

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

30 July 2020 (30.07.2020)



(10) International Publication Number

WO 2020/154632 A1

(51) International Patent Classification:

A01K 67/00 (2006.01) C12N 15/09 (2006.01)

A01K 67/027 (2006.01) C12N 15/85 (2006.01)

(21) International Application Number:

PCT/US2020/015008

(22) International Filing Date:

24 January 2020 (24.01.2020)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/796,437 24 January 2019 (24.01.2019) US

(71) Applicant: NORTH CAROLINA STATE UNIVERSITY [US/US]; 1021 Main Campus Drive, 2nd Floor, Raleigh, North Carolina 27606 (US).

(72) Inventor: PIEDRAHITA, Jorge A.; c/o North Carolina State University, 1021 Main Campus Drive, 2nd Floor, Raleigh, North Carolina 27606 (US).

(74) Agent: SCHLUETER, Peter J.; Casimir Jones, S.C., 2275 Deming Way, Suite 310, Middleton, Wisconsin 53562 (US).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, 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).

(54) Title: TRANSGENIC MODELS FOR STEM CELL THERAPIES

(57) Abstract: The present disclosure provides compositions and methods related to the generation and use of transgenic animal models having stem cell reporter systems. In particular, the present disclosure provides a novel transgenic animal model that expresses a nuclear-localized fluorescent reporter in cells endogenously expressing a leucine-rich repeat-containing G-protein coupled receptor (LGR) gene (e.g., LGR5 gene). Given the role of LGR genes in stem cell and cancer biology, the transgenic animal models provided herein are useful for a wide range of therapeutic and diagnostic purposes.



WO 2020/154632 A1

TRANSGENIC MODELS FOR STEM CELL THERAPIES

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to and the benefit of U.S. Provisional Patent Application No. 62/796,437 filed January 24, 2019, which is incorporated herein by reference in its entirety for all purposes.

GOVERNMENT FUNDING

[0002] This invention was made with government support under grant number OD019738 awarded by the National Institutes of Health. The government has certain rights in the invention.

FIELD

[0003] The present disclosure provides compositions and methods related to the generation and use of transgenic animal models having stem cell reporter systems. In particular, the present disclosure provides a novel transgenic animal model that expresses a nuclear-localized fluorescent reporter in cells endogenously expressing the leucine-rich repeat-containing G-protein coupled receptor 5 (LGR5) gene. Given the role of LGR5 in stem cell and cancer biology, the transgenic animal models provided herein are useful for a wide range of therapeutic and diagnostic purposes.

BACKGROUND

[0004] Leucine-rich repeat-containing G-protein coupled receptors (LGR) are expressed in a wide range of stem cells including, for example, skin, gut, lung, ovary, mammary gland, kidney, pancreas and liver. In particular, the LGR5 gene has been implicated in cancer and LGR5 expressing cells have been referred as the cancer stem cell. Being able to track and study LGR5 expressing cells will not only facilitate investigations into their normal development, but also help elucidate the function of LGRs during cell regeneration and repair, as well as their roles in cancer initiation and progression.

SUMMARY

[0005] Embodiments of the present disclosure include a non-human transgenic animal comprising a genome that expresses a nuclear-localized reporter gene in at least one cell type or tissue type that also expresses an endogenous leucine-rich repeat-containing G-protein coupled receptor (LGR) gene.

[0006] In some embodiments, the endogenous LRG gene is leucine-rich repeat-containing G-protein coupled receptor 4 (LGR4), LRG gene is leucine-rich repeat-containing G-protein coupled receptor 5 (LGR5), or LRG gene is leucine-rich repeat-containing G-protein coupled receptor 6 (LGR6).

[0007] In some embodiments, the nuclear-localized reporter gene is a fluorescent reporter. In some embodiments, the fluorescent reporter comprises at least one of GFP, eGFP, mCherry, CFP, BFP, YFP, eYFP, photoactivatable GFP, dsRed (*Discosoma* species fluorescent protein), mFruits (mutants of dsRed), TagRFPs (Evrogen), eqFP611 (isolated from sea anemone *Entacmaea quadricolor*), Dronpa (photoswitchable fluorescent protein), and EosFP (photoconvertable fluorescent protein).

[0008] In some embodiments, the nuclear-localized reporter gene comprises a gene encoding H2B. In some embodiments, the nuclear-localized reporter gene comprises a gene encoding an H2B-GFP fusion protein. In some embodiments, the nuclear-localized reporter gene comprises a gene encoding an LRG protein, or fragment thereof, fused to a nuclear-localized fluorescent reporter. In some embodiments, the nuclear-localized reporter gene comprises a gene encoding an LGR5-H2B-GFP fusion protein.

[0009] In some embodiments, the gene encoding the LGR5-H2B-GFP fusion protein comprises a gene encoding H2B-GFP downstream of the LGR5 ATG start site. In some embodiments, the gene encoding the LGR5-H2B-GFP fusion protein comprises a gene encoding H2B-GFP downstream of the LGR5 ATG start site, and upstream of intron 1 of LGR5. In some embodiments, the gene encoding the LGR5-H2B-GFP fusion protein does not contain an IRES site. In some embodiments, the gene encoding the LGR5-H2B-GFP fusion protein comprises one or more fragments of SEQ ID NO: 2.

[0010] In some embodiments, the at least one cell type or tissue type that expresses the endogenous LGR gene is a stem cell. In some embodiments, the at least one cell type or tissue type that expresses the endogenous LGR gene comprises skin, eye, inner ear, gastrointestinal track, uterus, ovary, prostate, mammary gland, kidney, liver, pancreas, cervix, and/or placenta cell types or tissue types.

[0011] Embodiments of the present disclosure also include a cell or cell line derived from the transgenic animal described above. In some embodiments, the cell line is a primary cell line or an immortalized cell line. In some embodiments, the cell is a stem cell. In some embodiments, the stem cell is capable of forming an organoid when cultured in a matrix.

[0012] Embodiments of the present disclosure also include a method of screening an intervention for a disease or condition. In accordance with these embodiments, the method

includes: a) contacting the transgenic animal of claim 1 or the cell or cell line of claim 15 with a candidate intervention; and b) determining the effect of said intervention on a disease or condition in the transgenic animal.

[0013] In some embodiments, the intervention is selected from the group consisting of a drug, a lifestyle change, an alternative medicine therapy, or a combination thereof. In some embodiments, the disease or condition is cancer. In some embodiments, the disease or condition is associated with stem cell function. In some embodiments, the disease or condition is an intestinal, hepatic, renal, lung, or skin disease or injury.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] FIG. 1: Representative schematic diagram of the LGR5 gene locus of a transgenic model of the present disclosure.

[0015] FIGS. 2A-2B: Representative diagram of the LGR5 genomic targeting strategy using CRISPR (FIG. 2A), and representative gels identifying two colonies as positive for the H2B-GFP knock-in (FIG. 2B).

[0016] FIGS. 3A-3B: Representative images of transgene expression in the duodenum (FIG. 3A) and the jejunum (FIG. 3B) of the transgenic lines of the present disclosure.

[0017] FIGS. 4A-4B: Representative images of transgene expression in the skin (FIG. 4A) and the biliary tree (FIG. 4B) of the transgenic lines of the present disclosure.

[0018] FIGS. 5A-5C: Representative images of the co-expression of LGR5-H2B-GFP transgene protein (FIG. 5A) and mRNA (FIG. 5B), merged in FIG. 5C.

[0019] FIGS. 6A-6C: Representative images of LGR5-H2B-GFP transgene expression in locations consistent with endogenous LGR5 expression in various stages of the hair growth cycle, including anagen (FIG. 6A), catagen (FIG. 6B), and telogen (FIG. 6C) stages.

[0020] FIGS. 7A-7C: Representative images of LGR5-H2B-GFP transgene expression in individual epidermal cells (FIGS. 7A-7B), which correlates with endogenous LGR5 mRNA expression (FIG. 7C).

[0021] FIGS. 8A-8D: Representative scatter plots from FACS experiments of GFP-positive stem cells expressing the LGR5-H2B-GFP transgene in liver (FIG. 8A), biliary tree (FIG. 8B), pancreas (FIG. 8C), and small intestine (FIG. 8D).

[0022] FIGS. 9A-9B: Representative bar graph results of organoid formation efficiency (FIG. 9A) and size (FIG. 9B) from LGR5-GFP-positive hair follicle stem cells.

[0023] FIGS. 10A-10B: Representative images demonstrate that LGR5-GFP-positive cells plated in Matrigel containing growth factors formed skin organoids (FIG. 10A) and biliary tree organoids (FIG. 10B).

[0024] FIG. 11: Representative images of LGR5-GFP (“hi”) organoids stained with various epidermal markers.

[0025] FIGS. 12A-12D: Representative images of low resolution cytoplasmic LGR5-GFP expression in a mouse model (FIGS. 12A-12B) as compared to high resolution nuclear LGR5-H2B-GFP expression in a porcine model (FIGS. 12C-12D); GFP expression is overlaid with DAPI staining in FIGS. 12B and 12D.

[0026] FIG. 13: Representative images of LGR5-GFP expression surrounding a wound, demonstrating increased cell proliferation.

[0027] FIGS. 14A-14C: Representative images of LGR5-GFP expression in hair germ cells, the primary proliferating hair germ cells in anagen (GFP expression is overlaid with Ki67 staining in FIG. 14B, and DAPI staining in FIG. 14C).

DETAILED DESCRIPTION

[0028] Embodiments of the present disclosure provide compositions and methods related to the generation and use of transgenic animal models having stem cell reporter systems. In particular, the present disclosure provides a novel transgenic animal model that expresses a nuclear-localized fluorescent reporter in cells endogenously expressing the leucine-rich repeat-containing G-protein coupled receptor 5 (LGR5) gene. Given the role of LGR5 in stem cell and cancer biology, the transgenic animal models provided herein are useful for a wide range of therapeutic and diagnostic purposes.

[0029] The use of a large animal model having anatomy and physiology more complementary to humans provides the ability to recapitulate human disease and overcomes many current limitations of rodent models. Thus, the transgenic porcine models provided herein are a powerful tool for the study of translational aspects of adult stem cells of the gut, skin, liver, and kidney, among other cells and tissues. These swine models can also be used to address questions regarding the role of stem cells in epithelial/mesenchymal replacement during homeostasis and intestinal, hepatic, renal, lung, and skin diseases/injuries, including cancer, and regenerative responses resulting from burns, wounds, radiation/chemotherapeutic or viral/bacterial damage.

[0030] The transgenic animal models described herein include a chromatin bound/nuclear-localized GFP fluorescent tag in the LGR5 locus that can facilitate the study of adult stem cells,

adult stem cell progenitors, and the stem-cell niche. LGR5 is a critical marker for multiple adult stem cells, including but not limited to, stem cells present in the skin, GI tract, liver and pancreas, as well as those present in lung regeneration, and mammary gland development and repair. Essentially almost all epithelial barriers involve an LGR5 progenitor cell as a repair stem cell. LGR5 has also been implicated in cancer, and it is upregulated in many cancers. LGR5-positive cells are often referred to as cancer stem cells, as they can be the initiator of many types of cancer. Thus, LGR5 is important from both a regenerative medicine standpoint, as well as oncology.

[0031] Previous investigations supported the use of IRES elements and cytoplasmic GFP in constructs used to generate LGR5 transgenic mice. Embodiments of the present disclosure used similar approaches, but these were only partially successful. Additionally, cytoplasmic GFP is very difficult to evaluate, both quantitatively and qualitatively, and it can be challenging to differentiate cytoplasmic GFP against a fluorescent background signal. Conversely, nuclear localization signals, such as H2B-GFP, are much easier to score and provide an enhanced degree of accuracy and spatial resolution. Also, the H2B nuclear localization tag (chromatin-specific) is particularly useful, as it allows for lineage tracing studies and investigations into chromosome dynamics by identifying the timing of when LGR5 gene expression is activated or deactivated.

[0032] More than ten transgenic mouse lines described in previous investigations using LGR transgenic models; however, none use H2B-GFP as a reporter and all include the use of an IRES. These existing mouse models have well characterized mosaic expression patterns, and cell ablation studies are limited due to the suboptimal penetrance of transgene expression in these models (see, *e.g.*, Stem Cell Reports; 3:234-241 (2014)).

[0033] Thus, in contrast to previous studies, embodiments of the present disclosure include at least two elements not previously identified as important for replicating LGR5 expression in a transgenic animal model—removing the IRES and switching to the chromatin bound GFP. Additionally, embodiments of the present disclosure include transgenic construct in which the transgene has been inserted into the 5' end of the transgene (*e.g.*, adjacent to the ATG start site). These features make the transgenic models of the present disclosure much easier to score and have led to the identification of unreported LGR5 cell populations in certain regions of GI tissues (see, *e.g.*, FIGS. 3 and 4), including cells that are not epithelial in nature, which may be more consistent with LGR expression in humans than other transgenic models.

[0034] As described further herein, embodiments of the present disclosure include a non-human transgenic animal comprising a genome that expresses a nuclear-localized reporter gene

in at least one cell type or tissue type that also expresses an endogenous leucine-rich repeat-containing G-protein coupled receptor (LGR) gene. In some embodiments, the endogenous LRG gene is LGR4, LGR5, or LRG6. LRG genes/proteins are a unique class of evolutionary conserved seven-transmembrane receptors characterized by a large extracellular region that includes multiple imperfect copies of a leucine-rich repeat protein interaction domain. The extracellular domain mediates ligand binding as a prerequisite to modulation of downstream intracellular signaling pathways via heterotrimeric G-proteins.

[0035] In some embodiments, the nuclear-localized reporter gene is a fluorescent reporter. In some embodiments, the fluorescent reporter comprises at least one of GFP, eGFP, mCherry, CFP, BFP, YFP, eYFP, photoactivatable GFP, dsRed (*Discosoma* species fluorescent protein), mFruits (mutants of dsRed), TagRFPs (Evrogen), eqFP611 (isolated from sea anemone *Entacmaea quadricolor*), Dronpa (photoswitchable fluorescent protein), and EosFP (photoconvertable fluorescent protein). Other reporters can also be used, as would be recognized by one of ordinary skill in the art based on the present disclosure.

[0036] In some embodiments, the nuclear-localized reporter gene comprises a gene encoding the histone H2B gene, a chromatin-specific nuclear-localization signal. In some embodiments, the nuclear-localized reporter gene comprises a gene encoding a nuclear-localization signal, such as but not limited to, a monopartite or bipartite nuclear localization signal of any other protein that is localized to the nucleus. Examples include the following: SV40 large T antigen, nucleoplasmin, Antennapedia Homeodomain Protein Antigens, Nuclear Ki-67 Antigen, Ku Autoantigen, snRNP Core Proteins, Ataxins, Ataxin-1, Ataxin-2, Ataxin-3, Ataxin-7, BRCA1 Protein, BRCA2 Protein, CCAAT-Enhancer-Binding Proteins, CCAAT-Binding Factor, CCAAT-Enhancer-Binding Protein-alpha, CCAAT-Enhancer-Binding Protein-beta, CCAAT, Enhancer-Binding Protein-delta, Transcription Factor CHOP, Y-Box-Binding Protein 1, CDX2 Transcription Factor, Chromosomal Proteins including histones, CCCTC-Binding Factor, Centromere Protein A, Centromere Protein B, High Mobility Group Proteins +, Methyl-CpG-Binding Protein 2, Minichromosome Maintenance Proteins +, Polycomb-Group Proteins +, SMARCB1 Protein, Telomere-Binding Proteins +, Tumor Suppressor p53-Binding Protein 1, Fanconi Anemia Complementation Group A Protein, Fanconi Anemia Complementation Group D2 Protein, Fanconi Anemia Complementation Group E Protein, Fanconi Anemia Complementation Group F Protein, Fanconi Anemia Complementation Group N Protein, Hepatocyte Nuclear Factors, Hepatocyte Nuclear Factor 1 +, Hepatocyte Nuclear Factor 3-alpha, Hepatocyte Nuclear Factor 3-beta, Hepatocyte Nuclear Factor 3-gamma, Hepatocyte Nuclear Factor 4, Hepatocyte Nuclear Factor 6, Host Cell Factor

C1, Immunoglobulin J Recombination Signal Sequence-Binding Protein, Inhibitor of Growth Protein 1, Karyopherins, alpha Karyopherins, beta Karyopherins, Mad2 Proteins, Mediator Complex, Cyclin C, Cyclin-Dependent Kinase 8, Mediator Complex Subunit 1, NF-kappa B, NF-kappa B p50 Subunit, NF-kappa B p52 Subunit, Transcription Factor RelA, Transcription Factor RelB, Nuclear Matrix-Associated Proteins, Heterogeneous-Nuclear Ribonucleoprotein U, Lamins +, Nuclear Receptor Coactivators, Mediator Complex Subunit 1, Nuclear Receptor Coactivator 1, Nuclear Receptor Coactivator 2, Nuclear Receptor Coactivator 3, Peroxisome Proliferator-Activated Receptor Gamma Coactivator 1-alpha, Nuclear Receptor Interacting Protein 1, Oncogene Protein p55(v-myc), Poly-ADP-Ribose Binding Proteins, Apoptosis Inducing Factor, Ataxia Telangiectasia Mutated Proteins, CCCTC-Binding Factor, Centromere Protein A, Centromere Protein B, Cyclin-Dependent Kinase Inhibitor p21, DNA-Activated Protein Kinase, DNA (Cytosine-5-)-Methyltransferase 1, DNA Topoisomerases, Type I, Heat Shock Transcription Factors, Histones, Ku Autoantigen, MRE11 Homologue Protein, Myristoylated Alanine-Rich C Kinase Substrate, NF-kappa B p52 Subunit, Nicotinamide-Nucleotide Adenylyltransferase, Nitric Oxide Synthase Type II, Telomerase, Tumor Suppressor Protein p53, Werner Syndrome Helicase, X-ray Repair Cross Complementing Protein 1, Xeroderma Pigmentosum Group A Protein, Proliferating Cell Nuclear Antigen, Promyelocytic Leukemia Protein, Protamines, Clupeine, Salmine, Proto-Oncogene Proteins c-fos, Proto-Oncogene Proteins c-jun, Proto-Oncogene Proteins c-mdm2, Proto-Oncogene Proteins c-myc, Proto-Oncogene Proteins c-rel, ran GTP-Binding Protein, Retinoblastoma Binding Proteins, E2F1 Transcription Factor, Retinoblastoma-Binding Protein 1, Retinoblastoma-Binding Protein 2, Retinoblastoma-Binding Protein 4, Retinoblastoma-Binding Protein 7, Retinoblastoma-Like Protein p107, Retinoblastoma-Like Protein p130, Retinoblastoma Protein, Silent Information Regulator Proteins, *Saccharomyces cerevisiae*, Thyroid Nuclear Factor 1, and Tumor Protein p73.

