



- (51) **International Patent Classification:**
A01K 67/027 (2006.01) *C12N 15/85* (2006.01)
- (21) **International Application Number:**
PCT/US2016/028165
- (22) **International Filing Date:**
18 April 2016 (18.04.2016)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
62/149,330 17 April 2015 (17.04.2015) US
- (71) **Applicant:** UNIVERSITY OF SOUTH FLORIDA
[US/US]; 3802 Spectrum Blvd., Suite 100, Tampa, FL
33612-9220 (US).
- (72) **Inventors:** WANG, Jia-Wang; 10429 Meadow Spring Dr,
Tampa, FL 33647 (US). LOCKEY, Richard F.; 2708
West Marlin Ave, Tampa, FL 33611 (US).
- (74) **Agents:** LADWIG, Glenn P. et al.; Saliwanchik, Lloyd &
Eisenschenk, PO Box 142950, Gainesville, FL 32614-2950
(US).
- (81) **Designated States** (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
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))
— with sequence listing part of description (Rule 5.2(a))



WO 2016/168848 A1

(54) **Title:** ANIMAL MODEL AND METHOD FOR STUDYING GENE-GENE INTERACTIONS

(57) **Abstract:** The invention concerns a non-human animal model useful for sensitively studying gene-gene interactions over a wide genetic background; methods for producing the animal model; and methods for studying gene-gene interactions using an animal model of the invention.

DESCRIPTION

ANIMAL MODEL AND METHOD FOR STUDYING GENE-GENE INTERACTIONS

5 CROSS-REFERENCE TO RELATED APPLICATION

The present application claims the benefit of U.S. Provisional Application Serial No. 62/149,330, filed April 17, 2015, which is hereby incorporated by reference herein in its entirety, including any figures, tables, nucleic acid sequences, amino acid sequences, or drawings.

10

BACKGROUND OF THE INVENTION

It is believed that many human common diseases result from the complex interplay of genetic and environmental factors, and individual gene contributions are usually small and can be masked by genetic and environmental variations, posing a great challenge in studying genetic interactions in human diseases. Although animal models are indispensable and powerful tools for studying these interactions, current animal models are not suitable to study genetic interactions. Techniques must be developed to rapidly, specifically and sensitively study the interplay of genetic factors in any laboratory animal models over a wide genetic background, essential to understand and treat most human diseases.

20

The genome-wide association studies (GWAS) show that many common diseases result from the complex interplay of many genes and environmental factors, and that individual gene contributions are small and can be masked by genetic and environmental variations [1]. The history of mouse genetics also makes it clear that the genetic background can cause considerable phenotypic variation in the mice with the same targeted gene. These variables pose a great challenge in studying human disease gene-gene interactions. GWAS have failed to establish common variant risk for the majority of common diseases [15,20]. Animal models are indispensable and powerful tools that have tremendously increased knowledge of gene function. However, animal techniques that can sensitively study gene-gene interactions over a wide genetic background required to decode complex human diseases have yet to be developed.

25

30

Current animal models are dependent on establishing a congenic strain and the presumption that animals within the strain are identical [6,7]. Unfortunately, these animals are still subject to genetic, epigenetic and environmental variations which can result in high background noise, low sensitivity, or even false data. Moreover, the lab mice have limited genetic diversity to recapitulate gene-gene interactions in human diseases.

Furthermore, most gene studies have focused on one gene and one strain. To recapitulate human disease gene-gene interactions, the targeted gene should be studied in a variety of inbred strains and wild mice to find interactors of the gene. However, the outbreeding will lose homogeneity, making it virtually impossible to study the gene, and a “congenic” strain from backcrossing has variations as stated above.

BRIEF SUMMARY OF THE INVENTION

To resolve problems associated with currently available animal models, this invention provides a novel animal model in which the nine genotypes of two genes are specifically labeled with four distinct detectable labels (also referred to as reporter molecules) such as fluorescent proteins in a single animal, using the clustered regularly interspaced short palindromic repeats (CRISPR) technique.

This model can be crossed with other strains or wild animals (*e.g.*, wild mice). As each genotype will have the other eight genotypes as controls, congenic strain establishment is not required, allowing a rapid (about two months) genetic background switch and the F1 animals can be directly used in study. Therefore, this technique, named “Genebow”, can be used to label cells with many signals (*e.g.*, colors) which may be measured or otherwise detected to study interactions of adjacent cells as in Brainbow. More importantly, Genebow can be used to study gene-gene interaction with high sensitivity, specificity due to the elimination of aforementioned variations, and over a wide genetic background required to model human disease gene-gene interactions.

As proof of principle, the lipopolysaccharide-responsive, beach and anchor containing (*Lrba*) and *Nfkb1* genes will be used to develop the Genebow model to study how the genetic interactions of the two genes affect the apoptosis of leukocytes. The mutations of LRBA cause immunodeficiency and autoimmunity with highly variable symptoms and LRBA interacts with many crucial regulators including NFkB *in vitro*. The

invention will permit study of how LRBA interacts with these regulators *in vivo* to better understand the immune system.

Mutations of LRBA paradoxically cause both immunodeficiency and autoimmunity. LRBA deficient patients have more CD20⁺ B cells responsive to anti-CD20 therapy [46-49], indicating increased B cell survival. LRBA knockdown (KD) protects Raji B cells from death induced by various stress stimuli (Figure 2D). However, LRBA has been suggested to be an oncogene, and LRBA KO decreases cell survival (Figure 2E) and LRBA KD causes more Raji cells death upon infections (Figure 2D, right panels). These contradictory results and the highly variable symptoms present in LRBA-deficient patients indicate complex interactions in multiple genes and environmental inputs. Indeed, LRBA interacts with multiple genes such as EGFR, NOTCH and RAS/MAPK [3-5], and LRBA KD increases phospho-NFkB levels and NFkB activity (Figures 2A and 2B). Correspondingly, the levels of TNF and IL10, NFkB target gene products, are increased (Figures 2C-1 and 2C-2). On another hand, NFkB potentially regulates LRBA [4,50]. NFkB is the central regulator of the immune system and plays a pivotal role in cell survival [4,51,52]. Without being limited by theory, the inventors propose that NFKB1 is an important modifier of LRBA and genetically interacts with the latter to affect cell survival, a key to understand the paradoxical association of immunodeficiency and autoimmunity, a unsolved fundamental question.

The CRISPR can generate up to 78% multiplex targeted mutations in mice in about three months [53-59,60]. It causes a DNA double strand break (DSB), which can be repaired by homology-directed-repair (HDR) [61,62]. A DNA fragment flanked by two homologous sequences can be inserted into the DSB cleaved site by homologous recombination, the efficiency of which can be 5000 times higher than traditional homologous recombination [63]. Due to these advantages, the CRISPR/HDR technique will be used to generate the Genebow mouse model (Figure 1). It requires a single guide RNA (sgRNA) and a target repair template for each gene.

BRIEF DESCRIPTION OF THE DRAWINGS

The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be

provided by the Patent and Trademark Office upon request and payment of the necessary fee.

Figure 1. Genebow: the nine genotypes of the two genes are labeled with four different FPs in a single mouse. Two transcription control FP cassettes (TCFP, shaded) and the Cre gene will be inserted into the gene *A*, gene *B* and the *Rosa26* loci, respectively by the CRISPR technique. FP2 and FP4 co-express with gene *A* and gene *B*, respectively, while FP1 and FP3 “co-express” with the STOP, which stops downstream gene expression. The Cre partial cleavage of the two loxP sites (triangle) will result in nine genotypes and cells with specific FP labeling. Genebow can also detect whether the promoter of a gene has activity. For example, if neither of the two FP genes knocked into the gene locus is expressed, it indicates that the promoter of the gene is not active. Oval represents cells. *A*, *B*, wild type alleles; *a*, *b*, knockout alleles. Arrows denote gene expression. X=No translation.

Figures 2A, 2B, 2C-1, 2C-2, 2D, and 2E. LRBA interacts with NFkB and regulates cell survival. Figure 2A: LRBA was stably knocked down in Raji cells by the short hairpin (sh) RNA technique in Raji B cells. NFkB activity was measured by an NFkB luciferase reporter. Figure 2B: The phospho-NFkB (pS529) levels were detected by Phosflow Kit. Figures 2C-1 and 2C-2: TNF α (Figure 2C-1) and IL10 (Figure 2C-2) proteins levels were measured by the human Th1/Th2 cytokine bead-array kit. Figure 2D: Apoptosis assay. Raji B cells were treated with various stress stimuli including proteasome inhibitors, serum starvation and bacterial infection. Apoptosis was detected by Annexin V-PE and 7-AAD. Figure 2E: LRBA (-/-) ES cell clones were obtained by culturing LRBA (+/-) ES cells with 1mg/ml G418 for ~ two weeks. Equal number of ES cells was plated into each well of a 96-well plate. The average cell numbers were counted from three wells at each time point.

Figures 3A-3C. Figure 3A: High throughput construction of target templates (SEQ ID NOs:8-11). The left and right arms (each 1 kb) that end or start with the sequences shown will be PCR amplified from the C57Bl/6J genomic DNA. The target template will be assembled by the enzymatic ligation assisted by nucleases (ELAN) via simultaneous ligation and digestion, in the presence of gel purified TCFP, both arms, BsmBI, with T4 DNA ligase and ATP added to the mixture following the protocol [2]. BsmBI will produce ends of the left and right arms compatible to the MluI and SalI stick

ends at the ends of the TCFP cassette. This one-tube preparative-scale assembly will produce enough DNA template for CRISPR targeting, which requires a 200 ng/ml DNA template. Thus, cloning is not necessary. Figures 3B and 3C: Excitation and emission spectra of the four FPs. Four excitation lasers (thick lines) and detection filter sets (open bars) are shown. The percentages are excitation efficiencies for each dye with the designed laser. There is no detection overlap between these FPs. The LIVE/DEAD Yellow and R-PE-Cy7 dyes, used for dead cell and annexin V labeling, also will not overlap with any FP after compensation. For clarity, their spectra are not shown.

Figures 4A-4I. Construction of CRISPR/HDR vectors and CRISPR cleavage test. The target vector is constructed from a 12 kb fragment subcloned from a C57BL/6J BAC clone using BAC subcloning kit (Gene Bridges). The TCFP cassette splits the genomic fragment into 3 kb left and 9kb right arms. Figure 4A: The *Lrba* target vector was digested with restriction enzymes. Figure 4B: PCR Genotyping test of *Lrba* KO using the primers from the TCFP cassette and the genomic DNA sequence outside of the short arm. Figure 4C: Flow cytometry detection of the three FP in H293 cells. Figure 4D: A secretable luciferase gene was inserted in place of *Lrba* through HR to functionally test the TCFP. The Luciferase assay was conducted in H293 cells transfected with the target vector with and without Dox. Without induction, the luciferase activity is negligible. Figure 4E: *SpeI* digestion of the PCR product using genomic DNA from mouse ES cells transfected with *Lrba* sgRNA/Cas9 vector. The white arrow indicates the mutated alleles (*SpeI* destroyed). The sgRNA target site in the PCR product, 713bp, contains a *SpeI* site. *SpeI* digestion produces 527bp and 186bp fragments in wt but cannot cut the CRISPR targeted allele. Figure 4F: Cas9/sgRNA can specifically digest the *Lrba* BAC subclone plasmid as expected. Portions of sequencing trace showing sequences of *Lrba*/SD/loxP (Figure 4G) (SEQ ID NO:12), STOP/loxP (Figure 4H) (SEQ ID NO:13), and Aqua/P2A/*Lrba* (Figure 4I) (SEQ ID NO:14).

Figures 5A-5C. Figure 5A: TetOn and inducible Cre system. The overexpression of gene A/B can be induced by Dox while TAM activates Cre to cleave the STOP and restore the expression of gene A/B. Figure 5B: Protocol for studying *Lrba* and *Nfkb1* genes. A concentration of 20 µg/ml Dox in 2% sucrose will be supplied in the drinking water for the Genebow mice >6 wks of age for three days (red line (rtTA)) to overexpress *Lrba* and *Nfkb1*. The withdrawal of Dox will turn off the expression of the two genes.

Administration of Tam by i.p. will activate Cre (green line). Peripheral blood will be collected as indicated from the mouse before and after each treatment and subjected to flow cytometry assay. ++wt, +-het, --ko, +++overexpression. B: bleeding. Figure 5C: Timeline.

5 **Figure 6.** Genebow: the nine genotypes of the two genes are labeled with two different FPs that localize to two different subcellular locations in a single mouse. Two transcription control FP cassettes (TCFP, shaded) and the Cre gene will be inserted into the gene *A*, gene *B* and the *Rosa26* loci, respectively by the CRISPR technique. The FP1 and FP4 are separately tagged with a nuclear localization sequence (N) and will be
10 localized in the nucleus. While the FP2 and FP3 are separately tagged (M) with a membrane localization sequence and will be localized in the cell membrane, FP2 and FP4 co-express with gene *A* and gene *B*, respectively, while FP1 and FP3 “co-express” with the STOP, which stops downstream gene expression. The Cre partial cleavage of the two loxP sites (triangle) will result in nine genotypes and cells with specific FP labeling,
15 which can be detected by confocal fluorescent microscopy and image flow cytometry. Oval represents cells. *A*, *B*, wild type alleles; *a*, *b*, knockout alleles. Arrows denote gene expression. X=No translation.

Figures 7A-7E. Comparison of Genebow and traditional method, and preliminary data. Figure 7A: Traditionally, cells from each mouse (genotype) have to be analyzed
20 separately, because if mixed, they cannot be distinguished. In contrast, Genebow can distinguish each genotype in a mixture. Consequently, e.g., to quantitate T cells of different genotypes, traditional method usually requires at least 27 mice and 27 tubes. While the Genebow requires only one mice and one tube. More importantly, the phenotyping sensitivity can be greatly increased, e.g., 1000 times (2/0.002. Bottom, an
25 example). Calculations [112] were based on hypothetical averages and variations.*Total number required to obtain statistically significant data ($p < 0.05$, power=80%). Nine genotypes are color-labeled. Wild type: uppercase, knockout: lowercase. Figure 7B: LRBA KD deregulates MAPKs. LRBA was stably knocked down (KD) in Raji cells by the short hairpin RNA technique. Cell lysates from WT or LRBA KD stable Raji cells
30 were analyzed by Western blot. Protein levels were quantified with Quantity One software. Phosphorylated (P-) protein levels were normalized to β -Actin. Figure 7C: NFkB activity was measured by an NFkB luciferase reporter. Figure 7D: NFkB binding

sites in human LRBA promoter and LRBA promoter assay (SEQ ID NOs:15-18). H1299 cells were transfected with the same amount of plasmid DNA expressing luciferase under LRBA promoter (LRBApLuci), NFkB p50 (P50) and RelB. DNA is an irrelevant plasmid DNA used to ensure equivalent amount of DNA in each transfection. RLU, relative luciferase unit. ****, $p < 5.8E-5$. Figure 7E: CRISPR knock-in in H1299 cells with short homologous arms (1kb) of LRBA. Primers from the genomic sequence outside of homologous arm and exogenous sequence in the knockin vector were used for the PCR screening. The predicted size of the PCR product for positive clones is 1.4 kb.

Figures 8A-8C. Principle of Genebow and Cloning strategy. Figure 8A: (SEQ ID NOs:19-30). Cre recombinase-mediated inversion. loxP (locus of crossover in P1) consists of two 13 bp inverted repeats flanking an 8 bp core sequence that determines its polarity (denoted by dashed arrows). Cre recombinase cleaves (small vertical arrows) the two inverted loxP sites. The cleavage produces two complementary protruding sequences: 5'-cataca at the ends of the intervening DNA, and 3'-tgtatg at the ends of the flanking DNA. The cleaved intervening DNA is ligated back in an inverted orientation. As the sequences of the two loxP sites and the size of the intervening DNA remain unchanged, the reaction is reversible and the reaction rates of forward and reverse reactions should be equal. In a balanced reaction, the numbers of the two end products, the original (i) and the inverted (ii), are equal. Figure 8B: Genotype-switching and labeling cassettes (GSLC). The first exon of *Lrba* is flanked with two inverted fluorescent protein genes. The green FP† and *Lrba* are coexpressed. Cre-mediated inversion inactivates the co-expression but activates red FP expression. The wt allele is converted to ko allele and the green FP is replaced by red FP. Correspondingly, as each cell has two alleles, ko cells (with two ko alleles) will be labeled as red FP, wt cells as green FP, and heterozygous (het) cells as red/green FPs, and can be distinguished by flow cytometry. As stated in Figure 8A, the numbers of wt and ko alleles should be equal. This is important, as Cre recombination will not be a variable confound the interpretations of results that are affected by genotype frequency. After the Cre reaction is balanced, Cre has to be removed. This can be easily done by stopping tamoxifen (TAM) treatment in the TAM-induced Cre system. Otherwise the cleavage/inversion will continue forever, and the wt/ko alleles in the cells will be switched back and forth. †mNeongreen (NeonG), mOrange2, mKate2, and TagBFP2 have relative brightnesses over 35.23 EGFP and mRFP1 have a relative

brightness of 36 and 13 have been successfully used to label cells [122]. Therefore, the inventors will use these four fluorescent proteins to label the two genes. The flow cytometer can detect 16 colors. Since four channels will be used by the four FPs, there are 12 channels available to conduct almost any multiplex flow cytometry that is routinely used in research. Figure 8C: Cloning strategy of donor vectors. The *Lrba* homologous arms (~2kb) will be PCR amplified and cloned into the pJet vector using the CloneJET PCR Cloning Kit. The GSLC1 (2kb) with short homologous sequences of *Lrba* at both ends will be synthesized by Genscript, and cloned into the pJetHR-*Lrba* vectors by using the CRISPR cloning technique [118].

Figures 9A-9C. Characterization of Cre-mediated inversion by restriction enzyme digestion, and Time Line (Figure 9C). The *ApaI/BstI* digestion will produce a 532 bp fragment when the intervening DNA is in original orientation (Figure 9A), or a 2444 bp fragment when the intervening DNA is in inverted orientation (Figure 9B). This assay will be used to analyze Cre-mediated inversion in bacteria detected by electrophoresis and in ES cells detected by Southern blot.

Figure 10. Comparison of Genebow and traditional method. Traditionally, cells from each mouse (genotype) have to be analyzed separately, because if mixed, they cannot be distinguished. In contrast, Genebow can distinguish each genotype in a mixture. Consequently, *e.g.*, to quantitate Tregs of different genotypes, the traditional method usually requires at least 27 mice and 27 tubes, while the Genebow requires only one mouse and one tube. More importantly, with greatly increased sample size and eliminated variations, the phenotyping sensitivity can be greatly increased. Nine genotypes are color-labeled. A=*Lrba*, B=*Nfkb1*, wild type: uppercase, knockout: lowercase.

Figures 11A-11E. LRBA interacts with important genes in regulating apoptosis and proliferation. LRBA was stably knocked down (KD) in Raji cells by the short hairpin RNA technique. Figure 11A: NFkB activity was measured by an NFkB luciferase reporter. The Phosphorylated (P-) NFkB (pS529) levels were detected by Phosflow Kit. Figure 11B: Apoptosis assay. Raji B cells were treated with PS-341, an NFkB inhibitor. Figure 11C: NFkB binding sites in human LRBA promoter and LRBA promoter assay (SEQ ID NOs:15-18). H1299 cells were transfected with the same amount of plasmid DNA expressing luciferase under LRBA promoter (LRBApLuci), NFkB p52 (P52) and

RelB. DNA is an irrelevant plasmid DNA used to ensure equivalent amount of DNA in each transfection. RLU, relative luciferase unit. ****, $p < 5.8E-5$. Figure 11D: LRBA^{-/-} ES cell clones were obtained by culturing LRBA^{+/-} ES cells with 1mg/ml G418 for ~ two weeks. Equal numbers of ES cells were plated into each well of a 96-well plate. The average cell numbers were counted from three wells at each time point. Figure 11E: LRBA KD deregulates MAPKs. Cell lysates from WT or LRBA KD stable Raji cells were analyzed by Western blot. Protein levels were quantified with Quantity One software. Phosphorylated (P-) protein levels were normalized to β -Actin.

Figures 12A-12F. CRISPR vector construction and CRISPR targeting of mouse embryonic stem (ES) cells and zygotes. The target vector was constructed from a 12.5 kb genomic fragment and has 3.5 kb left and 9 kb right arms flanking the transcription control and labeling cassette (TCLC). Figure 12A. Portions of sequencing trace show the sequences at the 5' junction (Figure 12A) (SEQ ID NO:31) and 3' junction (Figure 12B) (SEQ ID NO.14) of the TCLC cassette insertion. Figure 12C: PCR screening for *Lrba* knockin-positive ES clones. Primers from the genomic sequence outside of homologous arm and exogenous sequence in the knockin vector were used for the PCR screening. The predicted size of the PCR product is 3.5 kb. Figure 12D: Southern blot. EcoRV-digested genomic DNAs were hybridized with a 5' external probe with expected fragment sizes of 5 kb (wt) and 13 kb (targeted). 1. Linearized target vector (22 kb). 2-11: G418-positive clones. #4 and #9 incorrectly targeted. #6 Wt, other clones are correctly targeted heterozygotes. The results agree with the PCR results. Figure 12E: Nested PCR detection of correctly targeted blastocysts derived from zygotes injected with CRISPR reagents. One quarter of DNA from a blastocyst was used for the first PCR (Figure 12C). Figure 12F: CRISPR knock-in with short homologous arms (1kb). Primers were designed as in Figure 12C. The predicted size of the PCR product is 1.4 kb.

Figures 13A-13C. Principle of Genebow. Figure 13A: The structure of TCLCs is the same as that of Brainbow 3.1 [158], except a porcine teschovirus-1 2A (P2A[†]) peptide DNA sequence is inserted between the last FR gene and the targeted genes just before the translation codon to co-express the two genes in links. CFP=Aequamarine, which is superior to the popular form ECFP [166]; RFP=iRFP670, a near-infrared FP [167]. lox sites: loxP (P), loxN (N), lox2272 (2), lox 5171 (5). pA, SV40 PolyA sequence is used to stop the transcription of downstream genes. [†]Due to its small size, high self-cleavability,

and ability to produce an equal molar ratio of the two proteins, P2A is widely used to link two proteins to be co-expressed by a single promoter [168]. Figure 13B: Cre recombinase-mediated recombination solely removes DNA between two identical lox sites. The recombination between the two loxP sites turns on the FR1 $\ddagger$, producing the ko allele. While the recombination between the two Lox2272 sites turns on CFP, producing the wt allele. $\ddagger$ **Fluorescent proteins** are the most commonly used reporters as best exemplified by the Brainbow [158,162]. Aquamarine and iRFP670 will be used as their spectra are well separated and do not overlap with those of commonly used fluorophores. mRuby2, mKate2, Aqua and iRFP670 with a relative brightness of 46, 35, 23 and 13, respectively [169]. EGFP and mRFP1 with a relative brightness of 36 and 13 have been successfully used to label cells [170]. Therefore, the inventors will use four fluorescent proteins to label the two genes. The flow cytometer can detect 16 FCs. Since four channels will be used by the four FRs, there are 12 channels available to conduct almost any multiplex flow cytometry that is routinely used in research. Figure 13C: Comparison of phenotyping sensitivity. In this example, phenotyping sensitivity can be greatly increased by 1000 times (2/0.002) in the Genebow model. Calculations [171] were based on hypothetic averages and variations. *Total number required to obtain statistically significant data ($p < 0.05$, power=80%) for the 9 genotypes.

Figures 14A-14D. LRBA's role in the development and function of Tregs in autoimmune diabetes mouse models. Figure 14A: Cloning strategy of donor vectors. The homologous arms (~2kb) for each gene will be PCR amplified and cloned into the pJet vector. TCLC1 or 2 (~3kb) with short homologous sequences of *Lrba* or *Nfkb1* at both ends will be synthesized, and cloned into the corresponding vectors that have homologous sequences for targeting *Lrba* and *Nfkb1* by using the inventors' CRISPR cloning techniques [139]. Figure 14B: Generation of mouse model and mating scheme. Genebow mouse model will be generated by triplex CRISPR targeting. Foxp3.DTR+ mice express human diphtheria toxin receptor from the Foxp3 promoter from BAC transgenesis--without disrupting expression of the endogenous Foxp3 gene. Diphtheria toxin (DT) administration results in ablation of Treg cells. DT administration results in ablation of Treg cells 2 days after injection. Neonates injected with DT daily all die from autoimmune disease within 27 days of birth. *Lrba* and *Nfkb1* are 49 Mb away on chromosome 3, CreERT2/Rosa 26 on chromosome 6, and FoxP3GFP on chromosome X.

Figure 14C: Tregs play a major role in regulating the production of antibody and autoantibody. Tfrs are required to produce high affinity antibodies, and to limit autoantibody production. Depletion of Tfrs leads to autoantibody overproduction and antibodies with low affinity, resulting in autoimmune disease and immunodeficiency, respectively.

Figure 15. One-step generation of Genebow mouse model by CRISPR technique and subsequent mating strategy. Two transcription control and labeling cassettes (TCLCs) will be inserted into the genomic loci of *Lrba* and *Nfkb1*, respectively, and the CreERT2 into the Rosa26 locus by using the CRISPR technique. Thus, Triple CRISPR targeting will produce Genebow and CreERT2 transgene mice. These genes are on different chromosomes, allowing free segregation and combination of these genes (*Lrba* and *Nfkb1* are 49 Mb away on chromosome 3, CreERT2/Rosa 26 on chromosome 6, and FoxP3GFP on chromosome X).

Figures 16A-16E. Genebow mouse model and genetic interaction study. Figure 16A: a transcription control and labeling cassette (TCLC) is inserted into the genomic locus of *Lrba* by using the CRISPR technique. This insertion turns off *Lrba* as the TCLC contains a STOP cassette but turns on a red FR (RFR) as it places the RFR under the *Lrba* promoter. For clarity, only one pair of lox sites is shown. Figure 16B: tamoxifen (TAM) will be administered to neonates via lactating mother intraperitoneally injected with tamoxifen at optimal dosage and time. Figure 16C: TAM activates Cre recombinase, which removes the RFR-STOP between the two loxP sites (triangle) and turns on *Lrba* and switches the RFR to blue FR (BFR). Thus, the on and off of *Lrba* can be labeled separately with two FRs. Cre recombination will result in three genotypes: knockout (ko, red), heterozygote (het, blue/red) and wild type (wt, blue). Similarly, another gene can be labeled with two other FRs (purple/green). The 9 genotypes of two genes thus can be specifically labelled with four FRs. X=no expression. Figure 16D: Tregs with nine genotypes will be sorted and 3.6×10^5 of mixed (4×10^4 /genotype) Tregs will be injected ip into Treg-depleted non-obese diabetic (NOD) neonates. For depletion of Treg cells, Foxp3-DTR+ mice or control littermates will be injected ip with DT every other day for 10 days. Figure 16E: Tregs with nine genotypes will be sorted and 3.6×10^5 of Tregs for each genotype from Genebow mice will be injected ip into a Treg-depleted non-obese

diabetic (NOD) neonate, and then followed for manifestations of autoimmune disease up to 100 days.

BRIEF DESCRIPTION OF THE SEQUENCES

- 5 **SEQ ID NO:1:** *Mus musculus* LPS-responsive beige-like anchor (Lrba), transcript variant 2, mRNA. Accession number: NM_001077688.
- SEQ ID NO:2:** *Mus musculus* LPS-responsive beige-like anchor (Lrba), transcript variant 1, mRNA. Accession number: NM_030695.
- SEQ ID NO:3:** *Mus musculus* LPS-responsive beige-like anchor (Lrba), transcript variant
10 3, mRNA. Accession number: NM_001077687.
- SEQ ID NO:4:** *Homo sapiens* LPS-responsive vesicle trafficking, beach and anchor containing (LRBA), transcript variant 1, mRNA. Accession number: NM_001199282.
- SEQ ID NO:5:** *Homo sapiens* LPS-responsive vesicle trafficking, beach and anchor containing (LRBA), transcript variant 2, mRNA. Accession number: NM_006726.
- 15 **SEQ ID NO:6:** *Mus musculus* nuclear factor of kappa light polypeptide gene enhancer in B cells 1, p105 (Nfkb1), mRNA. Accession number: NM_008689.
- SEQ ID NO:7:** Human nuclear factor kappa-B DNA binding subunit (NF-kappa-B) mRNA, complete cds. Accession number: M58603.
- SEQ ID NOs:8-11:** target template sequences (Figure 3A).
- 20 **SEQ ID NO:12:** Lrba/SD/loxP portion of Lrba BAC subclone plasmid (Figure 4G).
- SEQ ID NO:13** STOP/loxP portion of Lrba BAC subclone plasmid (Figure 4H).
- SEQ ID NO:14:** Aqua/P2A/Lrba portion of Lrba BAC subclone plasmid (Figure 4I).
- SEQ ID NOs:15-18:** NFkB binding site of LRBA promoter (Figure 7D).
- SEQ ID NOs:19-30:** Locus of Cre recombinase-mediated crossover (Figure 8A).
- 25 **SEQ ID NO:31:** Sequence at 5' junction of TCLC cassette insertion (Figure 12A).

DETAILED DESCRIPTION OF THE INVENTION

An aspect of the invention concerns non-human animal models useful, for example, in studying gene-gene interactions, wherein the animal model comprises nine
30 genotypes of two genes specifically labeled with four distinct detectable labels. In some embodiments, the two genes comprise a wild-type first gene (A), a wild-type second gene

(B), an inactivated form of the first gene (a), and an inactivated form of the second gene (b), wherein the four distinct detectable labels comprise:

- a first label co-expressed with the wild-type first gene (A),
- a second label co-expressed with the inactivated form of the first gene (a),
- 5 a third label co-expressed with the wild-type second gene (B), and
- a fourth label co-expressed with the inactivated form of the second gene (b).

In some embodiments, the animal model comprises:

- (a) a first genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the wild-type second gene (BB), wherein the wild-type first gene (A) is co-expressed with the first label and the wild-type second gene (B) is co-expressed with the third label;

- (b) a second genotype comprising homozygous alleles for the wild-type first gene (AA), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the first gene (A) is co-expressed with the first label, the second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

- (c) a third genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the inactivated form of the second gene (bb), wherein the homozygous alleles for the wild-type first gene (AA) are co-expressed with the first label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

- (d) a fourth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the wild-type second gene (BB), wherein allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed with the second label, and the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

- (e) a fifth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, wherein the allele

for the wild-type second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(f) a sixth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the inactivated version of the second gene (bb), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(g) a seventh genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the wild-type second gene (BB), wherein the homozygous alleles for the inactivated version of the first gene (aa) are co-expressed with the second label, and wherein the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(h) an eighth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the one allele for the inactivated form of the second gene (b) is co-expressed with the fourth label; and

(i) a ninth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the inactivated version of the second gene (bb), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, and wherein the homozygous alleles for the inactivated version of the second gene (bb) is co-expressed with the fourth label.

Genes may be inactivated by various methods, such as insertion, deletion, substitution, and/or recombination.

The first and second genes may be any genes of interest. Any two genes that at least potentially interact with each other either, physically or genetically, are suitable to be selected and studied using this animal model. In some embodiments, the first gene is *Lrba* and the second gene is *NFkB*, or vice-versa. These two genes are used as examples herein. Gene pairs that are important in human diseases are particularly advantageous to

use in the animal model in place of Lrba and NFkB, include, for example, P53 and MDM2, NFkB and NOTCH1, EGFR and AKT1 gene pairs.

In some embodiments, the detectable labels are fluorescent labels or luminescent labels.

5 In some embodiments, the animal model is a rodent, such as a mouse. In some embodiments, the animal model is a primate.

Another aspect of the invention includes a method for studying gene-gene interaction, comprising: providing a non-human animal model of the invention; and analyzing the interaction of the two genes specifically labeled with four distinct
10 detectable labels. In some embodiments of the method, the analysis is carried out by analyzing the characteristics and/or behavior of the animal model or one or more cell populations of the animal model. In some embodiments, the analysis includes subjecting cells of the animal model to flow cytometry. In some embodiments of the method, the analysis is carried out in the presence and/or absence of an exogenous agent administered
15 to the animal model.

In some embodiments of the method, the analysis includes measuring or otherwise detecting a detectable label (*e.g.*, one label, some of the labels, or all of the labels) of one of the genes (the label signal) and, optionally, comparing the measured detectable label (signal) to that of the other gene. Depending upon the type of label, the signal may be
20 detected and optionally quantitatively measured using the appropriate modality for the type of label.

