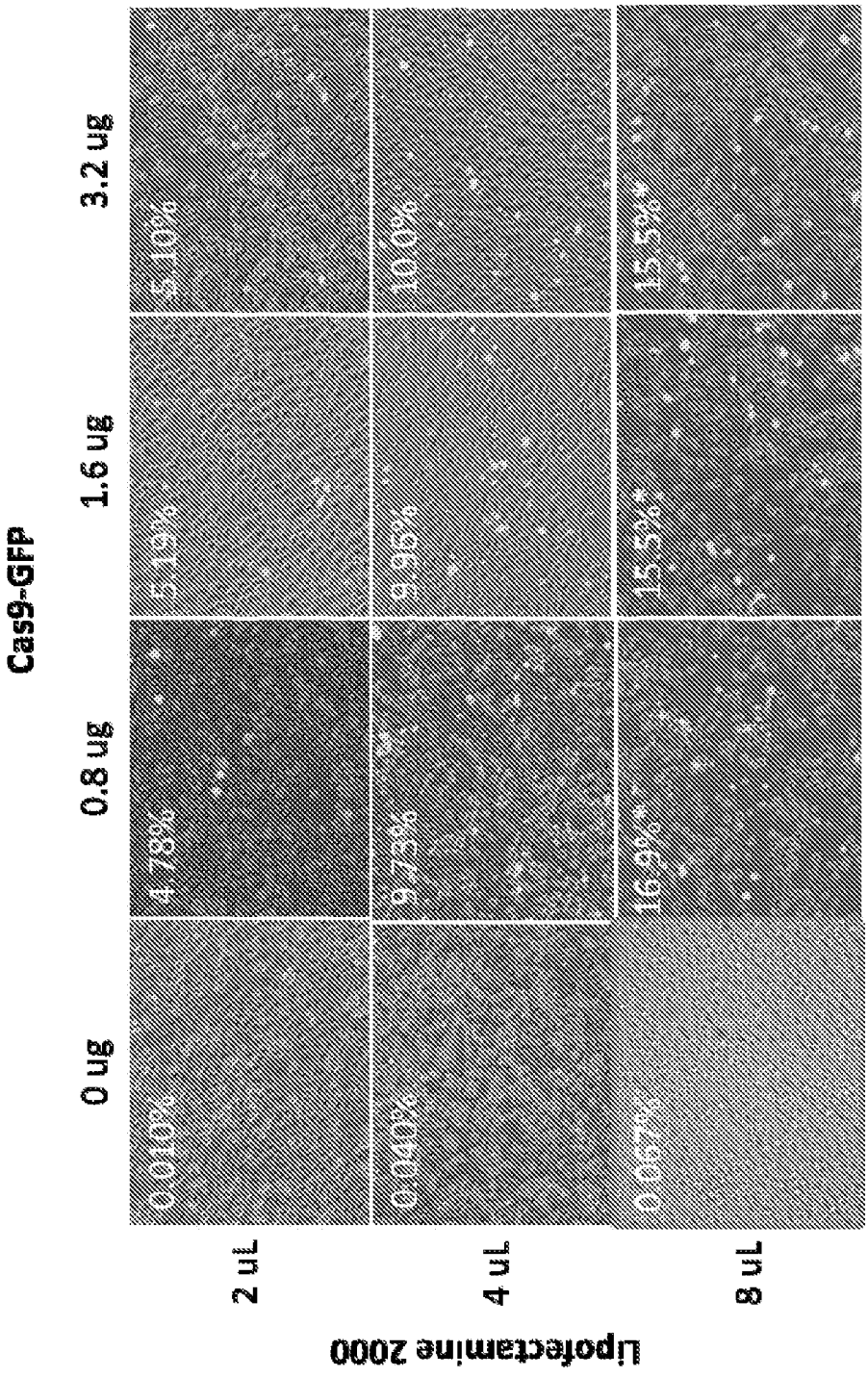

Cas9-GFP
0 ug
0.8 ug
1.6 ug
3.2 ug
2 uL
4 uL
8 uL
Lipofectamine 2000
0.011%
4.78%
5.19%
5.10%
0.041%
9.73%
9.96%
10.0%
0.067%
16.9%*
15.5%*
15.5%*
untreated
24 hours post transfection
*Reduced survival