[0037] In some embodiments, the nuclear-localized reporter gene comprises a gene encoding H2B fused to at least one fluorescent reporter gene. In some embodiments, the nuclear-localized reporter gene comprises a gene encoding an LRG protein, or fragment thereof, fused to a nuclear-localized fluorescent reporter. In some embodiments, the nuclear-localized reporter gene comprises a gene encoding an LGR5-H2B-GFP fusion protein, such that GFP is present and detectable in cells expressing endogenous LGR5. In some embodiments, the nuclear-localized reporter gene comprises a gene encoding an LRG4-H2B-GFP fusion protein, such that GFP is present and detectable in cells expressing endogenous LRG4. In some embodiments, the nuclear-localized reporter gene comprises a gene encoding

an LRG6-H2B-GFP fusion protein, such that GFP is present and detectable in cells expressing endogenous LRG6. In some embodiments, an LRG promoter (*e.g.*, LRG4 promoter, LGR5 promoter, or LRG6 promoter), or a fragment thereof, is present and drives expression of the transgene, independent of the presence or absence of one or more portions of the LRG gene itself (*e.g.*, introns, exons, or fragments thereof). In other embodiments, the LRG promoter, or a fragment thereof, is present and drives expression of the transgene, and expression of the transgene may involve one or more portions of the LRG gene itself (*e.g.*, introns, exons, or fragments thereof), which may act as enhancer elements (*e.g.*, an LRG gene, or a portion thereof, fused to a nuclear-localized fluorescent reporter gene, or a portion thereof).

[0038] In some embodiments, a gene encoding an LGR5-H2B-GFP fusion protein comprises a gene encoding H2B-GFP downstream of the LGR5 ATG start site. In some embodiments, the gene encoding the LGR5-H2B-GFP fusion protein comprises a gene encoding H2B-GFP downstream of the LGR5 ATG start site, and upstream of intron 1 of LGR5 (*e.g.*, within exon 1). In some embodiments, the gene encoding the LGR5-H2B-GFP fusion protein does not contain an IRES site. In some embodiments, the gene encoding the LGR5-H2B-GFP fusion protein comprises one or more fragments of SEQ ID NO: 2, such as one or more fragments of the sequence surrounding the genomic insertion site.

[0039] In some embodiments, the transgene used to generate the animal models of the present disclosure includes one or more fragments of SEQ ID NO: 2 (*e.g.*, one or more sequences surrounding the genomic insertion site) or sequences with at least 80% homology to these fragments of SEQ ID NO: 2. In some embodiments, the transgene used to generate the animal models of the present disclosure includes one or more fragments of SEQ ID NO: 2 (*e.g.*, one or more sequences surrounding the genomic insertion site) or sequences that are at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to one or more fragments of SEQ ID NO: 2. In some embodiments, variants of the transgene include one or more portions of an LRG coding region, and/or one or more portions of an LRG non-coding region (*e.g.*, comprising gene expression regulatory elements). In some embodiments, variants of the transgene include neither a portion of an LRG coding region, nor a portion of an LRG non-coding region.

[0040] In some embodiments, conservative or non-conservative substitutions are made. For example, it is contemplated that isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (*e.g.*, conservative mutations) will not have a major effect on the biological activity of the resulting molecule. Conservative replacements are those

that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic (aspartate, glutamate); (2) basic (lysine, arginine, histidine); (3) nonpolar (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan); and (4) uncharged polar (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine). Phenylalanine, tryptophan, histidine, and tyrosine are sometimes classified jointly as aromatic amino acids. In similar fashion, the amino acid repertoire can be grouped as (1) acidic (aspartate, glutamate); (2) basic (lysine, arginine, histidine), (3) aliphatic (glycine, alanine, valine, leucine, isoleucine, serine, threonine), with serine and threonine optionally be grouped separately as aliphatic hydroxyl; (4) aromatic (phenylalanine, tyrosine, tryptophan); (5) amide (asparagine, glutamine); and (6) sulfur containing (cysteine and methionine) (*e.g.*, Stryer ed., Biochemistry, pg. 17-21, 2nd ed, WH Freeman and Co., 1981). “Non-conservative” changes (*e.g.*, replacement of a glycine with a tryptophan) are also contemplated. Analogous minor variations can also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological activity can be found using computer programs (*e.g.*, LASERGENE software, DNASTAR Inc., Madison, WI.).

[0041] In some embodiments, at least one cell type or tissue type that expresses the endogenous LGR gene is a stem cell. In some embodiments, the at least one cell type or tissue type that expresses the endogenous LGR gene comprises skin, eye, inner ear, gastrointestinal track, uterus, ovary, prostate, mammary gland, kidney, liver, pancreas, cervix, and/or placenta cell types or tissue types. In some embodiments, one or more transgene-expressing cells of these tissue types can be isolated, and a cell line can be established. In some embodiments, the cell line is a primary cell line or an immortalized cell line.

[0042] Embodiments of the present disclosure also include a method of screening an intervention for a disease or condition. In accordance with these embodiments, the method can include contacting a transgenic animal described herein, or a cell or cell line of derived from a transgenic model described herein, with a candidate intervention. The method can include determining an effect of the intervention on a disease or condition in the transgenic animal (or cell derived therefrom). In some embodiments, the disease or condition is imparted to the transgenic animal via genetic or chemical means.

[0043] In some embodiments, the intervention is selected from the group consisting of a drug, a lifestyle change, an alternative medicine therapy, or a combination thereof. In some embodiments, the disease or condition is cancer. In some embodiments, the disease or

condition is associated with stem cell function. In some embodiments, the disease or condition is an intestinal, hepatic, renal, lung, or skin disease or injury.

[0044] Section headings as used in this section and the entire disclosure herein are merely for organizational purposes and are not intended to be limiting.

1. Definitions

[0045] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art. In case of conflict, the present document, including definitions, will control. Preferred methods and materials are described below, although methods and materials similar or equivalent to those described herein can be used in practice or testing of the present disclosure. All publications, patent applications, patents and other references mentioned herein are incorporated by reference in their entirety. The materials, methods, and examples disclosed herein are illustrative only and not intended to be limiting.

[0046] The terms “comprise(s),” “include(s),” “having,” “has,” “can,” “contain(s),” and variants thereof, as used herein, are intended to be open-ended transitional phrases, terms, or words that do not preclude the possibility of additional acts or structures. The singular forms “a,” “and” and “the” include plural references unless the context clearly dictates otherwise. The present disclosure also contemplates other embodiments “comprising,” “consisting of” and “consisting essentially of,” the embodiments or elements presented herein, whether explicitly set forth or not.

[0047] For the recitation of numeric ranges herein, each intervening number there between with the same degree of precision is explicitly contemplated. For example, for the range of 6-9, the numbers 7 and 8 are contemplated in addition to 6 and 9, and for the range 6.0-7.0, the number 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, and 7.0 are explicitly contemplated.

[0048] “Correlated to” as used herein refers to compared to.

[0049] As used herein, the term “animal” refers to any animal (*e.g.*, a mammal), including, but not limited to, humans, non-human primates, pigs, rodents (*e.g.*, mice, rats, etc.), flies, and the like.

[0050] As used herein, the term “non-human animals” refers to all non-human animals including, but are not limited to, vertebrates such as rodents, non-human primates, ovines, bovines, ruminants, lagomorphs, porcines, caprines, equines, canines, felines, aves, etc.

[0051] The term “transgene” as used herein refers to a foreign, heterologous, or autologous gene and/or fragment thereof that is placed into an organism (*e.g.*, by introducing the gene into

newly fertilized eggs or early embryos). The term “foreign gene” refers to any nucleic acid (*e.g.*, gene sequence) that is introduced into the genome of an animal by experimental manipulations and may include gene sequences found in that animal so long as the introduced gene does not reside in the same location as does the naturally-occurring gene.

[0052] As used herein, the term “transgenic animal” refers to any animal containing a transgene.

[0053] As used herein, the term “gene transfer system” refers to any means of delivering a composition comprising a nucleic acid sequence to a cell or tissue. For example, gene transfer systems include, but are not limited to, vectors (*e.g.*, retroviral, adenoviral, adeno-associated viral, and other nucleic acid-based delivery systems), microinjection of naked nucleic acid, polymer-based delivery systems (*e.g.*, liposome-based and metallic particle-based systems), biolistic injection, and the like. As used herein, the term “viral gene transfer system” refers to gene transfer systems comprising viral elements (*e.g.*, intact viruses, modified viruses and viral components such as nucleic acids or proteins) to facilitate delivery of the sample to a desired cell or tissue. As used herein, the term “adenovirus gene transfer system” refers to gene transfer systems comprising intact or altered viruses belonging to the family Adenoviridae.

[0054] As used herein, the term “site-specific recombination target sequences” refers to nucleic acid sequences that provide recognition sequences for recombination factors and the location where recombination takes place.

[0055] As used herein, the term “nucleic acid molecule” refers to any nucleic acid containing molecule, including but not limited to, DNA or RNA. The term encompasses sequences that include any of the known base analogs of DNA and RNA including, but not limited to, 4-acetylcytosine, 8-hydroxy-N6-methyladenosine, aziridinylcytosine, pseudoisocytosine, 5-(carboxyhydroxymethyl) uracil, 5-fluorouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxymethylaminomethyluracil, dihydrouracil, inosine, N6-isopentenyladenine, 1-methyladenine, 1-methylpseudouracil, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-methyladenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarbonylmethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, oxybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, N-uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, pseudouracil, queosine, 2-thiocytosine, and 2,6-diaminopurine.

[0056] The term “gene” refers to a nucleic acid (*e.g.*, DNA) sequence that comprises coding sequences for the production of a polypeptide, precursor, or RNA (*e.g.*, rRNA, tRNA). The polypeptide can be encoded by a full-length coding sequence or by any portion of the coding sequence so long as the desired activity or functional properties (*e.g.*, enzymatic activity, ligand binding, signal transduction, immunogenicity, etc.) of the full-length or fragment are retained. The term also encompasses the coding region of a structural gene and the sequences located adjacent to the coding region on both the 5' and 3' ends for a distance of about 1 kb or more on either end such that the gene corresponds to the length of the full-length mRNA. Sequences located 5' of the coding region and present on the mRNA are referred to as 5' non-translated sequences. Sequences located 3' or downstream of the coding region and present on the mRNA are referred to as 3' non-translated sequences. The term “gene” encompasses both cDNA and genomic forms of a gene. A genomic form or clone of a gene contains the coding region interrupted with non-coding sequences termed “introns” or “intervening regions” or “intervening sequences.” Introns are segments of a gene that are transcribed into nuclear RNA (hnRNA); introns may contain regulatory elements such as enhancers. Introns are removed or “spliced out” from the nuclear or primary transcript; introns therefore are absent in the messenger RNA (mRNA) transcript. The mRNA functions during translation to specify the sequence or order of amino acids in a nascent polypeptide.

[0057] As used herein, the term “heterologous gene” refers to a gene that is not in its natural environment. For example, a heterologous gene includes a gene from one species introduced into another species. A heterologous gene also includes a gene native to an organism that has been altered in some way (*e.g.*, mutated, added in multiple copies, linked to non-native regulatory sequences, etc). Heterologous genes are distinguished from endogenous genes in that the heterologous gene sequences are typically joined to DNA sequences that are not found naturally associated with the gene sequences in the chromosome or are associated with portions of the chromosome not found in nature (*e.g.*, genes expressed in loci where the gene is not normally expressed).

[0058] As used herein, the term “oligonucleotide,” refers to a short length of single-stranded polynucleotide chain. Oligonucleotides are typically less than 200 residues long (*e.g.*, between 15 and 100), however, as used herein, the term is also intended to encompass longer polynucleotide chains. Oligonucleotides are often referred to by their length. For example, a 24-residue oligonucleotide is referred to as a “24-mer.” Oligonucleotides can form secondary and tertiary structures by self-hybridizing or by hybridizing to other polynucleotides. Such

structures can include, but are not limited to, duplexes, hairpins, cruciforms, bends, and triplexes.

[0059] The term “isolated” when used in relation to a nucleic acid, as in “an isolated oligonucleotide” or “isolated polynucleotide” refers to a nucleic acid sequence that is identified and separated from at least one component or contaminant with which it is ordinarily associated in its natural source. Isolated nucleic acid is such present in a form or setting that is different from that in which it is found in nature. In contrast, non-isolated nucleic acids as nucleic acids such as DNA and RNA found in the state they exist in nature. For example, a given DNA sequence (*e.g.*, a gene) is found on the host cell chromosome in proximity to neighboring genes; RNA sequences, such as a specific mRNA sequence encoding a specific protein, are found in the cell as a mixture with numerous other mRNAs that encode a multitude of proteins. However, isolated nucleic acid encoding a given protein includes, by way of example, such nucleic acid in cells ordinarily expressing the given protein where the nucleic acid is in a chromosomal location different from that of natural cells, or is otherwise flanked by a different nucleic acid sequence than that found in nature. The isolated nucleic acid, oligonucleotide, or polynucleotide may be present in single-stranded or double-stranded form. When an isolated nucleic acid, oligonucleotide or polynucleotide is to be utilized to express a protein, the oligonucleotide or polynucleotide will contain at a minimum the sense or coding strand (*i.e.*, the oligonucleotide or polynucleotide may be single-stranded), but may contain both the sense and anti-sense strands (*i.e.*, the oligonucleotide or polynucleotide may be double-stranded).

[0060] As used herein, the term “purified” or “to purify” refers to the removal of components (*e.g.*, contaminants) from a sample. For example, antibodies are purified by removal of contaminating non-immunoglobulin proteins; they are also purified by the removal of immunoglobulin that does not bind to the target molecule. The removal of non-immunoglobulin proteins and/or the removal of immunoglobulins that do not bind to the target molecule results in an increase in the percent of target-reactive immunoglobulins in the sample. In another example, recombinant polypeptides are expressed in bacterial host cells and the polypeptides are purified by the removal of host cell proteins; the percent of recombinant polypeptides is thereby increased in the sample.

[0061] Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. For example, any nomenclatures used in connection with, and techniques of, cell and tissue culture, molecular biology, immunology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are

those that are well known and commonly used in the art. The meaning and scope of the terms should be clear; in the event, however of any latent ambiguity, definitions provided herein take precedent over any dictionary or extrinsic definition. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular.

2. Transgenic Reporter Systems and Methods of Use

[0062] Embodiments of the present disclosure include the use of the transgenic reporter systems provided herein, including the LGR5-H2B-GFP reporter system. In accordance with these embodiments, the methods include using the transgenic models of the present disclosure in the context of a disease or condition, such as a disease related to stem cell biology and/or cancer. For example, the transgenic models of the present disclosure can be used to investigate skin regeneration and/or wound healing (*e.g.*, as a result of a burn). In some embodiments, the LGR5-H2B-GFP reporter system can be used to track the progress of LGR5-expressing stem cells after the transgenic model (or cells isolated from the transgenic model) is subjected to an intervention involving damage to the skin. Similarly, the transgenic model can be used to investigate the efficacy of various therapies designed to reduce or treat the skin damage. For example, the H2B nuclear localization tag (chromatin-specific) is particularly useful, as it allows for lineage tracing studies and investigations into chromosome dynamics by identifying the timing of when LGR5 gene expression is activated or deactivated.

[0063] In some embodiments, the transgenic models of the present disclosure can be used to investigate cancer progression and cancer therapy, such as, but not limited to, ovarian cancer, colorectal cancer, hepatocellular carcinoma, basal cell carcinoma, breast cancer, and esophageal adenocarcinoma, among others. In some embodiments, the LGR5-H2B-GFP reporter system can be used to track the progress of LGR5-expressing stem cells after the transgenic model (or cells isolated from the transgenic model) is subjected to an intervention that causes cancer in the model (*e.g.*, genetic manipulation). Similarly, the transgenic model can be used to investigate the efficacy of various therapies designed to reduce or treat the cancer. As would be recognized by one of ordinary skill in the art based on the present disclosure, other diseases or conditions involving LGRs can be investigated using the transgenic models of the present disclosure.