In some embodiments of the method, the two genes comprise a wild-type first gene (A), a wild-type second gene (B), an inactivated form of the first gene (a), and an inactivated form of the second gene (b), wherein the four distinct detectable labels
25 comprise:

- a first label co-expressed with the wild-type first gene (A),
- a second label co-expressed with the inactivated form of the first gene (a),
- a third label co-expressed with the wild-type second gene (B), and
- a fourth label co-expressed with the inactivated form of the second gene (b); and

30 wherein the analyzing comprises measuring one or more of the first label, second label, third label, or fourth label.

In some embodiments of the method, the animal comprises:

(a) a first genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the wild-type second gene (BB), wherein the wild-type first gene (A) is co-expressed with the first label and the wild-type second gene (B) is co-expressed with the third label;

5 (b) a second genotype comprising homozygous alleles for the wild-type first gene (AA), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the first gene (A) is co-expressed with the first label, the second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

10 (c) a third genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the inactivated form of the second gene (bb), wherein the homozygous alleles for the wild-type first gene (AA) are co-expressed with the first label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

15 (d) a fourth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the wild-type second gene (BB), wherein allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed with the second label, and the homozygous alleles for the wild-type second
20 gene (BB) are co-expressed with the third label;

(e) a fifth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the
25 inactivated version of the first gene (a) is co-expressed second label, wherein the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(f) a sixth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the
30 inactivated version of the second gene (bb), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first

gene (a) is co-expressed second label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(g) a seventh genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the wild-type second gene (BB),
5 wherein the homozygous alleles for the inactivated version of the first gene (aa) are co-expressed with the second label, and wherein the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(h) an eighth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), one allele for the wild-type second gene (B), and one allele for the
10 inactivated form of the second gene (b), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the one allele for the inactivated form of the second gene (b) is co-expressed with the fourth label; and

(i) a ninth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the inactivated version of the second gene (bb), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, and wherein the homozygous alleles for the inactivated version of the second gene (bb) is co-expressed with the fourth label; and

20 wherein the analyzing includes measuring one or more of the first label, second label, third label, or fourth label (*i.e.*, one label, two labels, three labels, or all four labels).

Another aspect of the invention includes a method for studying gene-gene interaction *in vitro*, comprising: providing a population of cells (*ex vivo*) from the non-human animal model of the invention (the cells may be in isolated form, or as a tissue, for
25 example); and analyzing the interaction of the two genes specifically labeled with four distinct detectable labels. In some embodiments of the method, the analyzing comprises analyzing the characteristics and/or behavior of the cells. In some embodiments of the method, the analyzing comprises subjecting cells of the animal model to flow cytometry.

In some embodiments of the method, the analyzing is carried out in the presence
30 and/or absence of an exogenous agent administered to the animal *in vivo* or brought into contact with the cells *in vitro*.

In some embodiments of the method, the analyzing comprises measuring or otherwise detecting the detectable label of one of the genes and, optionally, comparing the measured detectable label to that of the other gene. Depending upon the type of label, the signal may be detected and optionally quantitatively measured using the appropriate modality for the type of label.

In some embodiments of the method, the two genes comprise a wild-type first gene (A), a wild-type second gene (B), an inactivated form of the first gene (a), and an inactivated form of the second gene (b), wherein the four distinct detectable labels comprise:

a first label co-expressed with the wild-type first gene (A),
a second label co-expressed with the inactivated form of the first gene (a),
a third label co-expressed with the wild-type second gene (B), and
a fourth label co-expressed with the inactivated form of the second gene (b); and
wherein said analyzing comprises measuring one or more of the first label, second label, third label, or fourth label (*i.e.*, one, two, three, or all four labels).

In some embodiments of the method, the animal comprises:

(a) a first genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the wild-type second gene (BB), wherein the wild-type first gene (A) is co-expressed with the first label and the wild-type second gene (B) is co-expressed with the third label;

(b) a second genotype comprising homozygous alleles for the wild-type first gene (AA), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the first gene (A) is co-expressed with the first label, the second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(c) a third genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the inactivated form of the second gene (bb), wherein the homozygous alleles for the wild-type first gene (AA) are co-expressed with the first label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(d) a fourth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the

wild-type second gene (BB), wherein allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed with the second label, and the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

5 (e) a fifth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, wherein the allele
10 for the wild-type second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(f) a sixth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the inactivated version of the second gene (bb), wherein the allele for the wild-type first gene
15 (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(g) a seventh genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the wild-type second gene (BB),
20 wherein the homozygous alleles for the inactivated version of the first gene (aa) are co-expressed with the second label, and wherein the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(h) an eighth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), one allele for the wild-type second gene (B), and one allele for the
25 inactivated form of the second gene (b), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the one allele for the inactivated form of the second gene (b) is co-expressed with the fourth label; and

30 (i) a ninth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the inactivated version of the second gene (bb), wherein the homozygous alleles for the inactivated version of the first gene (aa) is

co-expressed with the second label, and wherein the homozygous alleles for the inactivated version of the second gene (bb) is co-expressed with the fourth label; and

wherein the analyzing comprises measuring one or more of the first label, second label, third label, or fourth label (*i.e.*, one, two, three, or all four labels).

5

Animals

The non-human animal model is preferably a mammal. For example, the animal model may be a rodent or non-human primate.

10 In some embodiments, the animal is selected from the group consisting of a mouse, rat, guinea pig, hamster, gerbil, pig, cow, dog, wolf, coyote, jackel, and cat. In some embodiments, the animal model is a monkey or ape. In some embodiments, the animal model is a primate selected from the group consisting of a macaque, marmoset, tamarin, spider monkey, vervet monkey, squirrel monkey, and baboon. In some
15 embodiments, the animal model is an ape selected from the group consisting of a gorilla, chimpanzee, orangutan, and gibbon. The animal model may be a hybrid of two non-human animals (*e.g.*, dog-wolf).

The animal model may have any desired genetic background. The animals may be crossed with many strains, and the genes may be studied with a wide genetic background, which is desired to recapitulate the complexity of human disease. As
20 controls are not required, establishing isogenic strains are not necessary. The animal model of the invention can be crossed with many strains simultaneously and the function of a gene can be studied in a wide genetic background, which is desirable for recapitulating human complex diseases and can save time and resources.

The animal may be further modified at the genetic or epigenetic level so as to be
25 useful in modeling a particular disease, such as cancer, cardiovascular disease, a metabolic disease such as diabetes, or a monogenic disease. For example, the animal model may be further modified to model Down Syndrome, cystic fibrosis, cancer, glaucoma, type-I diabetes, type-II diabetes, epilepsy, heart disease, muscular dystrophy, or gynecological tumors.

30 Down Syndrome - One of the most common genetic birth defects in humans, occurring once in every 800 to 1,000 live births, Down syndrome results from an extra copy of chromosome 21, an abnormality known as trisomy. The Ts65Dn mouse,

developed at The Jackson Laboratory, mimics trisomy 21 and exhibits many of the behavioral, learning, and physiological defects associated with the syndrome in humans, including mental deficits, small size, obesity, hydrocephalus and thymic defects. This model represents the latest and best improvement of Down syndrome models to facilitate research into the human condition.

Cystic Fibrosis (CF) - The Cftr knockout mouse has helped advance research into cystic fibrosis, the most common fatal genetic disease in the United States today, occurring in approximately one of every 3,300 live births. Scientists now know that CF is caused by a small defect in the gene that manufactures CFTR, a protein that regulates the passage of salts and water in and out of cells. Studies with the Cftr knockout have shown that the disease results from a failure to clear certain bacteria from the lung, which leads to mucus retention and subsequent lung disease. These mice have become models for developing new approaches to correct the CF defect and cure the disease.

Cancer - The p53 knockout mouse has a disabled Trp53 tumor suppressor gene that makes it highly susceptible to various cancers, including lymphomas and osteosarcomas. The mouse has emerged as an important model for human Li-Fraumeni syndrome, a form of familial breast cancer.

Glaucoma - The DBA/2J mouse exhibits many of the symptoms that are often associated with human glaucoma, including elevated intraocular pressure. Glaucoma is a debilitating eye disease that is the second leading cause of blindness in the United States.

Type 1 Diabetes - This autoimmune disease, also known as Juvenile Diabetes, or Insulin Dependent Diabetes Mellitus (IDDM), accounts for up to 10 percent of diabetes cases. Non-obese Diabetic (NOD) mice are enabling researchers to identify IDDM susceptibility genes and disease mechanisms.

Type 2 Diabetes - A metabolic disorder also called Non-Insulin Dependent Diabetes Mellitus (NIDDM), this is the most common form of diabetes and occurs primarily after age 40. The leading mouse models for NIDDM and obesity research were all developed at The Jackson Laboratory: Cpefat, Lepob, Leprdb and tub.

Epilepsy - The “slow-wave epilepsy,” or swe, mouse is the only model to exhibit both of the two major forms of epilepsy: petit mal (absence) and grand mal (convulsive). It shows particular promise for research into absence seizures, which occur most often in children.

Heart Disease - Elevated blood cholesterol levels and plaque buildup in arteries within three months of birth (even on a low-fat diet) are characteristics of several experimental models for human atherosclerosis: the Apoe knockout mouse and C57BL/6J.

- 5 Muscular Dystrophy - The Dmd mdx mouse is a model for Duchenne Muscular Dystrophy, a rare neuromuscular disorder in young males that is inherited as an X-linked recessive trait and results in progressive muscle degeneration.

Ovarian Tumors - The SWR and SWXJ mouse models provide excellent research platforms for studying the genetic basis of ovarian granulosa cell tumors, a common and
10 very serious form of malignant ovarian tumor in young girls and post-menopausal women.

Reporters

The various genotypes are labeled with a detectable label (also referred to herein
15 as a reporter) in order to detect and track promoter activity and gene expression, *e.g.*, by flow cytometry. Thus, the reporter gene and the first gene and second gene are operably linked such that they are co-expressed.

Such detectable labels are known in the art, and include, for example, fluorescent reporter proteins encoded by fluorescent reporter genes. Preferably, each reporter can be
20 detected in a living animal. Thus, the amount, distribution, proliferation, movement, properties, and behavior of the labeled cells can thus be assessed and, optionally, monitored.

Exemplary reporters include light-emitting reporters, such as fluorescent and luminescent reporters. Polypeptides that result in the generation of light in a living
25 organism (bioluminescence) include, but are not limited to, various luciferases, green fluorescent protein (GFP), yellow fluorescent protein (YFP) and aequorin (Wilson and Hastings, *Annu. Rev. Cell Dev. Biol.*, 1998, 14:197-230). Fluorescence reporters have many diverse uses, the most common of which are for fluorescence microscopy and also for flow cytometry. In both cases internal expression of the fluorescence reporter, using a
30 reporter plasmid system, allows simple assessment of cell properties and/or behavior.

Luciferase is a luminescent molecule, and thus does not require excitation in order to generate light. It does typically require a substrate (*e.g.*, luciferin, an aldehyde or coelenterazine), an energy source (*e.g.*, ATP) and oxygen. In the case of bacterial

luciferases, the genes encoding the substrate can be supplied the same vector as the gene(s) encoding the luciferase enzyme, thus eliminating the need for exogenously-supplied substrate (see, *e.g.*, U.S. Patent No. 5,650,135).

5 In some embodiments, the reporter is selected from the group consisting of red fluorescent protein (RFP), green fluorescence protein (GFP), yellow fluorescence protein (YFP), and cyano fluorescence protein (CFP). Polynucleotide cassettes encoding such polypeptides may be transfected into the target site as extra-chromosomal genetic elements (*e.g.*, plasmids) or are stably incorporated into the genome (*e.g.*, “hopped” in using, for example, a transposon).

10 If the reporter is a light-emitting reporter, method of measurement include using a photon detection device, such as an intensified CCD camera, a cooled CCD camera, or any other photon detection device with a high sensitivity. However, other methods may be used. For example, a light-emitting reporter may also be detected using a sensitive luminometer; a radioactive reporter may be detected by counts, X-ray imaging
15 or scintillation.

The term “operatively linked” or “operably linked” refers to the connection of elements being a part of a functional unit such as a gene or an open reading frame (*e.g.*, encoding LRBA). Accordingly, by operatively linking a promoter to a nucleic acid sequence encoding a gene product such as a polypeptide the two elements becomes part
20 of the functional unit--a gene. The linking of the expression control sequence (promoter) to the nucleic acid sequence enables the transcription of the nucleic acid sequence directed by the promoter. By operatively linking two heterologous nucleic acid sequences encoding a polypeptide the sequences becomes part of the functional unit--an open reading frame encoding a protein or proteins comprising the amino acid sequences
25 encoded by the heterologous nucleic acid sequences. By operatively linking two coding sequences, the sequences can be co-expressed.

To alleviate the limitations of current animal model techniques, a method for studying gene-gene interactions in a single animal (*e.g.*, mouse) is provided (Figure 1): Genebow, coined from the Brainbow [8]. The method of the invention aims to achieve 1)
30 high sensitivity: phenotypes of the nine genotypes of the two genes from one mouse can be studied by flow cytometry in a single tube with high sensitivity due to the elimination of experimental variations; 2) congenic strain-independence: since within a single animal,

each genotype has the other eight genotypes as controls, establishing a congenic strain is not necessary; 3) high specificity: the gene expression can be turned on/off (*e.g.*, by the TetOn system). Phenotypes can be specifically attributed to the gene if they change in response to the switching gene expression on/off. The mutations of lipopolysaccharide-responsive, beach and anchor containing (LRBA) cause immunodeficiency and autoimmunity with highly variable symptoms and LRBA interacts with many crucial regulators including NFkB *in vitro* [3-5]. This animal model will permit study of how LRBA interacts with these regulators *in vivo* to better understand the immune system. As a proof of concept, *Lrba* and *Nfkb1* can be used for the model.

The Genebow is a revolutionary technique for rapid, specific and sensitive study of the gene-gene interaction in any animals over a wide genetic background, which is essential to recapitulate the genetic interactions of human diseases. The PI has cloned the LRBA gene, contributed fifteen GenBank entries and generated six mouse models including a conditional knockout (KO) model [9,10]. The PI has also developed a novel CRISPR cloning method [11]. The expertise will ensure the accomplishment of this proposal.

The GWA studies have demonstrated that many human common diseases, such as cancer, cardiovascular disease, diabetes and allergic disease, and even rare “monogenic” diseases, result from the complex interplay of many genes and environmental factors, and that individual gene contribution typically is small and can be masked by genetic and environmental variations [1,12-18]. The history of mouse genetics also makes it clear that the genetic background can cause considerable phenotypic variation in the mice with the same targeted gene [19]. These variables pose a great challenge in studying human disease gene-gene interactions. GWA studies have failed to establish common variant risk for the majority of common diseases [15,20]. Animal models are indispensable and powerful tools, tremendously increasing knowledge of gene function. However, animal techniques that can sensitively study gene-gene interactions over a wide genetic background required to decode complex human diseases have yet to be developed. Current animal models are dependent on establishing a congenic strain and on the presumptions that the animals within the strain are identical and that reproducible data can be obtained [6,7]. Unfortunately, these animals are still subject to genetic, epigenetic and environmental variations. These variations may stem from the “passenger DNA”

from the original targeted genome, epigenetic modifications of DNA, non-coding RNAs and RNA editing, and bacterial and viral infections [21-28]. All of these variations can play an important role in determining phenotype, which can result in high background noise, low sensitivity or even false data, *e.g.*, the phenotypes originally attributed to the targeted gene actually attribute to other genes [29-33]. Experimental protocol also can introduce significant variations [34-36]. Moreover, the genetic diversity of current mouse models is only a fraction of that presented in wild mice [37], as the classical inbred mouse strains are derived from the same ancestors. In addition, most gene-targeting studies have focused on one strain, which is analogous to only a single human in the sense of genetic diversity [38]. To recapitulate the complex human disease gene-gene interactions, a gene should be studied in a variety of inbred strains and wild mice to increase genetic diversity to find the genetic modifiers of the targeted gene [39,40]. However, there is no such technique available, since the outbreeding will lose homogeneity, making it virtually impossible to study the targeted gene, and the intra-“congenic” strain animals from backcrossing are not identical.

To address these limitations, the inventors provide a novel concept to study gene-gene interactions in a single animal (Figure 1) to achieve 1) high sensitivity: phenotypes of the nine genotypes from one mouse can be studied by flow cytometry in a single tube with high sensitivity due to the elimination of experimental variations and the power of flow cytometry, which has an analysis rate of up to 40,000 cells per second; 2) congenic strain-independence: since each genotype has the other eight genotypes as controls in a single mouse, establishment of a congenic strain is not necessary; 3) high specificity: the gene can be turned on/off by the TetOn system, allowing spatiotemporal control of gene expression. Phenotypes can be specifically attributed to the gene of interest if they change in response to switching the gene expression on/off. The nine-in-one model allows the rapid, specific and highly sensitive study of the genetic interaction of any two genes. This will form the base to study the gene-gene interactions in any animals, including non-human primates [41], on a wide genetic background, which is essential to recapitulate human diseases.

Brainbow is a process to label neuron cells with a random choice of up to 90 distinctive FP colors so that adjacent cells can be distinguished from each other to visually study neural connections in the brain [8,42]. The Genebow technique can be

used to label cells with many distinct labels (*e.g.*, colors) to study interactions of adjacent cells similar to Brainbow and other similar methods [43-45]. However, the Genebow labeling, unlike the other methods, is associated with the nine different genotypes and the promoter activities of the two genes. Therefore, it can be used to study gene-gene interaction with high sensitivity and specificity in a single animal. In addition, since each genotype has the other eight genotypes as controls in a single animal, there is no need for genotype control animals. Thus, a congenic strain establishment is not necessary. Any animal that has the genotype [*aabb*(rtTACreERT2)+/-, usually F2] can be used directly for experiments. This enables a rapid genetic background switch (~2 months) of the two targeted genes in many different inbred strains and wild animals (*e.g.*, wild mice). As stated above, different genetic backgrounds may help to identify new genetic modifiers of the two genes, from which a gene interaction networks can be discovered. This is virtually impossible to accomplish using the current existing techniques.

There are alternative ways in which the animal model can be produced, including alternative ways to switch the genes on and off. An alternative to the CRISPR method, the traditional knock-in method based on mouse embryonic stem cells can be used to obtain this model. Similar to the Cre-loxP technology, Flp-FRT or PhiC31 Integrase-mediated recombination can also be used in the place of Cre-loxP recombination to turn on gene expression. In order to produce other species of animal models, the species-specific sequences will be used to replace mouse sequences. Other potential modifications include, for example, fertilized eggs and foster mothers from that species.

As an alternative to the fluorescent proteins (FP), two Fluorogen activating peptide – FAP-tags®, a new class of small genetically encoded reporters that exhibit fluorescence (Green or red) only in the presence of micromolar concentrations of particular nontoxic soluble fluorogens, can be used in the place of two FPs (<http://spectragenetics.com/>). Therefore the two genes can be labeled with two fluorescent proteins and two FAP-tags. The FAP-tags can be used directly with Fluorogens to produce desired fluorescent colors. They also have Epitope Tag sequences that can be labeled with any colors, which gives the flexibility of accommodating multiplex fluorescence assay by flow cytometry and fluorescent microscopy. An LSRII flow cytometer may have the capacity of detecting 15 colors. The labeling of the two genes uses two fixed fluorescent colors and two other flexible colors. There are 11 color

available to label cell surface markers of interest to conduct almost any multiplex fluorescence flow cytometry routinely used in research (Figure 1).

An alternative labeling of the two genes is shown in Figure 6. Each gene is labeled with only one label (with one fluorescence gene). However, the subcellular localization of each fluorescence protein can be switched from cell membrane to nucleus or vice versa. Therefore, this two-fluorescence-gene-labeling is equivalent to the four-fluorescence-gene-labeling described previously.

A vast array of experiments can be carried out using the animal model and cells obtained therefrom. Since each genotype of the nine genotypes of two genes is specifically labelled, *e.g.*, with a color-code, it is not necessary to genotype cells in study. This will greatly facilitate the study of phenotype of the two genes. For example, lineage tracing of hematopoietic cells may be done: the 9 genotypes can be induced at early development of hematopoietic cells, *e.g.*, hematopoietic stem cells, then trace the development of each genotype using flow cytometry and fluorescent microscopy, and answer the questions whether different genotypes have different development consequences in terms of cell numbers, cell types (B, T lymphocytes, monocytes, dendritic cells) and cell subtypes (CD4, CD8 T cells, B1 and B2 B cells, *etc.*). It is also possible to study lymphocyte and stromal cell interaction using the animal model. Lymphocytes undergo massive cell death at multiple developmental stages in order to eliminate non- or self-reactive lymphocytes through positive and negative selection, in which stromal cells play an important role. Since both lymphocytes and stromal cells are labeled with fluorescent colors specific to the genotypes, the influence of the different genotype of the stromal cells on the selection of lymphocytes can be studied. In addition, it is possible to study the cell-cell interactions of hematopoietic cells. Similarly, the cell-cell interaction play an important role in the activation or inhibition of hematopoietic cells, such as the activation of B cells by T cells, the inhibition of T cell proliferation by regulatory T cells.

Example 1 describes how to construct two target vectors for the two genes with the universal TCFP cassettes, the third target vector with the inducible Cre, and the reverse tetracycline transactivator (rtTA) gene will be constructed. The functionality of all three will be tested *in vitro*.

Example 2 describes how to generate the mouse model and model functionality. The clustered regularly interspaced short palindromic repeats (CRISPR) technique, which can quickly generate mutated mice at high efficiency, will be used to generate the model. The functionality of the model in terms of gene expression control and FP labeling will be tested by studying the genetic interaction of the two genes on cell apoptosis.

Following are examples that illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

Example 1 – Targeting vector construction and functionality

1. Gene labeling strategy.

As shown in Figure 1, the wild type (WT) allele (*A*) will be labeled with FP1, the KO allele (*a*) with FP2. The cells with alleles *A* and *a* will have two colors, FP1 and FP2. Thus the three genotypes of one gene can be specifically labeled with two FPs. To this end, *FP2* and gene *A* are linked in frame with the porcine teschovirus-1 2A (P2A) sequence encoding a self-cleavable peptide, and they will be co-expressed. *FP1* is placed downstream of the promoter of the gene *A* and will be expressed if the promoter is active. *FP1* is followed by a STOP that serves as a stop signal for the transcription and translation of the downstream genes, *i.e.*, *FP2* and *A* genes. However, once the FP1 and the STOP are removed by the Cre/loxP system, FP2 and *A* will be expressed. Thus, FP1 (WT) and FP2 (KO) are exclusively expressed. Similarly, FP3 & 4 will be used to label gene *B* (Figure 1). Together, the nine genotypes of the two genes can be specifically labeled (Figure 1) and detected by flow cytometry.

2. Choosing florescent proteins (FPs).

Four bright FPs have been chosen for sensitive and specific detection: mNeonGreen, mRuby2, Aqua and iRFP670 with a relative brightness of 100, 46, 23 and 13, respectively [64]. EGFP and mRFP1 with a relative brightness of 36 and 13 have been successfully used to label cells [65]. These FPs do not overlap with each other in cytometry detection using the four lasers and detection filter sets as shown (Figures 3B and 3C). This ensures specific detection by flow cytometry (Figure 3C).

3. Cloning strategy.

For high throughput target construction, two universal TCFP cassettes will be constructed so the target repair template for any gene can be constructed as described in Figure 3A. Gibson cloning kit (NEB) will be used to facilitate the cloning of the two universal TCFP cassettes [66,67]. The two target repair templates for *Lrba* and *Nfkb1* will be assembled as described in Figure 3A. The TCFP cassette is composed of the following components that have been tested previously. (A) Splice sites (1, 2 and 6 in Figure 3A) for splicing out unwanted sequences that may impair translation of downstream coding sequences [1, splice donor (SD); 2, human β -globin (HBB) intron splice acceptor (SA) [44]; 6, rabbit β -globin SA [68]]. (B) FP reporter genes iRFP670 or mRuby2 (3); mNeonGreen or Aquamarine (Aqua) (7). iRFP670 is a near-infrared FP [69]. Aqua is superior to the popular form ECFP [70]. (C) STOP sequence (4) [71]. (D) Tetracycline responsive element (TRE) (5) [72-74]. (D) P2A (8). P2A peptide is used to link two proteins to be co-expressed by a single promoter [75]. Due to its small size (57 bp), high self-cleavability, and ability to produce an equal molar ratio of the two proteins, 2A is superior to IRES, which is large and causes differential expression of the two genes that it links [75].

4. Construction of the rtTA/CreERT2 target vector.

The tetracycline-inducible expression system has been used previously to turn on/off gene expression [76]. The Tet-On third generation tet-inducible gene expression system is 100-fold more Dox-sensitive than the original TetOn system. It has significantly reduced background, and also is widely cited [72-74]. The CreERT2 is a fusion protein, consisting of the loxP site-specific Cre recombinase linked to a triple mutant form of the ligand-binding domain of the human estrogen receptor, which does not bind its natural ligand but will bind the synthetic TAM. Cre-ERT2 is restricted to the cytoplasm and can only enter the nuclear compartment after exposure to TAM. Thus, it can be used to efficiently induce the recombinase activity by administration of TAM [77,78]. CreERT2 is currently the most successful CreER version [79]. The pMB80 (R26-CreER) (Addgene plasmid 12168) is a target vector for inserting CreERT2 into the Rosa26 locus [80]. The neomycin selection marker in this vector will be removed by Cre recombination *in vitro* and will be cleaved by a unique site enzyme PflFI just in the translation start codon ATG. Then, the rtTA-P2A (third generation TetOn gBlock) PCR fragment will be inserted and

fused with the CreERT2 by Gibson cloning, so that rtTA and Cre-ERT2 can be coexpressed.

5. The functionality of three target vectors will be validated before CRISPR injection.

5 Each TCFP target vector and the rtTA/CreERT2 will be inserted into the JM8A3-N1 ES cells (passage #17, the Mouse Biology Program) as previously with the CRISPR technique [9]. The corrected targeted ES clones will be expanded, mixed, and subjected to (1) Doxycycline (Dox) induction of gene expression and (2) Tam induced Cre cleavage of the STOP sequence and gene expression. The four FP will be detected by flow cytometry.

10 The gene expression of *Lrba* and *Nfkb1* will be detected by real time PCR and Western blot and compared to that from the WT ES cells.

6. Preliminary results and feasibility.

The inventors have constructed an *Lrba* target vector from a BAC clone using recombination-mediated genetic engineering techniques: homologous recombineering (HR) and Gibson cloning. The correct cloning of the target vector was characterized by restriction digestion (Figure 4A) and confirmed by sequencing (Figures 4G-4I). A PCR method for screening *Lrba* KO has been worked out (Figure 4B) and will be used to genotype mice generated by the CRISPR technique. The FPs can be readily detected (Figure 4C). A secretable luciferase gene was inserted in place of *Lrba* and used to show that the all-in-one Tet-inducible system in the vector functions as expected (Figure 4D). The single guide (sgRNA) sequence for *Lrba* was designed using the online CRISPR Design Tool and then cloned into the pX330 vector [81]. As the sgRNA sequence spans the insertion site, the target vector does not have the sequence and will not be cleaved by the Cas9 nuclease. The sgRNA binding sequence also has a *SpeI* site. Thus *SpeI* digestion was used to identify gene mutations (Figure 4E). The *in vitro* Cas9/sgRNA cleavage assay was also used to test the *Lrba* sgRNA (Figure 4F). Both assays indicate that the Cas9/sgRNA is working and *Lrba* locus has been successfully mutated in the ES cells. Similarly, the sgRNAs for *Nfkb1* and *Rosa26* will be designed and cloned. Others have been successful in inserting a 3.5kb loxP-STOP-TRE-loxP cassette and multiple functional lines have been generated using the traditional ko method [82].

The fragments for the two universal TCFP cassettes, the *Hbb* intron with SD and SA sequence and a loxP site, TRE-SA, mNeonGreen-2A, Aqua-2A and rtTA-2A with the

P2A sequence at their C-termini have been synthesized as gBlocks and cloned. The vectors containing mRuby2, iRFP670 and STOP were ordered from Addgene. All of these fragments have been confirmed through sequencing and their functionality has been tested. With high fidelity PCR, Gibson cloning techniques [66], and the availability of all of these fragments, it is expected that the three target templates will be successfully assembled and function properly *in vitro*. One potential concern is that although high fidelity DNA polymerase will be used in the PCR reaction, mutations may be introduced. To avoid this problem, several clones will be picked up and sequence all of the two universal cassettes and the homologous arms. The correct clones will then be chosen for the next step.

Example 2 – Animal model generation and model functionality

1. Generation of mouse model through CRISPR technology.

The mouse model will be generated as described previously [53,83]. Briefly, the T7 promoter will be added to the Cas9 coding sequence, the sgRNAs of Nfkb1 and Rosa26 by PCR using the primer pairs (Lrba sgRNA has been synthesized) [53]. The RNAs will be synthesized by T7 RNA polymerase *in vitro* and purified. Then, the Cas9 mRNA (100 ng/ml), and sgRNA (50 ng/ml) and 200 ng/ml target DNA template for each target will be injected into the cytoplasm of fertilized B57BL/6 eggs. Cytoplasm injection yields the highest number of positive pups with 89% homozygous KO efficiency [84]. The genomic DNA from targeted and control mice, age 8 to 12 days, will be extracted from clipped toes and used for PCR screening: The correct 5' and 3' end targeting will be confirmed by the PCR protocol [9] using the primers from the vector and the genomic DNA sequence outside of the left or right arms (Figure 4B). The mice that are correctly targeted as identified by PCR will be confirmed by Southern blot with the probes from mRuby2, iRFP670 and Cre genes in order to characterize the targeted Nfkb1, Lrba and Rosa26 loci, respectively as previously described [9].

2. Characterization of mouse model functionality.

Dox-induced Lrba gene expression: This model has the all-in-one TetOn system. The expression of the Lrba and Nfkb1 genes can be induced by Dox and detected by flow cytometry. The triple knockin mouse *aabbrrTAcreeERT2+/-* (*a*=Lrba KO, *b*=Nfkb1 KO) mice and C57BL/6 WT mice (6 wks of age) will be treated with or without Dox. Dox

binds to the rtTA and will activate *Lrba* and *Nfkb1* expression in the knockin mice in the presence of Dox.

Tam-induced Cre mosaic recombination: Nine genotypes of the two genes will be produced in a single *Lrba Nfkb1 rtTAcreERT2+/-* mouse by Tam-induced Cre mosaic recombination. It can be distinguished by flow cytometry (Figure 1) since Cre mediated recombination is normally incomplete [9,85]. The conditions (dosing and time) of Tam treatment will be optimized to obtain roughly equal numbers of the three B cell types for each gene, as determined by flow cytometry. Mice will receive 0, 3, or 9 mg of Tam (ip injection) for 1 to 5 consecutive days. Then, on day 24, the mice will be treated with Tam (i.p.) at the optimized condition to induce partial Cre recombination, producing nine types of peripheral blood cells [86]. On days 0, 10, 17 and 31, blood from the submandibular vein will be collected and subjected to flow cytometry (Figure 5B).

Flow cytometry: As mouse peripheral blood is limited, a no-lyse, no-wash staining flow cytometry technique will be utilized, which will use 20 μ L of whole blood for each analysis [87,88]. Cells will be stained with the anti-mouse CD45-PE-Cy7 antibody (Ebioscience) to distinguish the white blood cells from the much more abundant red blood cells and the LIVE/DEAD® Yellow dye, which will keep data set sizes manageable. Live CD45 positive cells will be gated and the nine populations will be analyzed. Although there is no overlapping between the four FPs and minimal overlapping with PE-Cy7 and the Yellow dye, compensation will be carried out by using the Raji B cells stained with the Yellow dye and anti-human CD45 PE-Cy7 separately, and the Raji B cells transfected with the four individual FP expression vectors.

Apoptosis assay: Apoptosis is central to the immune system. Useful immune cells are kept alive; unwanted or harmful immune cells are eliminated by apoptosis. Deregulated apoptosis can cause diseases. The best example is the primary immunodeficiency autoimmune lymphoproliferative syndrome (ALPS) resulting from defective apoptosis [89,90]. Both *LRBA* and *NFkB* can be both anti-apoptotic and proapoptotic. Both are involved in immunodeficiency and autoimmunity. Without being limited by theory, the inventors hypothesize that *Lrba* and *Nfkb1* genetically interact with each other to affect leukocyte survival. To test this hypothesis, flow cytometry will be carried out using Annexin V, biotin-X conjugate/R-PE/CY7 streptavidin conjugate and LIVE/DEAD® Yellow Dead Cell Stain Kit to stain the cells following a standard

cytometry protocol with red cell lysis, using 100 µl of blood from each mouse older than 6 months. The nine types of cells with nine different genotypes will be gated and analyzed as standard apoptosis assay. The concentration of total white blood cells from each bleeding will be measured by single bead-enhanced cytofluorimetry [91].