Figure 3

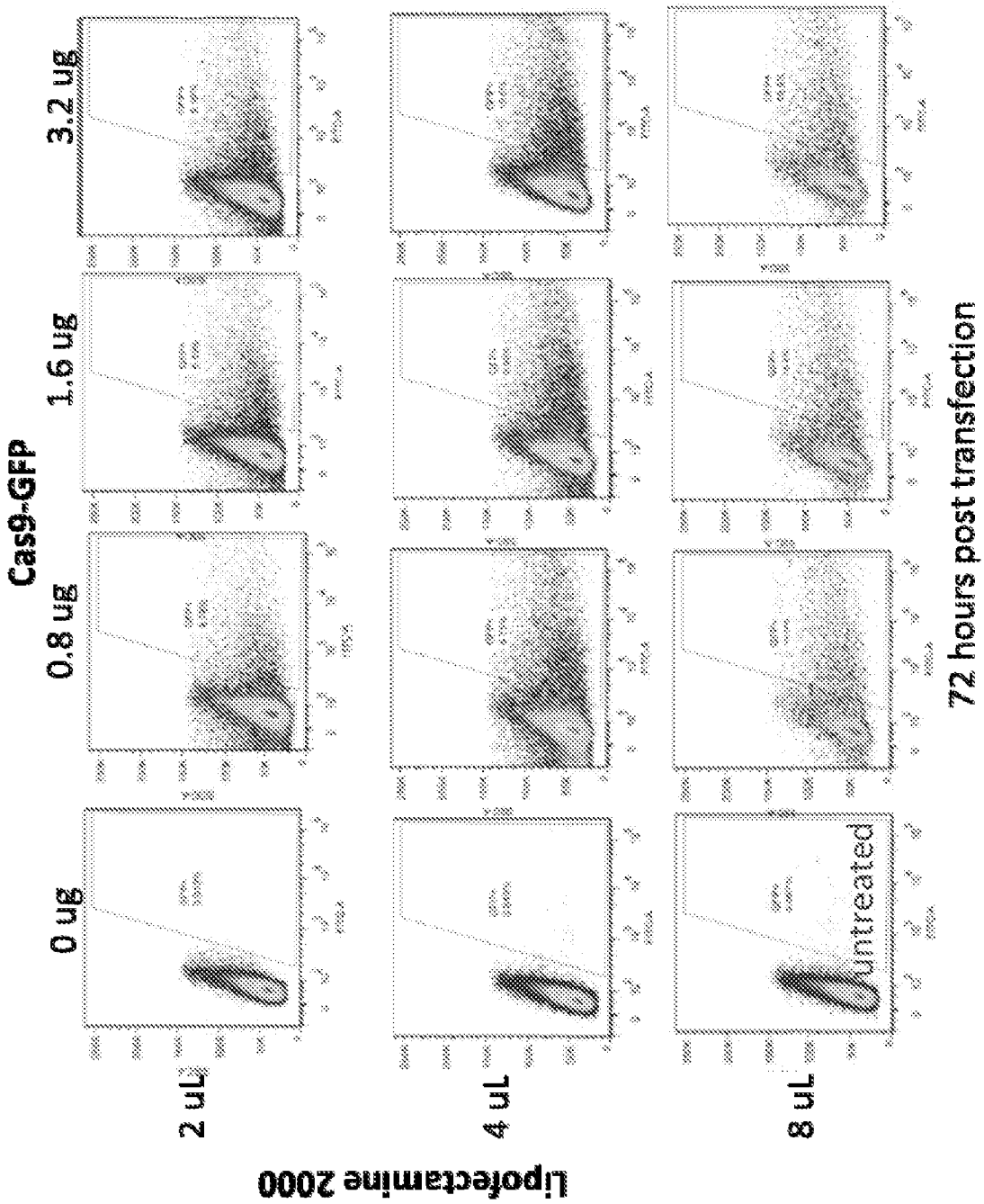

Cas9-GFP
0 ug
0.8 ug
1.6 ug
3.2 ug
2 uL
4 uL
8 uL
Lipofectamine 2000
Untreated
72 hours post transfection

Figure 4

Figure 5

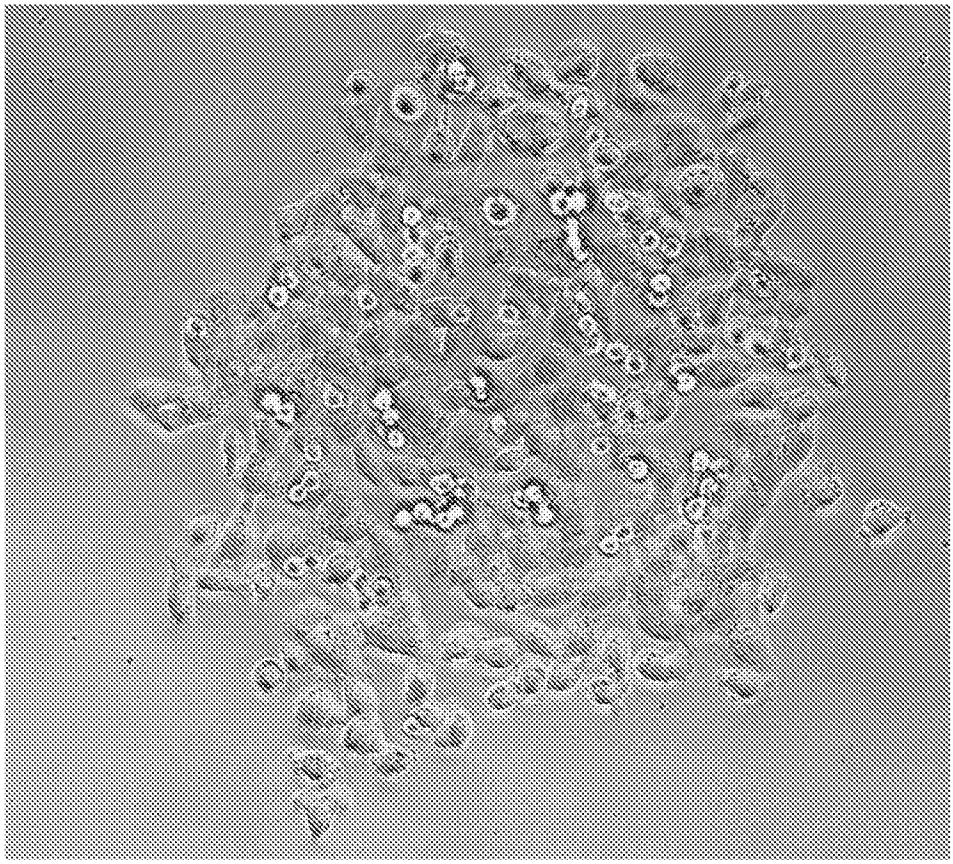


Figure 6

Human Codon Optimized Cas9:

gccaccATGGACAAGAAGTACTCCATTGGGCTCGATATCGGCACAAACAGCGTCGGCTGGGCCGTCATTA
CGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGTTCTGGGCAATACCGATCGCCACAGCATAAAGAA
GAACCTCATTGGCGCCCTCCTGTTCGACTCCGGGGAGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCA
CGGCGCAGATATACCCGCAGAAAGAATCGGATCTGCTACCTGCAGGAGATCTTTAGTAATGAGATGGCTA
AGGTGGATGACTCTTTCTTCCATAGGCTGGAGGAGTCCTTTTTGGTGGAGGAGGATAAAAAGCACGAGCG
CCACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCAACCATATATCATCTG
AGGAAGAAGCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCATATGA
TCAAATTTCGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAACAGCGATGTCGACAAACTCTT
TATCCAACTGGTTCAGACTTACAATCAGCTTTTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCC
AAAGCAATCCTGAGCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCCTGGGG
AGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGGCTGACCCCCAACTTTAAATCTAA
CTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAGACACCTACGATGATGATCTCGACAATCTG
CTGGCCCAGATCGGCGACCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTCTGC
TGAGTGATATTCTGCGAGTGAACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTA
TGATGAGCACCACCAAGACTTGACTTTGCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGAAGTACAAG
GAAATTTTCTTCGATCAGTCTAAAAATGGCTACGCCGGATACATTGACGGCGGAGCAAGCCAGGAGGAAT
TTTACAAATTTATTAAGCCCATCTTGGAAAAAATGGACGGCACCGAGGAGCTGCTGGTAAAGCTTAACAG
AGAAGATCTGTTGCGCAAACAGCGCACTTTCGACAATGGAAGCATCCCCCACCAGATTCACCTGGGCGAA
CTGCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTTGAAAGATAACAGGGAAAAGATTGAGA
AAATCCTCACATTTCGGATACCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTCGCGTGGAT
GACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGAGGAAGTCGTGGATAAGGGGGCCTCTGCC
CAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTGCCTAACGAAAAGGTGCTTCCTAAACACT
CTCTGCTGTACGAGTACTTCACAGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAG
AAAGCCAGCATTCCTGTCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGGAAA
GTTACCGTGAAACAGCTCAAAGAAGACTATTTCAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCG
GAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAATCATTAAAGACAAGGA
CTTCCTGGACAATGAGGAGAACGAGGACATTCTTGAGGACATTGTCCTCACCCTTACGTTGTTTGAAGAT
AGGGAGATGATTGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACGACAAAGTCATGAAACAGCTCA
AGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACTGATCAATGGGATCCGAGACAAGCAGAG
TGGAAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACTTCATGCAGTTGATCCAT
GATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAGTCTTCACG
AGCACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGGA
TGAACTCGTCAAAGTAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCGAGAGAACCAA
ACTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGG
GGTCCCAAATCCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTACCTGTACTA
CCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAACTGGACATCAATCGGCTCTCCGACTACGACGTG
GATCATATCGTGCCCCAGTCTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATA
AAAATAGAGGGAAGAGTGATAACGTCCCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCA
GCTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGATAATCTGACTAAGGCTGAACGAGGTGGCCTG
TCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCTTGTTGAGACACGCCAGATCACCAAGCACGTGG
CCCAAATTCTCGATTCACGCATGAACACCAAGTACGATGAAAATGACAAACTGATTCGAGAGGTGAAAGT
TATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGACTTTCAGTTTTATAAGGTGAGAGAGATC
AACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCACTGCACTTATCAAAAAATATC
CCAAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAATGATCGCAAAGTC
TGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTTTCAAGACC
GAGATTACACTGGCCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAA
TCGTGTGGGACAAGGGTAGGGATTTCGCGACAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGT
TAAAAAGACCGAAGTACAGACCGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCGACAAG
CTGATCGCACGCAAAAAAGATTGGGACCCCAAGAAATACGGCGGATTCGATTCTCCTACAGTCGCTTACA
GTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAAAAACTCAAAAGCGTCAAGGAACTGCTGGG
CATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACTTTCTCGAGGCGAAAGGATATAAA
GAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTGAAAACGGCCGGAAAC
GAATGCTCGCTAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATACGTTAATTT
CTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTTC

Figure 6 (Continued)

GTGGAACAACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCC
TCGCCGACGCTAACCTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCA
GGCAGAAAACATTATCCACTTGTTTACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCAAGTACTTCGAC
ACCACCATAGACAGAAAGCGGTACACCTCTACAAAGGAGGTCCTGGACGCCACACTGATTCATCAGTCAA
TTACGGGGCTCTATGAAACAAGAATCGACCTCTCTCAGCTCGGTGGAGACAGCAGGGCTGACCCCAAGAA
GAAGAGGAAGGTGTGA