[0064] In some embodiments, adeno-associated viruses designed to mutate the BRCA-1 and/or BRCA-2 gene and/or p53 gene can be injected into the mammary glands of transgenic animals. In addition, chemical induction methods can also be tested. If LGR5 is indeed a cancer stem cell, the tumors would be positive for transgene expression (*e.g.*, eGFP) and may or may

not be clonal. If tumors are induced, the animals would be used for studying tumor progression and identifying early stage biomarkers; studying tumor progression, biology of tumor growth, and transcriptomics of tumor to understand molecular mechanisms; and understanding the role of mesenchymal LGR5 cells in tumor progression in a tissue-specific manner.

[0065] In some embodiments, one or more transgene positive cells from the transgenic models provided herein can be isolated and cultured, either as primary cells or as immortalized cells. This facilitates *in vitro* experiments that are more difficult to conduct *in vivo*, such as experiments to identify modulators of a disease (*e.g.*, proteins, peptides, peptidomimetics, peptoids, small molecules or other drugs). In some embodiments, the transgenic models provided herein can be used to identify and isolate LGR5 stem cells from various tissues, for example, based on GFP expression from an LGR5 transgene (see Example 3). In some embodiments, these isolated stem cells can be cultured in a suitable matrix and under suitable conditions to form organoids, which can provide a 3D model that more accurately mimics tissue architecture than 2D cell culture.

[0066] The transgenic models of the present disclosure can be generated via a variety of methods. In some embodiments, CRISPR/Cas 9 systems are used to generate transgenic animals (see *e.g.*, Zhang F, Wen Y, Guo X (2014) *Human Molecular Genetics*. 23 (R1): R40–6). In some embodiments, embryonal cells at various developmental stages are used to introduce transgenes for the production of transgenic animals. Different methods are used depending on the stage of development of the embryonal cell. The zygote is the best target for micro-injection. The use of zygotes as a target for gene transfer has a major advantage in that in most cases the injected DNA will be incorporated into the host genome before the first cleavage (Brinster *et al.*, *Proc. Natl. Acad. Sci. USA* 82:4438-4442 (1985)). As a consequence, all cells of the transgenic non-human animal will carry the incorporated transgene. This will in general also be reflected in the efficient transmission of the transgene to offspring of the founder since 50% of the germ cells will harbor the transgene. Additional methods for generating transgenic animal are described, for example, in Palmiter, *Ann.Rev.Genet.* 20:465-99 (1986); Gordon, *Methods in enzymology*, vol.225; and Camper, *Biotechniques*. Vol 5. No 7. (1987).

[0067] In other embodiments, retroviral infection can be used to introduce transgenes into a non-human animal. In some embodiments, the retroviral vector is utilized to transfect oocytes by injecting the retroviral vector into the perivitelline space of the oocyte. In other embodiments, the developing non-human embryo can be cultured *in vitro* to the blastocyst stage. During this time, the blastomeres can be targets for retroviral infection (Janenich, *Proc.*

Natl. Acad. Sci. USA 73:1260 (1976)). Efficient infection of the blastomeres is obtained by enzymatic treatment to remove the zona pellucida (Hogan *et al.*, in *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1986)). The viral vector system used to introduce the transgene is typically a replication-defective retrovirus carrying the transgene (Jahner *et al.*, Proc. Natl. Acad. Sci. USA 82:6927 (1985)). Transfection is easily and efficiently obtained by culturing the blastomeres on a monolayer of virus-producing cells (Stewart, *et al.*, EMBO J., 6:383 (1987)). Additional means of using retroviruses or retroviral vectors to create transgenic animals known to the art involve the micro-injection of retroviral particles or mitomycin C-treated cells producing retrovirus into the perivitelline space of fertilized eggs or early embryos (PCT International Application WO 90/08832 (1990), and Haskell and Bowen, Mol. Reprod. Dev., 40:386 (1995)).

[0068] In other embodiments, the transgene is introduced into embryonic stem cells and the transfected stem cells are utilized to form an embryo. ES cells are obtained by culturing pre-implantation embryos *in vitro* under appropriate conditions (Evans *et al.*, Nature 292:154 (1981); Bradley *et al.*, Nature 309:255 (1984); Gossler *et al.*, Proc. Acad. Sci. USA 83:9065 (1986); and Robertson *et al.*, Nature 322:445 (1986)). Transgenes can be efficiently introduced into the ES cells by DNA transfection by a variety of methods known to the art including calcium phosphate co-precipitation, protoplast or spheroplast fusion, lipofection and DEAE-dextran-mediated transfection. Transgenes may also be introduced into ES cells by retrovirus-mediated transduction or by micro-injection. Such transfected ES cells can thereafter colonize an embryo following their introduction into the blastocoel of a blastocyst-stage embryo and contribute to the germ line of the resulting chimeric animal (for review, See, Jaenisch, Science 240:1468 (1988)). Prior to the introduction of transfected ES cells into the blastocoel, the transfected ES cells may be subjected to various selection protocols to enrich for ES cells which have integrated the transgene assuming that the transgene provides a means for such selection. Alternatively, the polymerase chain reaction may be used to screen for ES cells that have integrated the transgene. This technique obviates the need for growth of the transfected ES cells under appropriate selective conditions prior to transfer into the blastocoel.

[0069] In still other embodiments, homologous recombination is utilized to knock-out gene function or create deletion mutants (*e.g.*, truncation mutants). Methods for homologous recombination are described in, for example, U.S. Pat. No. 5,614,396.

3. Examples

[0070] It will be readily apparent to those skilled in the art that other suitable modifications and adaptations of the methods of the present disclosure described herein are readily applicable and appreciable, and may be made using suitable equivalents without departing from the scope of the present disclosure or the aspects and embodiments disclosed herein. Having now described the present disclosure in detail, the same will be more clearly understood by reference to the following examples, which are merely intended only to illustrate some aspects and embodiments of the disclosure, and should not be viewed as limiting to the scope of the disclosure. The disclosures of all journal references, U.S. patents, and publications referred to herein are hereby incorporated by reference in their entireties.

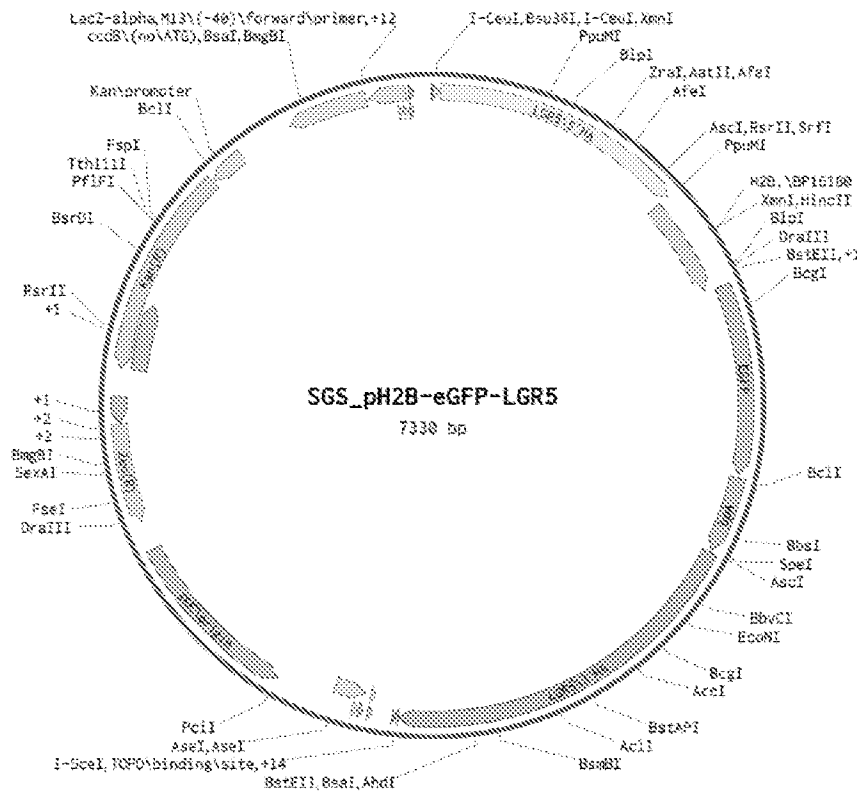
[0071] The present disclosure has multiple aspects, illustrated by the following non-limiting examples.

Example 1

[0072] **Generation of LGR5-H2B-GFP transgenic model.** As shown in FIG. 1, embodiments of the present disclosure include an H2B-GFP transgene inserted into a region of the endogenous LGR5 gene locus. In one embodiment, an H2B-GFP fusion construct was inserted into exon 1 of LGR5, downstream of the ATG start site and upstream of intron 1, and expression of the transgene is driven by the endogenous LGR5 promoter.

[0073] Shown below is a representative schematic diagram of a construct used to make the transgenic models, according to one embodiment of the present disclosure, including the corresponding insertion sites of the transgene.

SGS_pH2B-eGFP-LGR5
7330 bp



[0075] Previous attempts to generate an LGR transgenic reporter model used a Bacterial Artificial Chromosome containing all the LGR5 regulatory regions and a fluorescent tag inserted into a region of the gene. The BAC was introduced into cell by nucleofection and the resulting cells used for SCNT. Transgenic pigs containing the BAC were analyzed and none

expressed the reporter as expected. Further analysis indicated that the BAC had rearranged making the reporter inactive. Additionally, an approach was used that was previously used in mice (see, *e.g.*, Morita et al., Mol Cell Bio; 24:9736-9743 (2004)), which introduced an internal ribosomal entry site (IRES) linked to a cytoplasmic GFP and inserted this reporter into the 3' UTR after Exon 18. This reporter was introduced into the pig LGR5 locus using TALEN's induced homology directed repair and generated multiple offspring. Progeny was extensively analyzed, and it was determined that the reporter was active but was of little utility as it only mirrored the endogenous LGR5 pattern in some tissues.

[0076] Consequently, the H2B-GFP was knocked-in and simultaneously the opposite allele was corrected/protected using CRISPR-CAS HDR and a single strand DNA carrying the wild type LGR5 sequence. Essentially, this system was designed to turn the INDEL in the opposite allele into a wild type allele (FIG. 2A). Additionally, a screen identified two colonies containing the knock in (left panel) and one of them also carrying the corrected allele (lane 6), demonstrating the efficacy of this approach (FIG. 2B). This has been confirmed using sequencing and by generating additional colonies.

[0077] Therefore, as provided herein, the transgenic models of the present disclosure include several alterations from past models and from currently existing models, including, but not limited to: removing the IRES that contributed to gene silencing in some tissues; moving the reporter to the 5' end to create a LGR5-H2B-GFP fusion protein; and moving from the cytoplasmic to the porcine H2B-GFP chromatin/nuclear label.

Example 2

[0078] **Gene expression patterns in LGR5-H2B-GFP transgenic model.** As shown in FIGS. 3 and 4, transgene expression in these lines consistent with endogenous LGR5 expression, contained bright and easily scoreable nuclear-localized GFP expression in LGR5-expressing cells, including, for example, the duodenum, the jejunum, the skin and biliary tree. Expression in other cells and tissue types were also observed (data not shown).

[0079] LGR5-expressing stem cells play key roles in many organs; thus, each organ can be a potential clinical target. LGR5 expression has been reported in skin, eye, inner ear, gastrointestinal track, uterus, ovary, prostate, mammary gland, kidney, liver, pancreas, cervix, and/or placenta cell types or tissue types. And in each of these tissues there are associated disorders, some of which may be dependent on LGR5 cells (*e.g.*, cancer). In addition, LGR5 cells are activated upon injury so the transgenic models provided herein can be used to elucidate

how each tissue deals with injury, which will help determine the role of LGR5 cells and to find new targets of therapeutic development.

[0080] Finally, transgenic pig models have unique characteristics that make them more amenable to developing therapeutics that can be successful in a human population. From a regulatory FDA perspective, it may be possible to do both safety and efficacy testing in these animals and reduce the burden needed to initiate human clinical trials.

Example 3

[0081] As shown in FIGS. 5A-5C, experiments were conducted to investigate whether expression of endogenous LGR5 mRNA co-localizes with LGR5-H2B-GFP transgene protein expression. Results clearly demonstrate that endogenous LGR5 RNA transcripts (FIG. 5B) were co-expressed in the same location as H2B-GFP (FIG. 5A). RNA Scope for a pig-specific LGR5 probe shows that LGR5 RNA transcript can be detected in the same location that LGR5 driven H2B-GFP is expressed in the outer-root sheath of the hair follicle in the skin. (FIG. 5A, anti-GFP antibody; FIG. 5B, RNA-Scope (*in situ* hybridization) probe specific to porcine LGR5; FIG. 5C, merged images).

[0082] Experiments were also conducted to investigate whether endogenous LGR5-H2B-GFP transgene protein expression is consistent with endogenous LGR5 protein expression. As shown in FIGS. 6A-6C, results clearly demonstrate that LGR5-H2B-GFP transgene is expressed in locations consistent with endogenous LGR5 expression in various stages of the hair growth cycle, including during anagen (FIG. 6A), catagen (FIG. 6B), and telogen (FIG. 6C) stages. FIGS. 6A-6C include representative 20x confocal microscopy images of dorsal skin from a one-week old transgenic piglet. Green: nuclear H2B-GFP, blue: DAPI. LGR5 expressing cells (nuclear green) are found within the hair follicle, but not in any other location throughout the epidermis. This model faithfully recapitulates the expected position of LGR5 stem cells within the stages of the hair cycle.

[0083] As shown in FIGS. 7A-7C, LGR5-H2B-GFP transgene expression in individual epidermal cells also correlates with endogenous LGR5 mRNA expression. Single cell epidermis was isolated from LGR5-H2B-GFP samples and sorted based on GFP expression. Each sample was analyzed for levels of endogenous LGR5 mRNA expression by RT-qPCR. Values are normalized to GAPDH and RNA extracted from all components of skin in a delta-delta CT analysis. Additionally, FIGS. 8A-8D includes representative scatter plots from FACS sorting of GFP-positive stem cells expressing the LGR5-H2B-GFP transgene in liver (FIG. 8A), biliary tree (FIG. 8B), pancreas (FIG. 8C), and small intestine (FIG. 8D), which

demonstrates that LGR5-positive stem cells can be obtained from various tissues of the transgenic animals described further herein.

[0084] Because the LGR5 transgenic animals of the present disclosure facilitate the isolation of LGR5 stem cells, experiments were conducted to test the ability of these stem cells to form organoids. As shown in FIGS. 9A-9B, representative bar graph results demonstrate the ability to form organoids with efficiency (FIG. 9A) and proper size (FIG. 9B) from LGR5-GFP-positive hair follicle stem cells. Cells were isolated from LGR5-H2B-GFP transgene containing porcine epidermis and single cell sorted based on fluorescence. Cells were plated in organoid conditions in Matrigel containing growth factors including Wnt and R-spondin. After passaging, organoids derived from LGR5-H2BGFP-expressing epidermal cells were able to form organoids. Additionally, as shown in FIGS. 10A-10B, LGR5-GFP-positive cells plated in Matrigel containing growth factors formed skin organoids (FIG. 10A) and biliary tree organoids (FIG. 10B). Natural nuclear GFP expression in a fluorescent microscope is shown in the left image, brightfield images are in the center, and merged images are on the right. Images were obtained after 12 days in culture and display markers of proliferation, epidermis, and stem cells (data not shown but can be available upon request). Additionally, as shown in FIG. 11, day 42 organoids were characterized using various epidermal markers. Each panel shows a cryosection of an organoid stained with DAPI, the indicated marker (ACTIN, Loricrin, KI67cd200 k85, and krt14), or merged. These results demonstrate that organoids form cellular layers with positive marker expression in the epidermis toward the exterior of the organoid, whereas intracellular portions of the organoid contained keratinized cells (no nuclei and ECM).

Example 4

[0085] Experiments were also conducted to compare LGR5-GFP expression in an existing mouse model and a porcine model. FIGS. 12A-12D include representative images of LGR5-GFP expression in a mouse model (FIGS. 12A-12B) as compared to a porcine model (FIGS. 12C-12D). GFP expression is overlaid with DAPI staining in FIGS. 12B and 12D. Confocal microscopy of cryosections from telogen stage mouse skin (naturally synchronized) from LGR5-IRES-GFP model skin. GFP fluorescence is detected in a few cells at the base of the hair follicle (hair is auto-fluorescent). Cross section of hairs from LGR5-H2B-GFP porcine skin demonstrating several stages of the hair cycle (naturally asynchronous). Nuclear GFP allows for tracking of cells at single cell resolution.

[0086] In addition to successful LGR5-GFP expression in a porcine model, experiments were conducted to assess expression in the context of cell proliferation. As shown in FIG. 13,

hair follicles surrounding a wound displayed increased proliferation. LGR5-GFP⁺ hair follicle cells from skin within 1 cm of a chronic ulcer wound show increased staining with Ki-67, a nuclear marker for proliferation. These results demonstrate contribution of the LGR5-GFP stem cells toward re-epithelialization of wounds and the hyper-proliferative response of the LGR5-stem cell populations. Uninjured follicles show minimal co-localization with Ki-67 in the LGR5-GFP stem cells (shown is catagen stage uninjured hair follicle to match catagen injured follicle).