5 It is expected that the Genebow mouse model with three knockins (*aabbrrTAcereERT2+/-*) will be successfully generated, since CRISPR technique has high efficiency of multiplex targeting [53,55,57,83,92-95]. However, if the mouse with the desired genotype is not generated, the inventors will mate these mice that have one or two knock-ins to obtain the mice with the desired genotype. CRISPR targeting may induce
10 off-target mutations [53,96]. To detect any such potential mutations, the top 20 genome-wide off-target sites identified by the CRISPR Design tool will be amplified by PCR and sequenced. If mutations are detected, then the founders will be backcrossed with C57BL/6 mice multiple times until no mutations can be detected. It is expected that the apoptosis will be increased following the knockout of *Lrba* and *Nfkb1* genes on a dose-
15 dependent manner, *i.e.*, apoptosis rate: A+++B+++ (Dox-induced overexpression of the two genes, bleeding II) < AABB < (AABb, AaBB) < (AAbb, AaBb, aaBB) < (Aabb, aaBb) < aabb < A---B--- (There is no expression of the of the two genes in any cells in the mouse bleeding I & III). The prediction is based on that both *Lrba* and *Nfkb1* are mainly anti-apoptotic [4,47,51,97,98]. They may have additive anti-apoptotic effects and
20 complement each another. However, the two gene appear to have contradictory functions in apoptosis. They can be pro-apoptotic [52]. In either case, a pattern should be observed from the study of the nine genotypes plus one overexpression (A+++B+++) and one complete KO (A---B---) and will design the next experiments, *e.g.*, using MACS® Technology to isolate B or T cells to test the conditions (*e.g.*, cell type, stimuli) under
25 which *Lrba* or *Nfkb1* is anti- or pro-apoptosis. This may help to demystify the paradoxical association of immunodeficiency and autoimmunity, an unsolved fundamental question.

Supplemental Disclosure 1

30 Genetic interaction study in the intact animal provides the most compelling means to define genotype–phenotype relationship, but it is extremely challenging in human and animal models, since the variations between individuals can interfere with a gene's

contribution, which usually is small, to a phenotype. To overcome this barrier, the inventors propose that the nine genotypes of two genes are produced and respectively labeled with fluorescent proteins in a single animal (*e.g.*, a mouse) so that they can be analyzed in a single flow cytometer tube. Phenotyping sensitivity thus can be greatly increased with the increased sample size and eliminated variations. To provide a proof of principle for this strategy, the inventors propose inserting a genotype-switching and labeling cassettes into the genomic loci of the lipopolysaccharide-responsive beige-like anchor (*Lrba*). The inventors will also test whether *Lrba* expression can be switched on and off with equal probabilities and the three phenotypes of *Lrba* are specifically labeled with fluorescent colors. The invention presents a revolutionary methodology to efficiently conduct genetic interaction study in animal models to understand the etiology of complex human diseases.

The inventors have developed a simple but eloquent technique to conditionally knockout a gene and label the three genotypes, wild type, heterozygous and knockout, with specific fluorescent colors, in a single mouse. This can be done by knocking a genotype-switching and labeling cassette into the genomic loci of a gene. The knockin will not interrupt the expression of the targeted gene as the gene and one fluorescent proteins (FP) linked by a porcine teschovirus-1 2A (P2A) peptide gene will be co-expressed under the same endogenous promoter. Cre recombinase-mediated cleavage and ligation of the two inverted loxP sites will invert the intervening DNA. The inversion will inactivate both the targeted gene and the FP that is co-expressed with the targeted gene, but will activate another FP. As the sequences of the loxP site and the size of the intervening DNA remain unchanged, the reaction is reversible with equal reaction rates at both directions.

The intervening DNA between two inverted loxP sites can be inverted back and forth with equal probabilities by Cre recombination, resulting in equal numbers of two plasmid molecules, in which the intervening DNA has opposite orientations, in bacteria. This will be proved by transforming competent cells with the two plasmids resulted from the Cre recombination. Equal numbers of bacterial colonies will harbor the two plasmids, respectively. The results should also be true in mouse embryonic stem (ES) cells. The frequencies of the two orientations of the intervening DNA in the genome are equal. Furthermore, the two alleles of *Lrba* gene are respectively labeled with two FP: green FP

labels wild type (wt) allele, while red FP labels knockout (ko) allele. In other words the three genotypes of ES cells are specifically labeled with FPs: wt cells are labeled with green FP, ko cells are labeled with red FP, while heterozygous (het) cells are labeled with green FP and red FP. These cells in a mixture thus can be distinguished by flow cytometry without using traditional genotyping techniques. These results will demonstrate that Cre-mediated DNA inversion can be used to turn on and off gene expression with equal probabilities and the on/off status of gene expression can be specifically labeled with FPs. These results will provide proof of principle to develop Genebow models to conditionally produce equal frequencies of wt and ko alleles that are labeled with specific FPs, so that all the nine genotypes of two genes from a single mouse can be detected and assayed in a single flow cytometry tube. The two requirements, i.e. a large sample size and high phenotyping sensitivity required for genetic interaction study, can be fulfilled. By using the Genebow technique, genetic interaction study can be, for the first time, efficiently conducted in animal models at high resolution, critical to understanding the etiology of complex human diseases.

The mutation of the LRBA gene [99,100] causes severe autoimmune diseases, *e.g.*, inflammatory bowel diseases [101], type 1 diabetes and rheumatoid arthritis [102,103]. The underlying molecular mechanisms of LRBA deficiency are unknown. Moreover, LRBA deficiency lacks a clear genotype–phenotype relationship (GPr) as the clinical manifestations are highly variable. This indicates complex genetic interactions, *i.e.*, the phenotype of LRBA deficiency is dependent on the genetic interaction of LRBA with other genes. Indeed, LRBA interacts with multiple critical genes (Figures 7B-D) [100,104-106]. Therefore, the genetic interactions (GI) of LRBA with other critical genes must be studied to define a clear GPr to understand this complex human disease. The phenotype-based GI study in the intact animal provides the most compelling means to define GPr [107]. While a physical interaction study, based on two-hybrid, affinity capture-Western/mass spectrometry, is important, it does not link genotype to phenotype. However, GI study is extremely challenging in humans and animal models, since the variations between individuals can severely interfere with a gene's contribution to a phenotype, resulting in low phenotyping sensitivity. Consistently, the history of mouse genetics has demonstrated that “clear and consistent phenotypes are the exception rather than the rule”, *e.g.*, phenotypes cannot be detected in most gene knockouts [108]. An

impractically large sample size is required to obtain statistically significant data, *e.g.*, 500,000 subjects are required to achieve sufficient statistical power to detect non-additive effects of genetic variants [109]. To overcome this barrier, the inventors propose that the nine genotypes of two genes are produced and respectively labeled with four fluorescent proteins (FPs) in a single mouse and each cell will be studied as an individual analogous to a mouse in a traditional assay (Figure 7A). In this manner, large numbers of cells, *e.g.*, one mouse spleen contains $\sim 2 \times 10^6$ T regulatory cells and $\sim 2 \times 10^8$ splenocytes, with the nine different genotypes from a single mouse can be analyzed in a single tube with a flow cytometer, which can analyze up to 40,000 cells per second [110]. The background variations between these cells thus can be eliminated, because these cells are from the same mouse, thus have the same genetic background, epigenetic modifications and environmental inputs. With large sample size and eliminated variations, phenotyping sensitivity can be greatly increased (Figure 7A bottom). The two requirements, *i.e.*, a large sample size and high phenotyping sensitivity by knocking two genotype-switching and labeling cassettes (GSLC) separately into the genomic loci of two genes.

A GSLC is shown in Figure 8B. The knockin will not interrupt *Lrba* expression, as *Lrba* and the green FP linked by a porcine teschovirus-1 2A (P2A) peptide gene will be co-expressed under the *Lrba* promoter. P2A is widely used to link two proteins to be co-expressed by a single promoter. Cre recombinase-mediated cleavage and ligation of the two inverted loxP sites will invert the intervening DNA, switching on/off between the expression of the green FP and *Lrba* (wild type, wt, allele) and that of red FP (knockout, ko, allele) (Figure 8B). Therefore, ko cells will be labeled as red FP, wt cells will be labeled as green FP, and heterozygous (het) cells will be labeled as red/green FPs (Figure 8A-C). Similarly, *Nfkb1* can be labeled with two other FPs: ko (purple), het (purple/blue) and wt (blue) (Figure 7A, Figure 8B legend). The nine genotypes of the two genes thus can be produced and respectively labelled with four FPs (Figure 7A). The concept of the Genebow technique, *i.e.*, to visualize or “see” the genotypes of conditional gene knockouts, is highly innovative.

Construction of *Lrba* knockin vector.

The CRISPR system causes DNA double strand break (DSB), which can be repaired by homology-directed-repair pathways. A DNA fragment flanked by two homologous sequences can be efficiently inserted into the cleavage site [114]. The

inventors obtained correctly targeted mouse embryonic stem (ES) cells with high efficiency (67%) [115]. A knockin construct with short homologous arms was also tried and high knockin efficiency was obtained (Figure 7E). Therefore, the CRISPR targeting system will be used to knock the GSLC1 into the genomic loci of Lrba [20,21]. Two
5 vectors, *i.e.*, a donor vector and Cas9/sgRNA vector, are required for targeting each gene. The Cas9/sgRNA vector provides Cas9 endonuclease and a target site-specific sgRNA to cleave the target site. The donor vector contains the GSLC1 flanked by two homologous fragments serves as a CRISPR repair template to mediate the knockin of the GSLC1. A protocol was established to readily construct these vectors, and the Cas9/sgRNA vector
10 for Lrba is ready [115,118]. The cloning strategy of the donor vector is depicted in Figure 8C. Gene synthesis will be used for easy gene assembly to facilitate the cloning [118-120]. The correct cloning of the knockin target vector will be characterized by restriction digestion and confirmed by sequencing.

Functionality of Lrba knockin vector.

15 The knockin vector contains the GSLC, which is used to switch and label the genotypes. The two functions will be tested *in vitro*. The Cre recombinase-mediated inversion will be tested in bacteria. Plasmid 706-Cre (Gene Bridges GmbH) will be transformed into the bacteria containing pLrba-Donor following the manufacturer's protocol, and the bacteria that harbor the two plasmids will be cultured 12-24 hours at
20 30°C. DNA will be extracted and re-transform competent cells. 20 colonies will be picked up and cultured for DNA isolation. Cre recombinase-mediated inversion will be examined by restriction enzyme digestion of isolated DNA (Figures 9A-C). It is expected that the numbers of the colonies that harbor the original and inverted intervening DNA are equal, indicating that the rates of Cre-mediated reactions at both directions are equal. The
25 functionality of the CRISPR vectors will be validated in ES cells. The GSLC1 will be knocked into the Lrba genomic locus in ES cells as described [115]. Green FP positive clones will be picked up and screened by 5' and 3' PCR, and confirmed by Southern blot following the inventors' methods [115]. The correctly targeted ES clones will be infected with the recombinant retroviruses made from the plasmid MSCV.CreERT2.puro
30 (Addgene) and selected by puromycin for stable clones, which will be used for Tamoxifen (TAM)-induced Cre-cleavage assay. Gene expression will be directly detected by the expression of the fluorescent proteins by flow cytometry to detect

Cremediated genotype-switching and labeling. The correct recombination and the association between the genotypes and FP colors will also be confirmed by sorting the three types of cells and conducting PCR and Western blot. Sorted cells will also be expanded and further confirmed by Southern blot [115]. It is expected that the three
5 genotypes of *Lrba* will be generated and specifically labeled in ES cells with the ratios: 1wt:2het:1ko.

The TAM-inducible Cre recombination system allows optimization of the dose and time of TAM administration to have complete or near complete Cre recombination, which will be determined by flow cytometry. The condition (dosing and time) of TAM
10 treatment that achieves balanced Cre reaction producing equal numbers of wt and ko cells as determined by flow cytometry, will be determined empirically, using 0 to 2 μ M TAM for 12 to 72 hours [123,124]. Each experiment will be conducted in triplicate and repeat three times. The knockin vector may need to be modified using a CRISPR cloning technique [118] to obtain optimal genotype-switching and labeling results.

15 The principle of the Genebow technique is an improvement of Brainbow [111,120,125] and Cre/lox system [126], in that the Brainbow system only labels cells and traditional conditional knockout cannot “see” (visualize) a genotype. Genebow is a conditional gene knockout and genotype labeling technique.

The Genebow technique will allow researchers to efficiently conduct genetic
20 interaction studies, for the first time, in animal models, which is critical to understanding the etiology of complex human diseases including the diseases caused by LRBA deficiency. LRBA is a critical immune regulator and plays a fundamental role in the immune system [99,100]. However, LRBA is only one example of the genes that may be investigated using the invention. For example, similar experiments with *Nfkb1* can be
25 conducted, generating the Genebow animals models for *Lrba* and *Nfkb1* to study the genetic interaction of the two genes in animal models. LRBA and NFKB1 closely interact with each other (Figure 7A). NFKB1, among the genes that cause autoimmunity when mutated, has the most autoimmune problems (70%) that are found in LRBA deficiency [129].

30

REFERENCES

1. Manolio, T.A. & Collins, F.S. Genes, environment, health, and disease: facing up to complexity. *Hum Hered* **63**, 63-6 (2007).

2. Cost, G.J. Enzymatic ligation assisted by nucleases: simultaneous ligation and digestion promote the ordered assembly of DNA. *Nat Protoc* **2**, 2198-202 (2007).
3. Shamloula, H.K. *et al.* rugose (rg), a Drosophila A kinase anchor protein, is required for retinal pattern formation and interacts genetically with multiple signaling pathways. *Genetics* **161**, 693-710 (2002).
- 5 4. Wang, J.W. *et al.* Deregulated expression of LRBA facilitates cancer cell growth. *Oncogene* **23**, 4089-97 (2004).
5. Yatim, A. *et al.* NOTCH1 nuclear interactome reveals key regulators of its transcriptional activity and oncogenic function. *Mol Cell* **48**, 445-58 (2012).
- 10 6. Beck, J.A. *et al.* Genealogies of mouse inbred strains. *Nat Genet* **24**, 23-5 (2000).
7. Rogner, U.C. & Avner, P. Congenic mice: cutting tools for complex immune disorders. *Nat Rev Immunol* **3**, 243-52 (2003).
8. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56-62 (2007).
- 15 9. Wang, J.W. *et al.* Influence of SHIP on the NK repertoire and allogeneic bone marrow transplantation. *Science* **295**, 2094-7 (2002).
10. Wang, J.W., Howson, J., Haller, E. & Kerr, W.G. Identification of a novel lipopolysaccharide-inducible gene with key features of both A kinase anchor proteins and chs1/beige proteins. *J Immunol* **166**, 4586-95 (2001).
- 20 11. Wang, J.-W., Wang, A., Li, K., Jin, S. & Lockey, R.F. CRISPR/Cas9 nuclease combined with Gibson assembly for scarless cloning into large vectors. *Biotechniques*. 2014. **accepted for publication**(2014).
12. Taylor-Cousar, J.L. *et al.* Histo-blood group gene polymorphisms as potential genetic modifiers of infection and cystic fibrosis lung disease severity. *PLoS One* **4**, e4270 (2009).
- 25 13. Drumm, M.L. *et al.* Genetic modifiers of lung disease in cystic fibrosis. *N Engl J Med* **353**, 1443-53 (2005).
14. Fuchs, H. *et al.* Mouse phenotyping. *Methods* **53**, 120-35.
15. Belizario, J.E. The humankind genome: from genetic diversity to the origin of human diseases. *Genome* **56**, 705-16 (2013).
- 30 16. Arnedo, J. *et al.* Uncovering the Hidden Risk Architecture of the Schizophrenias: Confirmation in Three Independent Genome-Wide Association Studies. *Am J Psychiatry*.
17. Podjasek, J.C. & Abraham, R.S. Autoimmune cytopenias in common variable immunodeficiency. *Front Immunol* **3**, 189 (2012).
- 35 18. Wright, F.A. *et al.* Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2. *Nat Genet* **43**, 539-46 (2011).
19. Doetschman, T. Influence of genetic background on genetically engineered mouse phenotypes. *Methods Mol Biol* **530**, 423-33 (2009).
- 40 20. Kurtz, T.W. Genome-wide association studies will unlock the genetic basis of hypertension.: con side of the argument. *Hypertension* **56**, 1021-5 (2010).
21. Ermann, J. & Glimcher, L.H. After GWAS: mice to the rescue? *Curr Opin Immunol* **24**, 564-70 (2012).
- 45 22. Kado, S. *et al.* Intestinal microflora are necessary for development of spontaneous adenocarcinoma of the large intestine in T-cell receptor beta chain and p53 double-knockout mice. *Cancer Res* **61**, 2395-8 (2001).

23. Franklin, C.L. Microbial considerations in genetically engineered mouse research. *ILAR J* **47**, 141-55 (2006).
24. Treuting, P.M., Clifford, C.B., Sellers, R.S. & Brayton, C.F. Of mice and microflora: considerations for genetically engineered mice. *Vet Pathol* **49**, 44-63
5 (2012).
25. Watkins-Chow, D.E. & Pavan, W.J. Genomic copy number and expression variation within the C57BL/6J inbred mouse strain. *Genome Res* **18**, 60-6 (2008).
26. Bryant, C.D. *et al.* Behavioral differences among C57BL/6 substrains: implications for transgenic and knockout studies. *J Neurogenet* **22**, 315-31 (2008).
- 10 27. Saeed, S. *et al.* Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. *Science* **345**, 1251086 (2014).
28. Stevens, J.C., Banks, G.T., Festing, M.F. & Fisher, E.M. Quiet mutations in inbred strains of mice. *Trends Mol Med* **13**, 512-9 (2007).
- 15 29. Eisener-Dorman, A.F., Lawrence, D.A. & Bolivar, V.J. Cautionary insights on knockout mouse studies: the gene or not the gene? *Brain Behav Immun* **23**, 318-24 (2009).
30. Ridgway, W.M. A new tool for dissecting genetic control of type 1 diabetes. *Diabetes* **63**, 56-8 (2014).
- 20 31. Ridgway, W.M., Healy, B., Smink, L.J., Rainbow, D. & Wicker, L.S. New tools for defining the 'genetic background' of inbred mouse strains. *Nat Immunol* **8**, 669-73 (2007).
32. Keane, T.M. *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289-94 (2011).
- 25 33. Sundberg, J.P., Roopenian, D.C., Liu, E.T. & Schofield, P.N. The Cinderella effect: searching for the best fit between mouse models and human diseases. *J Invest Dermatol* **133**, 2509-13 (2013).
34. Greve, B. *et al.* High-grade loss of leukocytes and hematopoietic progenitor cells caused by erythrocyte-lysing procedures for flow cytometric analyses. *J Hematother Stem Cell Res* **12**, 321-30 (2003).
- 30 35. Einwallner, E. *et al.* Lysis matters: red cell lysis with FACS Lyse affects the flow cytometric enumeration of circulating leukemic blasts. *J Immunol Methods* **390**, 127-32 (2013).
36. Greve, B., Beller, C., Cassens, U., Sibrowski, W. & Gohde, W. The impact of erythrocyte lysing procedures on the recovery of hematopoietic progenitor cells in
35 flow cytometric analysis. *Stem Cells* **24**, 793-9 (2006).
37. Frazer, K.A. *et al.* A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* **448**, 1050-3 (2007).
38. Yang, H. *et al.* Subspecific origin and haplotype diversity in the laboratory mouse. *Nat Genet* **43**, 648-55 (2011).
- 40 39. Chalfin, L. *et al.* Mapping ecologically relevant social behaviours by gene knockout in wild mice. *Nat Commun* **5**, 4569 (2014).
40. Linder, C.C. Genetic variables that influence phenotype. *ILAR J* **47**, 132-40 (2006).
41. Niu, Y. *et al.* Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell* **156**, 836-43.
45
42. Cai, D., Cohen, K.B., Luo, T., Lichtman, J.W. & Sanes, J.R. Improved tools for the Brainbow toolbox. *Nat Methods* **10**, 540-7 (2013).

43. Espinosa, J.S., Tea, J.S. & Luo, L. Mosaic analysis with double markers (MADM) in mice. *Cold Spring Harb Protoc* **2014**, 182-9 (2014).
44. Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D. & Luo, L. Mosaic analysis with double markers in mice. *Cell* **121**, 479-92 (2005).
- 5 45. Loulier, K. *et al.* Multiplex cell and lineage tracking with combinatorial labels. *Neuron* **81**, 505-20 (2014).
46. Wang, J.-W. & Lockey, R.F. Lipopolysaccharide-responsive beige-like anchor (LRBA), a novel regulator of human immune disorders. *Austin Journal of Clinical Immunology* **1**, 9 (2014).
- 10 47. Lopez-Herrera, G. *et al.* Deleterious mutations in LRBA are associated with a syndrome of immune deficiency and autoimmunity. *Am J Hum Genet* **90**, 986-1001 (2012).
48. Burns, S.O. *et al.* LRBA gene deletion in a patient presenting with autoimmunity without hypogammaglobulinemia. *J Allergy Clin Immunol* **130**, 1428-32 (2012).
- 15 49. Alangari, A. *et al.* LPS-responsive beige-like anchor (LRBA) gene mutation in a family with inflammatory bowel disease and combined immunodeficiency. *J Allergy Clin Immunol* **130**, 481-8 e2 (2012).
50. Wang, J.-W. *et al.* LRBA Causes Immunodeficiency and Autoimmunity By Deregulating NFkB-Mediated Multiple Immune Effectors Critical For B Cell Activation. *Journal of Allergy and Clinical Immunology* **133**, AB251-AB251 (2014).
- 20 51. Baltimore, D. NF-kappaB is 25. *Nat Immunol* **12**, 683-5 (2011).
52. Hayden, M.S. & Ghosh, S. NF-kappaB, the first quarter-century: remarkable progress and outstanding questions. *Genes Dev* **26**, 203-34 (2012).
- 25 53. Yang, H. *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370-9 (2013).
54. Zhou, J. *et al.* Dual sgRNAs facilitate CRISPR/Cas9-mediated mouse genome targeting. *FEBS J* **281**, 1717-25 (2014).
55. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-8.
- 30 56. Yang, H. *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370-9.
57. Fujii, W., Kawasaki, K., Sugiura, K. & Naito, K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res* **41**, e187.
- 35 58. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-21 (2012).
59. Harrison, M.M., Jenkins, B.V., O'Connor-Giles, K.M. & Wildonger, J. A CRISPR view of development. *Genes Dev* **28**, 1859-1872 (2014).
- 40 60. Yasue, A. *et al.* Highly efficient targeted mutagenesis in one-cell mouse embryos mediated by the TALEN and CRISPR/Cas systems. *Sci Rep* **4**, 5705.
61. Bottcher, R. *et al.* Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells. *Nucleic Acids Res* **42**, e89 (2014).
- 45 62. Gratz, S.J. *et al.* Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics* **196**, 961-71 (2014).

63. Donoho, G., Jasin, M. & Berg, P. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol* **18**, 4070-8 (1998).
64. Dean, K.M. & Palmer, A.E. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat Chem Biol* **10**, 512-23 (2014).
65. Wan, Y.Y. & Flavell, R.A. Identifying Foxp3-expressing suppressor T cells with a bicistronic reporter. *Proc Natl Acad Sci U S A* **102**, 5126-31 (2005).
66. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-5 (2009).
67. Gibson, D.G., Smith, H.O., Hutchison, C.A., 3rd, Venter, J.C. & Merryman, C. Chemical synthesis of the mouse mitochondrial genome. *Nat Methods* **7**, 901-3.
68. Niwa, H., Yamamura, K. & Miyazaki, J. EFFICIENT SELECTION FOR HIGH-EXPRESSION TRANSFECTANTS WITH A NOVEL EUKARYOTIC VECTOR. *Gene* **108**, 193-199 (1991).
69. Shcherbakova, D.M. & Verkhusha, V.V. Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat Methods* **10**, 751-4 (2013).
70. Erard, M. *et al.* Minimum set of mutations needed to optimize cyan fluorescent proteins for live cell imaging. *Mol Biosyst* **9**, 258-67 (2013).
71. Sauer, B. Manipulation of transgenes by site-specific recombination: use of Cre recombinase. *Methods Enzymol* **225**, 890-900 (1993).
72. Zhou, X., Vink, M., Klaver, B., Berkhout, B. & Das, A.T. Optimization of the Tet-On system for regulated gene expression through viral evolution. *Gene Ther* **13**, 1382-90 (2006).
73. Loew, R., Heinz, N., Hampf, M., Bujard, H. & Gossen, M. Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnol* **10**, 81 (2010).
74. Kistner, A. *et al.* Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. *Proc Natl Acad Sci U S A* **93**, 10933-8 (1996).
75. Kim, J.H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**, e18556 (2011).
76. Ewald, D. *et al.* Time-sensitive reversal of hyperplasia in transgenic mice expressing SV40 T antigen. *Science* **273**, 1384-6 (1996).
77. Aghajani, K., Keerthivasan, S., Yu, Y. & Gounari, F. Generation of CD4CreER(T2)) transgenic mice to study development of peripheral CD4-T-cells. *Genesis* **50**, 908-13.
78. Metzger, D., Clifford, J., Chiba, H. & Chambon, P. Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. *Proc Natl Acad Sci U S A* **92**, 6991-5 (1995).
79. Feil, S., Valtcheva, N. & Feil, R. Inducible Cre mice. *Methods Mol Biol* **530**, 343-63 (2009).
80. McLaughlin, M.E. *et al.* The Nf2 tumor suppressor regulates cell-cell adhesion during tissue fusion. *Proc Natl Acad Sci U S A* **104**, 3261-6 (2007).
81. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-23.
82. Tanaka, K.F. *et al.* Flexible Accelerated STOP Tetracycline Operator-knockin (FAST): a versatile and efficient new gene modulating system. *Biol Psychiatry* **67**, 770-3 (2010).

83. Yang, H., Wang, H. & Jaenisch, R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat Protoc* **9**, 1956-68 (2014).
84. Horii, T. *et al.* Validation of microinjection methods for generating knockout mice by CRISPR/Cas-mediated genome engineering. *Sci Rep* **4**, 4513.
- 5 85. Leneuve, P. *et al.* Cre-mediated germline mosaicism: a new transgenic mouse for the selective removal of residual markers from tri-lox conditional alleles. *Nucleic Acids Res* **31**, e21 (2003).
86. Hayashi, S. & McMahon, A.P. Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene
- 10 87. activation/inactivation in the mouse. *Dev Biol* **244**, 305-18 (2002).
87. Weaver, J.L., Broud, D.D., McKinnon, K. & Germolec, D.R. Serial phenotypic analysis of mouse peripheral blood leukocytes. *Toxicol Mech Methods* **12**, 95-118 (2002).
88. Weaver, J.L., McKinnon, K. & Germolec, D.R. Phenotypic analysis using very
- 15 89. small volumes of blood. *Curr Protoc Cytom* **Chapter 6**, Unit 6 30 (2010).
89. Todoric, K., Koontz, J.B., Mattox, D. & Tarrant, T.K. Autoimmunity in immunodeficiency. *Curr Allergy Asthma Rep* **13**, 361-70 (2013).
90. Mackay, I.R., Leskovsek, N.V. & Rose, N.R. The odd couple: a fresh look at autoimmunity and immunodeficiency. *J Autoimmun* **35**, 199-205 (2010).
- 20 91. Montes, M., Jaensson, E.A., Orozco, A.F., Lewis, D.E. & Corry, D.B. A general method for bead-enhanced quantitation by flow cytometry. *J Immunol Methods* **317**, 45-55 (2006).
92. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-23 (2013).
- 25 93. Kabadi, A.M., Ousterout, D.G., Hilton, I.B. & Gersbach, C.A. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Res* (2014).
94. Ma, S. *et al.* CRISPR/Cas9 mediated multiplex genome editing and heritable mutagenesis of BmKu70 in *Bombyx mori*. *Sci Rep* **4**, 4489 (2014).
- 30 95. Sakuma, T., Nishikawa, A., Kume, S., Chayama, K. & Yamamoto, T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci Rep* **4**, 5400 (2014).
96. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822-6 (2013).
- 35 97. Andres, S.A., Brock, G.N. & Wittliff, J.L. Interrogating differences in expression of targeted gene sets to predict breast cancer outcome. *BMC Cancer* **13**, 326 (2013).
98. Beg, A.A. & Baltimore, D. An essential role for NF-kappaB in preventing TNF-alpha-induced cell death. *Science* **274**, 782-4 (1996).
- 40 99. Wang, J.W., Howson, J., Haller, E. & Kerr, W.G. Identification of a novel lipopolysaccharide-inducible gene with key features of both A kinase anchor proteins and chs1/beige proteins. *J Immunol* **166**, 4586-95 (2001).
- 45 100. Wang, J.W., Gamsby, J.J., Highfill, S.L., Mora, L.B., Bloom, G.C., Yeatman, T.J., Pan, T.C., Ramne, A.L., Chodosh, L.A., Cress, W.D., Chen, J. & Kerr, W.G. Deregulated expression of LRBA facilitates cancer cell growth. *Oncogene* **23**, 4089-97 (2004).

101. Serwas, N.K., Kansu, A., Santos-Valente, E., Kuloglu, Z., Demir, A., Yaman, A., Yaneth Gamez Diaz, L., Artan, R., Sayar, E., Ensari, A., Grimbacher, B. & Boztug, K. Atypical Manifestation of LRBA Deficiency with Predominant IBD-like Phenotype. *Inflamm Bowel Dis* **21**, 40-7 (2015).
- 5 102. Wang, J.-W. & Lockey, R.F. Lipopolysaccharide-responsive beige-like anchor (LRBA), a novel regulator of human immune disorders. *Austin Journal of Clinical Immunology* **1**, 9 (2014).
103. Charbonnier, L.M., Janssen, E., Chou, J., Ohsumi, T.K., Keles, S., Hsu, J.T., Massaad, M.J., Garcia-Lloret, M., Hanna-Wakim, R., Dbaiibo, G., Alangari, A.A., Al Sultan, A., Al-Zahrani, D., Geha, R.S. & Chatila, T.A. Regulatory T-cell deficiency and immune dysregulation, polyendocrinopathy, enteropathy, X-linked-like disorder caused by loss-of-function mutations in LRBA. *J Allergy Clin Immunol* **135**, 217-27 (2015).
- 10 104. Schreiber, S.L., Preiss, A., Nagel, A.C., Wech, I. & Maier, D. Genetic screen for modifiers of the rough eye phenotype resulting from overexpression of the Notch antagonist hairless in *Drosophila*. *Genesis* **33**, 141-52. (2002).
105. Shamloula, H.K., Mbogho, M.P., Pimentel, A.C., Chrzanowska-Lightowlers, Z.M., Hyatt, V., Okano, H. & Venkatesh, T.R. rugose (rg), a *Drosophila* A kinase anchor protein, is required for retinal pattern formation and interacts genetically with multiple signaling pathways. *Genetics* **161**, 693-710 (2002).
- 20 106. de Souza, N., Vallier, L.G., Fares, H. & Greenwald, I. SEL-2, the *C. elegans* neurobeachin/LRBA homolog, is a negative regulator of lin-12/Notch activity and affects endosomal traffic in polarized epithelial cells. *Development* **134**, 691-702 (2007).
- 25 107. Ramsdell, F. & Ziegler, S.F. FOXP3 and scurfy: how it all began. *Nat Rev Immunol* **14**, 343-9 (2014).
108. Pearson, H. Surviving a knockout blow. *Nature* **415**, 8-9 (2002).
109. Zuk, O., Hechter, E., Sunyaev, S.R. & Lander, E.S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* **109**, 1193-8 (2012).
- 30 110. cytometer, L.f. Flow cytometer can analyze up to 40,000 cells per second flow.ucsf.edu/selfanalysis-bd-lsriifortessa (2015).
111. Cai, D., Cohen, K.B., Luo, T., Lichtman, J.W. & Sanes, J.R. Improved tools for the Brainbow toolbox. *Nat Methods* **10**, 540-7 (2013).
- 35 112. Boston-University-and-Boston-Medical-Center. Sample Size Calculations. bu.edu/orccommittees/iacuc/policies-and-guidelines/sample-size-calculations/.
113. Kim, J.H., Lee, S.R., Li, L.H., Park, H.J., Park, J.H., Lee, K.Y., Kim, M.K., Shin, B.A. & Choi, S.Y. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**, e18556 (2011).
- 40 114. Donoho, G., Jasin, M. & Berg, P. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol* **18**, 4070-8 (1998).
- 45 115. Wang, B., Li, K., Wang, A., Reiser, M., Saunders, T.L., Lockey, R.F. & Wang, J.W. Highly efficient CRISPR/HDR-mediated knockin in mouse embryonic stem cells and zygotes. *BioTechniques*. **59**, 201-208 (2015).

116. Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L. & Jaenisch, R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370-9.
117. Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F. & Jaenisch, R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-8.
118. Wang, J.W., Wang, A., Li, K., Wang, B., Jin, S., Reiser, M. & Lockey, R.F. CRISPR/Cas9 nuclease cleavage combined with Gibson assembly for seamless cloning. *BioTechniques* **58**, 161-70 (2015).
119. IDTDNA. gBlocks® Gene Fragments Protocol: Gibson Assembly® Method. idtdna.com/pages/docs/default-source/user-guides-and-protocols/gibsonassembly.pdf?sfvrsn=12 (2012).
120. Loulier, K., Barry, R., Mahou, P., Le Franc, Y., Supatto, W., Matho, K.S., Ieng, S., Fouquet, S., Dupin, E., Benosman, R., Chedotal, A., Beaurepaire, E., Morin, X. & Livet, J. Multiplex cell and lineage tracking with combinatorial labels. *Neuron* **81**, 505-20 (2014).
121. Dean, K.M. & Palmer, A.E. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat Chem Biol* **10**, 512-23 (2014).
122. Wan, Y.Y. & Flavell, R.A. Identifying Foxp3-expressing suppressor T cells with a bicistronic reporter. *Proc Natl Acad Sci U S A* **102**, 5126-31 (2005).
123. Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L. & Smith, A. Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230-4 (2007).
124. Endoh, M., Endo, T.A., Endoh, T., Fujimura, Y., Ohara, O., Toyoda, T., Otte, A.P., Okano, M., Brockdorff, N., Vidal, M. & Koseki, H. Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* **135**, 1513-24 (2008).
125. Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R. & Lichtman, J.W. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56-62 (2007).
126. Kratochwil, C.F. & Rijli, F.M. The Cre/Lox system to assess the development of the mouse brain. *Methods Mol Biol* **1082**, 295-313 (2014).
127. Wang, J.W., Howson, J.M., Ghansah, T., Despots, C., Ninos, J.M., May, S.L., Nguyen, K.H., Toyama-Sorimachi, N. & Kerr, W.G. Influence of SHIP on the NK repertoire and allogeneic bone marrow transplantation. *Science* **295**, 2094-7 (2002).
128. van de Ven, A.A. & Warnatz, K. The autoimmune conundrum in common variable immunodeficiency disorders. *Curr Opin Allergy Clin Immunol* **15**, 514-24 (2015).
129. Wang, J.W., Qi, Y.P., Huang, Y.X. & Li, S.D. Nucleotide sequence of a 1446 base pair SalI fragment and structure of a novel early gene of Leucania separata nuclear polyhedrosis virus. *Arch Virol* **140**, 2283-91 (1995).
130. Kerr, W.G. & Wang, J.-W. LPS-responsive chs1/beige-like anchor gene and therapeutic applications thereof. *US patent 7,704,963* (2010).
131. Reiser, M.A., Li, K., Lockey, R.F. & Wang, J.W. Lipopolysaccharide Responsive Beige-Like Anchor Subcellular Localization Involving in Vesicle Trafficking Responsive to Lipopolysaccharide. *Austin Journal of Clinical Immunology* **1**, 8 (2014).

132. Wang, J.-W., Wang, A., Li, K., Jin, S. & Lockey, R.F. CRISPR/Cas9 nuclease combined with Gibson assembly for seamless cloning into large vectors. *Biotechniques*. (2015).
- 5 133. Tu, Z., Ninos, J.M., Ma, Z., Wang, J.W., Lemos, M.P., Despons, C., Ghansah, T., Howson, J.M. & Kerr, W.G. Embryonic and hematopoietic stem cells express a novel SH2-containing inositol 5'-phosphatase isoform that partners with the Grb2 adapter protein. *Blood* **98**, 2028-38 (2001).

Supplemental Disclosure 2

10 Mutation of the LRBA gene causes primary immunodeficiency and severe autoimmune diseases, including several human common diseases such as inflammatory bowel diseases (IBD), type 1 diabetes, and rheumatoid arthritis. The molecular mechanism of how LRBA deficiency causes such a wide spectrum of diseases is unclear. Accordingly, there is no causal treatment available for LRBA deficiency. Tregs are
15 indispensable for controlling various autoimmune diseases. LRBA regulates two important Treg regulators, and Treg is defective in some patients. Thus, it is hypothesized that LRBA is required for the development and function of Treg, and that deregulated Treg plays a major role in causing the wide spectrum of autoimmune manifestations. However, LRBA deficiency lacks a clear genotype-phenotype relationship (GPr) as the
20 manifestations are highly variable, indicating complex genetic interactions (GI). In other words, the phenotype of LRBA deficiency is modified by other genes. In this sense, LRBA deficiency is a polygenic disease, and the genetic interactions of LRBA with other critical genes must be studied to define a clear GPr. The phenotype-based GI study in the intact animal provides the most compelling means to define GPr, but it is extremely
25 challenging in humans and animal models, since the variations from genetic background, epigenetic modifications and environmental factors can interfere with or mask a gene's contribution to a phenotype, resulting in low phenotyping sensitivity. An impractically large sample size and highly sensitive phenotyping are required to obtain statistically significant data.

30 To overcome this barrier, the inventors propose studying each cell as an individual analogous to a mouse in a traditional assay and to label each genotype with a specific, distinct label (*e.g.*, color), so that large number of cells (up to several million) with nine different genotypes in a single animal can be analyzed in a single flow cytometry tube. In this manner, the background variation between cells should be eliminated because the
35 cells from the same mouse have the same genetic background, epigenetic modifications

and environmental input. Phenotyping sensitivity thus can be greatly increased with the greatly increased sample size and eliminated variation. To test the hypothesis, the inventors will develop a mouse model, named “Genebow”, by using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) cloning and mouse model generation techniques. The Genebow model will be used to determine how the genetic interaction of *Lrba* and *Nfkb1* regulates the development and function of Treg. This will demonstrate that GI study, for the first time, can be efficiently conducted in animal models at high resolution, and provide high resolution data to determine whether LRBA plays, modified by NFKB1, a critical role in regulating Tregs to suppress autoimmunity, and to provide deep insight into how LRBA deficiency causes such a wide spectrum of highly variable clinical manifestations. It will also shed light on the etiology of these human common diseases including some top 10 deadliest diseases that are associated with LRBA deficiency.

The study of genetic interaction plays a key role in defining GPr in complex human diseases, but it is extremely challenging in humans and mice, because an impractically large sample size and highly sensitive phenotyping are often required to obtain statistically significant data due to variations between individuals. A variation-free technique which can fulfill the two requirements, *i.e.*, a large sample size and high phenotyping sensitivity, for genetic interaction study will be developed, so that genetic interaction study can be efficiently conducted in animal models for the first time. As an exemplification, this technique can be used to determine how the genetic interaction of two critical immune genes in suppressing autoimmunity at high resolution, and such high resolution results may provide deep insight into the etiology of these complex human autoimmune diseases that are associated with LRBA deficiency.

The mutation of the LRBA gene causes primary immunodeficiency and a myriad of autoimmune problems.¹ The molecular and cellular mechanisms of how LRBA deficiency causes such a wide spectrum of highly variable clinical manifestations are unknown, and thus there is no causal treatment available for LRBA deficiency. LRBA regulates two important regulators of Tregs, the cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) and programmed death 1 (PD-1). LRBA deficiency also causes the immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX)-like syndrome.² IPEX is caused by deficiency of Tregs. Tregs are indispensable for

controlling various autoimmune diseases. The inventors thus hypothesized that LRBA is required for the development and function of Tregs, and that defective Tregs play a major role in LRBA deficiency. A goal is to delineate the molecular mechanisms of LRBA's role in the regulation of Tregs. However, LRBA deficiency lacks a clear GPr as the clinical manifestations are highly variable. This indicates complex genetic interactions (GI), *i.e.*, the phenotype of LRBA deficiency is dependent on the genetic interaction of LRBA with other genes, *e.g.*, NFKB1. In this sense, LRBA deficiency is a polygenic disease, and the genetic interactions of LRBA with other critical genes must be studied to define a clear GPr to understand this complex human disease.

However, GI study is extremely challenging in humans and mice since that variations from genetic background, epigenetic modifications and environmental factors can interfere with or mask a gene's contribution, usually small, to a phenotype, greatly reducing the phenotyping sensitivity. Consistently, most mouse gene knockouts do not have a detectable phenotype. An impractically large sample size and highly sensitive phenotyping are required to obtain statistically significant data. For example, 500,000 subjects might be necessary to achieve sufficient statistical power to detect nonadditive effects of genetic variants.³ To fulfil the two requirements, the inventors propose to study each cell as an individual analogous to a mouse in a traditional assay, so that large numbers of cells (~2x 10⁶ Tregs/spleen) in a single animal (*e.g.*, a single mouse) can be analyzed in a single flow cytometry tube with greatly increased phenotyping sensitivity as a result of eliminated variations (Figure 10, Figure 13C). To this end, the nine genotypes of two genes will be generated and respectively labelled with four fluorescent reporters (FRs) (Figure 13A) in a single mouse, so that the phenotypes of all genotypes of the two genes can be studied by flow cytometry in a single tube. As indicated above, this variation-free phenotyping technique was named "Genebow". LRBA and NFKB1 are closely related, *e.g.*, among the genes that cause autoimmunity when mutated, NFKB1 has the most autoimmune problems that are found in LRBA deficiency. To test the inventors' hypothesis, a Genebow mouse model will be generated for the two genes to study their genetic interaction.

The first objective is to generate a Genebow mouse model in which *Lrba* and *Nfkb1* expression can be turned on/off in a spatiotemporal and trackable manner. By using the CRISPR cloning and mouse model generation techniques, two transcription

control and labeling cassettes (TCLC) will be separately inserted into the genomic loci of *Lrba* and *Nfkb1*. The second objective is to determine how the genetic interaction of *Lrba* and *Nfkb1* regulates the development and function of Tregs. Reduced cell number and suppressive function of Tregs are two causes of the autoantibody overproduction.

- 5 Autoantibodies cause autoimmunity in many autoimmune diseases. The working hypothesis is that deregulation of *Lrba* and *Nfkb1* genetic interaction impairs the development and function of Tregs, and results in reduced cell number and ability to suppress autoimmunity.

If successful, it will demonstrate that a genetic interaction study of two genes, for
10 the first time, can be efficiently conducted in animal models at high resolution, and provide high resolution data to determine whether LRBA plays, modified by NFKB1, a critical role in regulating Tregs to suppress autoimmunity, and to provide deep insight into how LRBA deficiency causes such a wide spectrum of highly variable clinical manifestations. It will also shed light on the etiology of these human common diseases,
15 including some top 10 deadliest diseases, that are associated with LRBA deficiency.

Mutation of LRBA causes primary immunodeficiency (PID) and autoimmune diseases including IPEX-like syndrome,² IBD,⁸ autoimmune lymphoproliferative syndrome (ALPS),⁹ rheumatoid arthritis,^{2,10} autoimmune hemolytic anemia (AHA), immune thrombocytopenia (ITP), and endocrinopathies (type 1 diabetes, hepatitis, growth
20 hormone deficiency and hypothyroidism).^{1,11} The clinical presentations of LRBA deficiency are highly variable.^{9,12} Some patients have early onset PID, while others initially present with autoimmunity without PID.⁸ The molecular mechanism of how LRBA deficiency causes such a wide spectrum of highly variable manifestations is unclear. Thus, there is no causal treatment available for LRBA deficiency. LRBA
25 deficiency decreases CTLA-4¹³ but increases PD-1.² CTLA-4 is required for the suppressive function of Tregs, while PD-1 is a negative regulator of Tregs. Accordingly, Treg is reduced in 16 out of 22 LRBA-deficient patients.¹ A significant number of IPEX-like syndromes might result from LRBA deficiency.¹⁴ Treg is indispensable for controlling autoimmune diseases as demonstrated by the severity and intensity of
30 autoimmune disease in IPEX patients and in scurfy mice. However, it is unknown whether the defective Tregs causes the autoimmune diseases in LRBA deficiency, because their suppressive ability has not been studied in vivo. It is also unknown whether

LRBA deficiency directly impairs Tregs, since Tregs have only been studied in the context of all cells with LRBA mutation. In this proposal, we will test our hypothesis that LRBA is required for the development and function of Treg, and that deregulated Tregs play a major role in autoimmunity by transferring Genebow cells into an autoimmune diabetic mouse model. Furthermore, LRBA deficiency lacks a clear GPr as the clinical manifestations are highly variable, indicating complex GI. In other words, the phenotype is dependent on the genetic interaction of LRBA with other genes. Indeed, LRBA closely interacts with NFkB and other critical genes (Figs 11A-E).¹⁵⁻¹⁸ In this sense, the LRBA deficiency is a polygenic disease, and the genetic interactions of LRBA with other critical genes must be studied to define a clear GPr.

Mouse models.

The phenotype-based GI study in the intact animal (*‘in vivo veritas’*) provides the most compelling means to define the GPr. While a physical interaction study, based on two-hybrid, affinity capture-Western/mass spectrometry, is important, it does not link genotype to phenotype. GI study has been efficiently conducted in yeast but is extremely challenging in humans and animal models. This is because the variations from genetic background, epigenetic modifications,¹⁹ infections, *etc.*,²⁰⁻²² can interfere with or mask a gene's contribution to a phenotype, resulting in low phenotyping sensitivity or even false data.^{23,24} An impractically large number of mice is required to obtain meaningful data in GI study. For example, 500,000 subjects might be necessary to achieve sufficient statistical power to detect non-additive effects of genetic variants.³ Since LRBA deficiency is rare, such a human study would be impossible. To fulfil the two requirements, the inventors propose to generate and label the nine genotypes of two genes with four FRs (Figure 13A) in a single animal (a mouse), so that a large number of immune cells with nine different genotypes can be studied in a single flow cytometry tube (Figure 10). With large sample size (*e.g.*, ~2x 10⁶ Tregs and ~2x 10⁸ splenocytes/spleen) and eliminated variations, phenotyping sensitivity will be greatly increased (Figures 13A-C).

The principle of Genebow is based on, but an improvement over, the Brainbow and Cre/lox system. The mosaic analysis with double markers (MADM) system can be used to produce three genotypes with distinct fluorescent colors in a mouse.²⁶⁻²⁸ However, MADM can only label the genotypes of a gene that is mutated and distal to the insertion

site of the fluorescent genes on the same chromosome. It cannot be used to label two genes at the same time. Brainbow only labels cells.^{25,29} In contrast, Genebow can be used to label the genotypes of two genes. The advantages of Genebow are large sample size and high sensitivity, and thus it can be broadly used to obtain high resolution data of any immunophenotyping based on single cells. Another premise of Genebow is that the competitive assay with different genotypes has been widely used in research, *e.g.*, transplantation of mixed bone marrows and yeast fitness assay.

By studying each cell as an individual analogous to a mouse in a traditional assay, Genebow can greatly increase the sample size (up to several million) and eliminate background variations. Therefore, this greatly increases phenotyping sensitivity, making it possible, for the first time, to efficiently conduct GI study in animal models at very high resolution. While traditional GI study methods based on one-genotype-one-animal are much less powerful; 2) Genebow will be used to clearly determine whether the genetic interaction of LRBA and NFKB1 plays a critical role in regulating the development and autoimmune-suppression of Tregs *in vivo*, and thus will result in a clear GPr. These GI studies that require such large sample sizes cannot be conducted using current techniques.

PRELIMINARY RESULTS

1. LRBA interacts with MAPKs and NFkB (Figures 11A-E).

Figure 11A shows that LRBA knockdown (KD) increases NFkB activity. Consistently, phospho (P)-NFkB and cell survival (Figure 11B) are increased, in agreement with NFkB being a major cell survival regulator. Figure 11B shows that the inhibition of NFkB by PS-341 promotes apoptosis, which can be rescued by LRBA KD, suggesting that LRBA KD functions against PS-341 by increasing NFkB activity. LRBA promoter has several NFkB binding sites conserved in mouse. Figure 11C shows that NFkB increases LRBA promoter activity. *Lrba* is required for cell proliferation. The results in Figure 11D demonstrate that LRBA interacts with important genes in regulating apoptosis and proliferation. Figure 11E shows that LRBA KD downregulates P-ERK but upregulates P-p38, P-JNK and correspondingly MAPK substrate P-MNK1. Complementing this data, MAPK (ERK orthologue) is decreased but JNK increased upon rugose (LRBA orthologue) mutation.³⁰

2. Construction of target vector and single guide RNA (sgRNA)/Cas9 vector (Figures 12A-F).

The methods have been described in the inventors' recent publication, which is incorporated herein by reference in its entirety.⁷ The one-step generation of a mouse model by CRISPR requires a sgRNA/Cas9 vector and a target repair template vector for each gene. The Lrba target vector was constructed from a BAC clone using homologous recombineering (HR). The subsequent engineering of the vector included multiple rounds of HR and the inventors' CRISPR cloning method.⁶ The correct cloning of the target vector was characterized by restriction digestion and then confirmed by sequencing (Figures 12A-B). The gRNA sequence was cloned into the pX330 vector.³¹ The inventors then conducted CRISPR-mediated knockin in mouse ES cells and zygotes.³² The inventors found that 67% of G418-positive ES clones are specifically targeted (Figure 12C). Southern blot (D) confirmed the PCR results. The inventors also obtained high targeting efficiency (32%) and survival rate (87% of two cell eggs, 57% of blastocysts) of zygotes injected with only 1/5 of the normal concentrations of CRISPR reagents (Figure 12E). The inventors also attempted a knockin construct with short homologous arms and obtained high knockin efficiency (Figure 12F).

Production of the nine genotypes of two genes specifically labeled (the Genebow principle).

TCLC1 (Figure 13A) will be knocked into the genomic loci of Lrba. Cre recombinase-mediated recombination turns on red FR1 (knockout, ko, allele) or blue CFP (wild type, wt, allele) (Figure 13B). The frequencies of the ko and wt alleles should be equal, as Cre stochastically chooses each lox pair. As each cell has two alleles, ko cells (with two ko alleles) will be labeled as red fluorescent color (FC), wt cells as blue FC, and heterozygous (het) cells as red/blue FCs. Thus wt and ko cells can be produced in equal number. Similarly, Nfkb1 can be labeled with two other FRs: ko (purple), het (purple/green) and wt (green) by knocking in TCLC2. The nine genotypes of the two genes thus can be produced and respectively labelled with four FRs (Figure 10).

Genebow can produce high resolution phenotyping data required for GI study.

Statistically, phenotyping sensitivity is dependent positively on sample size but inversely on variation. The inventors propose to study each cell as an individual,

analogous to a mouse in a traditional assay, to greatly increase the sample size (up to several million: $\sim 2 \times 10^6$ Tregs/spleen). This can be done by producing all genotypes of two genes in a single animal (*e.g.*, a single mouse) and analyzing cells with these genotypes in one vessel (*e.g.*, flow cytometry tube). Thus, there are no background variations between the cells, because the cells from the same mouse should have the same genetic background, epigenetic modifications and environmental input. With greatly increased sample size and zero variation, high resolution (*e.g.*, 1000 times higher than current methods) phenotyping data can be obtained (Figure 13C).

To generate a Genebow mouse model in which *Lrba* and *Nfkb1* expression can be turned on/off in a spatiotemporal and trackable manner. The CRISPR system is currently used to quickly generate mouse models (~ 3 months) with high efficiency (up to 78%) and $\sim 100\%$ germline transmission.³⁹⁻⁴⁴ It causes DNA double strand break (DSB), which can be repaired by homology-directed-repair pathways. A DNA fragment flanked by two homologous sequences can be inserted into the DSB with high efficiency.⁴⁵ By using this technique, the inventors obtained correctly targeted ES cells and zygotes at high efficiency (67% and 32%, respectively).⁷ Therefore, the inventors will take advantage of the efficient multiplex targeting CRISPR system to knock the two TCLCs into the genomic loci of *Lrba* and *Nfkb1*, and the CreERT2 into the *Rosa26* genomic locus. (Figure 13 A).^{41,42}

Construction of CRISPR targeting vectors.

Two vectors, *i.e.*, a donor vector and Cas9/sgRNA vector, are required for targeting each gene. The Cas9/sgRNA vector provides Cas9 endonuclease and a target site-specific sgRNA to cleave the target site. The donor vector contains a TCLC flanked by two homologous fragments and serves as a CRISPR repair template to mediate the knockin of the TCLC. The inventors have established a protocol to readily construct these vectors.⁶ *Lrba* & *Nfkb1*: The Cas9/sgRNA vector for *Lrba* is ready and a Cas9/sgRNA vector for *Nfkb1* will be cloned as described.⁷ The detailed cloning strategy of donor vectors is shown in Figure 14A. CreERT2: CreERT2 consists of the Cre linked to the ligand-binding domain of the estrogen receptor. Cre recombinase activity thus can be efficiently induced by Tamoxifen (TAM) to mediate recombination between two lox sites (Figure 13B).⁴⁶⁻⁴⁸ Although transgenic mice that harbor this transgene are available, they

are not on a NOD background, which is required for this study. The sgRNA vector for this knockin is ready, and the donor vector pMB80 has been obtained from Addgene.⁴⁹ The functionality of the CRISPR vectors for the genes will be validated in ES cells. These vectors will be tested in ES cells before zygote injection and optimized as necessary using the inventors' CRISPR vector modification technique.⁶ Briefly, triple knockins of the two TCLCs and CreERT2 will be obtained by using the CRISPR technique (Figures 12A-F).⁴ The corrected targeted ES clones, confirmed by 5' and 3' PCR and Southern blot, will be subjected to TAM-induced Cre-cleavage. Gene expression will be detected by the expression of fluorescent proteins by flow cytometry directly.

Generation of mouse model through CRISPR technology.

The T7 promoter will be added to the sgRNAs of Nfkb1 or Rosa26 by PCR.^{6,39} The RNAs will be synthesized by T7 RNA polymerase in vitro and purified.⁶ Then, the Cas9 mRNA (TriLink)/sgRNA/plasmid donor (1 ng/μl/0.5 ng/μl/2 ng/μl) for each gene will be injected into the pronucleus of fertilized eggs. The concentrations of CRISPR reagents are only 1/5 of the amounts described,⁵⁰ but they produce high targeting efficiency and low toxicity (Figure 12E). The genomic DNA from targeted and control mice (age 8 to 12 days) will be extracted from clipped toes and used to PCR-identify mice that have Lrba, Nfkb1, and Rosa26 correctly targeted at 5' and 3' ends (Figure 12C),⁴ which will be confirmed by Southern blot with the probes from CFP or RFP (Figure 12D).⁴

Expected Outcomes and Alternative Strategy: 1) The inventors expect to obtain the correctly triple-targeted mice. The three loci have been targeted by traditional and/or CRISPR methods and disruption of any of them is not lethal in humans and/or mice.⁵¹ The Rosa26 is a safe harbor for knockin.⁵² Although TCLC is large (~3 kb), the inventors have already inserted large (7.4 and 5.8kb) fragments into the Lrba locus at high efficiency (Figures 12A-F). Others have also inserted ~6 kb fragments.^{39,53} The off-target mutations induced by CRISPR targeting are low or undetectable in mice.^{54,55} Furthermore, the inventors will use the new high-fidelity CRISPR-Cas9 nuclease to eliminate off-targeting.⁵⁶ Triple knockin should not be a concern as nine target loci were simultaneously edited in mouse zygotes.⁵⁷ If it is, single and/or double knockins, probably

as side products of the triple targeting, will be obtained separately and then bred into one triple knockin.

5 Determination of how the genetic interaction of Lrba and Nfkb1 regulates the development and function of Treg.

Many autoimmune diseases are caused by autoantibodies. Tregs play a determinant role in controlling autoimmunity,⁵⁸ likely by preventing autoantibody production, as Treg deficiency results in overproduction of autoantibodies. Treg is downregulated in 73% of LRBA deficiency, and autoantibodies are overproduced in 10 LRBA-deficient patients. Thus, Tregs are likely defective and play a key role in LRBA deficiency and it is important to study how Treg development and function are regulated by LRBA. Furthermore, highly variable clinical presentations indicate background genes have a strong influence on the overall penetrance of the phenotype. LRBA deficiency is modified by other genes to the extent that one subject is asymptomatic. One of the most 15 important modifiers could be NFKB1. LRBA and NFKB1 are closely related, *e.g.*, among the genes that cause autoimmunity when mutated, NFKB1 has the most autoimmune problems that are found in LRBA deficiency, and NFKB1 also likely regulates Tregs. It would be intriguing to study the genetic interaction of the two genes. Thus, an objective of this study was to determine whether Lrba plays a critical role, modified by Nfkb1, in 20 autoimmune diseases by regulating Treg development and function. The inventors' hypothesis is that deregulation of Lrba and Nfkb1 genetic interaction impairs the development and function of Tregs, resulting in reduced cell number and ability to suppress autoimmunity. To test this hypothesis, Tregs with nine different genotypes from Genebow mice will be transferred to an autoimmune mouse model to study how the 25 genotype affects the development and suppressive function of Tregs. It is the inventors' expectation that the phenotypic differences of the nine genotypes will be detected at high resolution to determine whether LRBA plays, modified by NFKB1, a critical role in regulating Tregs to suppress autoimmunity. It will demonstrate that genetic interaction study, for the first time, can be efficiently conducted in animal models.

30 LRBA and NFKB1 interact with each other (Figures 11A-E). NFKB1, among the genes that cause autoimmunity when mutated, has the most autoimmune problems that are found in LRBA deficiency. Mutation of NFKB2 or several upstream genes of NFKB

also cause immunodeficiency and autoimmunity.⁵⁹ LPS upregulates both *Lrba* and *Nfkb1*. Both genes interact with *Notch1* and *MAPK*.¹⁶ *LRBA-NFKB1* fusion is present in cancers,⁶⁰ reminiscent of *BCR-ABL* fusion in leukemia. NF- κ B family member *Rel* is required to turn on *FoxP3*, the most important transcription factor for Tregs, and *NFKB1* p105 is a repressor of *Rel*, suggesting that *NFKB1* may regulate Tregs. Therefore, *Lrba* and *Nfkb1* likely interact with each other to regulate Tregs and play a major role in autoimmunity.

Testing whether the genetic interaction of *Lrba* and *Nfkb1* is required for Treg development.

The majority of Tregs are generated in thymus and matured into functional Tregs in the periphery.⁵⁸ One of the most important Treg subsets is T follicular Treg (Tfr). Tfr cells promote the production of high affinity antibodies but suppress autoantibody production. Tfr deficiency results in overproduction of autoantibodies. Consistently, circulating Tfrs were reduced and autoantibodies were overproduced in *LRBA*-deficient patients.² However, the frequency of recent thymic emigrants (RTEs) was normal,² it may suggest that the development of RTEs into Tfrs is impaired. Developmental defects usually result in no or fewer mature cells. Cell number is one of the most important parameters of deregulated immune cells. Because Tfr is one of the end products of Treg development, defective Tfrs indicate defect(s) in one or more developmental stages of Tregs.

Experimental design: As shown in Figures 14A-D, tregs will be isolated by cell sorting from T cells enriched by autoMACS as described.⁶¹ T cells will be purified from the spleens of 12-day-old neonates treated with TAM* via negative selection of non-T cells. Cells will be stained with fluorochrome-conjugated CD4, CD3. *Foxp3* is genetically tagged with EGFP. Tregs (*Foxp3*+CD4+CD3+) with nine genotypes will be sorted and 3.6×10^5 of mixed (4×10^4 /genotype) Tregs will be injected i.p. into Treg-depleted non-obese diabetic (NOD) neonates at 12 days of age as described.⁶¹ For depletion of Treg cells, *Foxp3*-DTR+ mice will be injected i.p. with DT or PBS every other day from day 0-10.⁶¹ Finally, treated mice will be killed at 25 days and splenocytes will be stained with fluorochrome-conjugated antibodies of CD4, PD1, CXCR5 (the follicle homing receptor for both B and T cells), followed by standard flow cytometry†. Tfr cells are

Foxp3+CD4+PD1+CXCR5+. The following four parameters will be analyzed: total cell numbers, percent of total CD4 T cells, percent of FoxP3+CD4 Treg cells, and percent of CD4+CXCR5+ (follicular T) cells. Total Tfr numbers indicate relative differentiation/expansion. Tfr percent of total CD4 T cells provides an indication of relative Tfr cell development compared with total CD4 T cells. Tfr percent of FoxP3+CD4 Treg indicates Tfr cell development versus expansion of precursor Treg cells. Tfr percent of CD4+CXCR5+ cells indicates the Tfr/Tfh ratio, which serves as an important way to gauge the GC reaction. GFP levels represent FoxP3 expression level in Tregs.

*TAM-induced Cre recombination. The TAM-inducible Cre system allows optimizing the dose and time of TAM administration to have complete or near complete Cre recombination, which will be determined by flow cytometry of Tregs. If the recombination is completed, RFP+ (wt) or mKate2 + (ko) Tregs should be a half of total Tregs (all Tregs are labeled with EGFP). TAM will be administered to neonates via lactating mother intraperitoneally (ip) injected with tamoxifen at a dosage of 0, 50, 100, 200 mg/kg body weight, once per day from P3 to P5.

†Flow cytometry. The LSR-II flow cytometer has an analysis rate of up to 40,000 cells per second and the capacity to measure 16 cell markers. 1) To exclude cell aggregates, two sequential gates of scatter width vs height signals will be applied. The singlet population will then be gated by forward scatter vs side scatter to exclude dead cells and debris. "Live cells" will be first gated using a contour plot and then switched to a dot plot for easy monitoring of acquisition. 2) A FMO-control (Fluorescence Minus One) is a control sample composed of all antibody labels except one, and it will be used as a negative control in place of an isotype control for that antibody staining.

Determination of whether the genetic interaction of Lrba and Nfkb1 is required for the autoimmune-suppression function of Tregs.

Some LRBA-deficient patients have reduced Tfrs, defective Tregs with decreased suppressive ability *in vitro* and autoantibody overproduction,² it is unknown whether the defective Tregs cause the autoimmune diseases, because their suppressive ability *in vivo* has not been studied. It is also unknown whether NFKB1 deficiency causes autoimmunity by causing defective Tregs, although it likely does as NFKB1 closely interacts with LRBA and causes similar manifestations of autoimmunity when mutated. The inventors

hypothesize that the deregulated GI of *Lrba* and *Nfkb1* causes defective Tregs, which plays a major role in causing a variety of autoimmune diseases. To test this hypothesis, the autoimmune-suppression ability of Tregs with 9 different genotypes will be compared in an autoimmune diabetes mouse model (100% spontaneously develop into diabetes).

5 Depletion of Treg greatly accelerates this process.

Experimental design: The following experiments will be conducted as described.⁶¹ Tregs will be isolated as described above with the exception that the Tregs mouse donors will be 12 and 19 days of age. Treg-depleted NOD neonates (*Foxp3-DTR+*) at 12 and 19 days of age will be injected ip with 3.6×10^5 of Tregs[†] for each genotype from Genebow mice of the same age, and then followed for manifestations of autoimmune disease up to 100 days. The reason same age Tregs are to be injected into the same age of mice is that recent thymic emigrant Tregs take about 3 weeks to be matured into Tfrs. Blood glucose levels will be measured every 3 days after cell transfer for 30 days. The measurements for each group at every time point will be statistically analyzed. Other readouts are organ infiltration, weight loss, and death.⁶¹

[†]Pooled Tregs: For testing whether the genetic interaction of *Lrba* and *Nfkb1* is required for Treg development 3.6×10^5 per Treg injection is desired. Three injections require 1.08×10^6 . It may be possible to obtain such amount of Tregs from one mouse. If not, sorted Tregs will be pooled from several mice. For determining whether the genetic interaction of *Lrba* and *Nfkb1* is required for the autoimmune suppression function of Tregs, 3.6×10^5 per Treg injection for each genotype is desired. For 9 injections (each genotype per injection) in triplicate will require 10×10^6 Tregs. The inventors thus will pool sorted Tregs from ~10 mice. To ensure that each mouse has the same contribution to the pooled Tregs, the inventors will mix equal numbers of Tregs for each genotype from each mouse. In this manner, the pooled Tregs for each genotype will have the same genetic background, epigenetic modifications and environmental input as they were from one mouse.