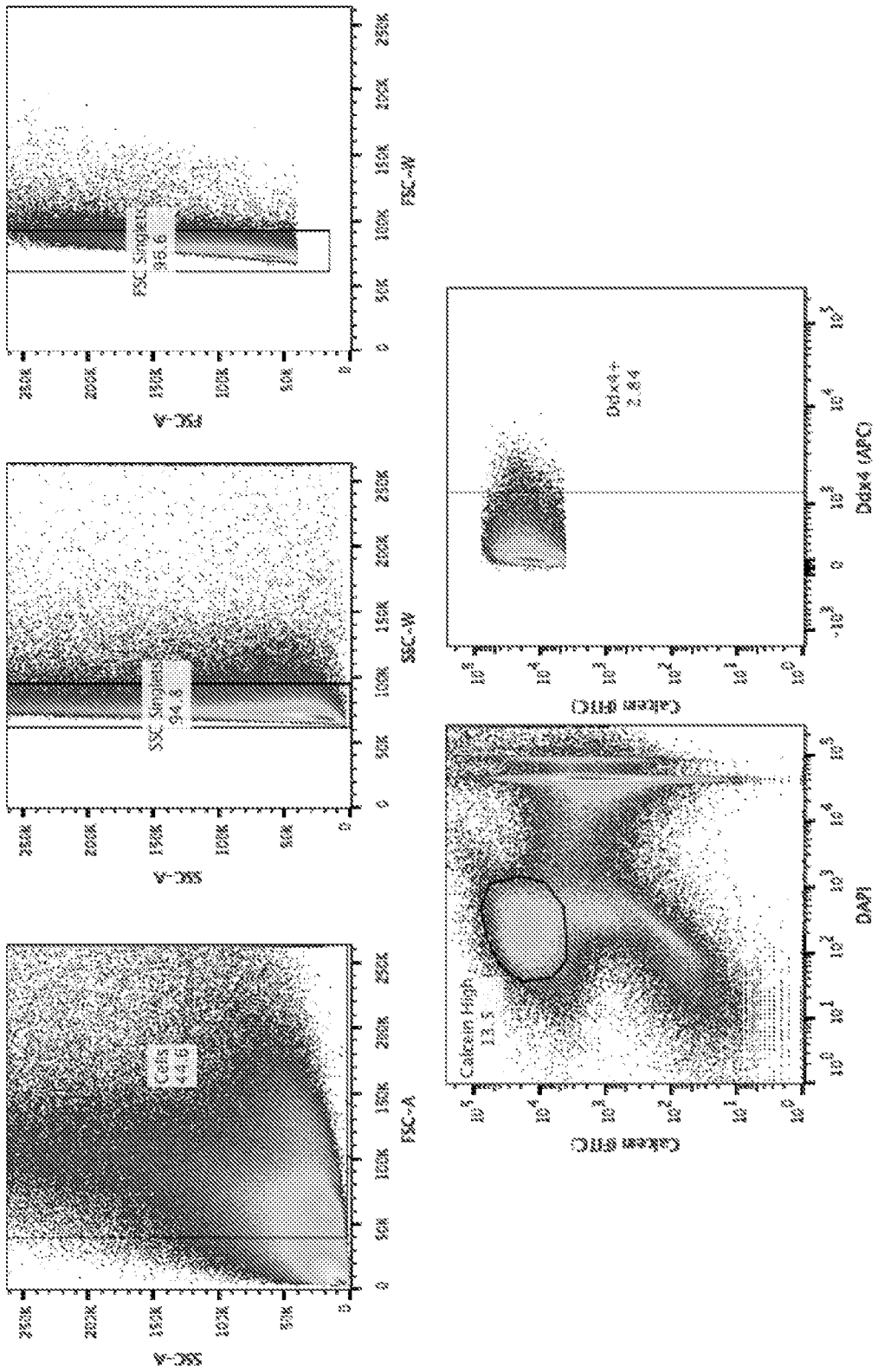

Cells
44.8
FSC-A
SSC-A
SSC Singlets
64.8
SSC-W
FSC Singlets
86.6
FSC-W
Calcein High
13.5
Calcein FITC
DAPI
Ddx4+
2.84
Ddx4 (APC)

Figure 7

Figure 8

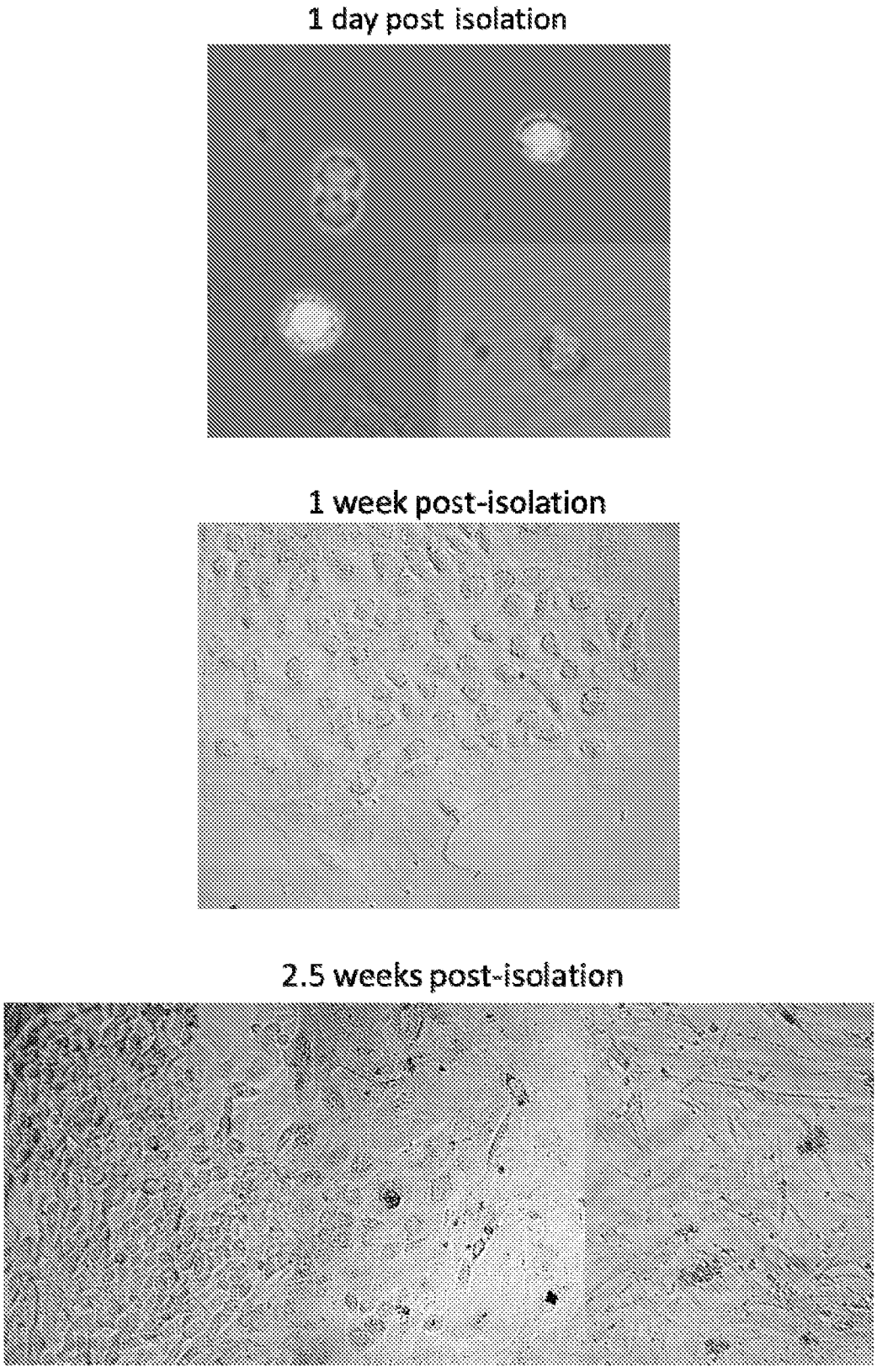
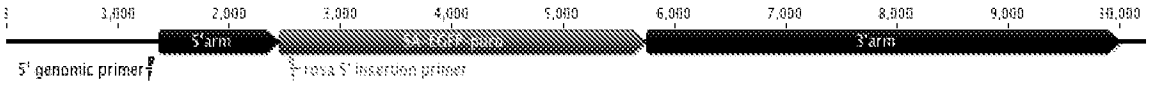