[0087] As shown in FIGS. 14A-14C, LGR5-GFP expression was also found in hair germ cells, the primary proliferating hair germ cells in anagen (GFP expression is overlaid with Ki67 staining FIG. 14B, and DAPI staining in FIG. 14C). Representative images of the base of an anagen stage hair follicle shows stem cell proliferation and H2B-GFP dilution. 20x confocal microscopy shows the hair follicle expressing natural GFP (FIG. 14A), co-stained with the nuclear marker for proliferation Ki-67 (FIG. 14B), and DAPI (FIG. 14C). Scale bar indicates 100 μ M. Nuclear GFP allows for single cell resolution and tracking of cell behavior.

CLAIMS

What is claimed is:

1. A non-human transgenic animal comprising a genome that expresses a nuclear-localized reporter gene in at least one cell type or tissue type that also expresses an endogenous leucine-rich repeat-containing G-protein coupled receptor (LGR) gene.
2. The transgenic animal of claim 1, wherein the endogenous LRG gene is leucine-rich repeat-containing G-protein coupled receptor 4 (LGR4), LRG gene is leucine-rich repeat-containing G-protein coupled receptor 5 (LGR5), or LRG gene is leucine-rich repeat-containing G-protein coupled receptor 6 (LGR6).
3. The transgenic animal of claim 1, wherein the nuclear-localized reporter gene is a fluorescent reporter.
4. The transgenic animal of claim 3, wherein the fluorescent reporter comprises at least one of GFP, eGFP, mCherry, CFP, BFP, YFP, eYFP, photoactivatable GFP, dsRed (*Discosoma* species fluorescent protein), mFruits (mutants of dsRed), TagRFPs (Evrogen), eqFP611 (isolated from sea anemone *Entacmaea quadricolor*), Dronpa (photoswitchable fluorescent protein), and EosFP (photoconvertable fluorescent protein).
5. The transgenic animal of claim 1, wherein the nuclear-localized reporter gene comprises a gene encoding H2B.
6. The transgenic animal of claim 5, wherein the nuclear-localized reporter gene comprises a gene encoding an H2B-GFP fusion protein.
7. The transgenic animal of claim 1, wherein the nuclear-localized reporter gene comprises a gene encoding an LRG protein, or fragment thereof, fused to a nuclear-localized fluorescent reporter.
8. The transgenic animal of claim 7, wherein the nuclear-localized reporter gene comprises a gene encoding an LGR5-H2B-GFP fusion protein.

9. The transgenic animal of claim 8, wherein the gene encoding the LGR5-H2B-GFP fusion protein comprises a gene encoding H2B-GFP downstream of the LGR5 ATG start site.
10. The transgenic animal of claim 8, wherein the gene encoding the LGR5-H2B-GFP fusion protein comprises a gene encoding H2B-GFP downstream of the LGR5 ATG start site, and upstream of intron 1 of LGR5.
11. The transgenic animal of claim 8, wherein the gene encoding the LGR5-H2B-GFP fusion protein does not contain an IRES site.
12. The transgenic animal of claim 8, wherein the gene encoding the LGR5-H2B-GFP fusion protein comprises one or more fragments of SEQ ID NO: 2.
13. The transgenic animal of claim 1, wherein the at least one cell type or tissue type that expresses the endogenous LGR gene is a stem cell.
14. The transgenic animal of claim 1, wherein the at least one cell type or tissue type that expresses the endogenous LGR gene comprises skin, eye, inner ear, gastrointestinal track, uterus, ovary, prostate, mammary gland, kidney, liver, pancreas, cervix, and/or placenta cell types or tissue types.
15. A cell or cell line derived from the transgenic animal of claim 1.
16. The cell or cell line of claim 15, wherein the cell line is a primary cell line or an immortalized cell line.
17. The cell or cell line of claim 15, wherein the cell is a stem cell.
18. An organoid derived from the stem cell of claim 17.
19. A method of generating an organoid comprising culturing the stem cell of claim 17 in a matrix under conditions sufficient to promote growth and/or proliferation of the stem cell.

20. A method of screening an intervention for a disease or condition, the method comprising:
- a) contacting the transgenic animal of claim 1 or the cell or cell line of claim 15 with a candidate intervention; and
 - b) determining the effect of said intervention on a disease or condition in the transgenic animal.
21. The method of claim 20, wherein the intervention is selected from the group consisting of a drug, a lifestyle change, an alternative medicine therapy, or a combination thereof.
22. The method of claim 20, wherein the disease or condition is cancer.
23. The method of claim 20, wherein the disease or condition is associated with stem cell function.
24. The method of claim 20, wherein the disease or condition is an intestinal, hepatic, renal, lung, or skin disease or injury.

FIG. 1
New Targeting Strategy

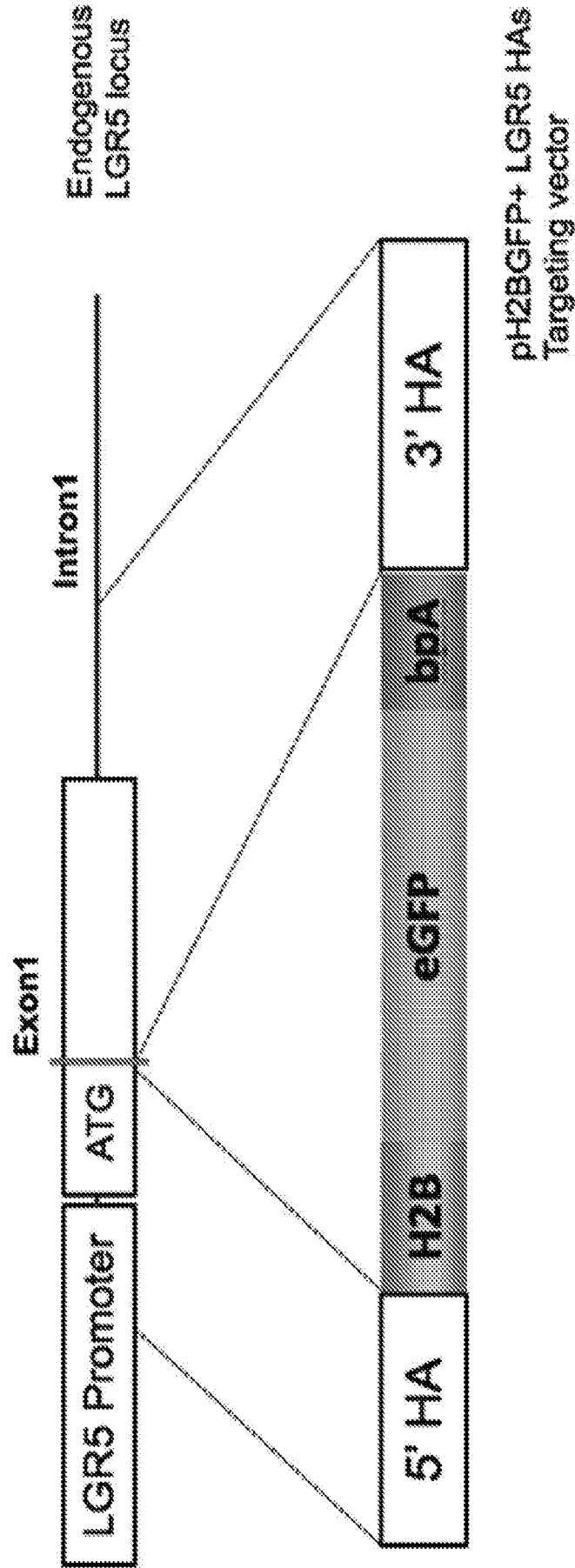
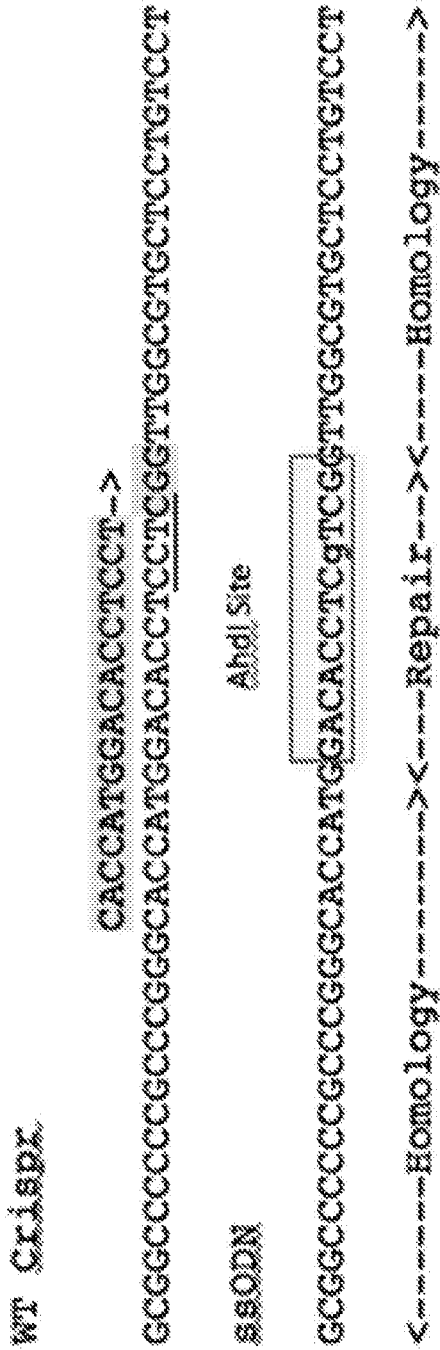


FIG. 2A

ssODN Design for Repair of the wild type allele

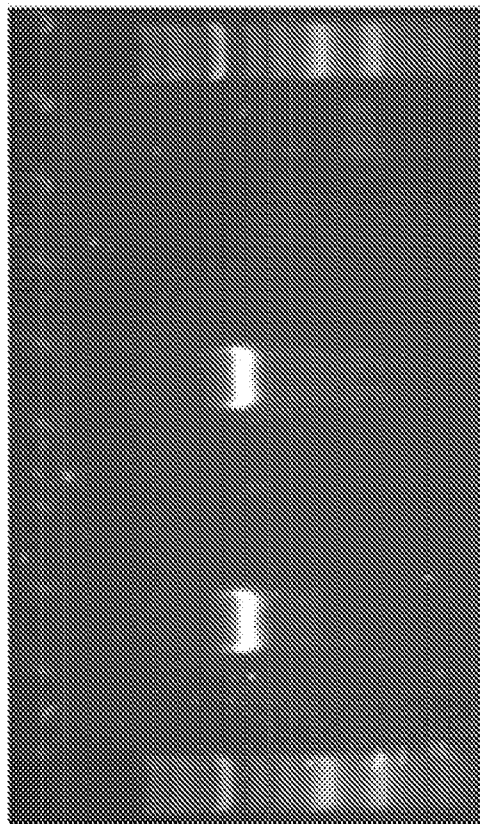


Repair by co-transfection of targeting vector and repair ssODN

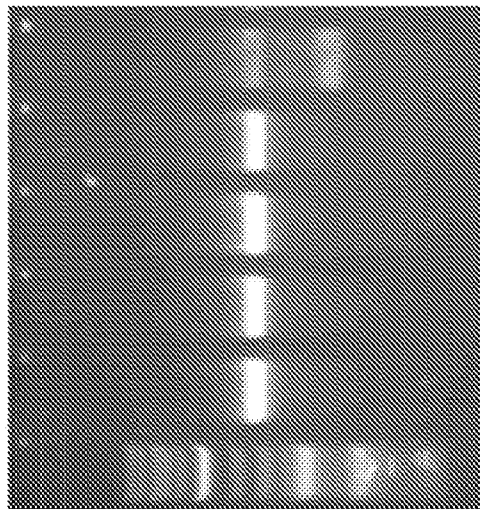
FIG. 2 (CONT'D)