Expected Outcomes, and Alternative Strategy. The inventors expect that 1) knockout (ko) of either or both genes will result in reduced Tfr number and suppression, and 2) dbko will have much fewer Tfr cells and less potent suppression than either single ko, and the reduced Tfr number and suppression are not additive results of two single knockouts. If both are true, it will demonstrate that both genes are required for Treg

development and function, and that the two genes genetically interact with each other, and NFKB1 would modify some phenotypes of LRBA deficiency, resulting in variable manifestations. In addition, the inventors expect that any small phenotypic differences between the nine genotypes will be detected by the Genebow model as it is highly sensitive. The Tfr number and suppressive ability of genotypes may follow the order: AABB > (AABb, AaBB) > (AAbb, AaBb, aaBB) > (Aabb, aaBb) > aabb (A=Lrba; B=Nfkb1). Thus the phenotypic difference between a mutant genotype and WT should be caused by that mutant genotype. Moreover, as each gene has two alleles, the dose effects of the two genes alone or in combination on Treg development and function should be detectable. Consequently, the inventors should be able to define a clear GPr at high resolution. These expectations are based, first, on that apoptosis plays a critical role in Treg development. Apoptosis is required for the positive and negative selection of Tregs in the thymus to develop into Tregs with appropriate affinity to self-antigens. In the periphery, Tregs are further shaped and maintained by the self-antigens through apoptosis. Second, there are more apoptosis in LRBA-deficient Tregs. The inventors also found LRBA downregulation induces apoptosis, and LRBA is overexpressed in several cancers. Others found that LRBA overexpression is a molecular signature of both mortality and recurrence risks for breast cancers. NFKB1 is an important cell survival regulator and is required for lymphocyte development,⁶² both genes likely exert their roles on apoptosis mainly by intrinsic mechanism. Since all genotypes are in the same mouse, Genebow also can be used to study the extrinsic function of a gene by transferring cells with different combinations of genotypes into wild type mice will allow us to distinguish cell intrinsic function from the extrinsic effects of other cells with the mutation. Alternative strategy: 1) Both genes likely exert their roles on the development and function of Tregs mainly intrinsically, as they may function mainly on intrinsic apoptosis of Tregs. However, even though there are some extrinsic influence, with such high sensitivity of Genebow assay, it should be possible to detect the difference caused by different genotypes. Further, if either of the genes functions extrinsically on apoptosis and confounds the interpretation of results, the inventors will omit Tregs with some genotypes in the Treg injection and should be able to determine what phenotype is caused by the extrinsic effects. It is possible that defects in other cells, such as B and dendritic cells, may contribute to the defective Tregs. If this is the case, the inventors will conduct the

above assay with Tregs combined with other type of cells (*e.g.*, B cells) with different genotypes.

Vertebrate Animal Subjects.

5 Animal model generation. Due to its small size, rapid breeding and low costs, mouse is the choice for this proposal. Type I diabetes is characterized by destruction of the beta cells (insulin-producing cells) relatively early in life, it is thought by autoantibodies against the beta cells. The nonobese diabetic mouse (NOD) is considered to be a good animal model of autoimmune type I diabetes. Two transcription control and
10 labeling cassette (TCLC) will be respectively inserted into the genomic loci of LRBA and NFkB1 by Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) technique. The correctly targeted mice identified will be used as the founders to establish the “Genebow” colony.

Study 1: Animal model generation. Both males and females (NOD/ShiLtJ) will be
15 used to establish the colony of the Genebow mouse model from the founders (could be both sex) generated by CRISPR/ homology-directed-repair (HDR). During breeding, unwanted mice also will be produced and will be included in the animal number as breeding. The colonies will be established as shown in Figure 15 and Table 1.

Table 1. Mice required to establish the colonies

Strain name/	Nickname	Genotype	Animal number	Ages	Sex	Usage
Founders	Triple knockin (Tkn)	Triple knockin (Tkn: Lrba, Nfkb1, Cre/ERT2)	20	2-3 month	Both	Breeding
NOD.Foxp3GFP	F3P	Foxp3GFP	4		Both	Tregs are EGFP+
NOD.Tkn.Foxp3P	TknF3P	Tkn(Lrba, Nfkb1, CreERT2+/-)/FoxP3GFP	100	2-3 month	Both	Breeding for neonates
NOD/ShiLtJ*	NOD	Wild type	4	2-3 month	Both	Breeding. Mated with the founders
NOD.Foxp3DTR/	F3D	Foxp3-DTR	4	2-3 month	Both	Tregs can be depleted by administration of Diphtheria Toxin
Total number			132			

*NOD/ShiLtJ, an autoimmune diabetes mouse model, which 100% spontaneously develop into diabetes and depletion of Treg greatly accelerates this process.

- 5 Study 2: Characterization of mouse model functionality. As the inventors conduct experiments with a single mouse that has leukocytes with the nine genotypes: Lrba (wt, het and ko) X Nfkb1 (wt, het and ko), because each genotype has the other eight genotypes as controls in a single mouse, there is no need for control mice for each genotype. Other exogenous protein controls such as Cre background control are not
- 10 required either, as all cells have the same proteins expressed. Moreover, as flow cytometry can analyze up to millions of cells with multiplex ability, small difference in phenotype can be detected with high confidence (low p values), allowing high sensitivity of phenotyping required to detect phenotypes caused by less penetrance (more dependent on genetic background, typical for genes in mutagenic diseases) of the targeted gene,
- 15 detecting more phenotypes. Therefore, the number of mice can be greatly reduced

Table 2. Mice required to conduct genetic interaction assay of *Lrba* and *Nfkb1* genes

Group	Nickname	Ages	Sex	Animal number	Usage and Treatment
1	TknF3P	Neonates -several months	Both	6* 3*3 = 54 [Animals(Both sex) x 3 (TAM doses) x 3 (experimental repeats)]	Optimizing TAM treatment to have complete or near complete Cre recombination. TAM will be administered to neonates via lactating mother intraperitoneally injected with tamoxifen at a dosage of 0, 100 mg/kg body weight, once per day from P3 to P5.
2	TknF3P	Neonates - several months	Both	6* 3*3 = 54 [6 Animals (Both sex) x 3 (with and without TAM) x 3 (experimental repeats)]	Treg donors†. Using optimal TAM treatments obtained from group 1. TAM treatment control included.
3	F3D	Neonates - several months	Both	6*3*3 = 54 [Animals (Both sex) x (with or without Treg transfer) x 3 (experimental repeats)]	Recipients for Treg transfer. Aim 2.1 , Experimental group. Tregs and DT control (PBS) included.
4	F3D	Neonates - several months	Both	54*1*3 = 162 [Animals (Both sex) x (with or without Treg transfer) x 3 (experimental repeats)]	Recipients for Treg transfer. Aim 2.2 , Experimental group. Tregs and DT control (PBS) included.
Total number				324	

5 Required: 324

+ Breeding: 276

Total = 600

Total animal numbers: 900.

†Pooled Tregs: For testing whether the genetic interaction of *Lrba* and *Nfkb1* is required

- 10 for Treg development, 3.6×10^5 per Treg injection is desired. Three injections require 1.08×10^6 . The inventors may be able to obtain such amount of Tregs from one mouse. If not, the inventors will pool sorted Tregs from several mice. For determination of whether the genetic interaction of *Lrba* and *Nfkb1* is required for the autoimmune suppression

function of Tregs, 3.6×10^5 per Treg injection for each genotype is desired. For 9 injections (each genotype per injection) in triplicate will require 10×10^6 Tregs. The inventors thus will pool sorted Tregs from ~10 mice. To ensure that each mouse has the same contribution to the pooled Tregs, the inventors will mix equal numbers of Tregs for each genotype from each mouse. In this manner, the pooled Tregs for each genotype will have the same genetic background, epigenetic modifications and environmental input as they were from one mouse.

All patents, patent applications, provisional applications, and publications referred to or cited herein are incorporated by reference in their entirety, including all figures and tables, to the extent they are not inconsistent with the explicit teachings of this specification.

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims. In addition, any elements or limitations of any invention or embodiment thereof disclosed herein can be combined with any and/or all other elements or limitations (individually or in any combination) or any other invention or embodiment thereof disclosed herein, and all such combinations are contemplated with the scope of the invention without limitation thereto.

REFERENCES FOR SUPPLEMENTAL DISCLOSURE 2

1. Gamez-Diaz, L., August, D., Stepensky, P., Revel-Vilk, S., Seidel, M.G., Noriko, M., Morio, T., Worth, A.J., Blessing, J., Van de Veerdonk, F., Feuchtinger, T., Kanariou, M., Schmitt-Graeff, A., Jung, S., Seneviratne, S., Burns, S., Belohradsky, B.H., Rezaei, N., Bakhtiar, S., Speckmann, C., Jordan, M. & Grimbacher, B. The extended phenotype of LPS-responsive beige-like anchor protein (LRBA) deficiency. *J Allergy Clin Immunol* 137, 223-30 (2016).
2. Charbonnier, L.M., Janssen, E., Chou, J., Ohsumi, T.K., Keles, S., Hsu, J.T., Massaad, M.J., Garcia-Lloret, M., Hanna-Wakim, R., Dbaibo, G., Alangari, A.A., Alsultan, A., Al-Zahrani, D., Geha, R.S. & Chatila, T.A. Regulatory T-cell deficiency and immune dysregulation, polyendocrinopathy, enteropathy, X-linked-like disorder caused by loss-of-function mutations in LRBA. *J Allergy Clin Immunol* 135, 217-27 (2015).

3. Zuk, O., Hechter, E., Sunyaev, S.R. & Lander, E.S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A* 109, 1193-8 (2012).
4. Wang, J.W., Howson, J.M., Ghansah, T., Despons, C., Ninos, J.M., May, S.L.,
5 Nguyen, K.H., Toyama-Sorimachi, N. & Kerr, W.G. Influence of SHIP on the NK repertoire and allogeneic bone marrow transplantation. *Science* 295, 2094-7 (2002).
5. Wang, J.W., Howson, J., Haller, E. & Kerr, W.G. Identification of a novel lipopolysaccharide-inducible gene with key features of both A kinase anchor proteins and chs1/beige proteins. *J Immunol* 166, 4586-95 (2001).
- 10 6. Wang, J.W., Wang, A., Li, K., Wang, B., Jin, S., Reiser, M. & Lockey, R.F. CRISPR/Cas9 nuclease cleavage combined with Gibson assembly for seamless cloning. *BioTechniques* 58, 161-70 (2015).
7. Wang, B., Li, K., Wang, A., Reiser, M., Saunders, T.L., Lockey, R.F. & Wang,
15 J.W. Highly efficient CRISPR/HDR-mediated knockin in mouse embryonic stem cells and zygotes. *BioTechniques*. 59, 201-208 (2015).
8. Serwas, N.K., Kansu, A., Santos-Valente, E., Kuloglu, Z., Demir, A., Yaman, A., Yaneth Gamez Diaz, L., Artan, R., Sayar, E., Ensari, A., Grimbacher, B. & Boztug, K. Atypical Manifestation of LRBA Deficiency with Predominant IBD-like Phenotype. *Inflamm Bowel Dis* 21, 40-7 (2015).
- 20 9. Seidel, M.G., Hirschmugl, T., Gamez-Diaz, L., Schwinger, W., Serwas, N., Deutschmann, A., Gorkiewicz, G., Zenz, W., Windpassinger, C., Grimbacher, B., Urban, C. & Boztug, K. Long-term remission after allogeneic hematopoietic stem cell transplantation in LPS-responsive beige-like anchor (LRBA) deficiency. *J Allergy Clin Immunol* (2014).
- 25 10. Wang, J.-W. & Lockey, R.F. Lipopolysaccharide-responsive beige-like anchor (LRBA), a novel regulator of human immune disorders. *Austin Journal of Clinical Immunology* 1, 9 (2014).
11. Alkhairy, O.K., Abolhassani, H., Rezaei, N., Fang, M., Andersen, K.K.,
30 Chavoshzadeh, Z., Mohammadzadeh, I., El-Rajab, M.A., Massaad, M., Chou, J., Aghamohammadi, A., Geha, R.S. & Hammarstrom, L. Spectrum of Phenotypes Associated with Mutations in LRBA. *J Clin Immunol* 36, 33-45 (2016).
12. Alangari, A., Alsultan, A., Adly, N., Massaad, M.J., Kiani, I.S., Aljebreen, A., Raddaoui, E., Almomen, A.K., Al-Muhsen, S., Geha, R.S. & Alkuraya, F.S. LPS-responsive beige-like anchor (LRBA) gene mutation in a family with
35 inflammatory bowel disease and combined immunodeficiency. *J Allergy Clin Immunol* 130, 481-8 e2 (2012).
13. Lo, B., Zhang, K., Lu, W., Zheng, L., Zhang, Q., Kanellopoulou, C., Zhang, Y., Liu, Z., Fritz, J.M., Marsh, R., Husami, A., Kissell, D., Nortman, S., Chaturvedi, V., Haines, H., Young, L.R., Mo, J., Filipovich, A.H., Bleesing, J.J., Mustillo, P., Stephens, M., Rueda, C.M., Chougnet, C.A., Hoebe, K., McElwee, J., Hughes, J.D., Karakoc-Aydiner, E., Matthews, H.F., Price, S., Su, H.C., Rao, V.K., Lenardo, M.J. & Jordan, M.B. AUTOIMMUNE DISEASE. Patients with LRBA deficiency show CTLA4 loss and immune dysregulation responsive to abatacept therapy. *Science* 349, 436-40 (2015).
- 40 14. Schreiner, F., Plamper, M., Dueker, G., Schoenberger, S., Gamez-Diaz, L., Grimbacher, B., Hilger, A.C., Gohlke, B., Reutter, H. & Woelfle, J. Infancy-onset

- T1DM, short stature and severe immunodysregulation in two siblings with a homozygous LRBA-mutation. *J Clin Endocrinol Metab*, jc20153382 (2016).
15. Schreiber, S.L., Preiss, A., Nagel, A.C., Wech, I. & Maier, D. Genetic screen for modifiers of the rough eye phenotype resulting from overexpression of the Notch antagonist hairless in *Drosophila*. *Genesis* 33, 141-52. (2002).
 - 5 16. Shamloula, H.K., Mbogho, M.P., Pimentel, A.C., Chrzanowska-Lightowlers, Z.M., Hyatt, V., Okano, H. & Venkatesh, T.R. rugose (rg), a *Drosophila* A kinase anchor protein, is required for retinal pattern formation and interacts genetically with multiple signaling pathways. *Genetics* 161, 693-710 (2002).
 - 10 17. de Souza, N., Vallier, L.G., Fares, H. & Greenwald, I. SEL-2, the *C. elegans* neurobeachin/LRBA homolog, is a negative regulator of lin-12/Notch activity and affects endosomal traffic in polarized epithelial cells. *Development* 134, 691-702 (2007).
 - 15 18. Wang, J.W., Gamsby, J.J., Highfill, S.L., Mora, L.B., Bloom, G.C., Yeatman, T.J., Pan, T.C., Ramne, A.L., Chodosh, L.A., Cress, W.D., Chen, J. & Kerr, W.G. Deregulated expression of LRBA facilitates cancer cell growth. *Oncogene* 23, 4089-97 (2004).
 - 20 19. Bryant, C.D., Zhang, N.N., Sokoloff, G., Fanselow, M.S., Ennes, H.S., Palmer, A.A. & McRoberts, J.A. Behavioral differences among C57BL/6 substrains: implications for transgenic and knockout studies. *J Neurogenet* 22, 315-31 (2008).
 - 20 20. Kado, S., Uchida, K., Funabashi, H., Iwata, S., Nagata, Y., Ando, M., Onoue, M., Matsuoka, Y., Ohwaki, M. & Morotomi, M. Intestinal microflora are necessary for development of spontaneous adenocarcinoma of the large intestine in T-cell receptor beta chain and p53 double-knockout mice. *Cancer Res* 61, 2395-8 (2001).
 - 25 21. Franklin, C.L. Microbial considerations in genetically engineered mouse research. *ILAR J* 47, 141-55 (2006).
 22. Treuting, P.M., Clifford, C.B., Sellers, R.S. & Brayton, C.F. Of mice and microflora: considerations for genetically engineered mice. *Vet Pathol* 49, 44-63 (2012).
 - 30 23. Ridgway, W.M., Healy, B., Smink, L.J., Rainbow, D. & Wicker, L.S. New tools for defining the 'genetic background' of inbred mouse strains. *Nat Immunol* 8, 669-73 (2007).
 24. Sundberg, J.P., Roopenian, D.C., Liu, E.T. & Schofield, P.N. The Cinderella effect: searching for the best fit between mouse models and human diseases. *J Invest Dermatol* 133, 2509-13 (2013).
 - 35 25. Cai, D., Cohen, K.B., Luo, T., Lichtman, J.W. & Sanes, J.R. Improved tools for the Brainbow toolbox. *Nat Methods* 10, 540-7 (2013).
 26. Espinosa, J.S., Tea, J.S. & Luo, L. Mosaic analysis with double markers (MADM) in mice. *Cold Spring Harb Protoc* 2014, 182-9 (2014).
 - 40 27. Hippenmeyer, S., Johnson, R.L. & Luo, L. Mosaic analysis with double markers reveals cell-type-specific paternal growth dominance. *Cell Rep* 3, 960-7 (2013).
 28. Liu, C., Sage, J.C., Miller, M.R., Verhaak, R.G., Hippenmeyer, S., Vogel, H., Foreman, O., Bronson, R.T., Nishiyama, A., Luo, L. & Zong, H. Mosaic analysis with double markers reveals tumor cell of origin in glioma. *Cell* 146, 209-21 (2011).
 - 45 29. Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R. & Lichtman, J.W. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* 450, 56-62 (2007).

30. Wech, I. & Nagel, A.C. Mutations in rugose promote cell type-specific apoptosis in the *Drosophila* eye. *Cell Death Differ* 12, 145-52 (2005).
31. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. & Zhang, F. Multiplex genome engineering using
5 CRISPR/Cas systems. *Science* 339, 819-23.
32. Yang, H., Wang, H. & Jaenisch, R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat Protoc* 9, 1956-68 (2014).
33. Erard, M., Fredj, A., Pasquier, H., Beltolngar, D.B., Bousmah, Y., Derrien, V., Vincent, P. & Merola, F. Minimum set of mutations needed to optimize cyan
10 fluorescent proteins for live cell imaging. *Mol Biosyst* 9, 258-67 (2013).
34. Shcherbakova, D.M. & Verkhusha, V.V. Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat Methods* 10, 751-4 (2013).
35. Kim, J.H., Lee, S.R., Li, L.H., Park, H.J., Park, J.H., Lee, K.Y., Kim, M.K., Shin, B.A. & Choi, S.Y. High cleavage efficiency of a 2A peptide derived from porcine
15 teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* 6, e18556 (2011).
36. Dean, K.M. & Palmer, A.E. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat Chem Biol* 10, 512-23 (2014).
37. Wan, Y.Y. & Flavell, R.A. Identifying Foxp3-expressing suppressor T cells with a
20 bicistronic reporter. *Proc Natl Acad Sci U S A* 102, 5126-31 (2005).
38. Boston-University-and-Boston-Medical-Center. Sample Size Calculations. <http://www.bu.edu/orcccommittees/iacuc/policies-and-guidelines/sample-size-calculations/>.
39. Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L. & Jaenisch, R. One-
25 step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370-9 (2013).
40. Zhou, J., Wang, J., Shen, B., Chen, L., Su, Y., Yang, J., Zhang, W., Tian, X. & Huang, X. Dual sgRNAs facilitate CRISPR/Cas9-mediated mouse genome targeting. *FEBS J* 281, 1717-25 (2014).
- 30 41. Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F. & Jaenisch, R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910-8.
42. Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L. & Jaenisch, R. One-
35 step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370-9.
43. Fujii, W., Kawasaki, K., Sugiura, K. & Naito, K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res* 41, e187.
44. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. & Charpentier, E. A
40 programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-21 (2012).
45. Donoho, G., Jasin, M. & Berg, P. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol* 18, 4070-8 (1998).
- 45 46. Aghajani, K., Keerthivasan, S., Yu, Y. & Gounari, F. Generation of CD4CreER(T2)) transgenic mice to study development of peripheral CD4-T-cells. *Genesis* 50, 908-13.

47. Metzger, D., Clifford, J., Chiba, H. & Chambon, P. Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. *Proc Natl Acad Sci U S A* 92, 6991-5 (1995).
48. Feil, S., Valtcheva, N. & Feil, R. Inducible Cre mice. *Methods Mol Biol* 530, 343-63 (2009).
49. McLaughlin, M.E., Kruger, G.M., Slocum, K.L., Crowley, D., Michaud, N.A., Huang, J., Magendantz, M. & Jacks, T. The Nf2 tumor suppressor regulates cell-cell adhesion during tissue fusion. *Proc Natl Acad Sci U S A* 104, 3261-6 (2007).
50. Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F. & Jaenisch, R. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* 153, 910-918 (2013).
51. Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M., Graham, D.B., Jhunjhunwala, S., Heidenreich, M., Xavier, R.J., Langer, R., Anderson, D.G., Hacohen, N., Regev, A., Feng, G., Sharp, P.A. & Zhang, F. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 159, 440-55 (2014).
52. Sadelain, M., Papapetrou, E.P. & Bushman, F.D. Safe harbours for the integration of new DNA in the human genome. *Nat Rev Cancer* 12, 51-8 (2012).
53. Kimura, Y., Hisano, Y., Kawahara, A. & Higashijima, S. Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. *Sci Rep* 4, 6545 (2014).
54. Inui, M., Miyado, M., Igarashi, M., Tamano, M., Kubo, A., Yamashita, S., Asahara, H., Fukami, M. & Takada, S. Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. *Sci Rep* 4, 5396 (2014).
55. Mashiko, D., Fujihara, Y., Satouh, Y., Miyata, H., Isotani, A. & Ikawa, M. Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. *Sci Rep* 3, 3355 (2013).
56. Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z. & Joung, J.K. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490-5 (2016).
57. Sakurai, T., Kamiyoshi, A., Kawate, H., Mori, C., Watanabe, S., Tanaka, M., Uetake, R., Sato, M. & Shindo, T. A non-inheritable maternal Cas9-based multiple-gene editing system in mice. *Sci Rep* 6, 20011 (2016).
58. Cebula, A., Seweryn, M., Rempala, G.A., Pabla, S.S., McIndoe, R.A., Denning, T.L., Bry, L., Kraj, P., Kisielow, P. & Ignatowicz, L. Thymus-derived regulatory T cells contribute to tolerance to commensal microbiota. *Nature* 497, 258-62 (2013).
59. Cunningham-Rundles, C. Human B cell defects in perspective. *Immunol Res* 54, 227-32 (2012).
60. Teles Alves, I., Hartjes, T., McClellan, E., Hiltemann, S., Bottcher, R., Dits, N., Temanni, M.R., Janssen, B., van Workum, W., van der Spek, P., Stubbs, A., de Klein, A., Eussen, B., Trapman, J. & Jenster, G. Next-generation sequencing reveals novel rare fusion events with functional implication in prostate cancer. *Oncogene* 34, 568-77 (2015).
61. Yang, S., Fujikado, N., Kolodin, D., Benoist, C. & Mathis, D. Immune tolerance. Regulatory T cells generated early in life play a distinct role in maintaining self-tolerance. *Science* 348, 589-94 (2015).

62. Gerondakis, S., Fulford, T.S., Messina, N.L. & Grumont, R.J. NF-kappaB control of T cell development. *Nat Immunol* 15, 15-25 (2014).
63. Tanaka, K.F., Ahmari, S.E., Leonardo, E.D., Richardson-Jones, J.W., Budreck, E.C., Scheiffele, P., Sugio, S., Inamura, N., Ikenaka, K. & Hen, R. Flexible Accelerated STOP Tetracycline Operator-knockin (FAST): a versatile and efficient new gene modulating system. *Biol Psychiatry* 67, 770-3 (2010).
64. Baudouin, S.J., Gaudias, J., Gerharz, S., Hatstatt, L., Zhou, K., Punnakal, P., Tanaka, K.F., Spooren, W., Hen, R., De Zeeuw, C.I., Vogt, K. & Scheiffele, P. Shared synaptic pathophysiology in syndromic and nonsyndromic rodent models of autism. *Science* 338, 128-32 (2012).
65. Ingenious-Targeting-Laboratory. Flexible Accelerated STOP Tetracycline Operator-knockin (FAST): a versatile and efficient new gene modulating system. <http://www.genetargeting.com/products-and-services/types-of-mouse-models/f-a-s-t/>.

Exemplified Embodiments

Examples of embodiments of the invention include, but are not limited to:

Embodiment 1. A non-human animal model comprising nine genotypes of two genes specifically labeled with four distinct detectable labels.

5 Embodiment 2. The non-human animal model of embodiment 1, wherein the two genes comprise a wild-type first gene (A), a wild-type second gene (B), an inactivated form of the first gene (a), and an inactivated form of the second gene (b), wherein the four distinct detectable labels comprise:

- 10 a first label co-expressed with the wild-type first gene (A),
- a second label co-expressed with the inactivated form of the first gene (a),
- a third label co-expressed with the wild-type second gene (B), and
- a fourth label co-expressed with the inactivated form of the second gene (b).

Embodiment 3. The non-human animal model of embodiment 2, wherein the animal comprises:

15 (a) a first genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the wild-type second gene (BB), wherein the wild-type first gene (A) is co-expressed with the first label and the wild-type second gene (B) is co-expressed with the third label;

20 (b) a second genotype comprising homozygous alleles for the wild-type first gene (AA), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the first gene (A) is co-expressed with the first label, the second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

25 (c) a third genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the inactivated form of the second gene (bb), wherein the homozygous alleles for the wild-type first gene (AA) are co-expressed with the first label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

30 (d) a fourth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the wild-type second gene (BB), wherein allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is

co-expressed with the second label, and the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(e) a fifth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, wherein the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(f) a sixth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the inactivated version of the second gene (bb), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(g) a seventh genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the wild-type second gene (BB), wherein the homozygous alleles for the inactivated version of the first gene (aa) are co-expressed with the second label, and wherein the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(h) an eighth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the one allele for the inactivated form of the second gene (b) is co-expressed with the fourth label; and

(i) a ninth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the inactivated version of the second gene (bb), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, and wherein the homozygous alleles for the inactivated version of the second gene (bb) is co-expressed with the fourth label.

Embodiment 4. The non-human animal model of embodiment 2 or 3, wherein the first gene is *Lrba* and the second gene is *NFkB*, or vice-versa.

Embodiment 5. The non-human animal model of any preceding embodiment, wherein the detectable labels are fluorescent labels or luminescent labels.

5 Embodiment 6. The non-human animal model of any preceding embodiment, wherein the animal model is a rodent.

Embodiment 7. The non-human animal model of any preceding embodiment, wherein the animal model is a mouse.

Embodiment 8. The non-human animal model of any one of embodiments 1 to 5,
10 wherein the animal model is a primate.

Embodiment 9. The non-human animal model of any preceding embodiment, wherein the animal model is further modified at the genetic or epigenetic level.

Embodiment 10. The non-human animal model of any preceding embodiment, wherein the animal model is further modified at the genetic or epigenetic level so as to be
15 useful in modeling a disease.

Embodiment 11. The non-human animal model of embodiment 10, wherein the disease is selected from the group consisting of cancer, cardiovascular disease, metabolic disease, monogenic disease.

Embodiment 12. The non-human animal model of embodiment 10, wherein the
20 disease is selected from the group consisting of Down Syndrome, cystic fibrosis, glaucoma, type-I diabetes, type-II diabetes, epilepsy, heart disease, muscular dystrophy, and gynecological cancer.

Embodiment 13. A composition comprising a population of cells from the non-human animal of embodiment 1.

25 Embodiment 14. A method for studying gene-gene interaction, comprising:
 providing a non-human animal model of any one of embodiments 1 to 12;
 analyzing the interaction of the two genes specifically labeled with four
distinct detectable labels.

Embodiment 15. The method of embodiment 14, wherein said analyzing
30 comprises analyzing the characteristics and/or behavior of the animal model or one or more cell populations of the animal model.

Embodiment 16. The method of embodiment 14, wherein said analyzing comprises subjecting cells of the animal model to flow cytometry.

Embodiment 17. The method of embodiment 14, wherein said analyzing is carried out in the presence and/or absence of an exogenous agent.

5 Embodiment 18. The method of embodiment 14, wherein said analyzing comprises measuring the detectable label of one of the genes and, optionally, comparing the measured detectable label to that of the other gene.

Embodiment 19. The method of embodiment 14, wherein the two genes comprise a wild-type first gene (A), a wild-type second gene (B), an inactivated form of the first
10 gene (a), and an inactivated form of the second gene (b), wherein the four distinct detectable labels comprise:

 a first label co-expressed with the wild-type first gene (A),
 a second label co-expressed with the inactivated form of the first gene (a),
 a third label co-expressed with the wild-type second gene (B), and
15 a fourth label co-expressed with the inactivated form of the second gene (b); and
 wherein said analyzing comprises measuring one or more of the first label, second
label, third label, or fourth label.

Embodiment 20. The method of embodiment 19, wherein the animal comprises:

 (a) a first genotype comprising homozygous alleles for the wild-type first gene
20 (AA) and homozygous alleles for the wild-type second gene (BB), wherein the wild-type
first gene (A) is co-expressed with the first label and the wild-type second gene (B) is co-
expressed with the third label;

 (b) a second genotype comprising homozygous alleles for the wild-type first gene
(AA), one allele for the wild-type second gene (B), and one allele for the inactivated form
25 of the second gene (b), wherein the first gene (A) is co-expressed with the first label, the
second gene (B) is co-expressed with the third label, and wherein the allele for the
inactivated form of the second gene (b) is co-expressed with the fourth label;

 (c) a third genotype comprising homozygous alleles for the wild-type first gene
(AA) and homozygous alleles for the inactivated form of the second gene (bb), wherein
30 the homozygous alleles for the wild-type first gene (AA) are co-expressed with the first
label, and wherein the homozygous alleles for the inactivated form of the second gene
(bb) are co-expressed with the fourth label;

(d) a fourth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the wild-type second gene (BB), wherein allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed with the second label, and the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(e) a fifth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, wherein the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(f) a sixth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the inactivated version of the second gene (bb), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(g) a seventh genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the wild-type second gene (BB), wherein the homozygous alleles for the inactivated version of the first gene (aa) are co-expressed with the second label, and wherein the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(h) an eighth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the one allele for the inactivated form of the second gene (b) is co-expressed with the fourth label; and

(i) a ninth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the inactivated version of the second gene (bb), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, and wherein the homozygous alleles for the inactivated version of the second gene (bb) is co-expressed with the fourth label; and
5 wherein said analyzing comprises measuring one or more of the first label, second label, third label, or fourth label.

Embodiment 21. A method for studying gene-gene interaction, comprising:
providing a population of cells from the non-human animal model of
10 embodiment 1; and
analyzing the interaction of the two genes specifically labeled with four distinct detectable labels.

Embodiment 22. The method of embodiment 21, wherein said analyzing comprises analyzing the characteristics and/or behavior of the cells.

15 Embodiment 23. The method of embodiment 21, wherein said analyzing comprises subjecting cells of the animal model to flow cytometry.

Embodiment 24. The method of embodiment 21, wherein said analyzing is carried out in the presence and/or absence of an exogenous agent administered to the animal *in vivo* or brought into contact with the cells *in vitro*.

20 Embodiment 25. The method of embodiment 21, wherein said analyzing comprises measuring the detectable label of one of the genes and, optionally, comparing the measured detectable label to that of the other gene.