1 day post isolation
1 week post-isolation
2.5 weeks post-isolation

Figure 9



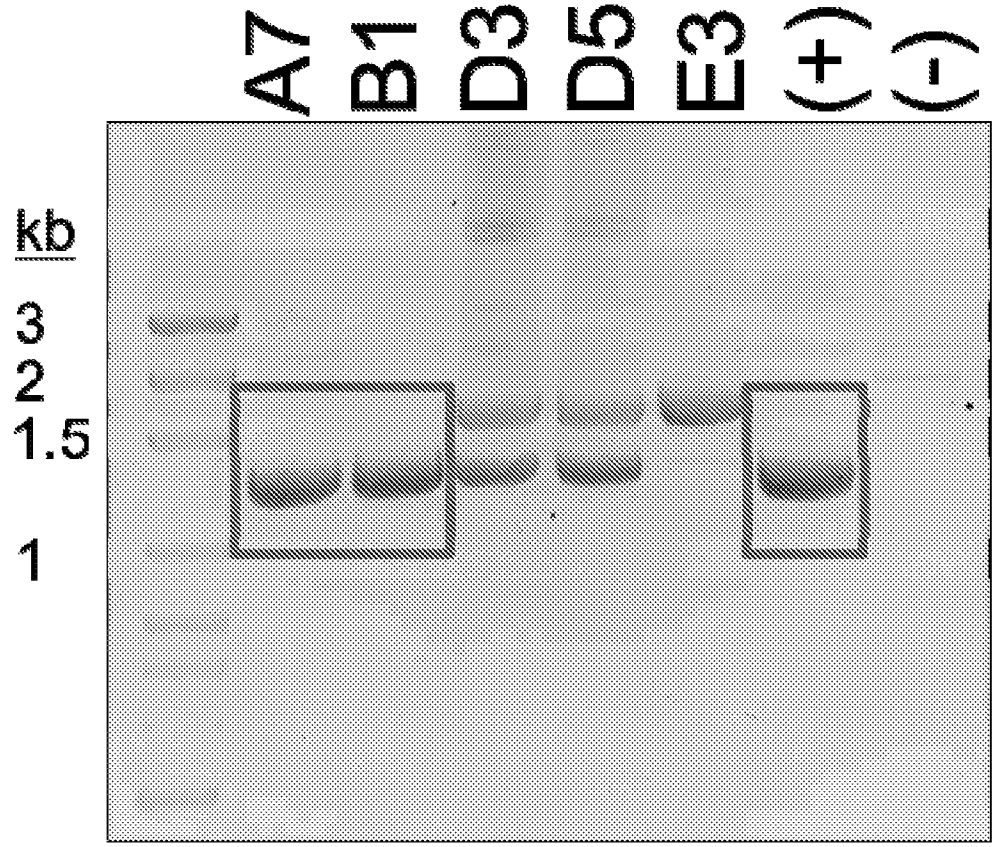
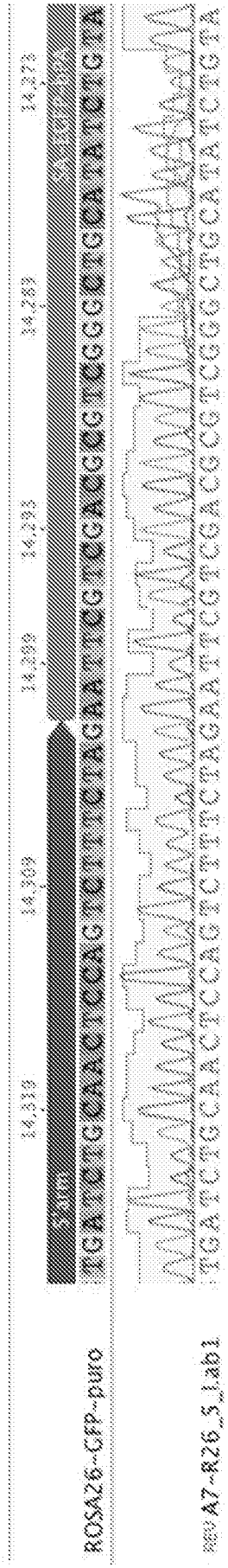


Figure 10


14,213
14,309
14,209
14,293
14,283
14,273
ROSA26-GFP-puro
TGATCCTGCAACTCCAGTCTTTCTAGAATTCGTCGACGGCGTCCGGCCTGCATATCCTGTA
REV A7-R26_5_lab1
TGATCCTGCAACTCCAGTCTTTCTAGAATTCGTCGACGGCGTCCGGCCTGCATATCCTGTA