FIG. 2B

KI Screen



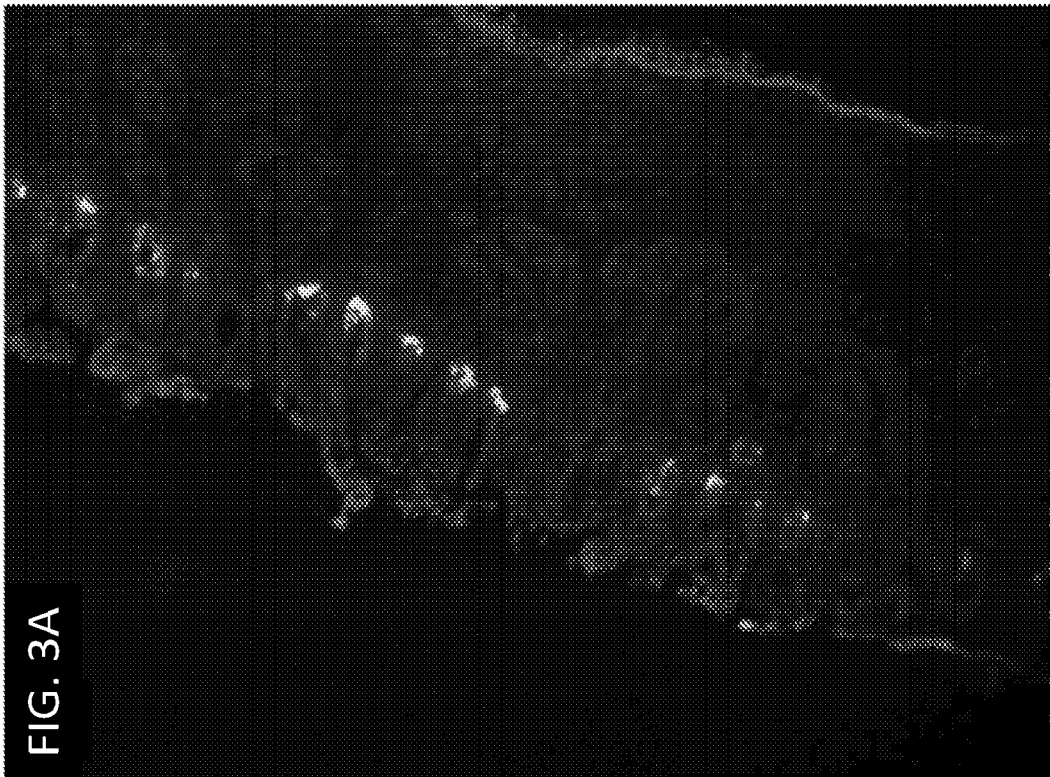
AhdI Digestion



FIGS. 3A-B

Duodenum

FIG. 3A



Jejunum

FIG. 3B



FIGS. 4A-B

FIG. 4B

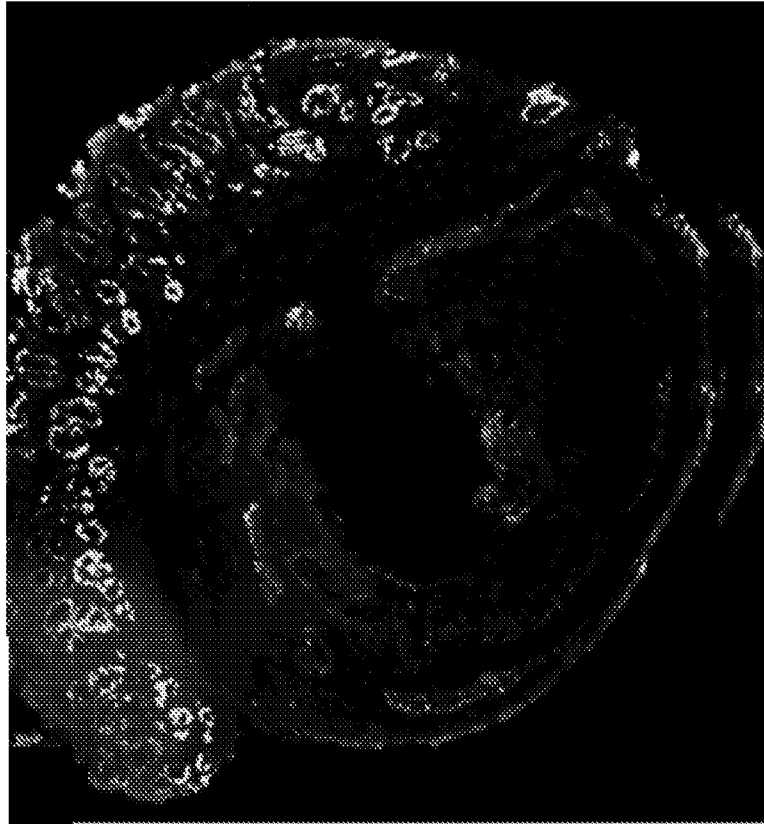


FIG. 4A



FIG. 5

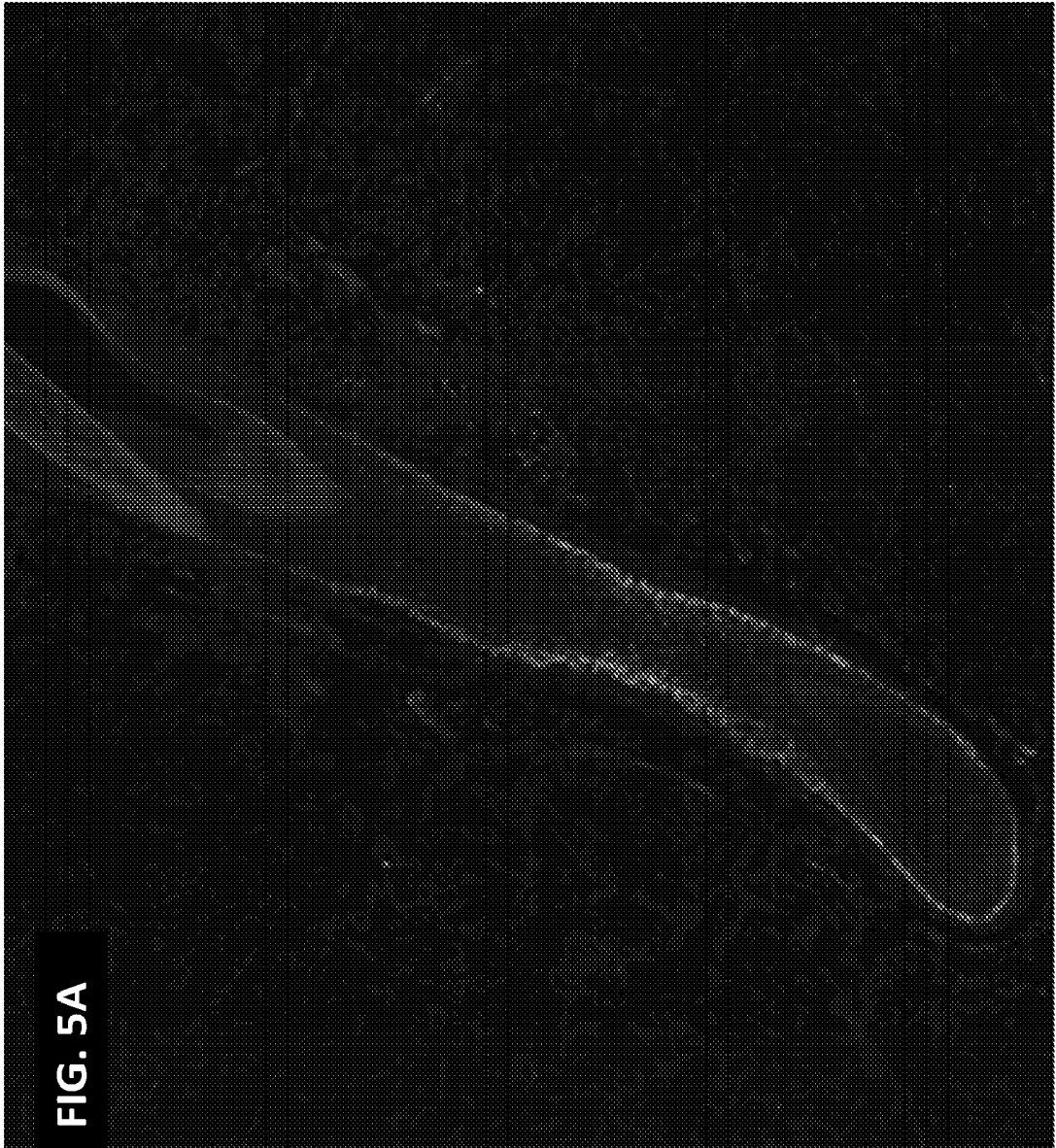


FIG. 5 (CONT'D)

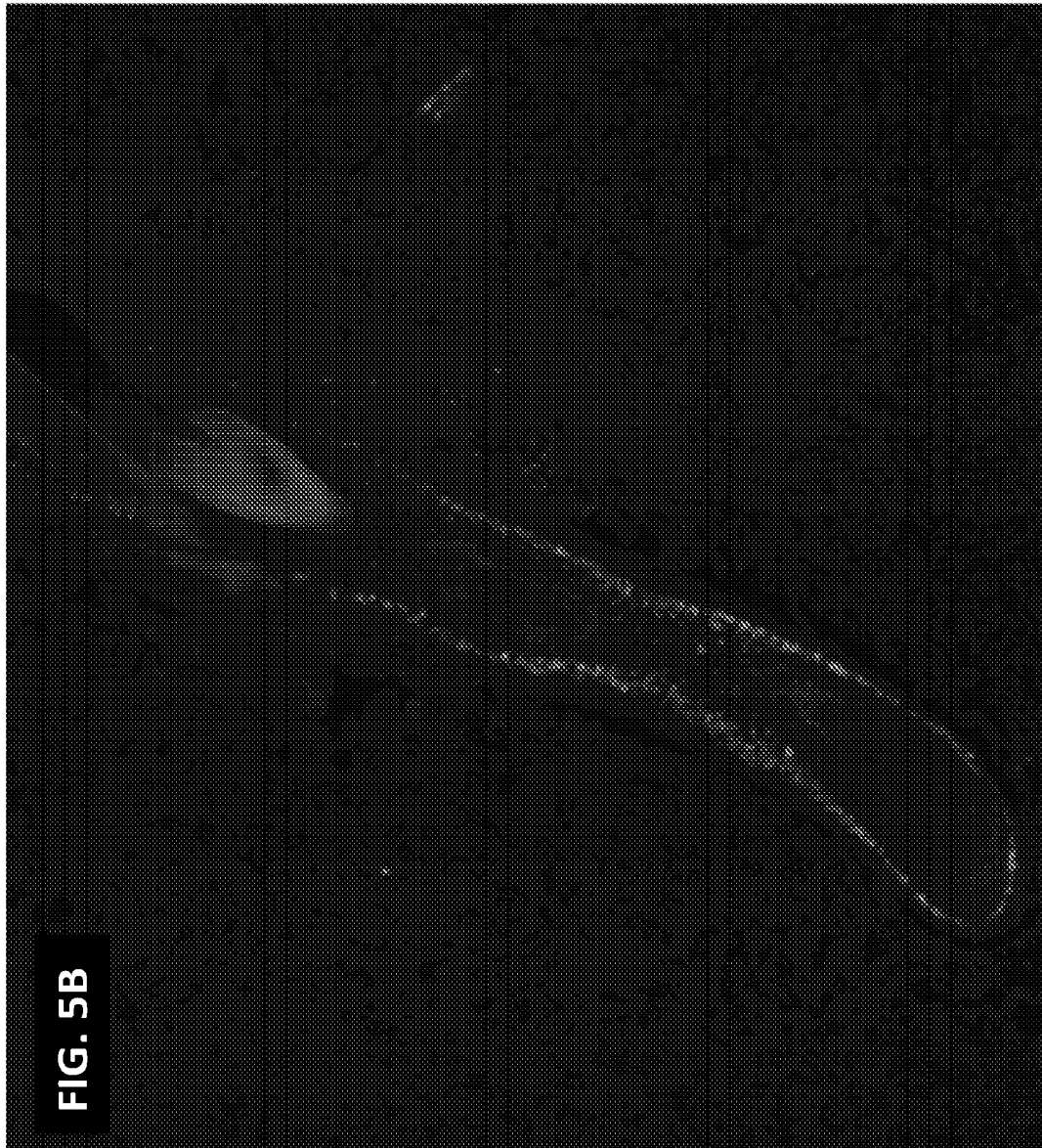


FIG. 5 (CONT'D)

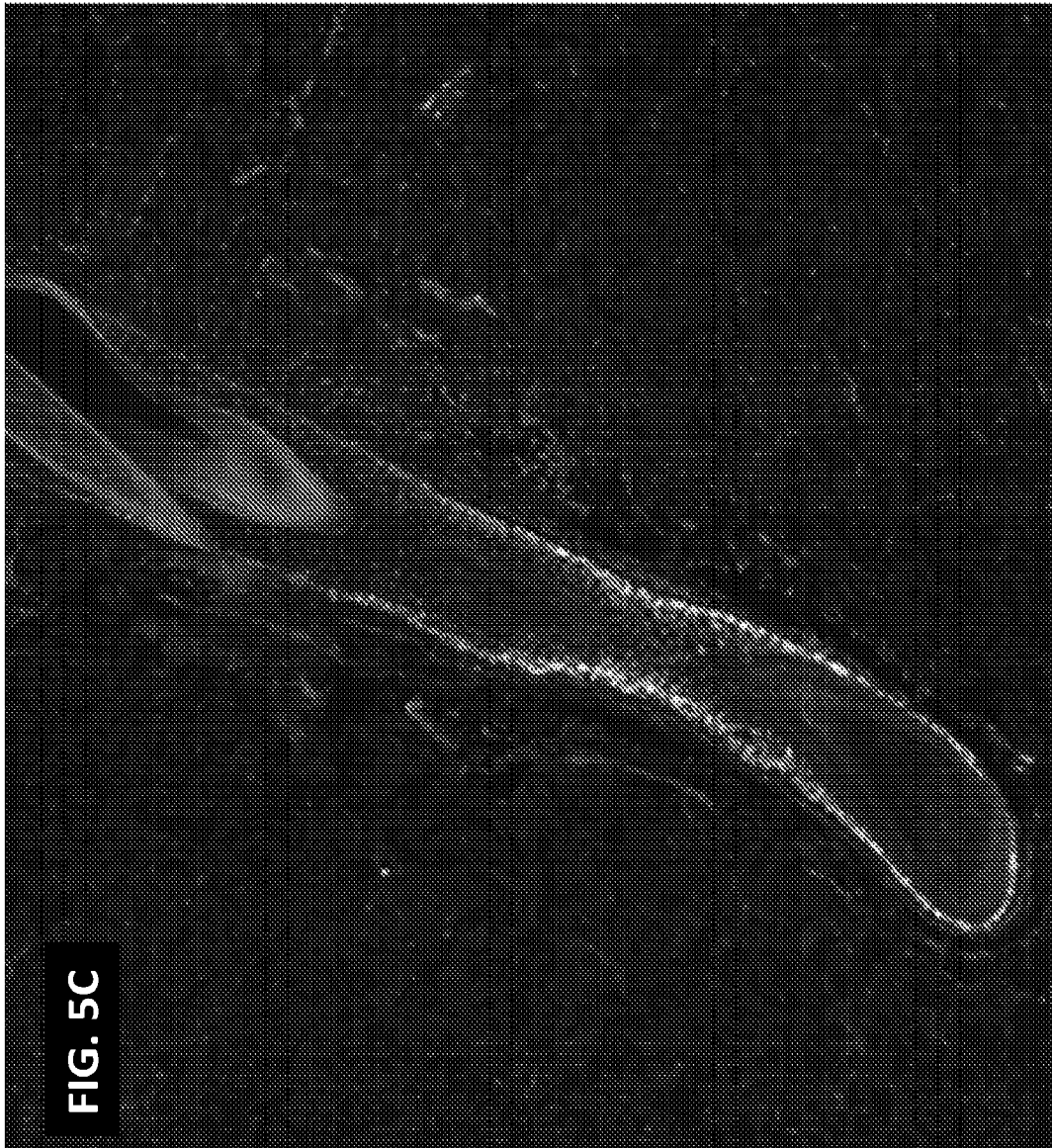


FIG. 6
Anagen (growth)

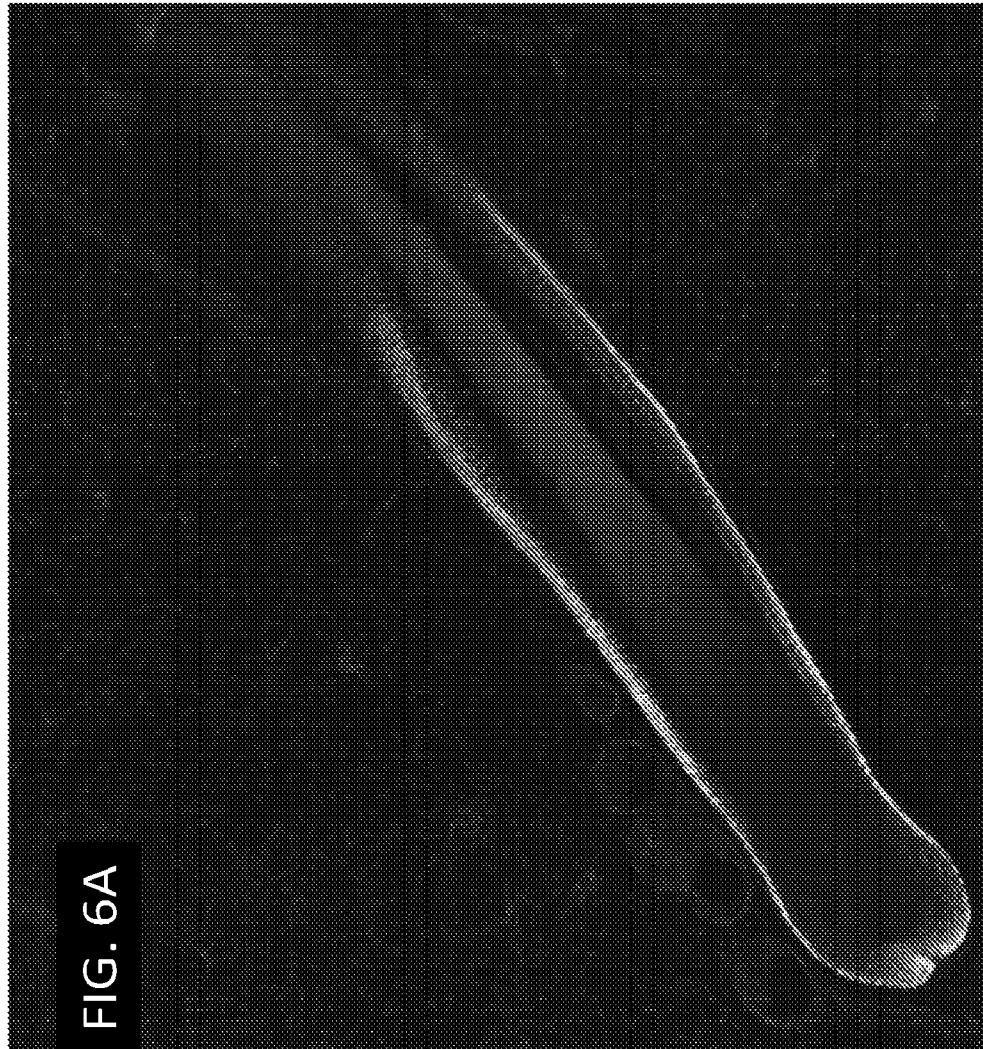


FIG. 6 (CONT'D)
Catagen (destruction)

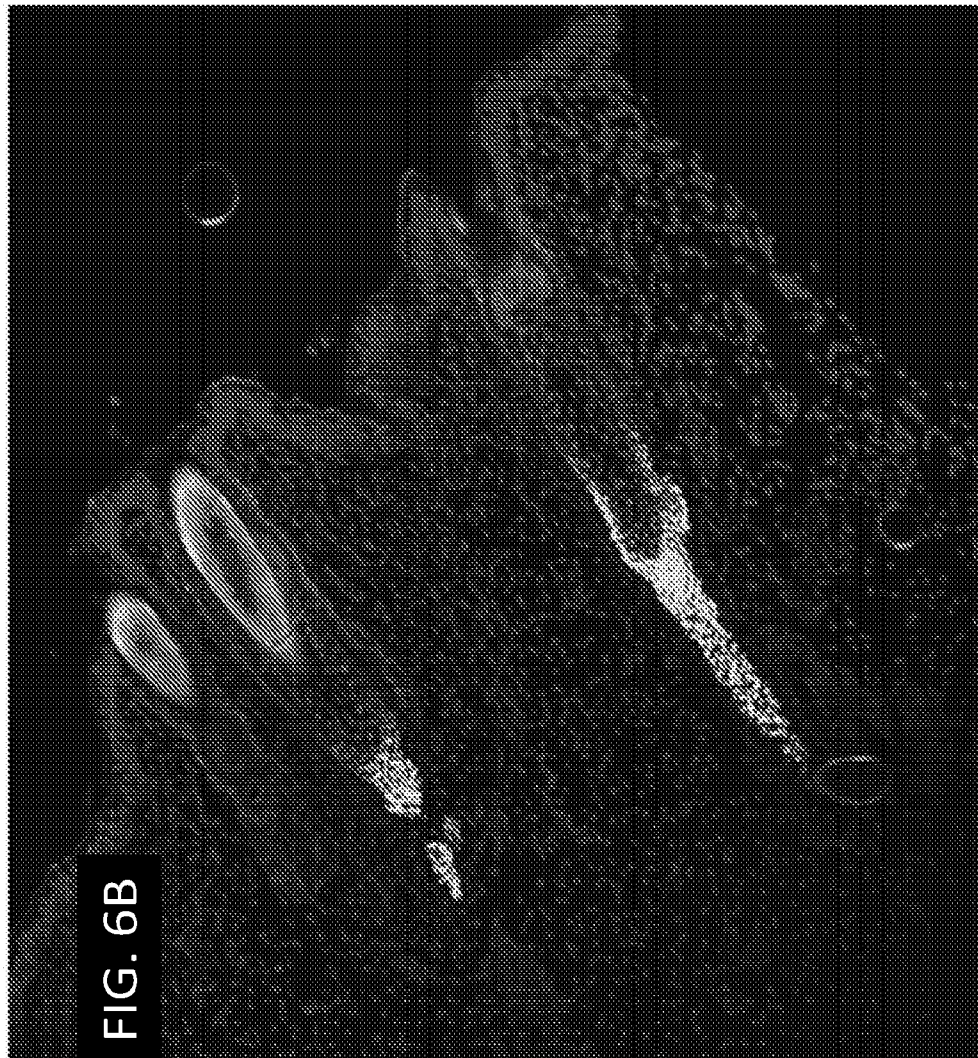


FIG. 6 (CONT'D)
Telogen (rest)