Embodiment 26. The method of embodiment 21, wherein the two genes comprise a wild-type first gene (A), a wild-type second gene (B), an inactivated form of the first gene (a), and an inactivated form of the second gene (b), wherein the four distinct
25 detectable labels comprise:

a first label co-expressed with the wild-type first gene (A),
a second label co-expressed with the inactivated form of the first gene (a),
a third label co-expressed with the wild-type second gene (B), and
30 a fourth label co-expressed with the inactivated form of the second gene (b); and
wherein said analyzing comprises measuring one or more of the first label, second label, third label, or fourth label (*i.e.*, one, two, three, or all four labels).

Embodiment 27. The method of embodiment 19, wherein the animal comprises:

(a) a first genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the wild-type second gene (BB), wherein the wild-type first gene (A) is co-expressed with the first label and the wild-type second gene (B) is co-expressed with the third label;

(b) a second genotype comprising homozygous alleles for the wild-type first gene (AA), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the first gene (A) is co-expressed with the first label, the second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(c) a third genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the inactivated form of the second gene (bb), wherein the homozygous alleles for the wild-type first gene (AA) are co-expressed with the first label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(d) a fourth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the wild-type second gene (BB), wherein allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed with the second label, and the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(e) a fifth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, wherein the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(f) a sixth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the inactivated version of the second gene (bb), wherein the allele for the wild-type first gene

(A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

5 (g) a seventh genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the wild-type second gene (BB), wherein the homozygous alleles for the inactivated version of the first gene (aa) are co-expressed with the second label, and wherein the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

10 (h) an eighth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the one allele for the inactivated form of the second gene (b) is co-expressed with the fourth
15 label; and

(i) a ninth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the inactivated version of the second gene (bb), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, and wherein the homozygous alleles for the
20 inactivated version of the second gene (bb) is co-expressed with the fourth label; and

wherein said analyzing comprises measuring one or more of the first label, second label, third label, or fourth label (*i.e.*, one, two, three, or all four labels).

CLAIMS

We claim:

1. A non-human animal model comprising nine genotypes of two genes specifically labeled with four distinct detectable labels.

2. The non-human animal model of claim 1, wherein the two genes comprise a wild-type first gene (A), a wild-type second gene (B), an inactivated form of the first gene (a), and an inactivated form of the second gene (b), wherein the four distinct detectable labels comprise:

- a first label co-expressed with the wild-type first gene (A),
- a second label co-expressed with the inactivated form of the first gene (a),
- a third label co-expressed with the wild-type second gene (B), and
- a fourth label co-expressed with the inactivated form of the second gene (b).

3. The non-human animal model of claim 2, wherein the animal comprises:

(a) a first genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the wild-type second gene (BB), wherein the wild-type first gene (A) is co-expressed with the first label and the wild-type second gene (B) is co-expressed with the third label;

(b) a second genotype comprising homozygous alleles for the wild-type first gene (AA), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the first gene (A) is co-expressed with the first label, the second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(c) a third genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the inactivated form of the second gene (bb), wherein the homozygous alleles for the wild-type first gene (AA) are co-expressed with the first label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(d) a fourth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the wild-type

second gene (BB), wherein allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed with the second label, and the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(e) a fifth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, wherein the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(f) a sixth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the inactivated version of the second gene (bb), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(g) a seventh genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the wild-type second gene (BB), wherein the homozygous alleles for the inactivated version of the first gene (aa) are co-expressed with the second label, and wherein the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(h) an eighth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the one allele for the inactivated form of the second gene (b) is co-expressed with the fourth label; and

(i) a ninth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the inactivated version of the second gene (bb), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-

expressed with the second label, and wherein the homozygous alleles for the inactivated version of the second gene (bb) is co-expressed with the fourth label.

4. The non-human animal model of claim 2, wherein the first gene is *Lrba* and the second gene is *NFkB*, or vice-versa.

5. The non-human animal model of claim 1, wherein the detectable labels are fluorescent labels or luminescent labels.

6. The non-human animal model of claim 1, wherein the animal model is a rodent.

7. The non-human animal model of claim 1, wherein the animal model is a mouse.

8. The non-human animal model of claim 1, wherein the animal model is a primate.

9. The non-human animal model of claim 1, wherein the animal model is further modified at the genetic or epigenetic level.

10. The non-human animal model of claim 1, wherein the animal model is further modified at the genetic or epigenetic level so as to be useful in modeling a disease.

11. The non-human animal model of claim 10, wherein the disease is selected from the group consisting of cancer, cardiovascular disease, metabolic disease, and monogenic disease.

12. The non-human animal model of claim 10, wherein the disease is selected from the group consisting of Down Syndrome, cystic fibrosis, glaucoma, type-I diabetes, type-II diabetes, epilepsy, heart disease, muscular dystrophy, and gynecological cancer.

13. A composition comprising a population of cells from the non-human animal of claim 1.

14. A method for studying gene-gene interaction, comprising:
providing a non-human animal model of claim 1; and
analyzing the interaction of the two genes specifically labeled with four distinct detectable labels.

15. The method of claim 14, wherein said analyzing comprises analyzing the characteristics and/or behavior of the animal model or one or more cell populations of the animal model.

16. The method of claim 14, wherein said analyzing comprises subjecting cells of the animal model to flow cytometry.

17. The method of claim 14, wherein said analyzing is carried out in the presence and/or absence of an exogenous agent.

18. The method of claim 14, wherein said analyzing comprises measuring the detectable label of one of the genes and, optionally, comparing the measured detectable label to the measured label of the other gene.

19. The method of claim 14, wherein the two genes comprise a wild-type first gene (A), a wild-type second gene (B), an inactivated form of the first gene (a), and an inactivated form of the second gene (b), wherein the four distinct detectable labels comprise:

a first label co-expressed with the wild-type first gene (A),
a second label co-expressed with the inactivated form of the first gene (a),
a third label co-expressed with the wild-type second gene (B), and
a fourth label co-expressed with the inactivated form of the second gene (b); and
wherein said analyzing comprises measuring one or more of the first label, second label, third label, or fourth label.

20. The method of claim 19, wherein the animal comprises:

(a) a first genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the wild-type second gene (BB), wherein the wild-type first gene

(A) is co-expressed with the first label and the wild-type second gene (B) is co-expressed with the third label;

(b) a second genotype comprising homozygous alleles for the wild-type first gene (AA), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the first gene (A) is co-expressed with the first label, the second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(c) a third genotype comprising homozygous alleles for the wild-type first gene (AA) and homozygous alleles for the inactivated form of the second gene (bb), wherein the homozygous alleles for the wild-type first gene (AA) are co-expressed with the first label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(d) a fourth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the wild-type second gene (BB), wherein allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed with the second label, and the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(e) a fifth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, wherein the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the allele for the inactivated form of the second gene (b) is co-expressed with the fourth label;

(f) a sixth genotype comprising one allele for the wild-type first gene (A), one allele for the inactivated version of the first gene (a), and homozygous alleles for the inactivated version of the second gene (bb), wherein the allele for the wild-type first gene (A) is co-expressed with the first label, the allele for the inactivated version of the first gene (a) is co-expressed second label, and wherein the homozygous alleles for the inactivated form of the second gene (bb) are co-expressed with the fourth label;

(g) a seventh genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the wild-type second gene (BB), wherein the homozygous alleles for the inactivated version of the first gene (aa) are co-expressed with the second label, and wherein the homozygous alleles for the wild-type second gene (BB) are co-expressed with the third label;

(h) an eighth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), one allele for the wild-type second gene (B), and one allele for the inactivated form of the second gene (b), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, the allele for the wild-type second gene (B) is co-expressed with the third label, and wherein the one allele for the inactivated form of the second gene (b) is co-expressed with the fourth label; and

(i) a ninth genotype comprising homozygous alleles for the inactivated version of the first gene (aa), and homozygous alleles for the inactivated version of the second gene (bb), wherein the homozygous alleles for the inactivated version of the first gene (aa) is co-expressed with the second label, and wherein the homozygous alleles for the inactivated version of the second gene (bb) is co-expressed with the fourth label; and

wherein said analyzing comprises measuring one or more of the first label, second label, third label, or fourth label.

21. A method for studying gene-gene interaction, comprising:

providing a population of cells from the non-human animal model of claim 1; and

analyzing the interaction of the two genes specifically labeled with four distinct detectable labels.

1/27

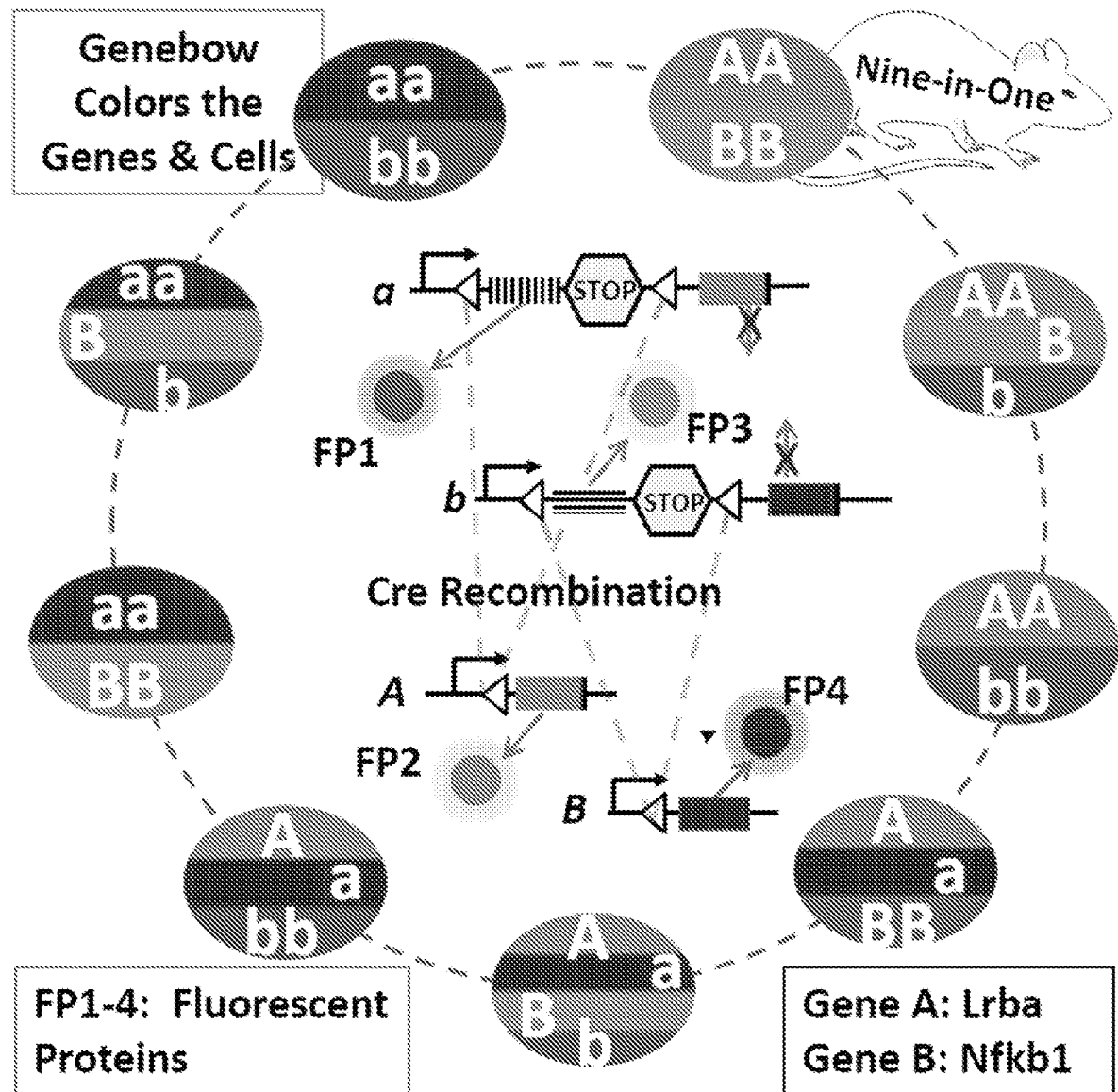
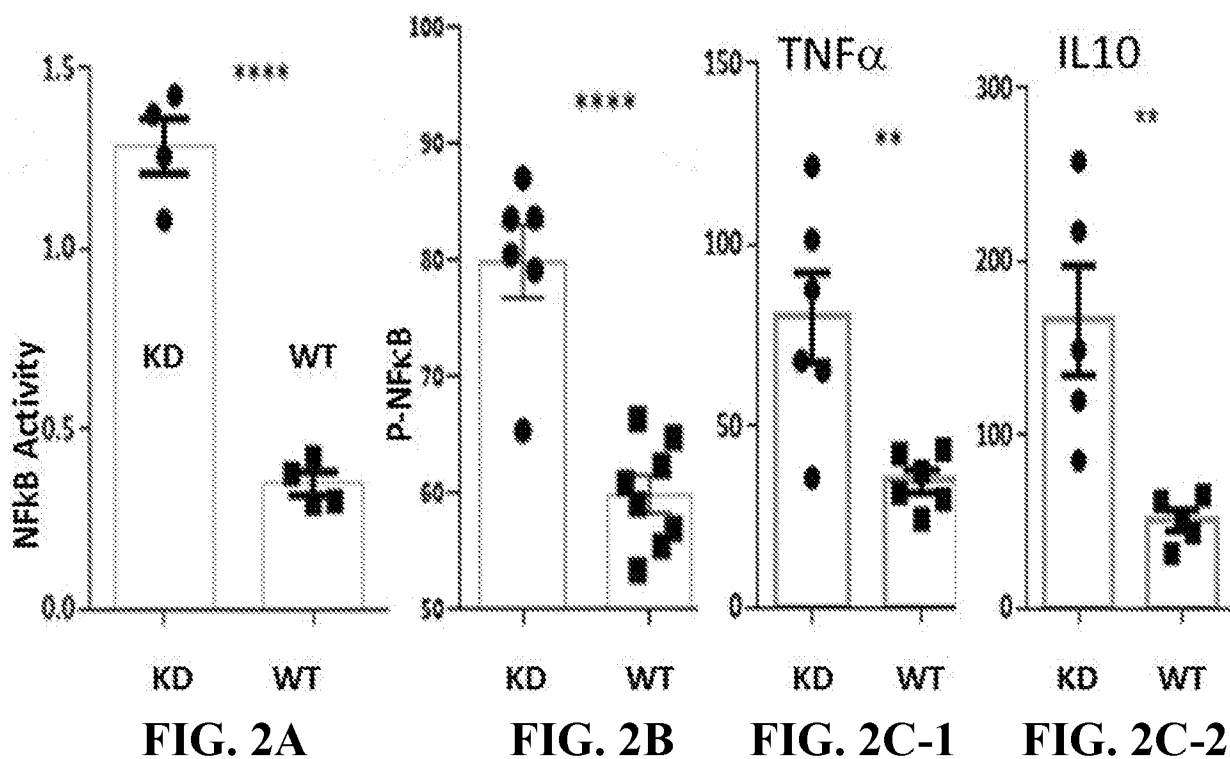


FIG. 1

2/27



PS-341 Serum Starvation Infection

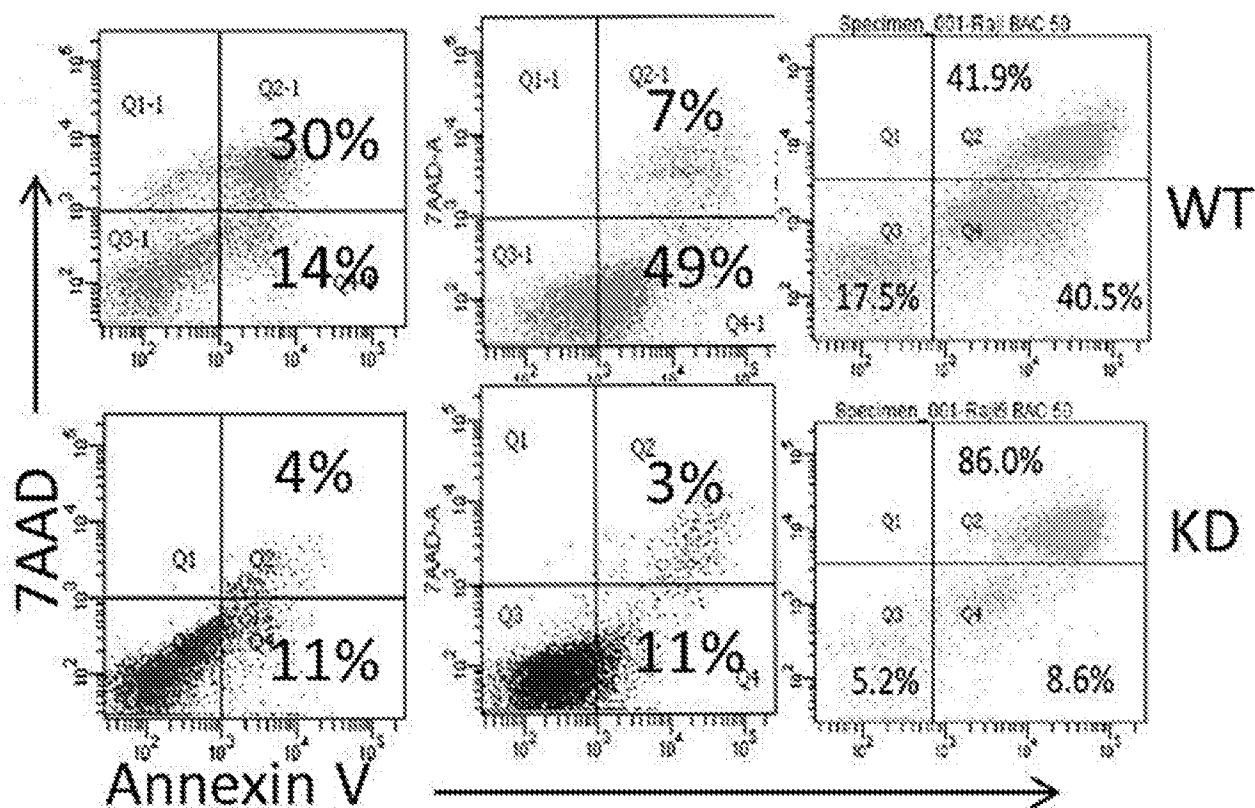


FIG. 2D

3/27

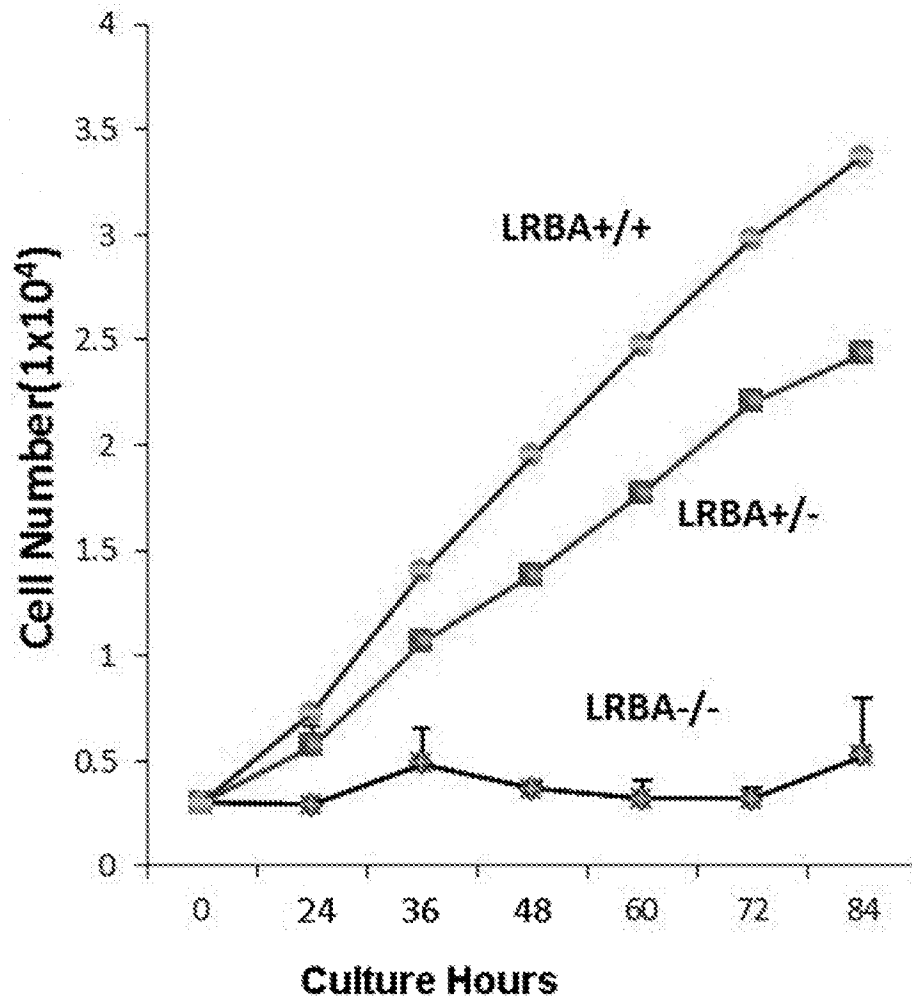


FIG. 2E

4/27

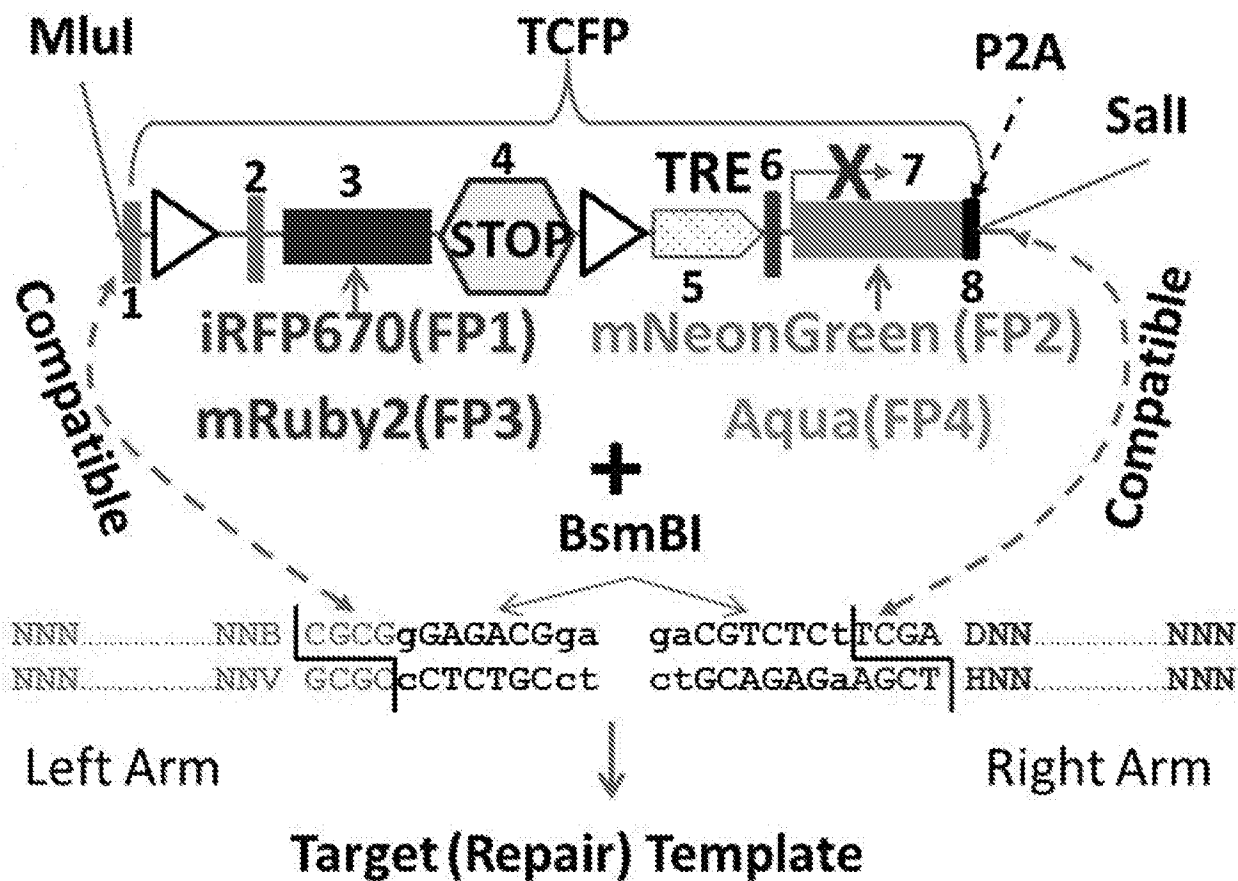


FIG. 3A

5/27

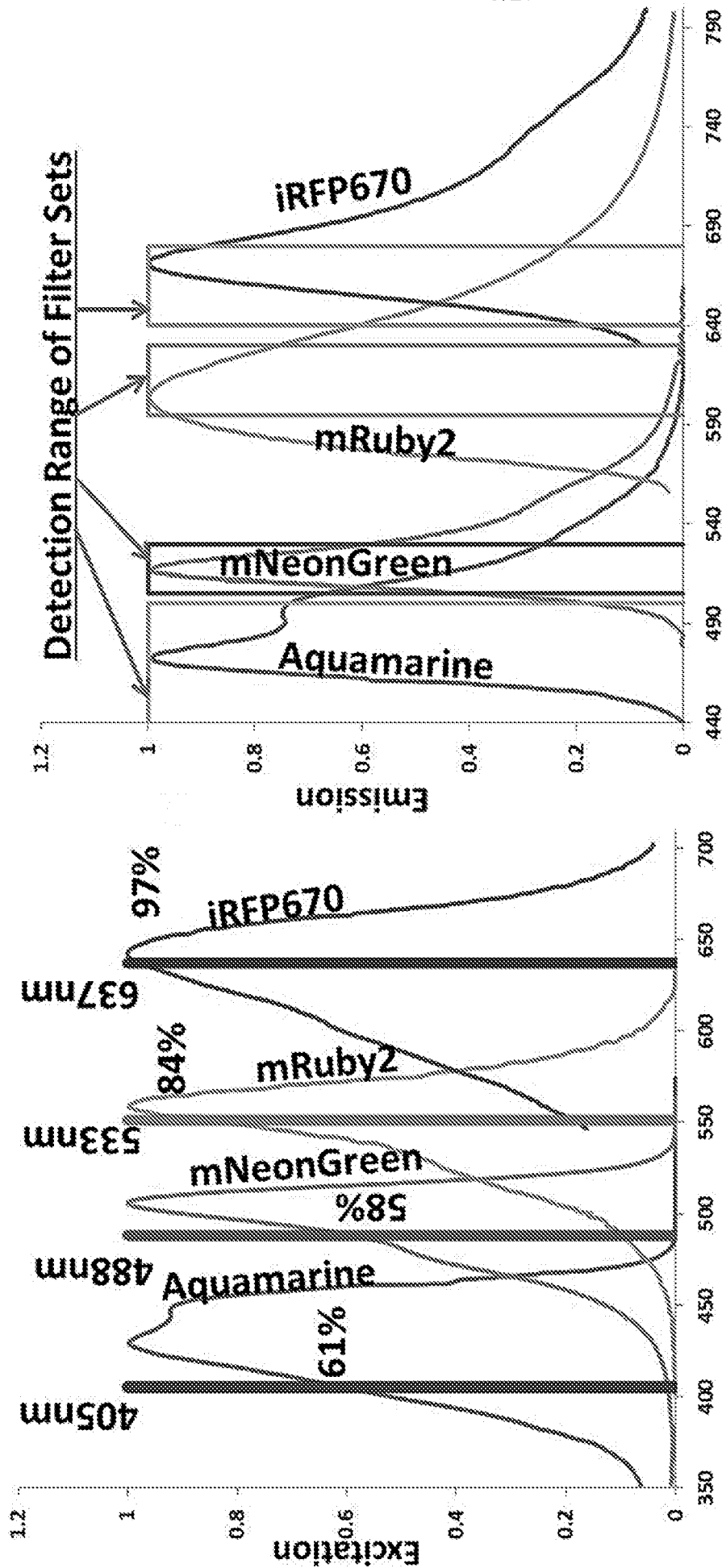
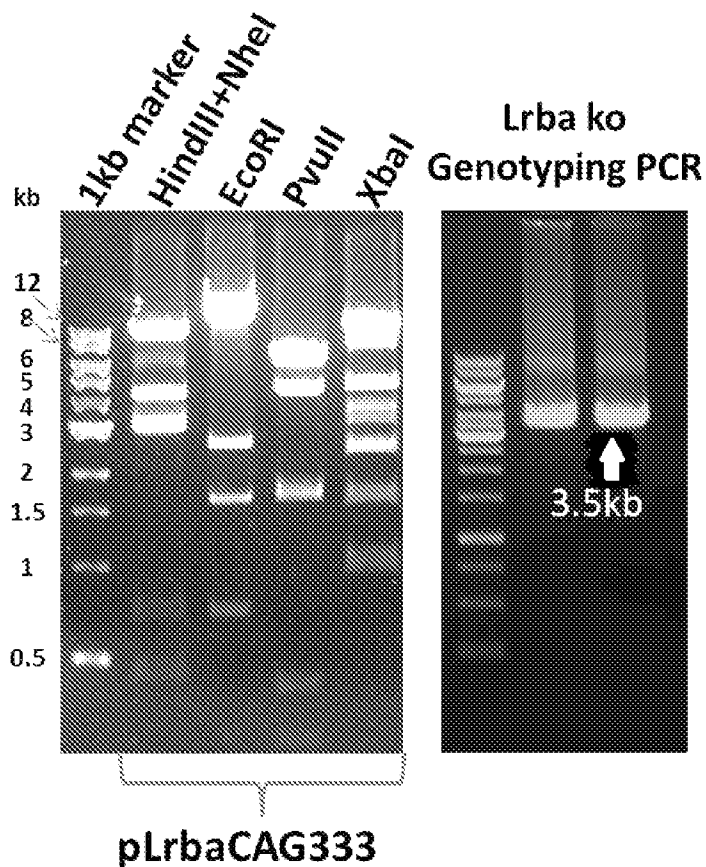