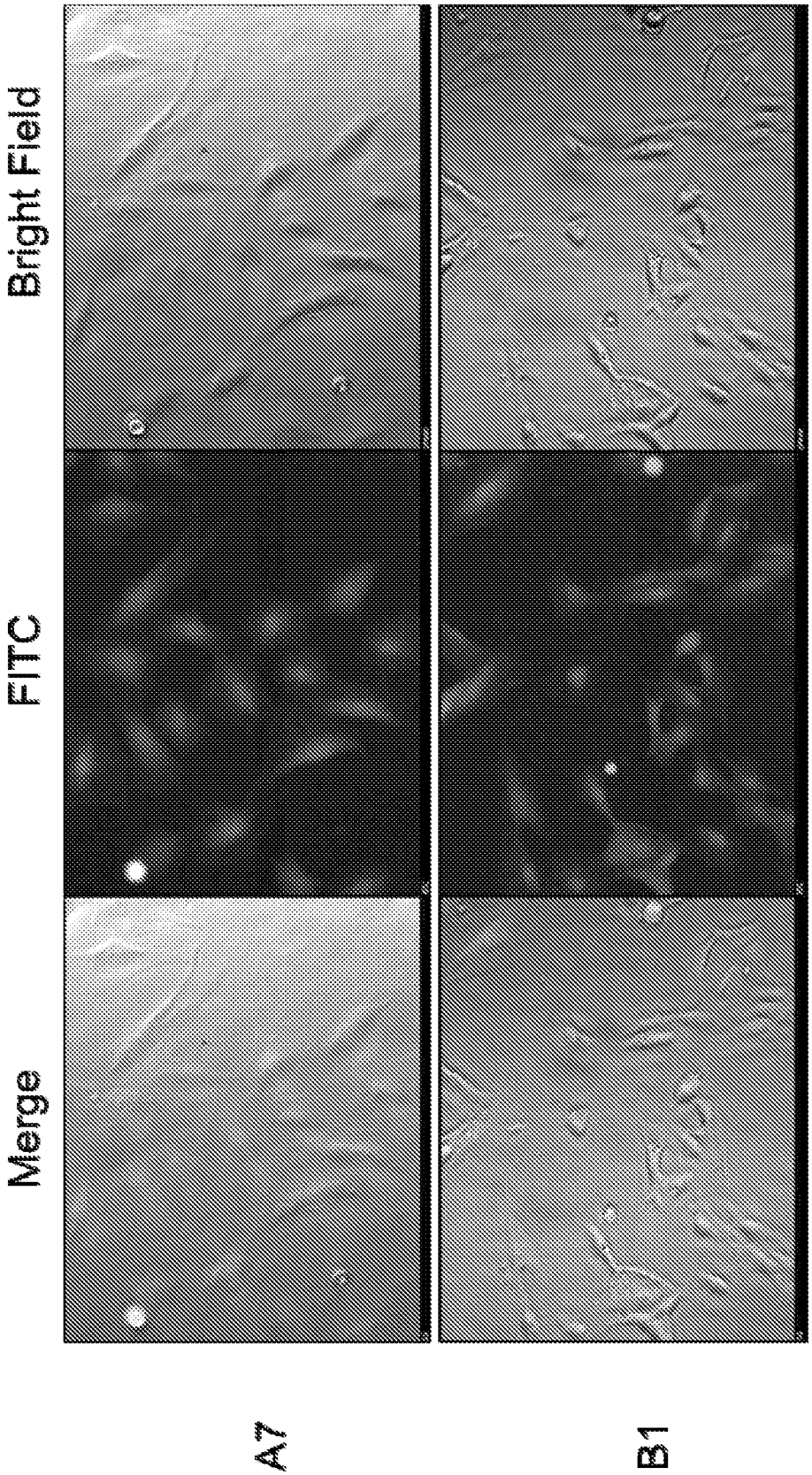


Figure 11