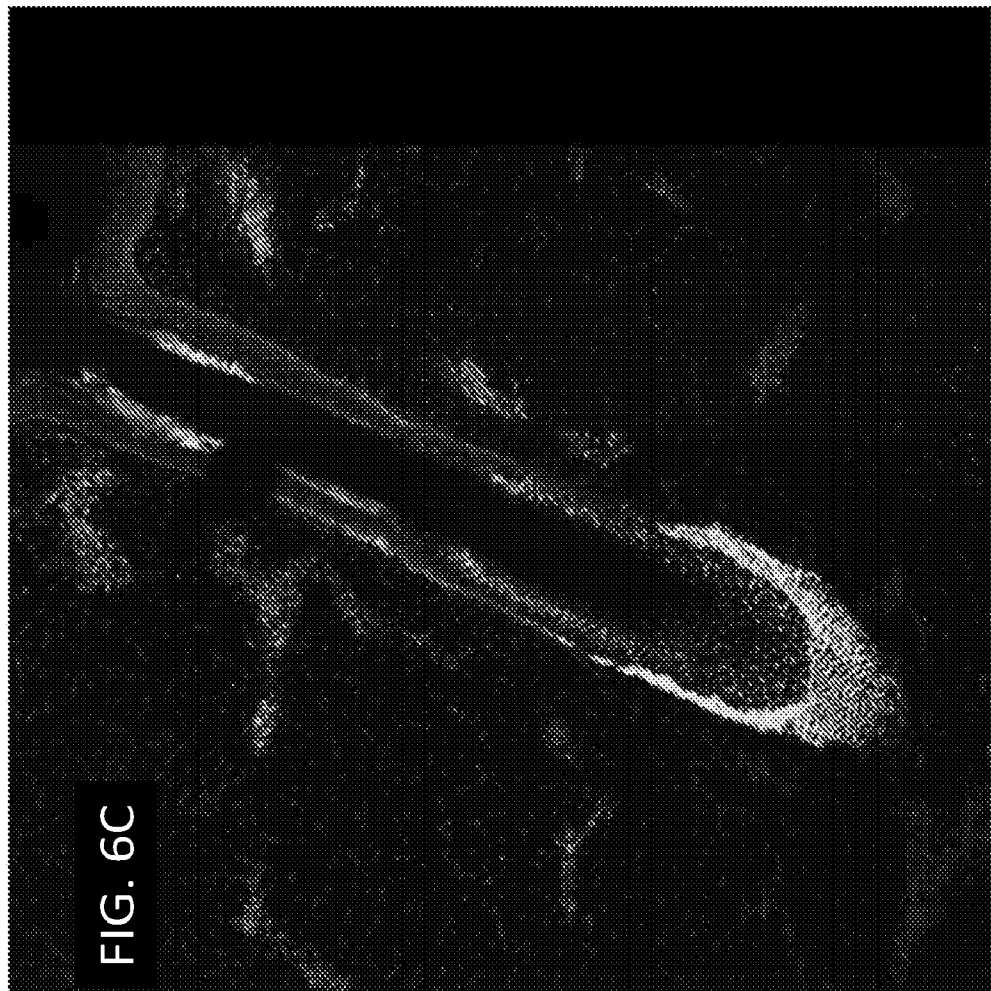


FIG. 7

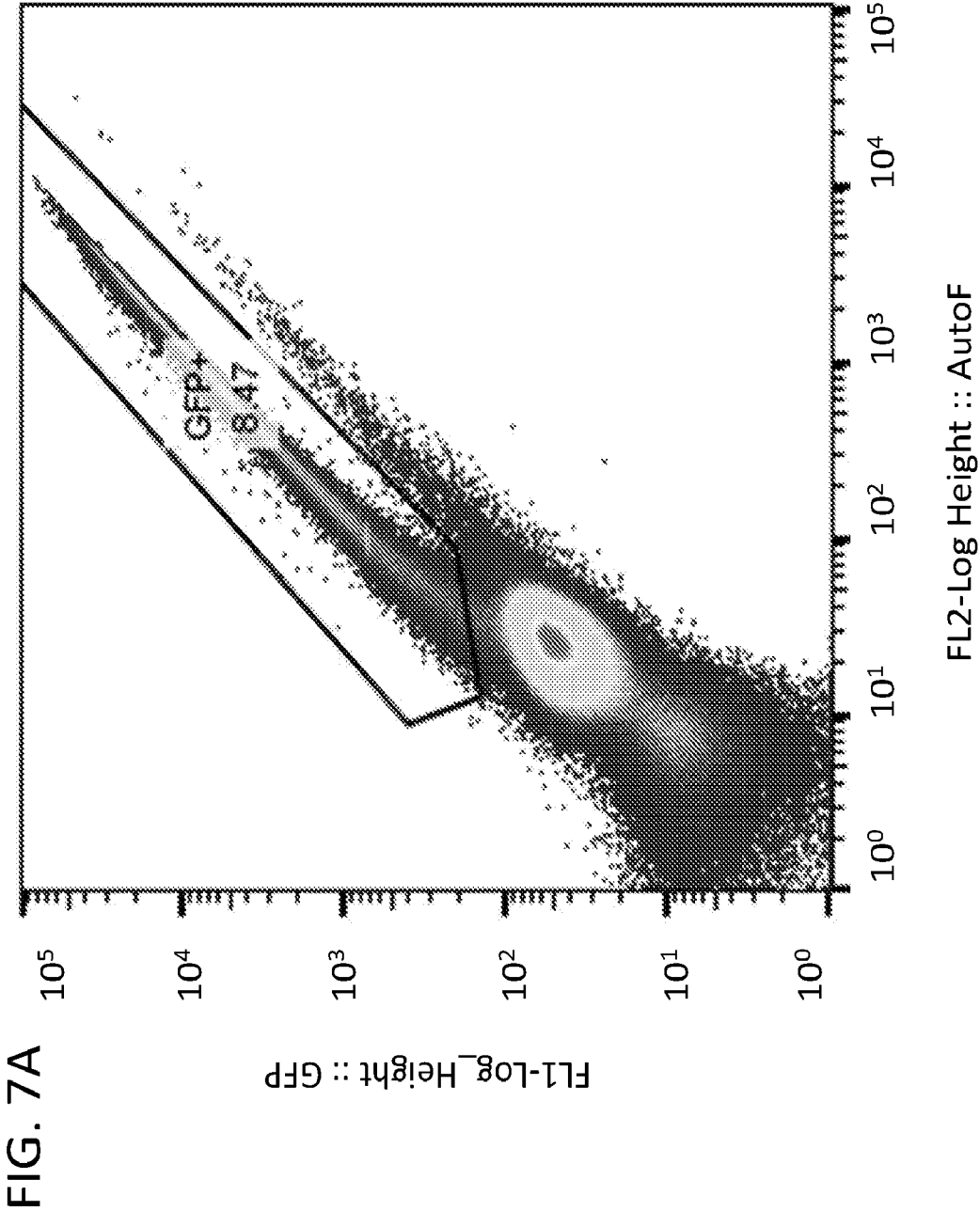


FIG. 7 (CONT'D)

FIG. 7B

Sample_10 (G2: GFP+ & Singlets & Scatter)

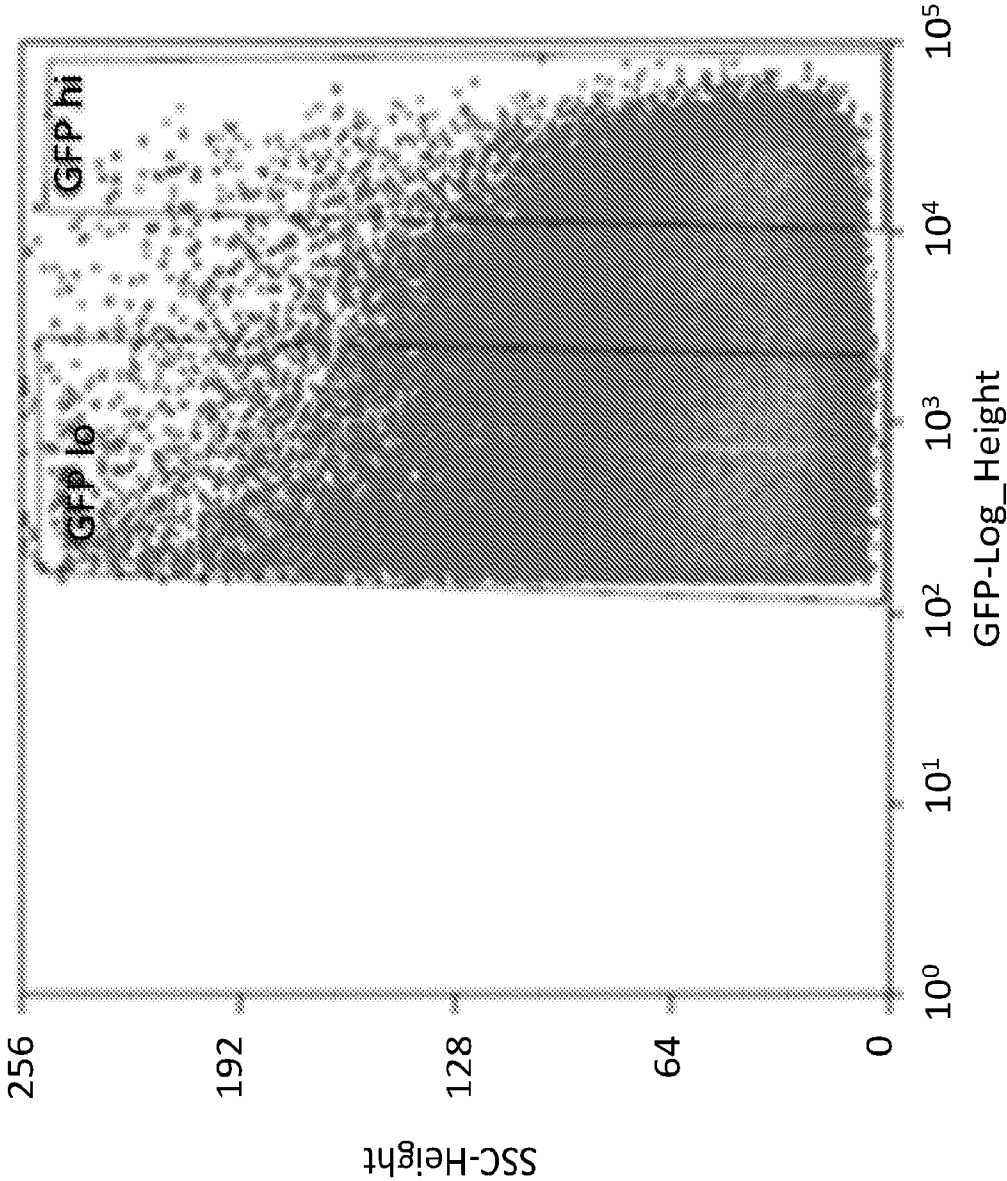


FIG. 7 (CONT'D)

Relative LGR5 expression

FIG. 7C



FIG. 8

FIG. 8A Sample_17 (G1: Singlets & Scatter)

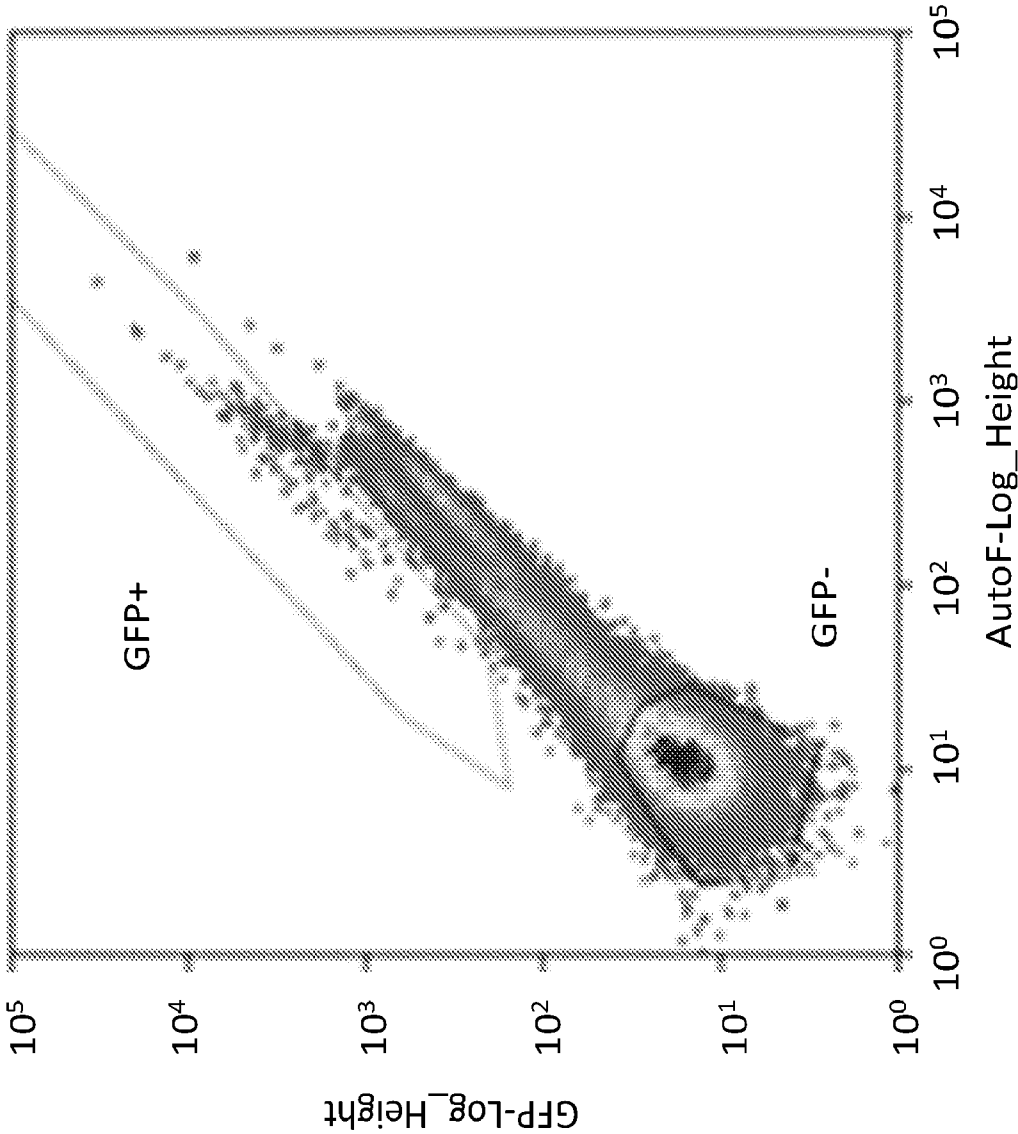


FIG. 8 (CONT'D)

FIG. 8B Sample_17 (G1: Singlets & Scatter)

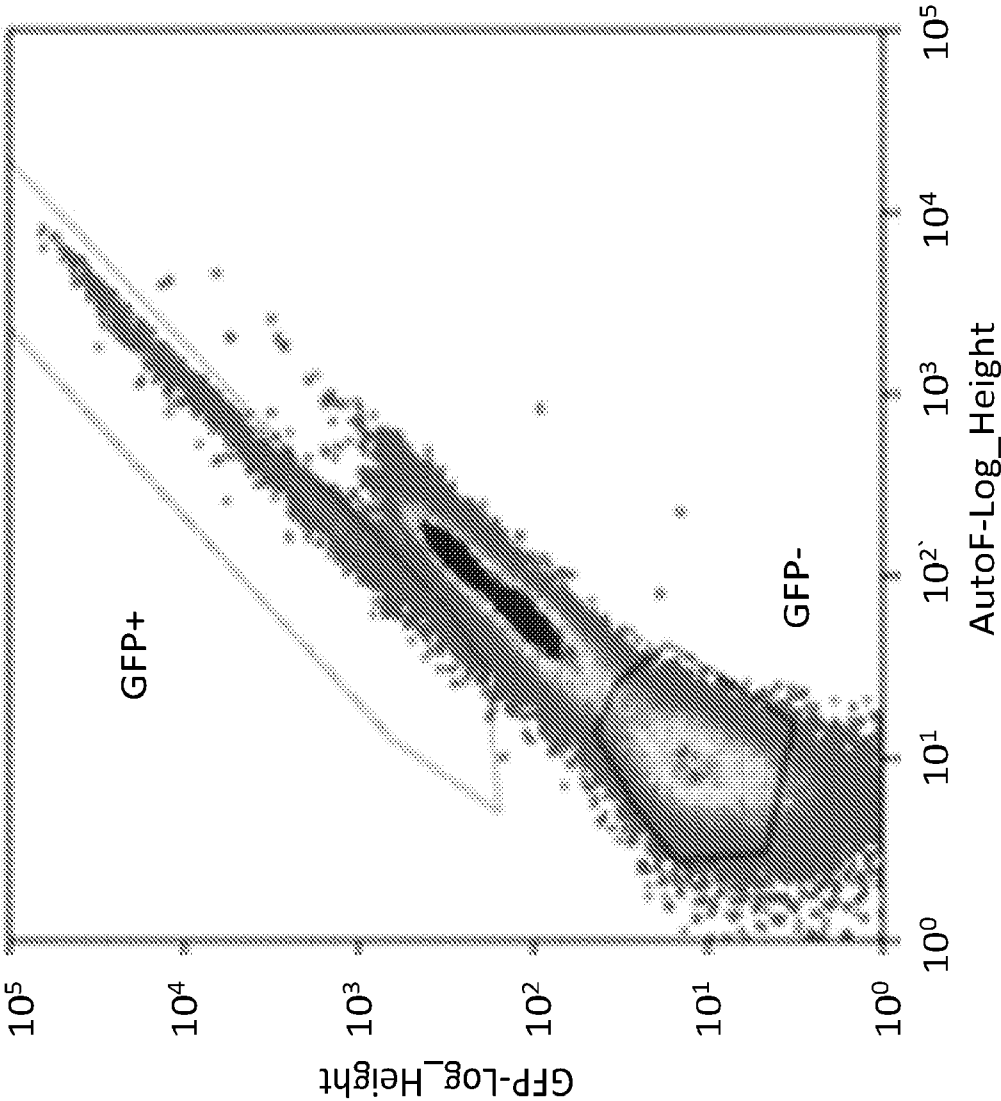


FIG. 8 (CONT'D)