FIG. 3C

FIG. 3B

6/27



pLrbaCAG333

FIG. 4A

FIG. 4B

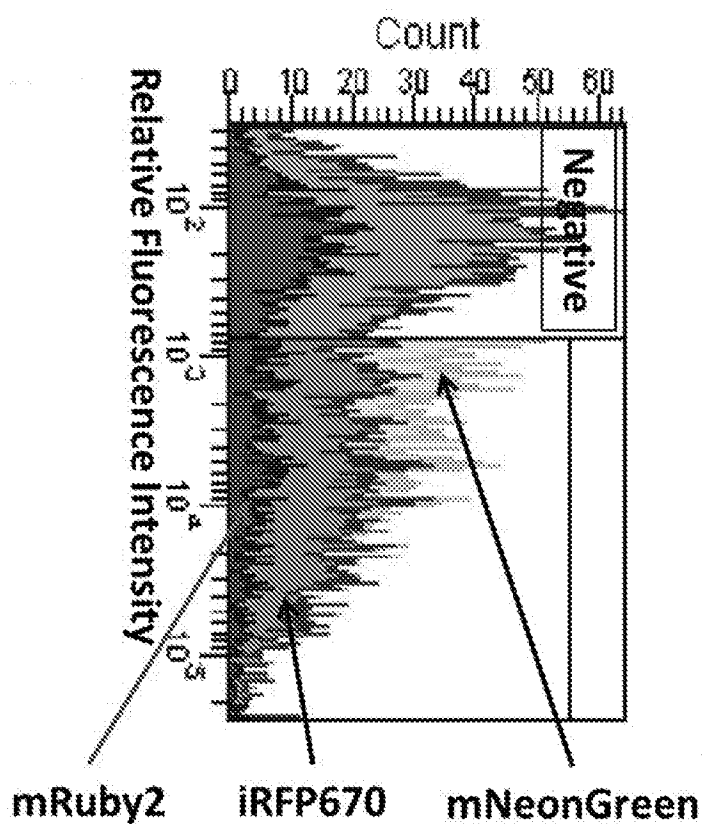


FIG. 4C

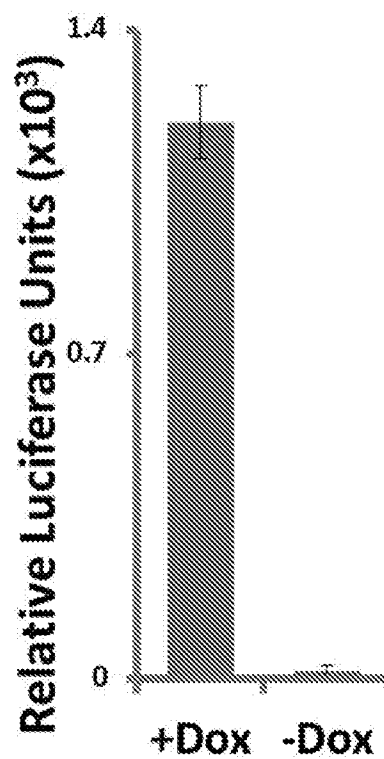


FIG. 4D

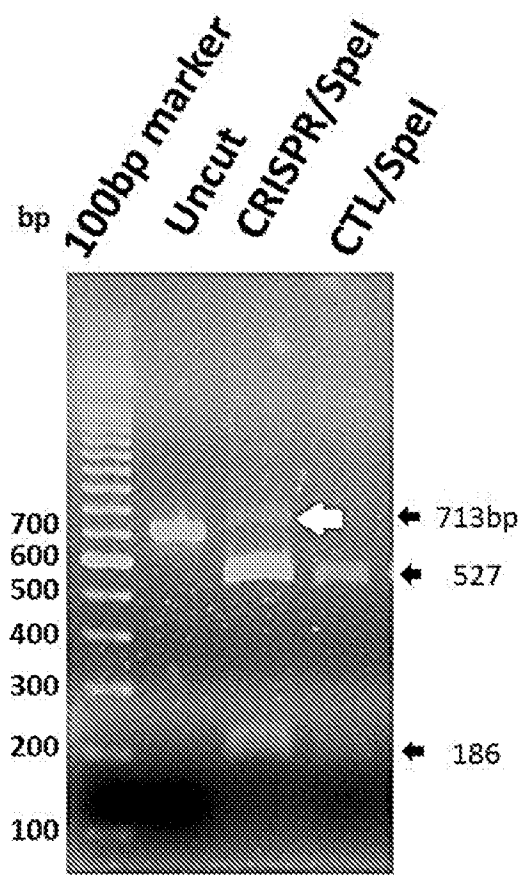


FIG. 4E

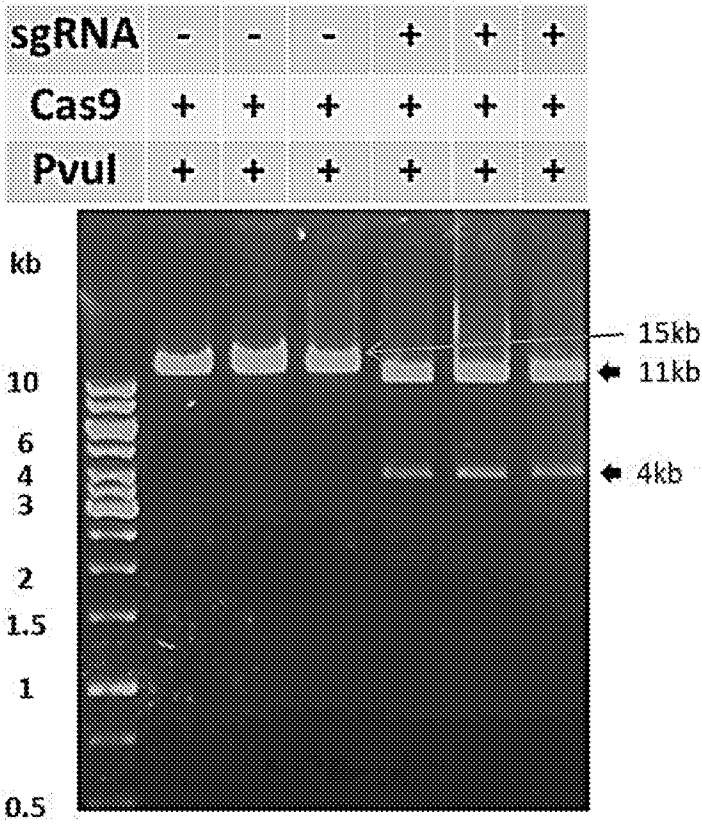


FIG. 4F

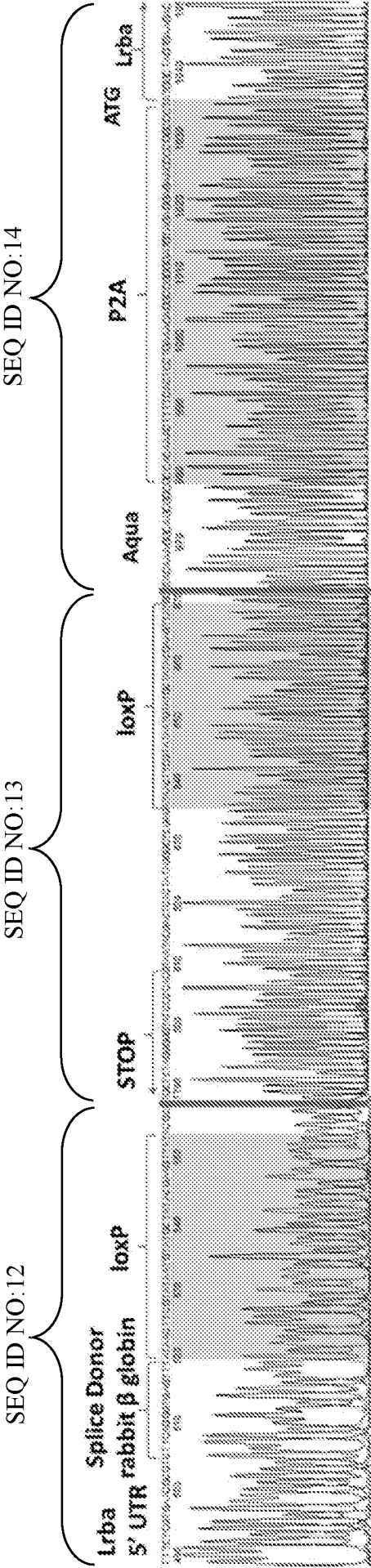


FIG. 4I

FIG. 4H

FIG. 4G

9/27

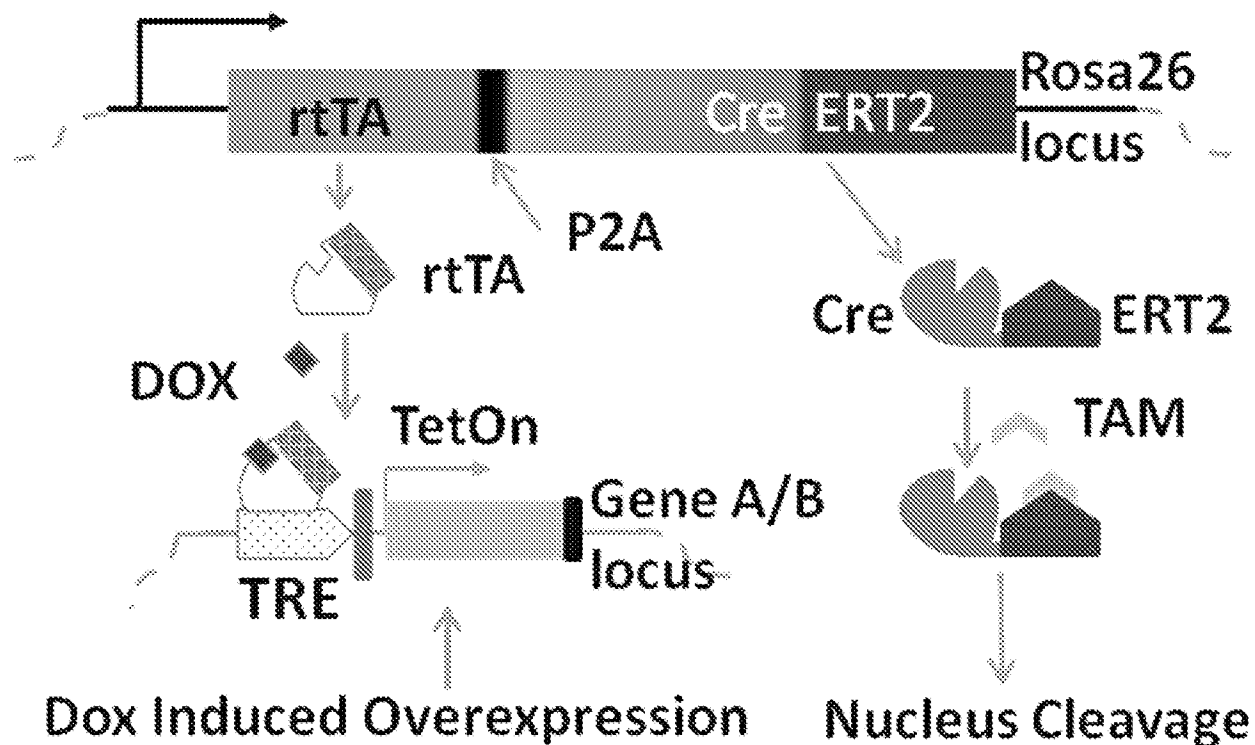


FIG. 5A

Days	0	7-10	10	17	24-31	31
Treatment	B I	DOX	B II	B III	TAM	B IV
Lrba	-	+++		-	++, +-, --	
Nfkb1	-	+++		-	++, +-, --	

FIG. 5B

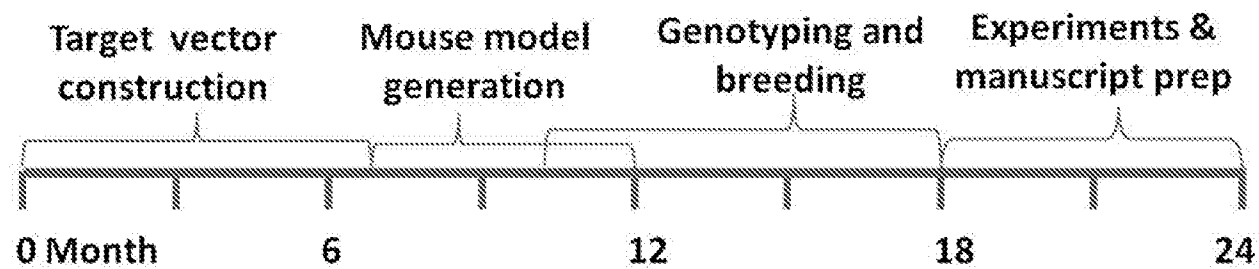


FIG. 5C

10/27

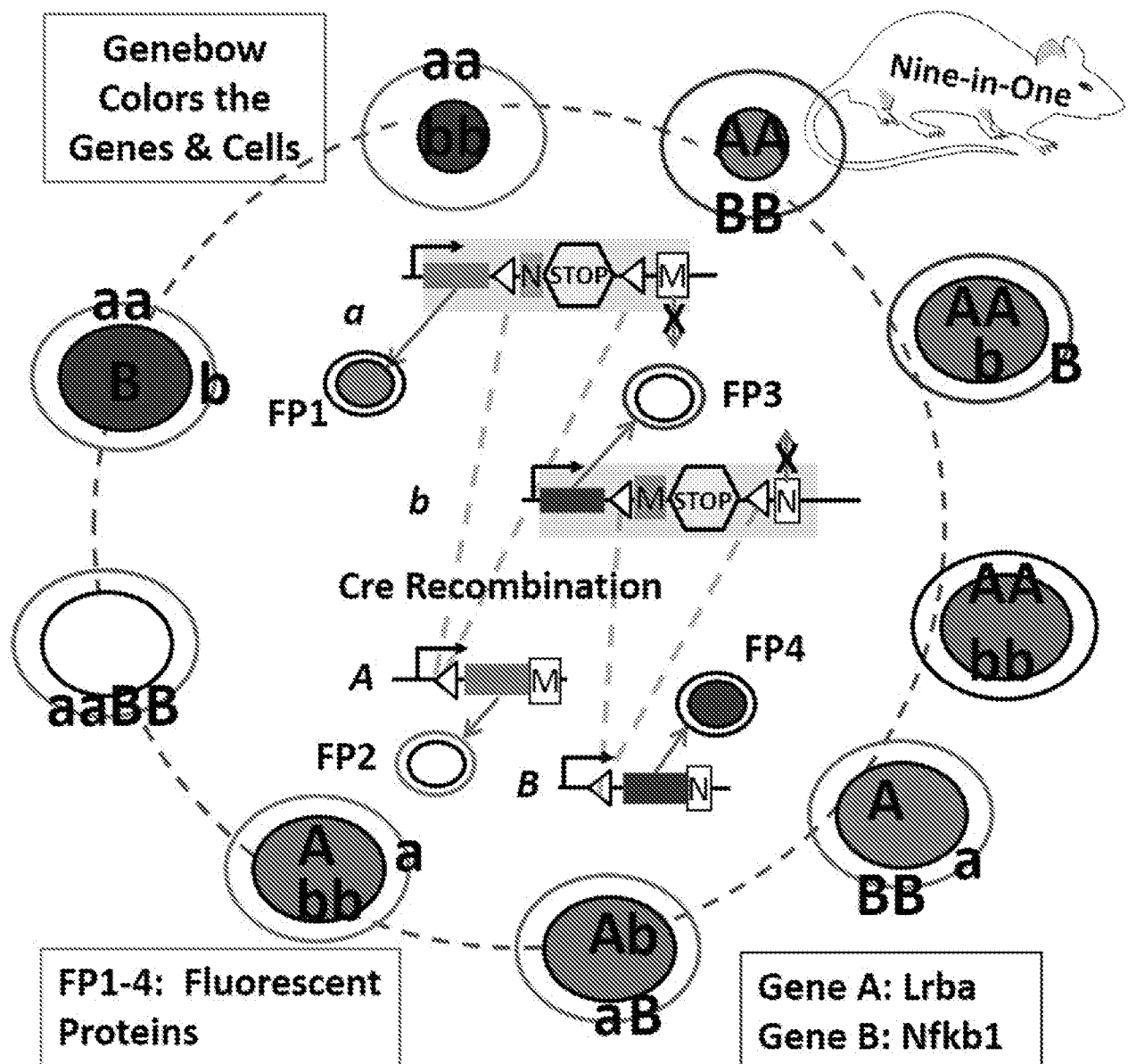


FIG. 6

11/27

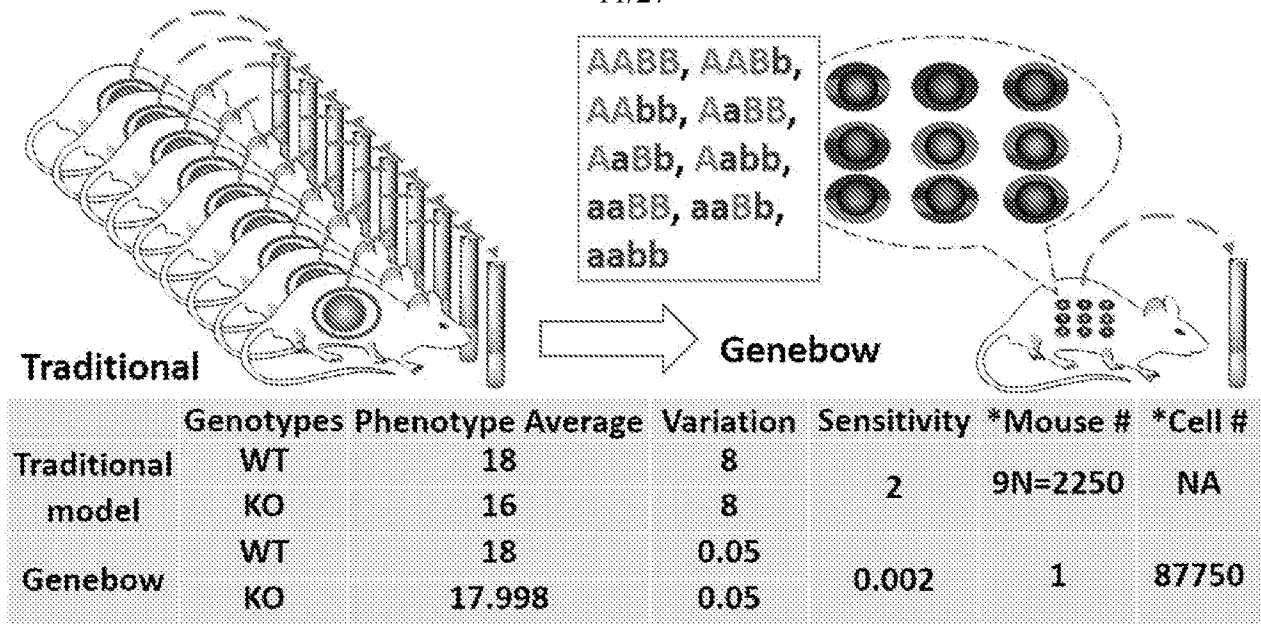


FIG. 7A

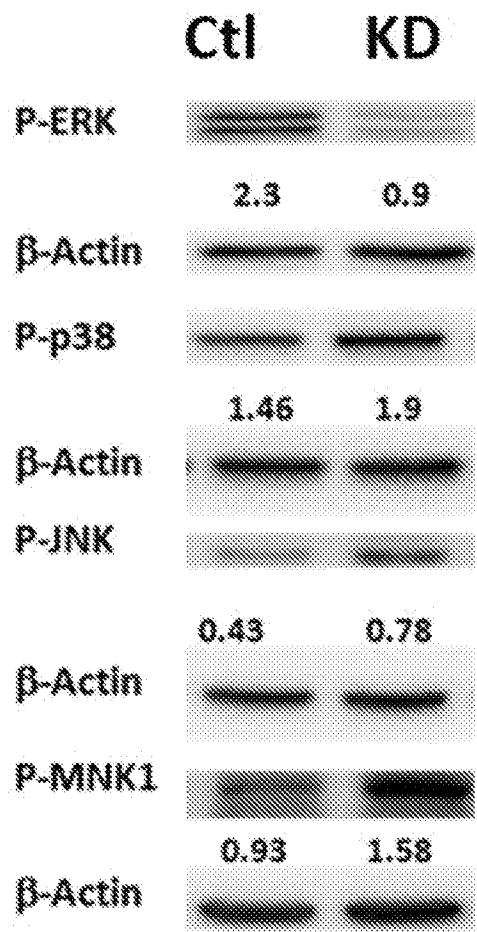


FIG. 7B

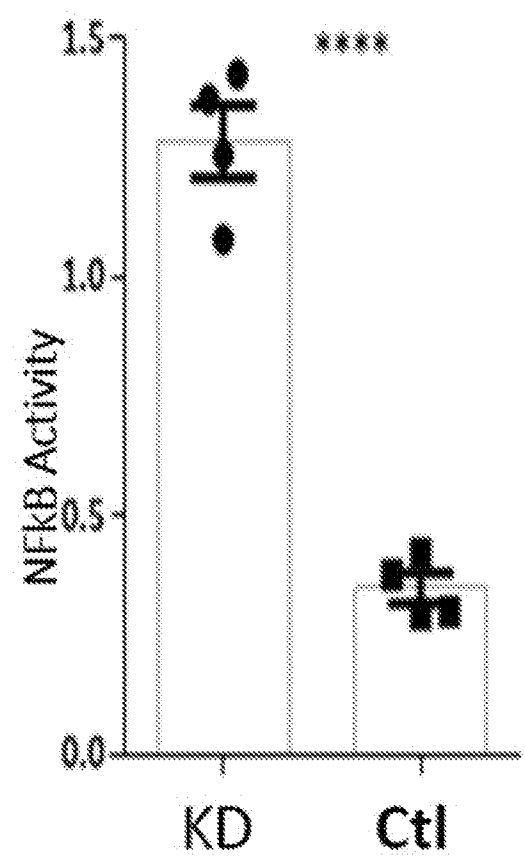


FIG. 7C

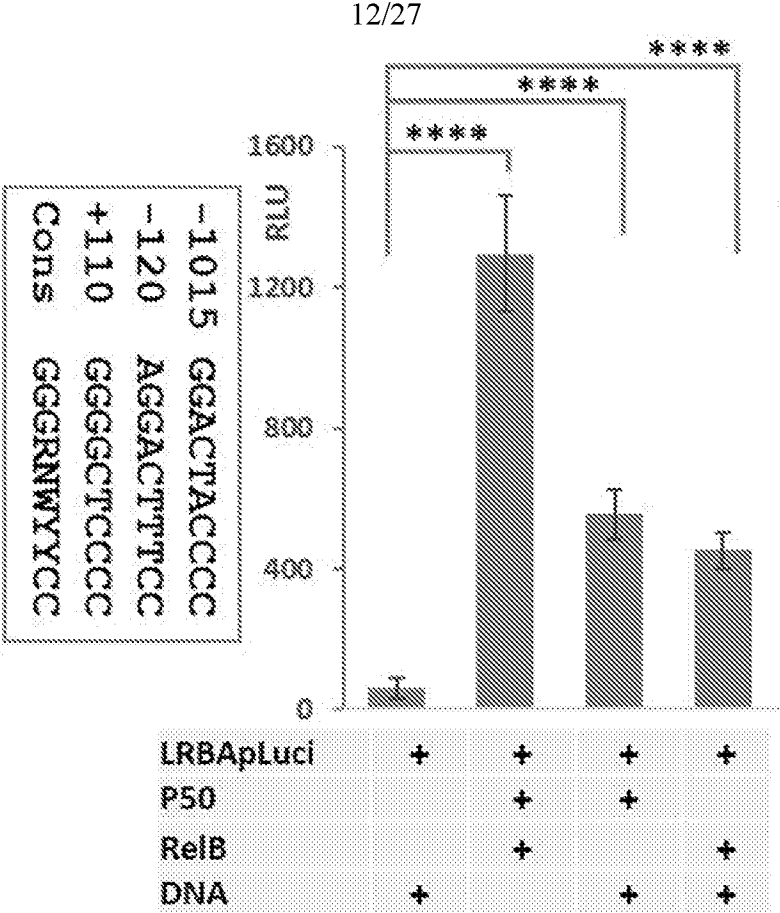


FIG. 7D

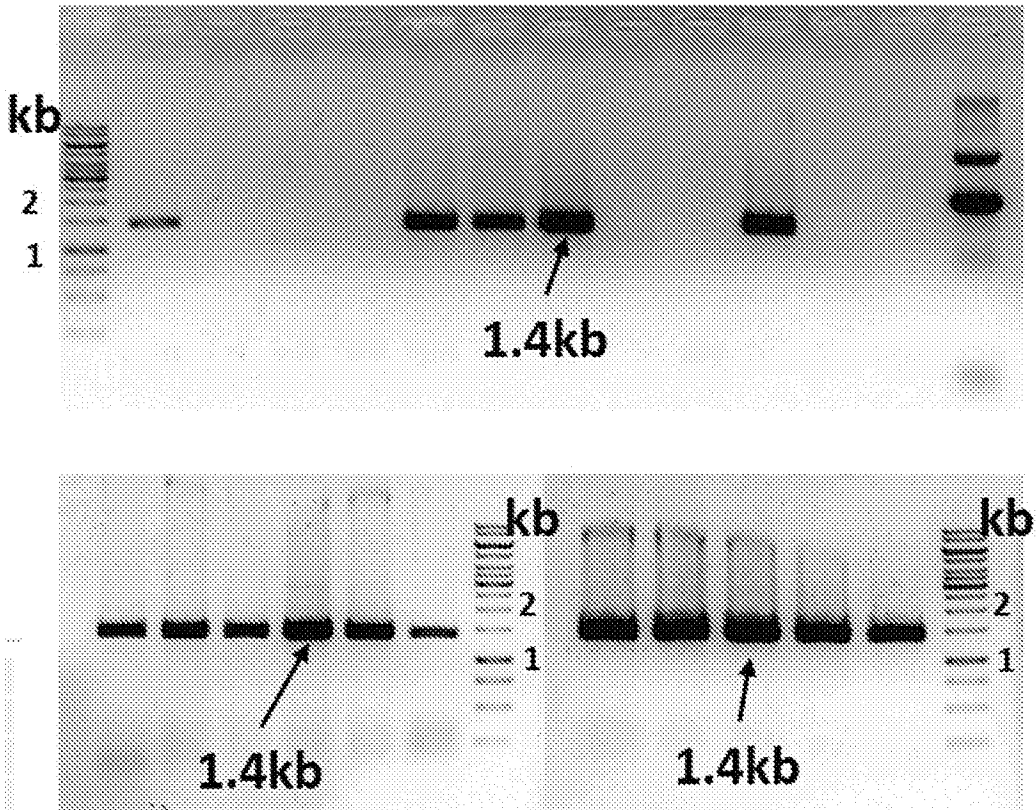


FIG. 7E

14/27

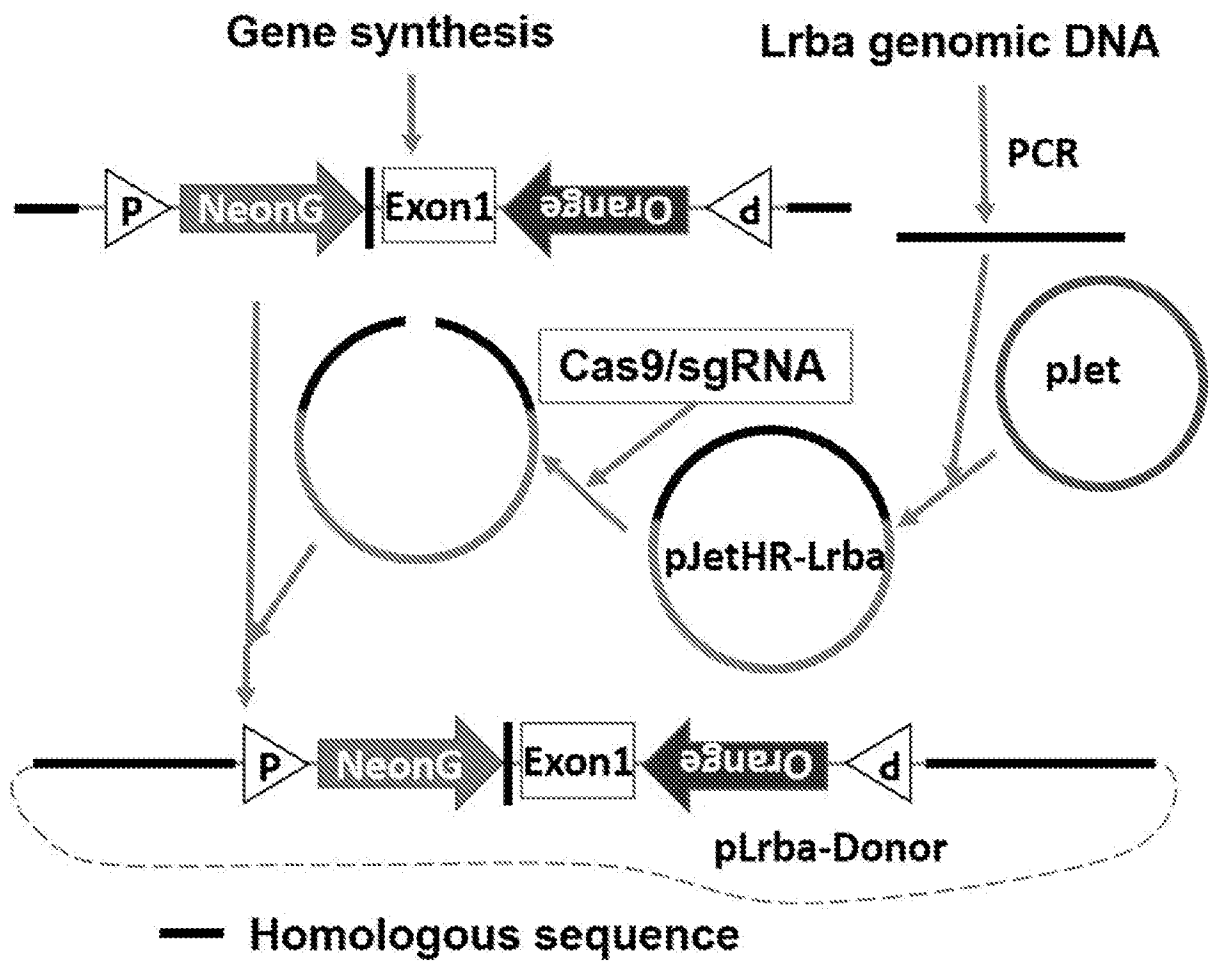


FIG. 8C

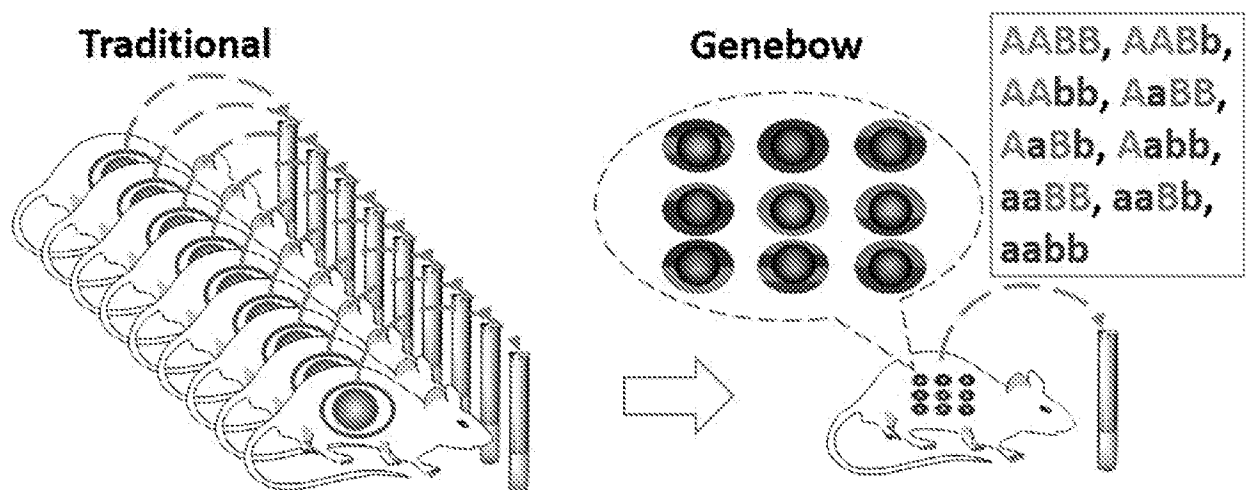
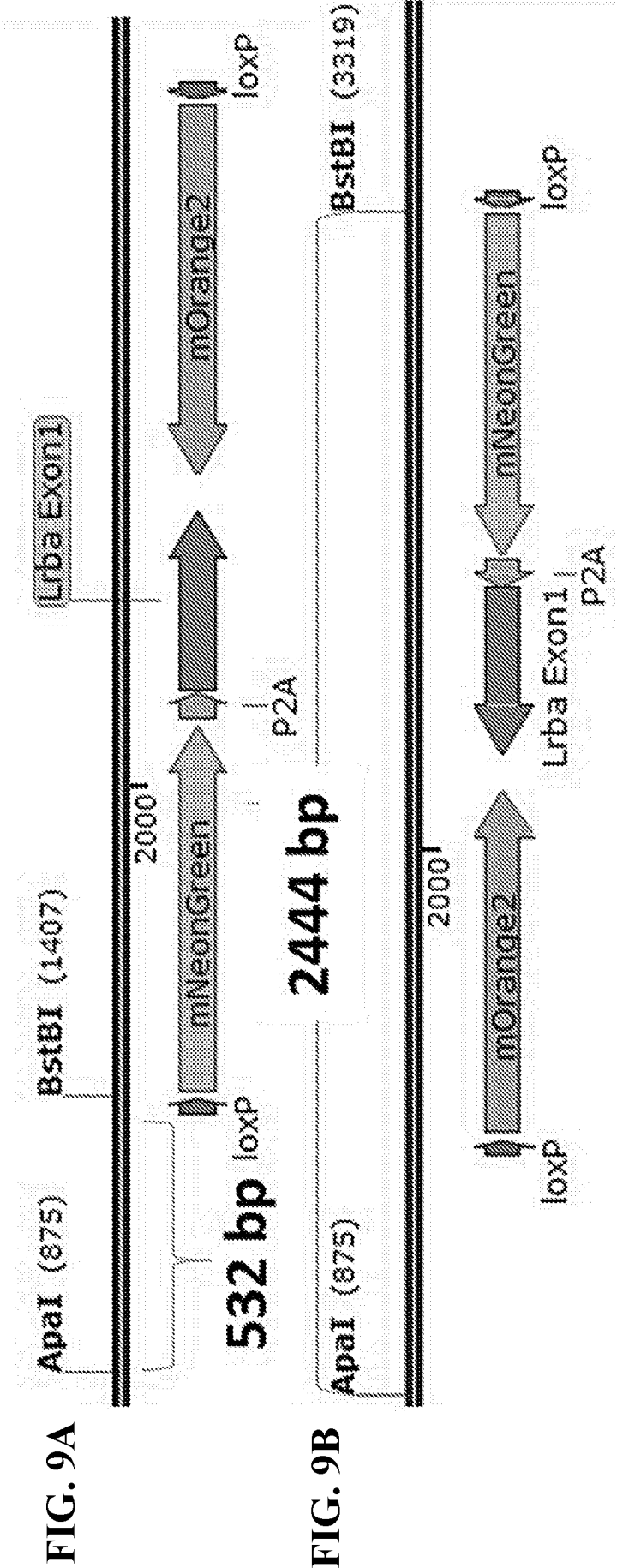