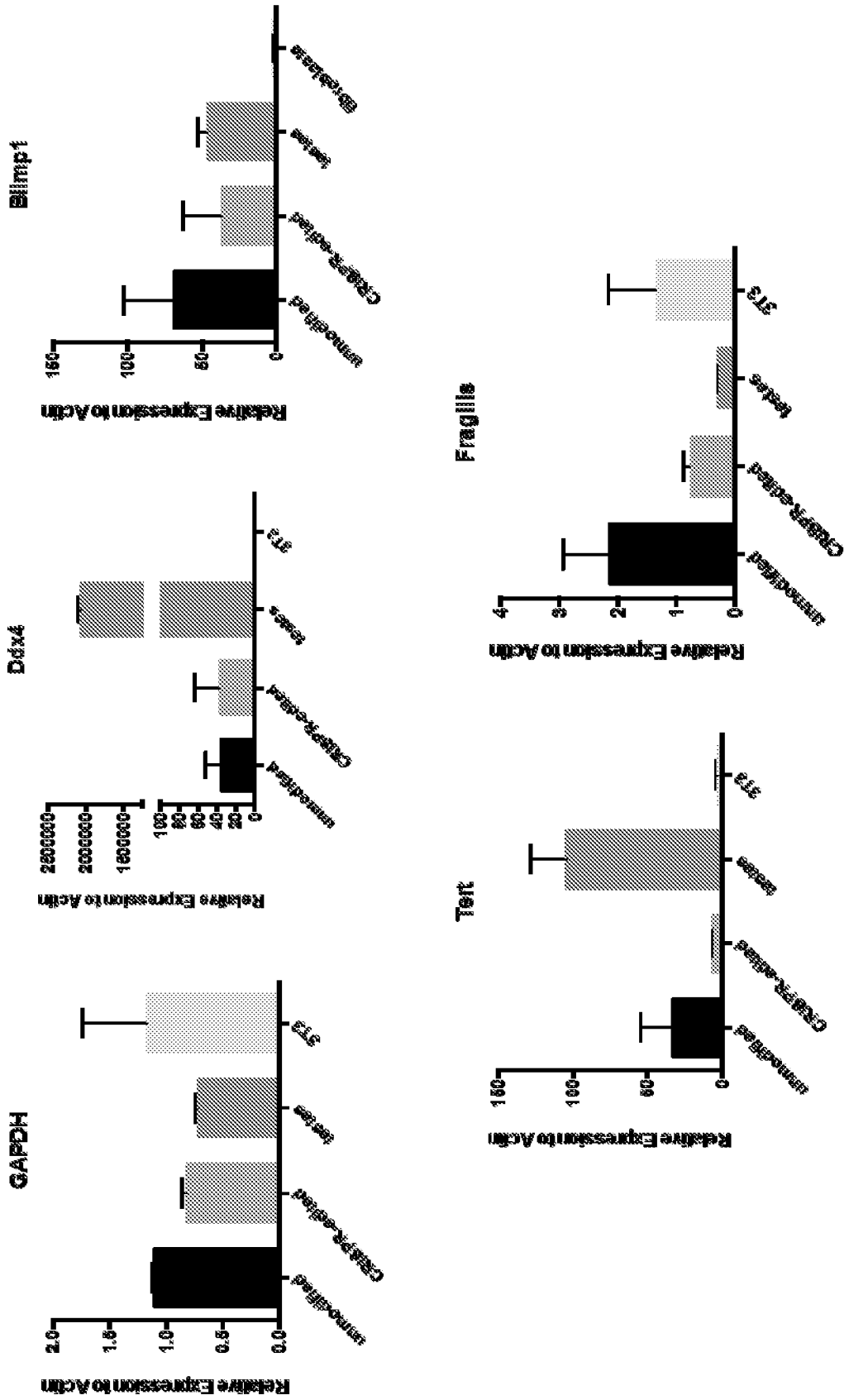
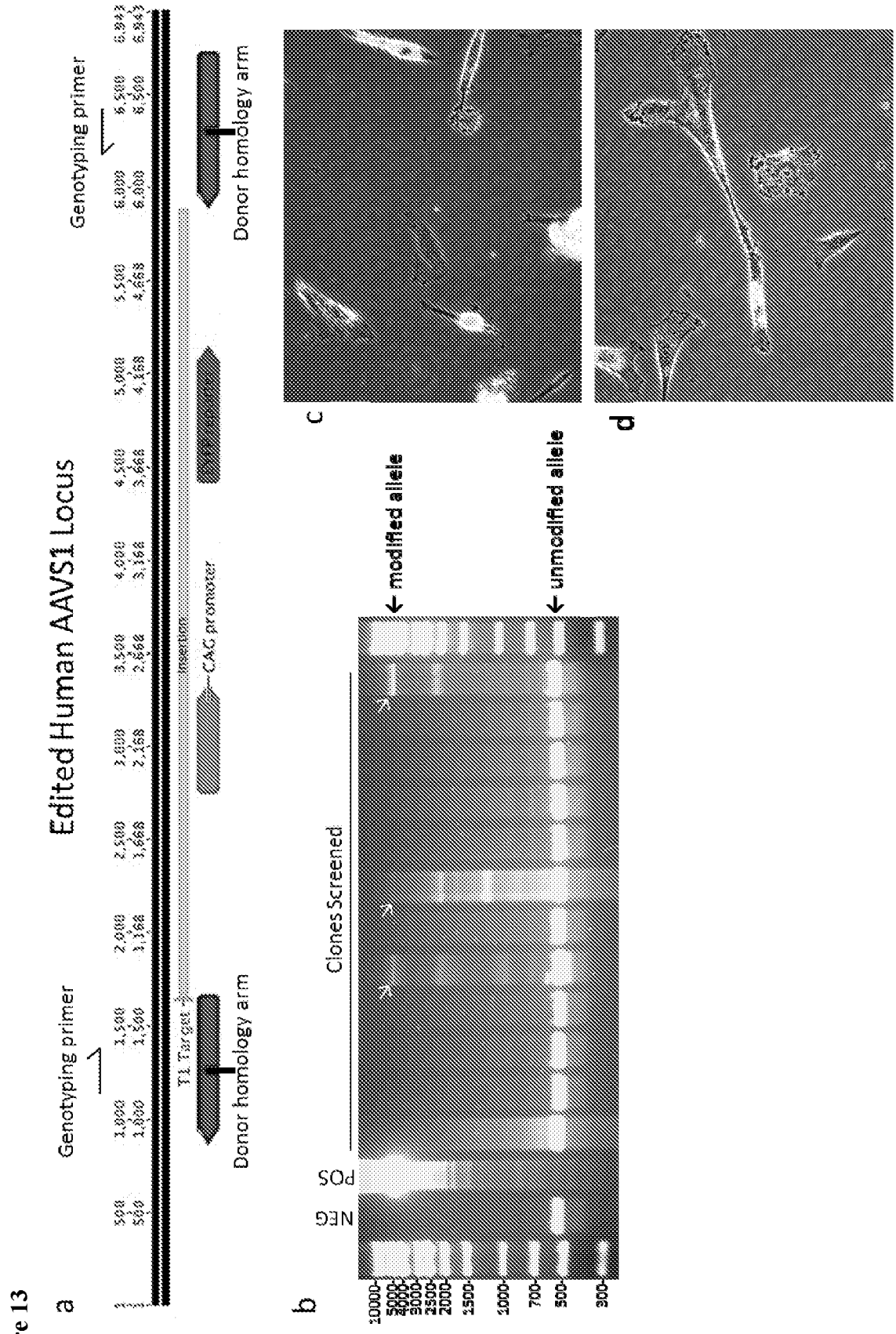