Sample_13 (G1: Singlets & Scatter)

FIG. 8C

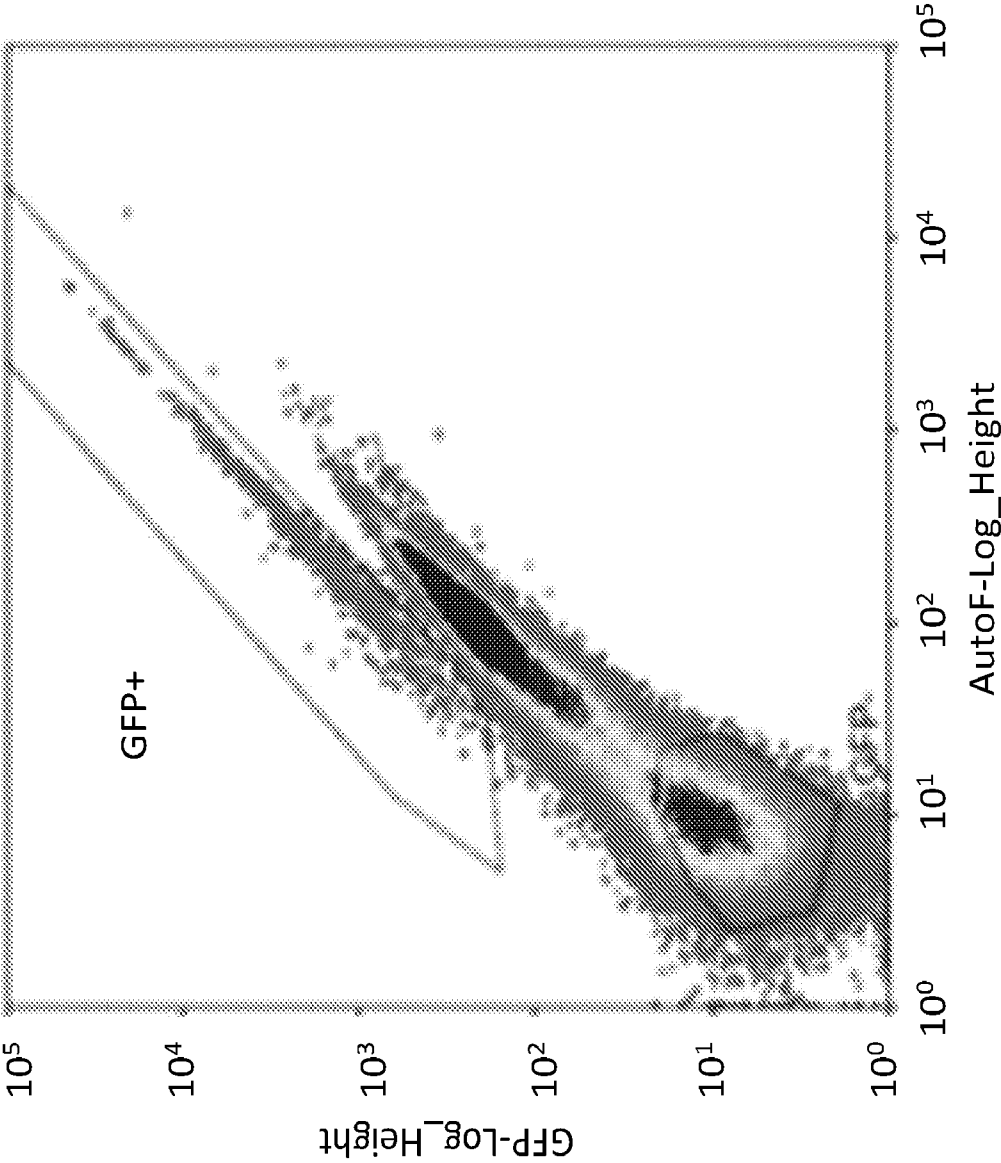


FIG. 8 (CONT'D)

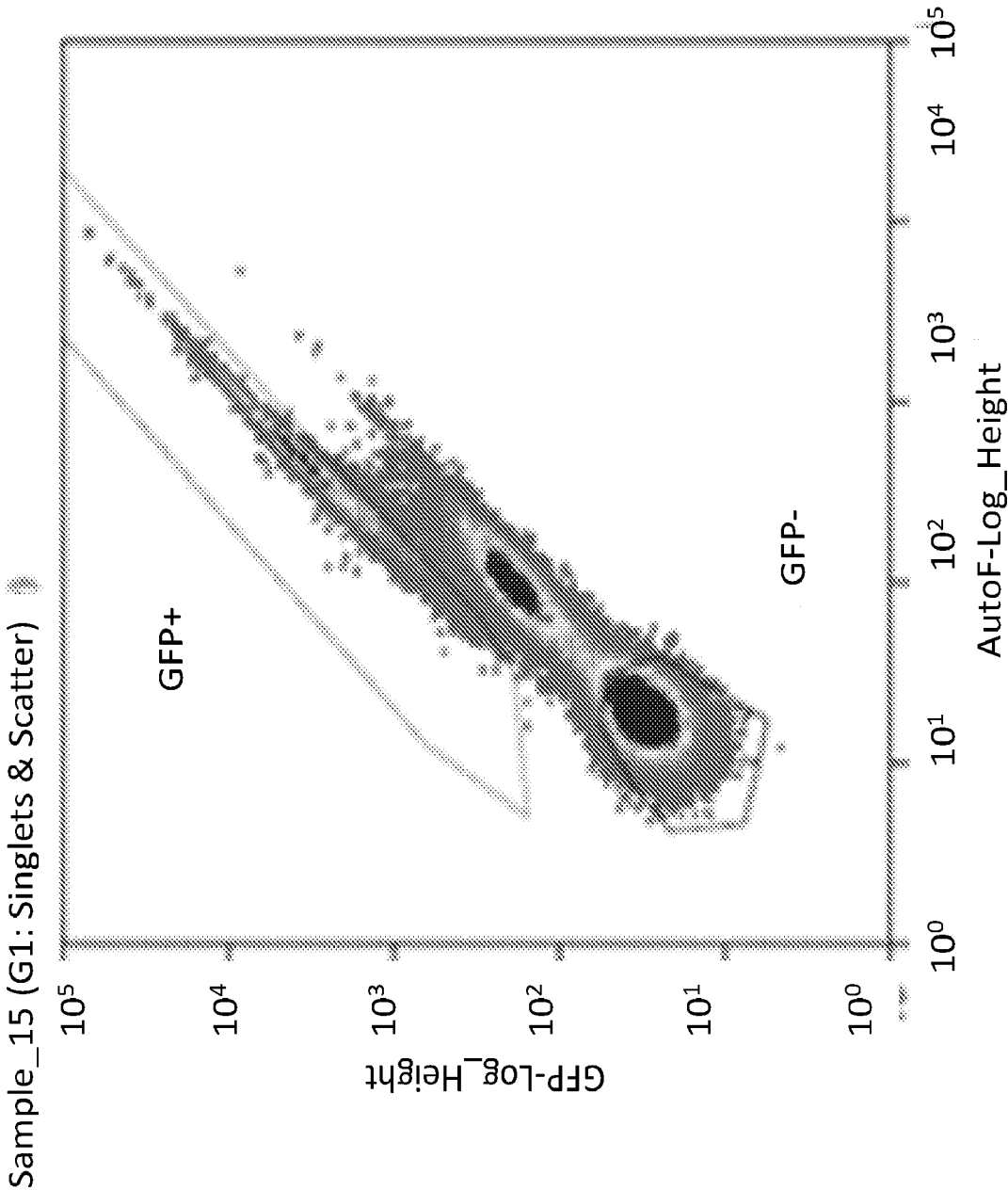


FIG. 8D

FIG. 9

FIG. 9A

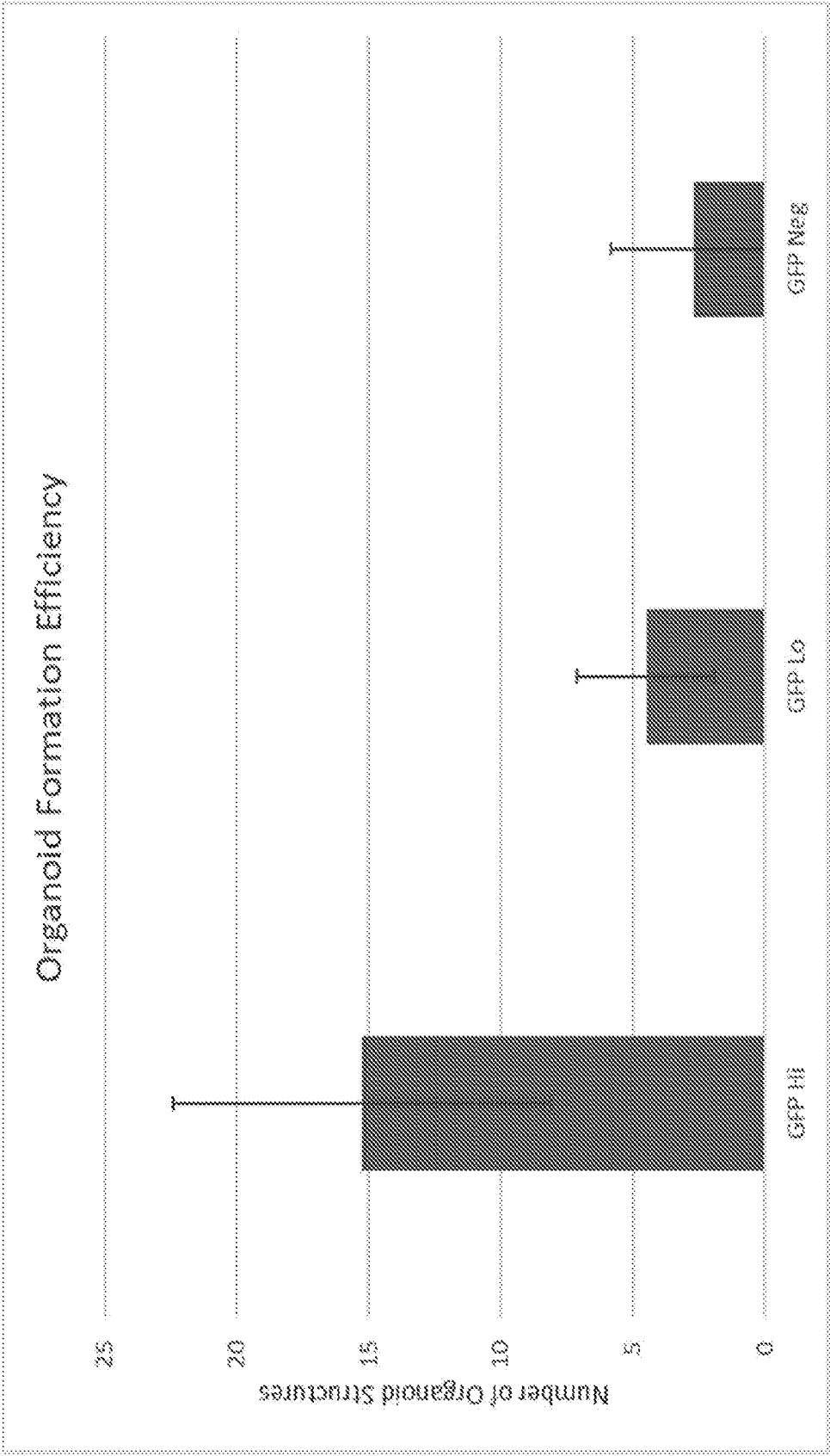


FIG. 9 (CONT'D)

FIG. 9B

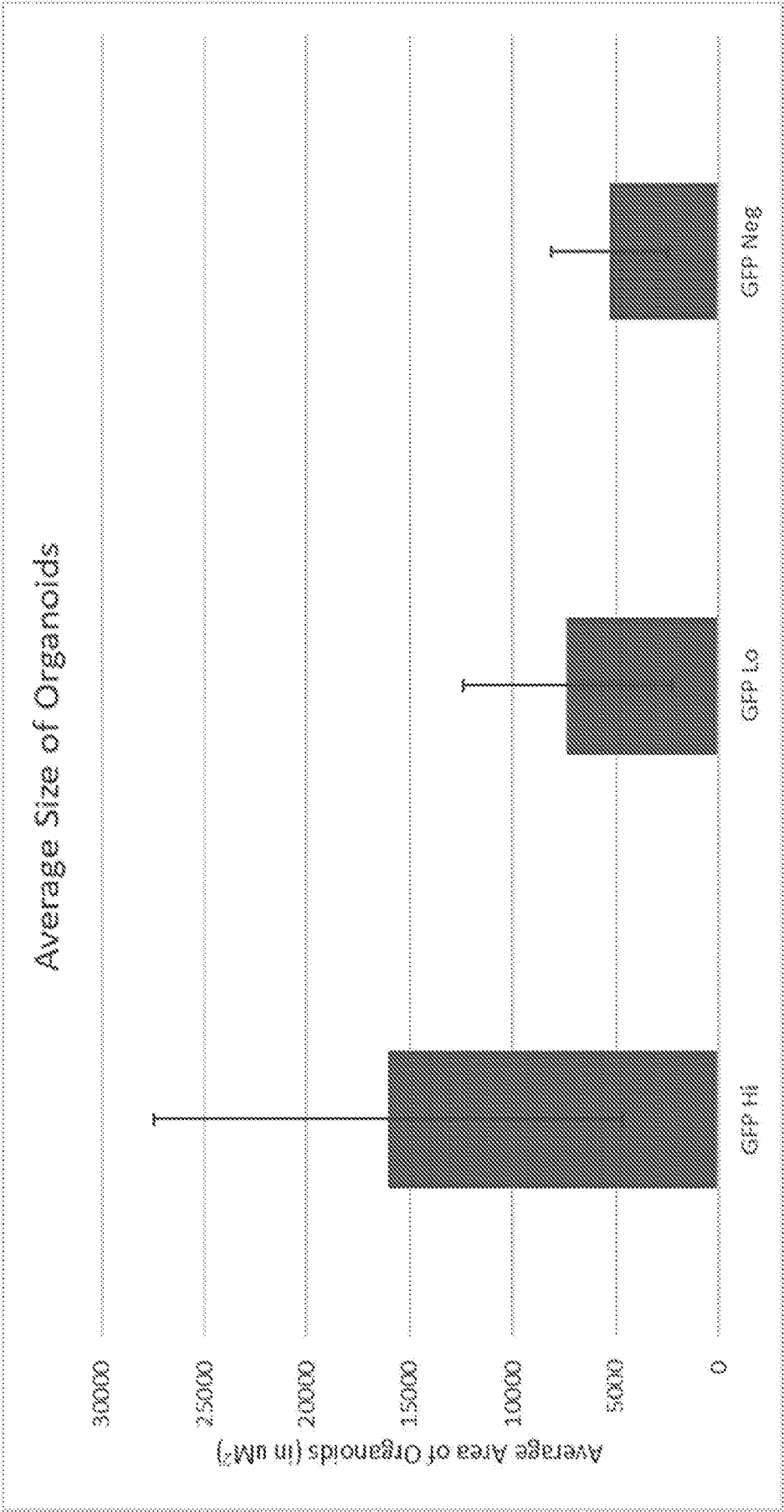


FIG. 10

FIG. 10A

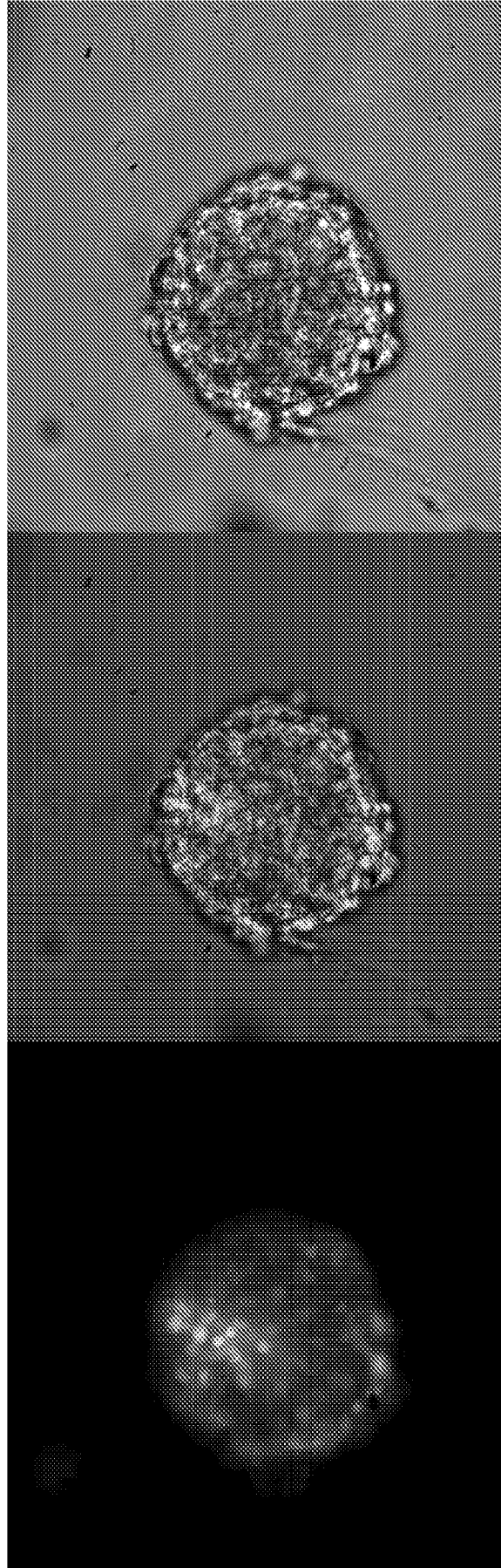
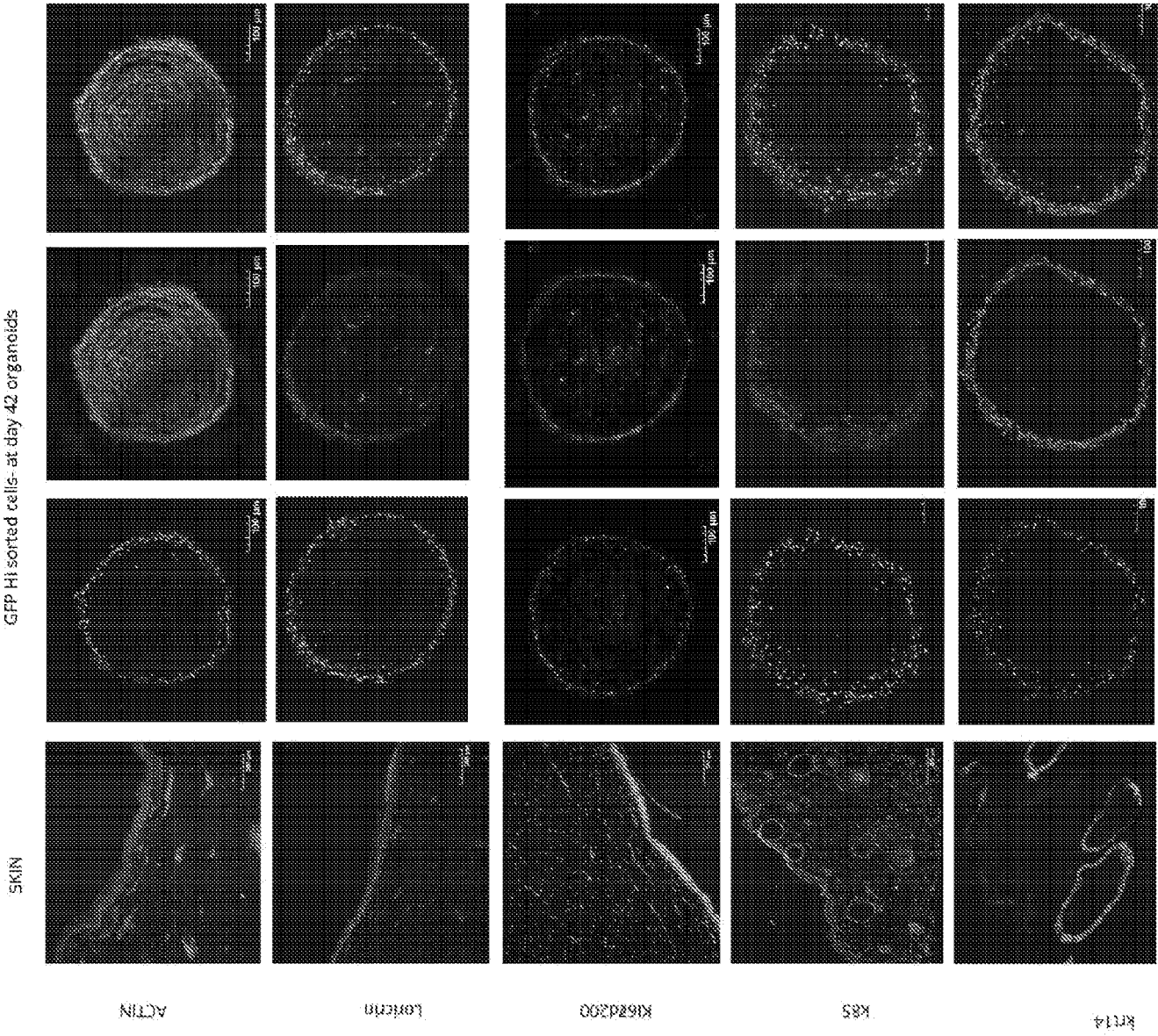
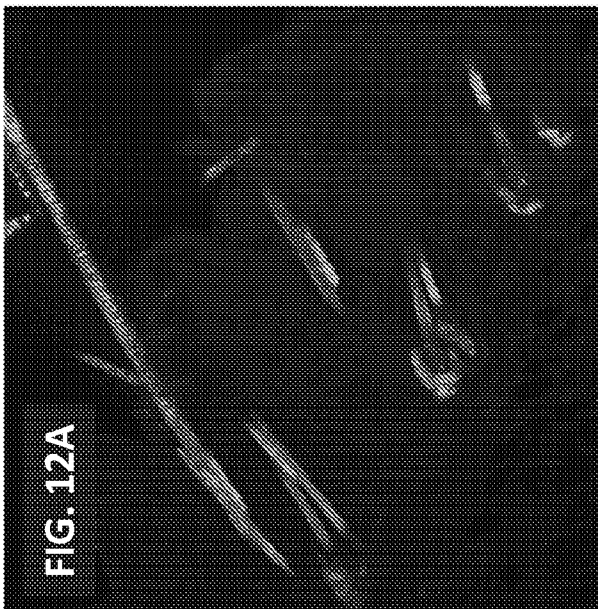
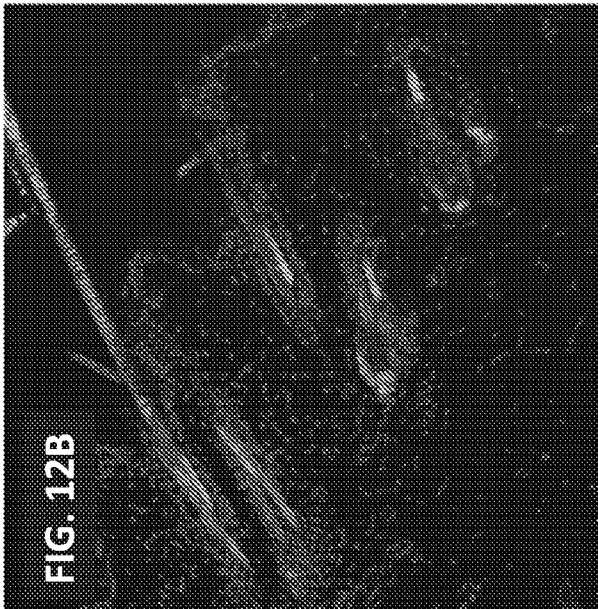


FIG. 10 (CONT'D)

FIG. 10B







FIGS. 12A-12D

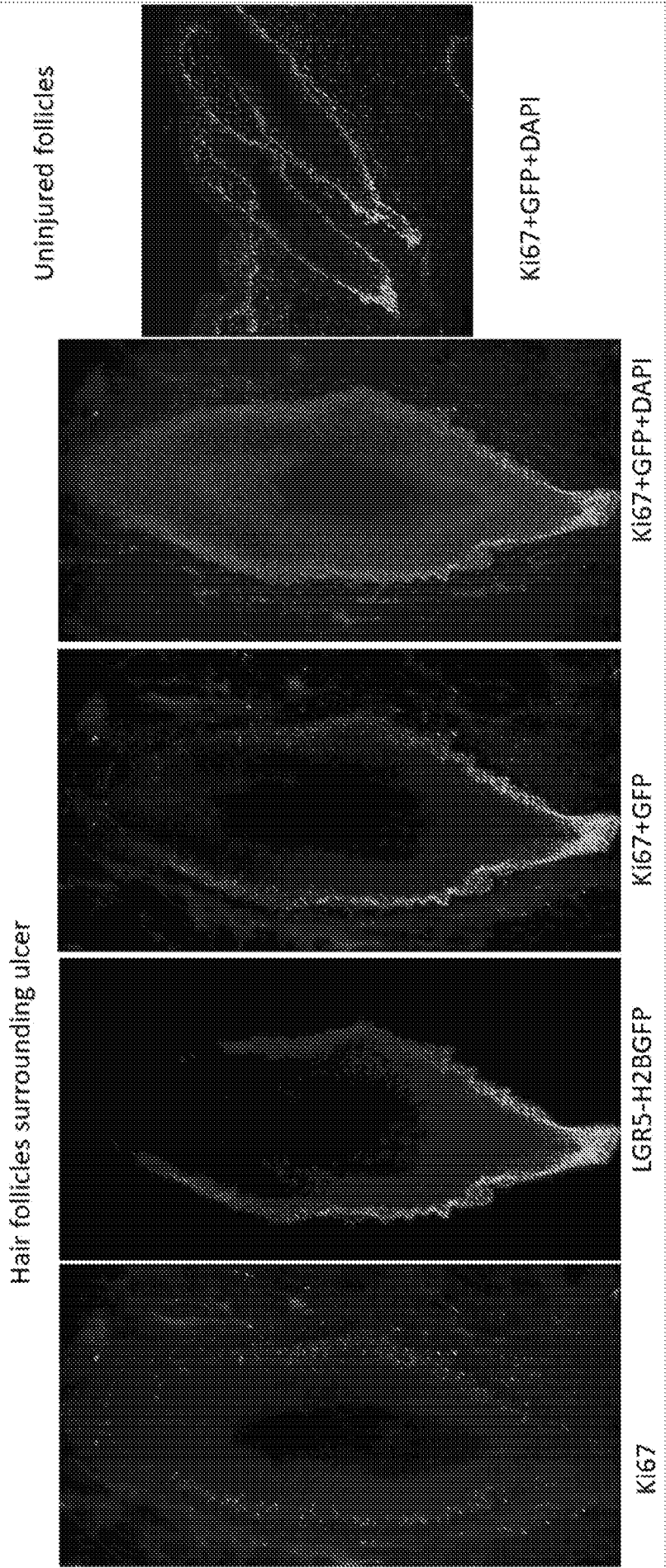


FIG. 13

FIG. 14A

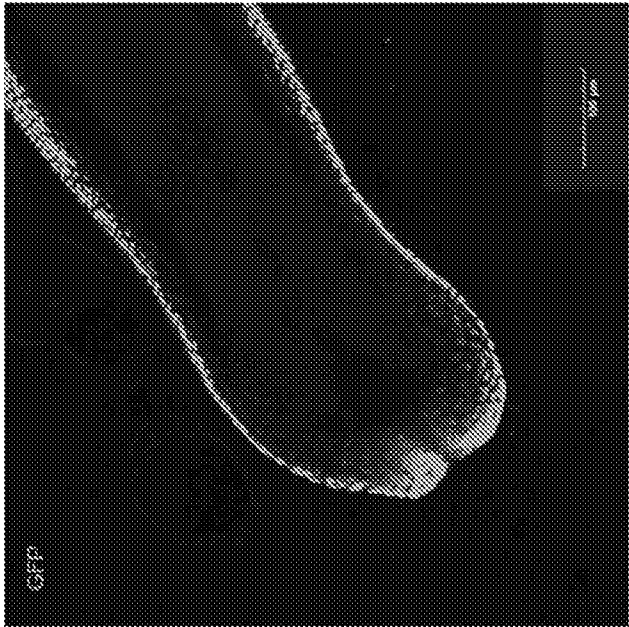


FIG. 14B



FIG. 14C



FIGS. 14A-14C

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 20/15008

A. CLASSIFICATION OF SUBJECT MATTER

IPC - A01K 67/00, A01K 67/027, C12N 15/09, C12N 15/85 (2020.01)

CPC - C12N 15/8509, A01K 67/0275, A01K 2217/05

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)

See Search History document

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

See Search History document

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

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ---- Y	US 2018/0221441 A1 (KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN) 09 August 2018 (09.08.2018) Abstract; para [0006-0007], [0035], [0155], [0164-0167], [0170-0171], [0252], [0324], [0329], [0339], [0388], [0404], [0399]	1-4, 7, 13-17 ----- 5-6, 8-12
Y	US 2018/0353652 A1 (THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL) 13 December 2018 (13.12.2018) para [0121], [0247]	5-6, 8-12
Y	GenBank Accession No: CU463987.2, 25 November 2007, [Retrieved 17 April 2020]. Retrieved from the internet <URL: https://www.ncbi.nlm.nih.gov/nuccore/CU463987.2/ > entire document	12

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

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

"D" document cited by the applicant in the international application

"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

17 April 2020

Date of mailing of the international search report

16 JUN 2020

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Lee Young

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 20/15008

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

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

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

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

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

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

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

Group I, claims 1-17, directed to a non-human transgenic animal, or a cell thereof.

Group II, claims 18-19, directed to an organoid or a method of generating an organoid.

Group III, claims 20-24, directed to a method of screening an intervention for a disease or condition.

-----Please see Supplemental Box-----

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-17

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/US 20/15008

Continuation of Box No. III Observations where unity of invention is lacking:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-17, directed to a non-human transgenic animal, or a cell thereof.

Group II, claims 18-19, directed to an organoid or a method of generating an organoid.

Group III, claims 20-24, directed to a method of screening an intervention for a disease or condition.

The inventions listed as Groups I-III do not relate to a single special technical feature under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special technical features:

Group I has the special technical feature of a composition consisting of a non-human transgenic animal, or a cell thereof, that is not required by Group II or III.

Group II has the special technical feature of a composition consisting of an organoid or a method of generating an organoid, that is not required by Group I or III.

Group III has the special technical feature of contacting a transgenic animal or cell(s) therefrom with a candidate intervention and determining the effect of said intervention, that is not required by Group I or II.

Common technical features:

Groups I-III share the common technical feature of a non-human transgenic animal comprising a genome that expresses a nuclear-localized reporter gene in at least one cell type or tissue type that also expresses an endogenous leucine-rich repeat-containing G-protein coupled receptor (LGR) gene, and a cell or cell line derived from the transgenic animal

However, this shared technical feature does not represent a contribution over prior art, because this shared technical feature is previously made obvious by US 2018/0221441 A1 to Koninklijke Nederlandse Akademie van Wetenschappen. (hereinafter "KNA").

KNA teaches a non-human transgenic animal comprising a genome that expresses reporter gene in at least one cell type or tissue type that also expresses an endogenous leucine-rich repeat-containing G-protein coupled receptor (LGR) gene (para [0165] "a recombinant animal comprising a first reporter gene under control of a Lgr5 or Lgr6 promoter such that the reporter gene product is expressed in cells expressing Lgr5 or Lgr6"; [0171] - "Said second reporter gene is preferably inserted into a genomic region that provides robust, prolonged expression of the transgene after activation, such as the ROSA locus if the transgenic animal is a mouse."; para [0388] - "Lgr6-EGFP-Ires-CreERT2 mice were generated by homologous recombination in embryonic stem cells targeting an EGFP-Ires-CreERT2 cassette to the ATG of Lgr6."; para [0404] - "The analysis of endogenous Lgr5 expression was determined in L8 cells."; para [0252] - "In yet another embodiment, the invention provides (a collection of) cancer stem cells comprising Lgr5 and/or 6 embedded in their cell membrane"). KNA does not specifically teach the reported is a nuclear localized. However, KNA does teach detection of the reporter in the nucleus of cells (para [0339] - "26 days old Lgr6-EGFP-Ires-CreERT2 mice (early anagen) were obtained and stained for nuclear DNA (Topro) and EGFP visualized using confocal microscopy (A-C)."). Use of nuclear localized reporters was well known in the art of biology at the time in order to measure expression of a gene in the nucleus of a transgenic cell. Therefore, it would have been obvious to one of ordinary skill in the art to have used a nuclear localized reporter in the KNA methods, in order to better detect expression of LGR due to transgene expression in the nucleus of a transgenic cell generated by KNA methods.

KNA further teaches a cell or cell line derived from the transgenic animal of claim I (para [0155] - "The invention further provides isolated, genomically modified stem cells obtainable by a method of the invention."; para [0252] - "In yet another embodiment, the invention provides (a collection of) cancer stem cells comprising Lgr5 and/or 6 embedded in their cell membrane").

As the technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Therefore, Group I-III inventions lack unity under PCT Rule 13 because they do not share the same or corresponding special technical feature.