GAPDH
Relative Expression to Actin
2.0
1.5
1.0
0.5
0.0
Uninoculated
CRISPR-edited
Vectors
3T3
Ddx4
Relative Expression to Actin
2500000
2000000
1500000
100
80
60
40
20
0
Uninoculated
CRISPR-edited
Vectors
3T3
Blimp1
Relative Expression to Actin
150
100
50
0
Uninoculated
CRISPR-edited
Vectors
Embryo tissue
Fragilis
Relative Expression to Actin
4
3
2
1
0
Uninoculated
CRISPR-edited
Vectors
3T3
Tert
Relative Expression to Actin
150
100
50
0
Uninoculated
CRISPR-edited
Vectors
3T3

Figure 12

Figure 13


a
Edited Human AAVS1 Locus
Genotyping primer
Genotyping primer
T1 Target
Donor homology arm
Donor homology arm
CAG promoter
b
Clones Screened
NEG
POS
modified allele
unmodified allele
10000
5000
4000
3000
2500
2000
1500
1000
700
500
300
c
d

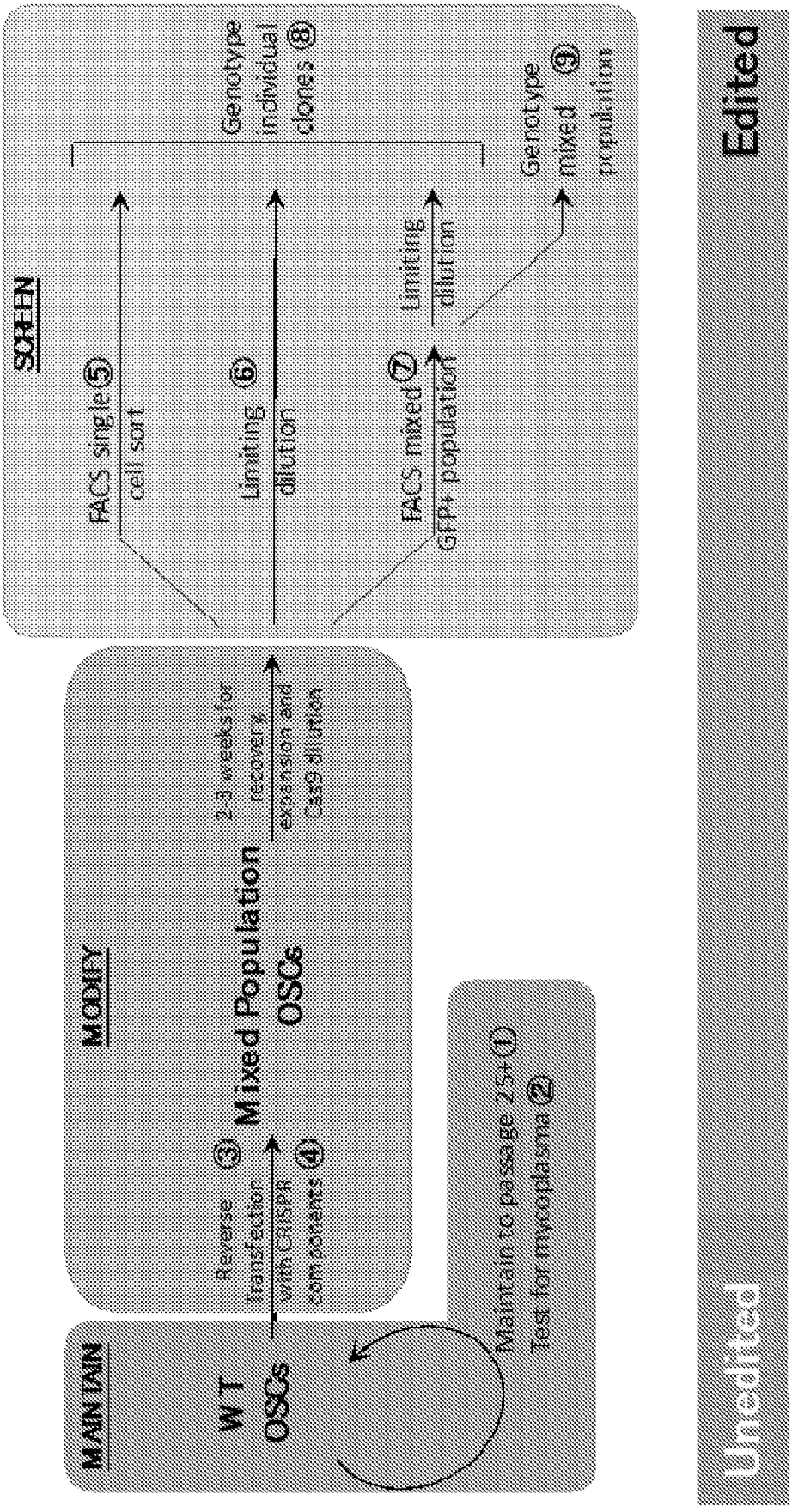

MAINTAIN
WT OSCs
Maintain to passage 25+ ①
Test for mycoplasma ②
MODIFY
Reverse Transfection ③ with CRISPR components ④
Mixed Population OSCs
2-3 weeks for recovery, expansion and Cas9 dilution
SCREEN
FACS single ⑤ cell sort
Limiting ⑥ dilution
FACS mixed ⑦ GFP+ population
Limiting dilution
Genotype individual clones ⑧
Genotype mixed ⑨ population
Unedited
Edited

Figure 14

b

a

Figure 15

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2016/060315

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/90 C12N5/075 C12N5/076 C12N9/22 C12N15/113
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N C07K

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 2015/054315 A1 (UNIV NORTHEASTERN [US]) 16 April 2015 (2015-04-16) claims 25-29, 32-37; example 1	1-4,6-8, 23,24 5,12,13, 22,29, 31,38

X	WO 2013/192316 A1 (RECOMBINETICS INC [US]) 27 December 2013 (2013-12-27) pages 27-29,39; claims 1-7,20,21,37-44; examples 19,20	1-3,6,7, 9,10, 14-18, 21, 23-27, 30,32, 34-37, 39-42

	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 March 2017	17/03/2017

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Deleu, Laurent

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2015/035034 A1 (MICE WITH HORNS LLC [US]) 12 March 2015 (2015-03-12)	1-4,6,7, 9-11,17, 18, 23-28, 32,33, 35,36, 39-42
Y	page 31; example 1	5,12,13, 29,31

Y	Susan M. Byrne ET AL: "Genome Editing in Human Stem Cells" In: "METHODS IN ENZYMOLOGY", 2014, ACADEMIC PRESS, US, XP055242837, ISSN: 0076-6879 vol. 546, pages 119-138, DOI: 10.1016/B978-0-12-801185-0.00006-4, page 10 - page 12	5,12,13, 29,31

Y	K. DUDA ET AL: "High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases or CRISPR/Cas9 nickase pairs", NUCLEIC ACIDS RESEARCH, vol. 42, no. 10, 21 April 2014 (2014-04-21), pages e84-e84, XP055352337, ISSN: 0305-1048, DOI: 10.1093/nar/gku251 page 11 - page 13	5,12,13, 29,31

Y	WANG SI ET AL: "Eliminate mitochondrial diseases by gene editing in germ-line cells and embryos", PROTEIN & CELL, SPRINGER ASIA, BEIJING, CN, vol. 6, no. 7, 17 June 2015 (2015-06-17), pages 472-475, XP035507312, ISSN: 1674-800X, DOI: 10.1007/S13238-015-0177-X [retrieved on 2015-06-17] page 472 - page 475	19,20

Y	WO 2013/002880 A1 (GEN HOSPITAL CORP [US]; HARVARD COLLEGE [US]; TILLY JONATHAN LEE [US];) 3 January 2013 (2013-01-03) claim 54	19,20

	-/--	

2

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	OZLEM BINGOL OZAKPINAR: "Ovarian stem cells: From basic to clinical applications", WORLD JOURNAL OF STEM CELLS, vol. 7, no. 4, 26 May 2015 (2015-05-26), page 757, XP055352649, CN ISSN: 1948-0210, DOI: 10.4252/wjsc.v7.i4.757 page 766	22,38
Y,P	CN 105 602 935 A (NIE LINGYUN) 25 May 2016 (2016-05-25) figures 2-4	19

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/060315

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

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

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

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

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

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

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

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

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

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2016/060315

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2015054315 A1	16-04-2015	CA 2932581 A1	16-04-2015
		CN 105916977 A	31-08-2016
		EP 3055408 A1	17-08-2016
		US 2016237402 A1	18-08-2016
		WO 2015054315 A1	16-04-2015
WO 2013192316 A1	27-12-2013	AR 091482 A1	04-02-2015
		AU 2013277214 A1	22-01-2015
		CA 2877297 A1	27-12-2013
		CN 105101787 A	25-11-2015
		EP 2863736 A1	29-04-2015
		JP 2015519923 A	16-07-2015
		KR 20150029715 A	18-03-2015
		RU 2015101740 A	20-08-2016
		WO 2013192316 A1	27-12-2013
WO 2015035034 A1	12-03-2015	US 2015064149 A1	05-03-2015
		WO 2015035034 A1	12-03-2015
WO 2013002880 A1	03-01-2013	AU 2012276038 A1	20-02-2014
		BR 112013033797 A2	14-02-2017
		CA 2847292 A1	03-01-2013
		CN 103827293 A	28-05-2014
		EA 201490050 A1	30-07-2014
		EP 2726601 A1	07-05-2014
		JP 2014520526 A	25-08-2014
		US 2013059384 A1	07-03-2013
		US 2016024527 A1	28-01-2016
		WO 2013002880 A1	03-01-2013
CN 105602935 A	25-05-2016	NONE	