FIG. 10



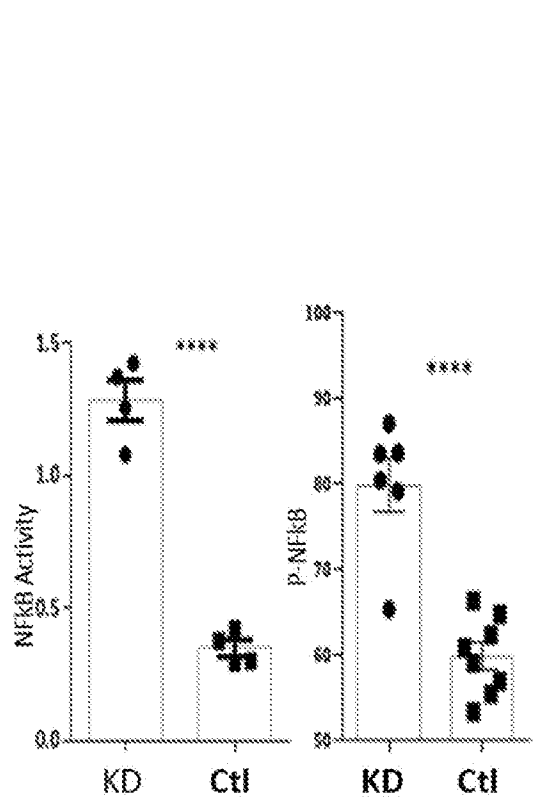


FIG. 11A

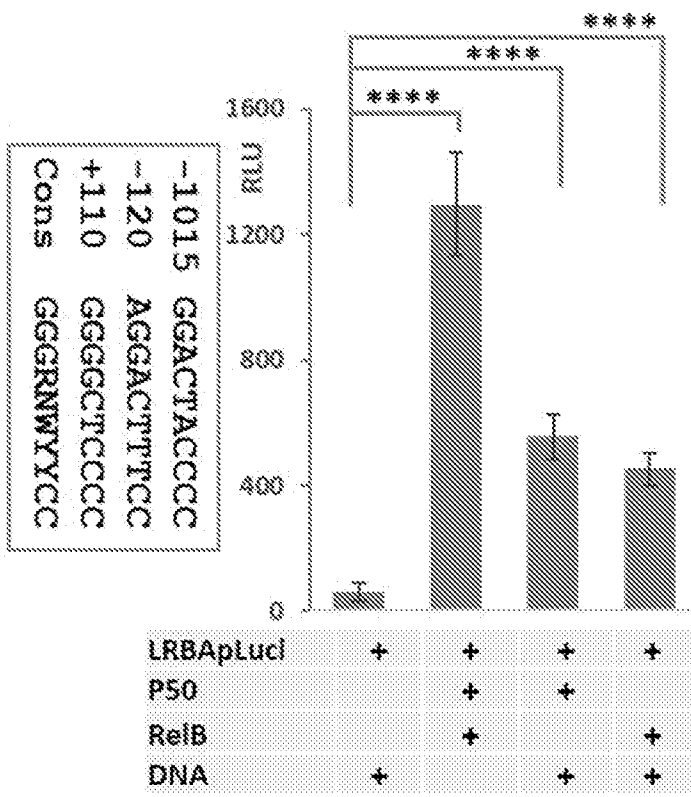


FIG. 11C

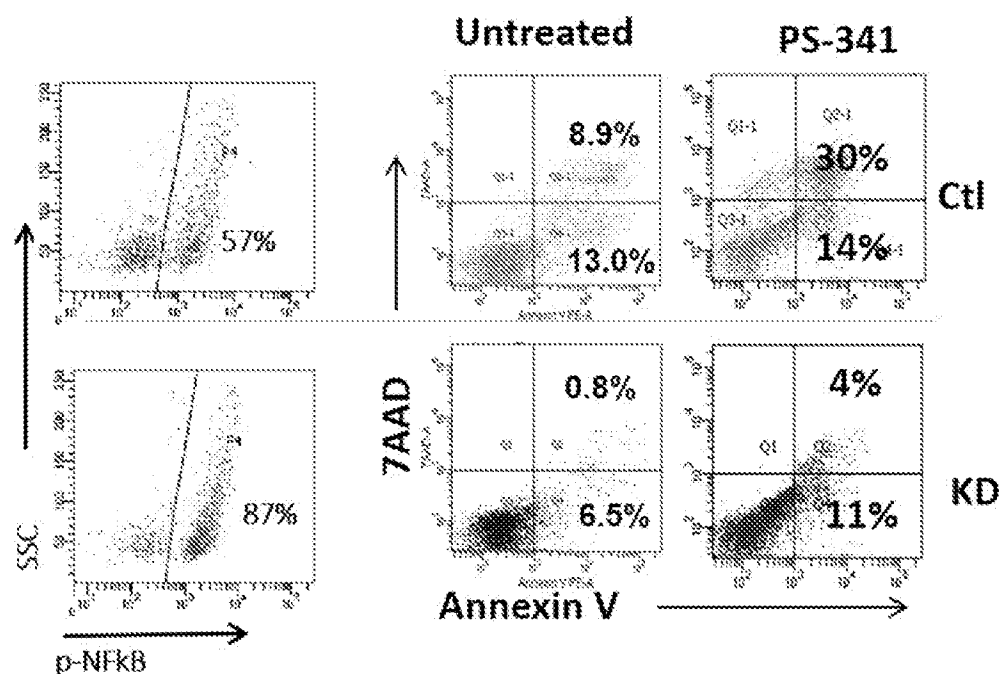


FIG. 11B

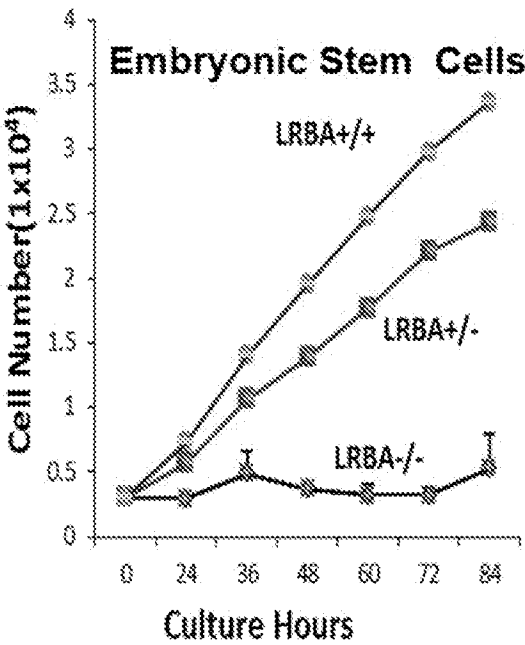


FIG. 11D

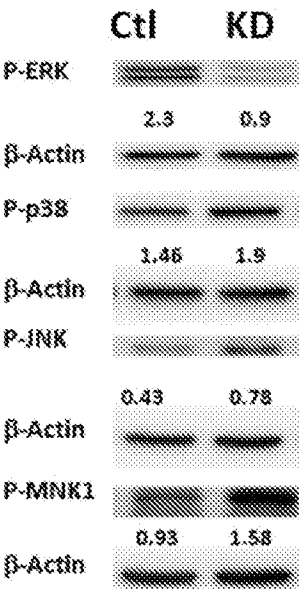


FIG. 11E

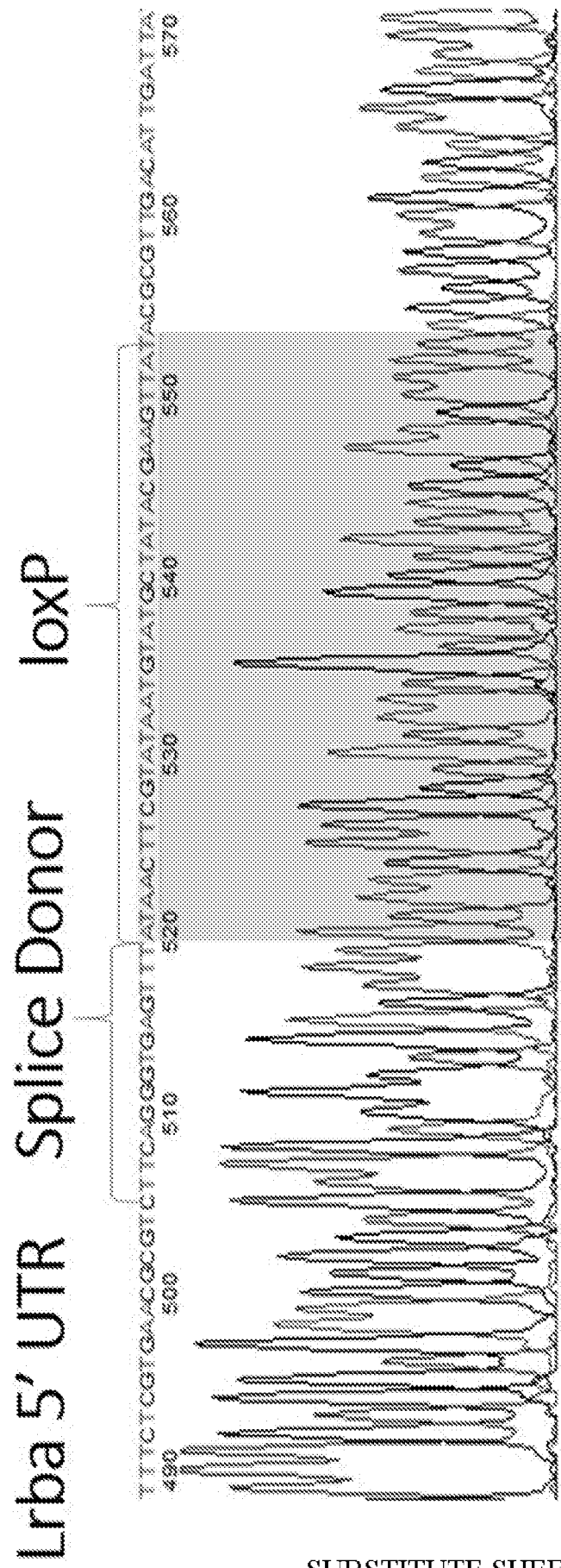


FIG. 12A

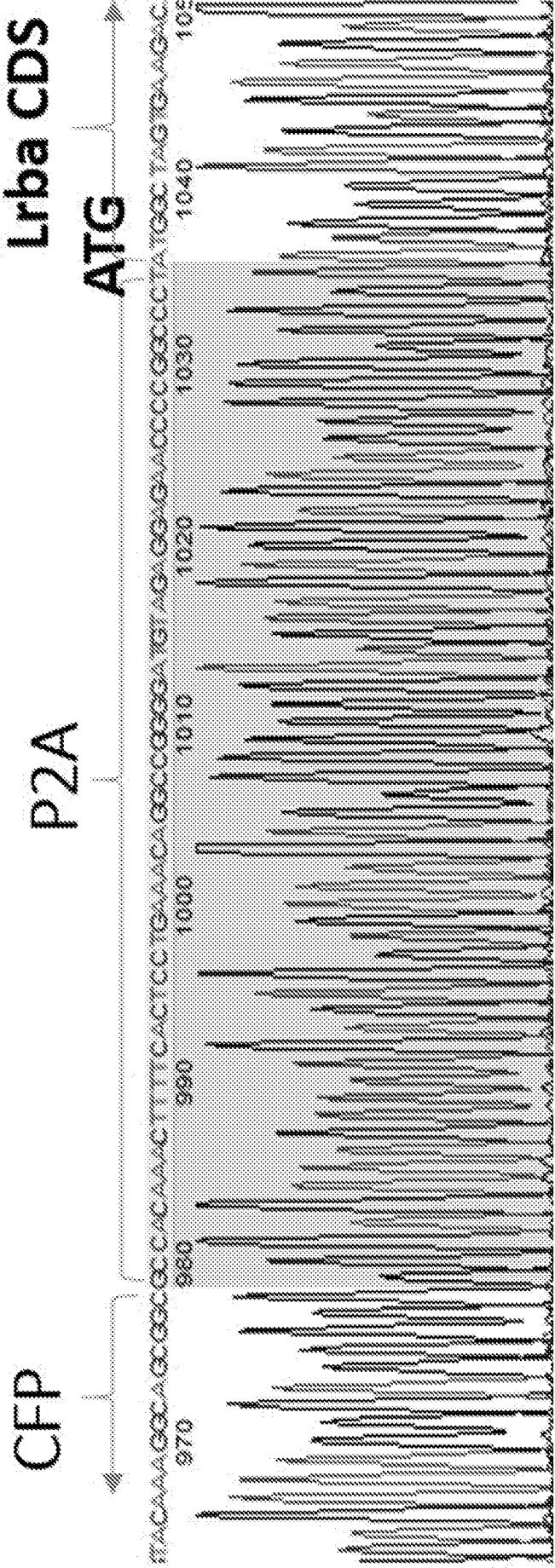


FIG. 12B

20/27

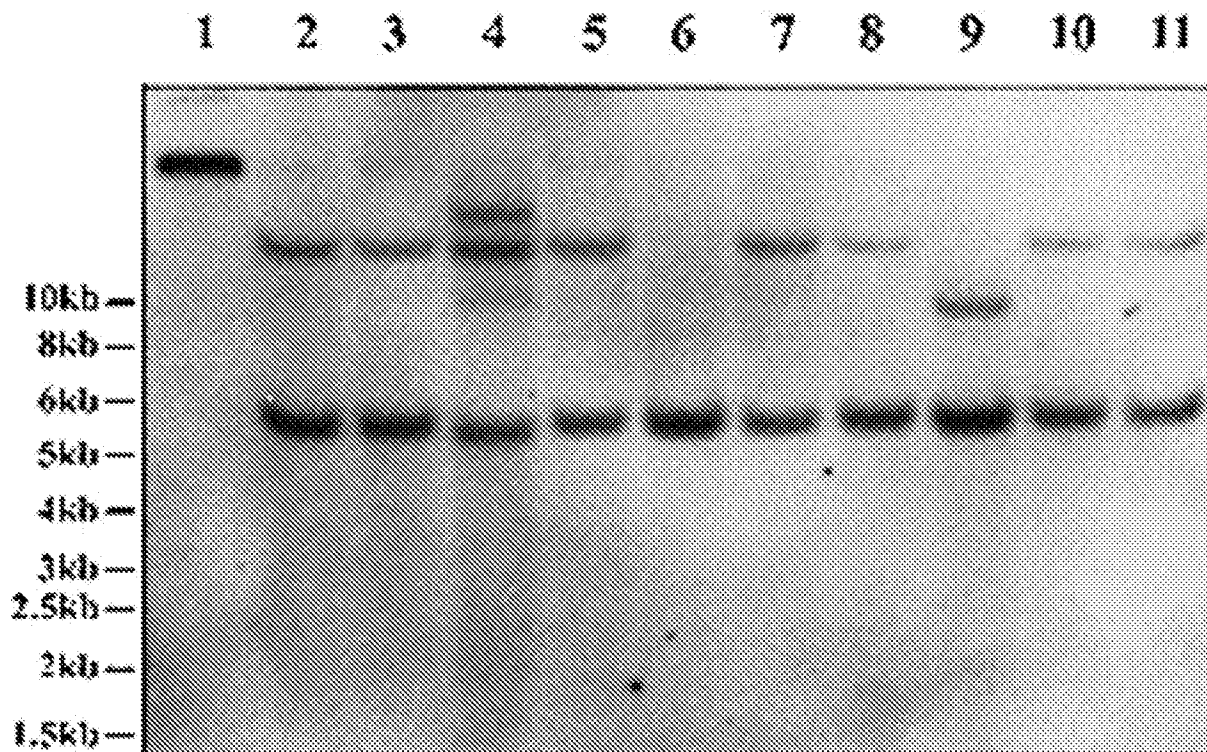


FIG. 12D

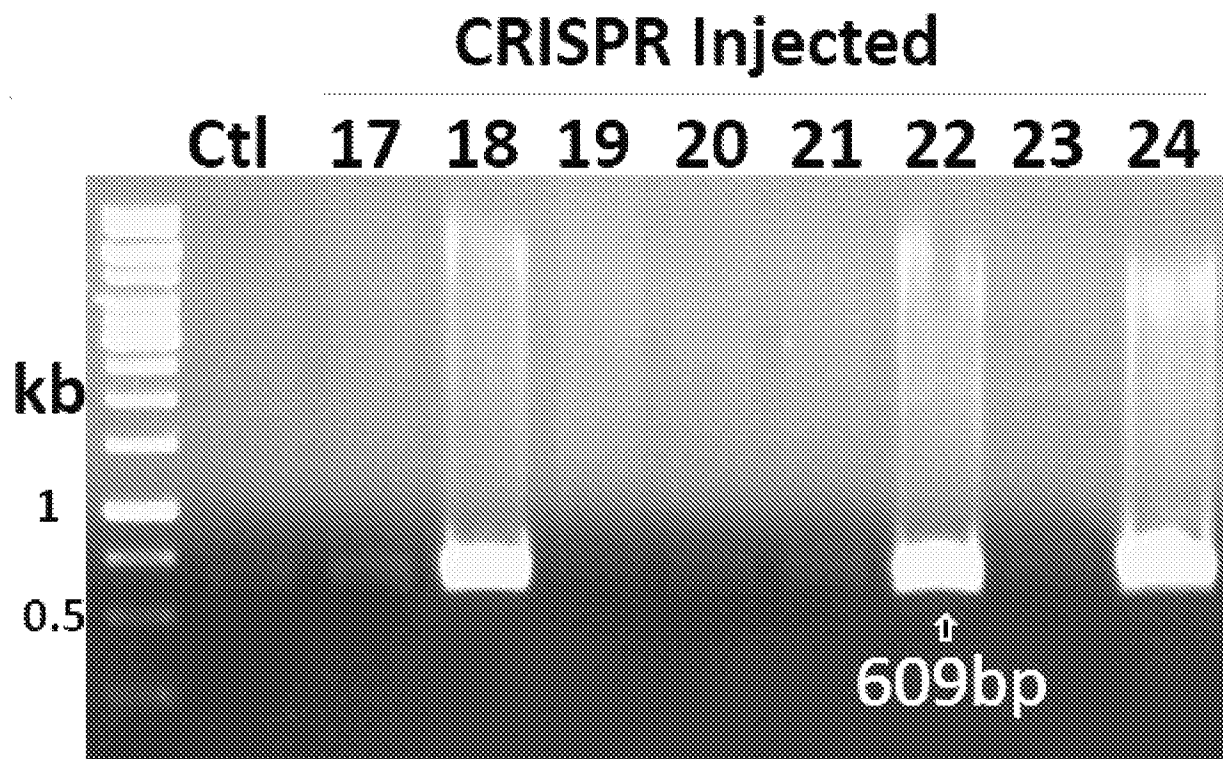


FIG. 12E

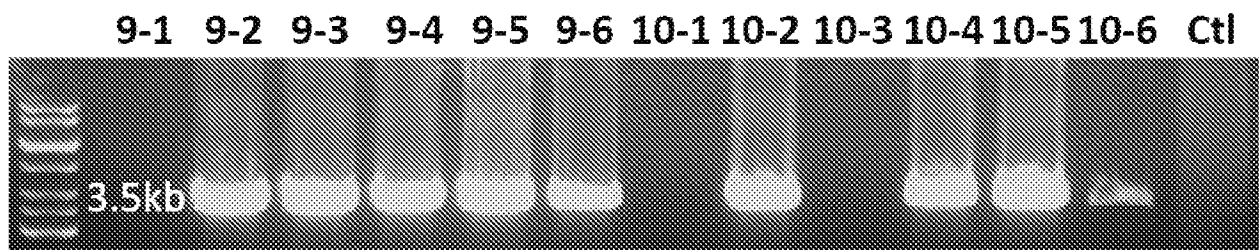


FIG. 12C

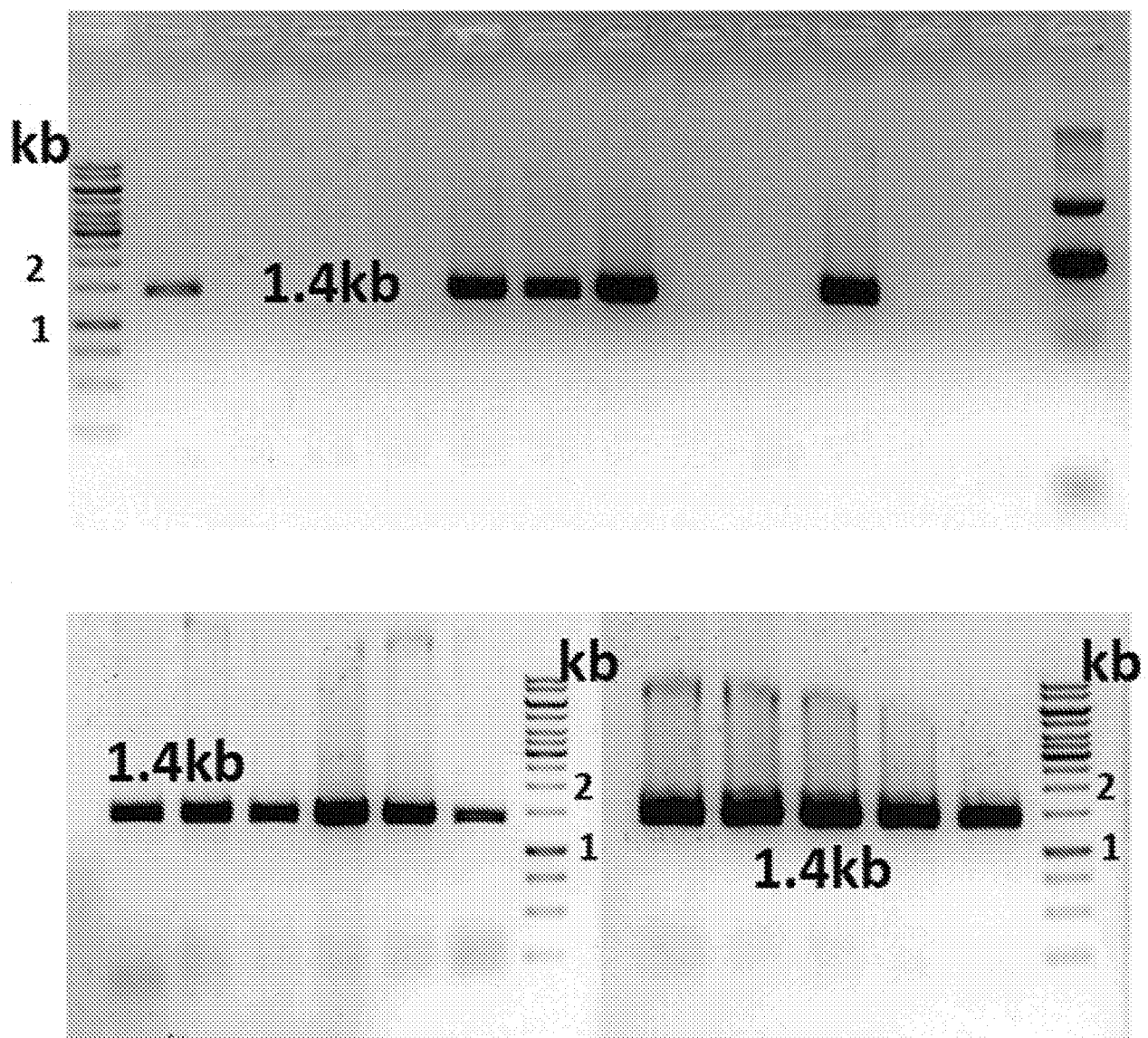
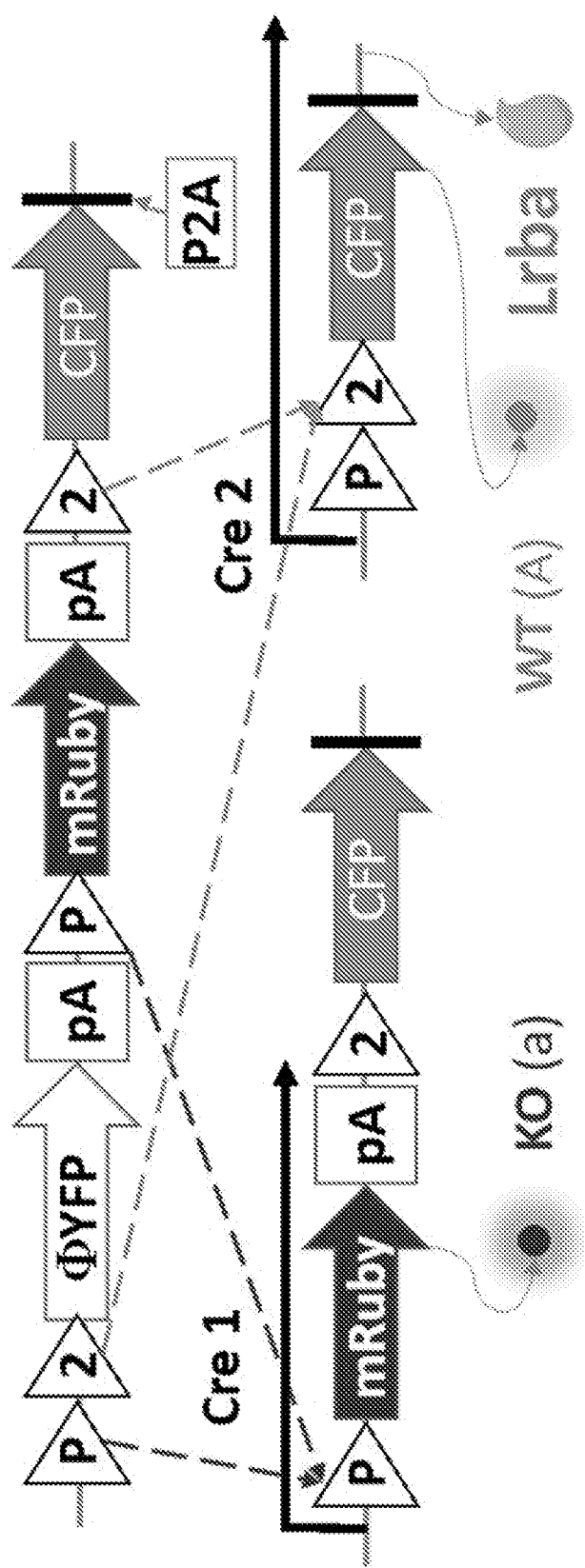


FIG. 12F

FIG. 13B



23/27

Comparison of phenotyping sensitivity

		Phenotype Average	Variation	Sensitivity	*Mouse #	*Cell #
T	WT	18	8	2	2250	NA
	KO	16	8			
G	WT	18	0.05	0.002	1	87750
	KO	17.998	0.05			

T=traditional methods; G= Genebow

FIG. 13C

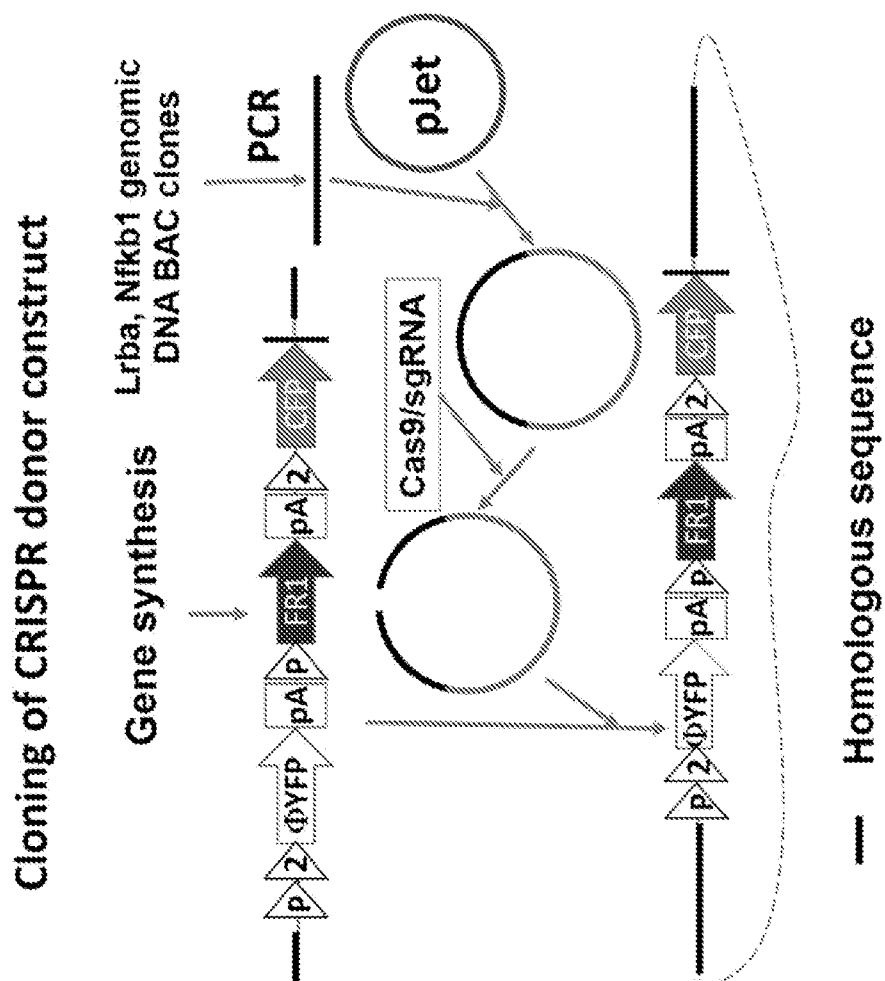


FIG. 14A

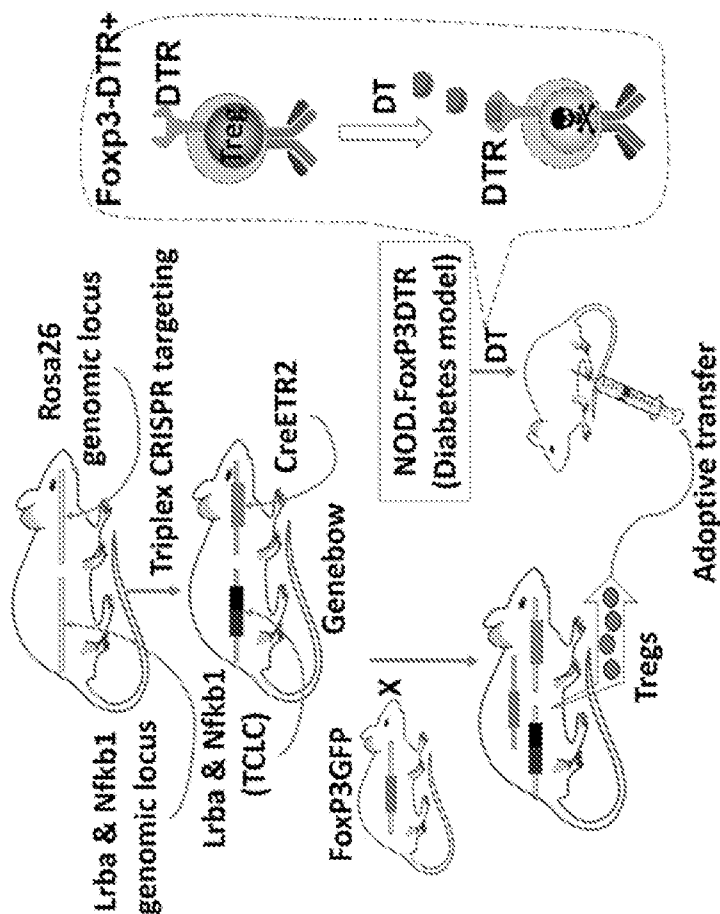
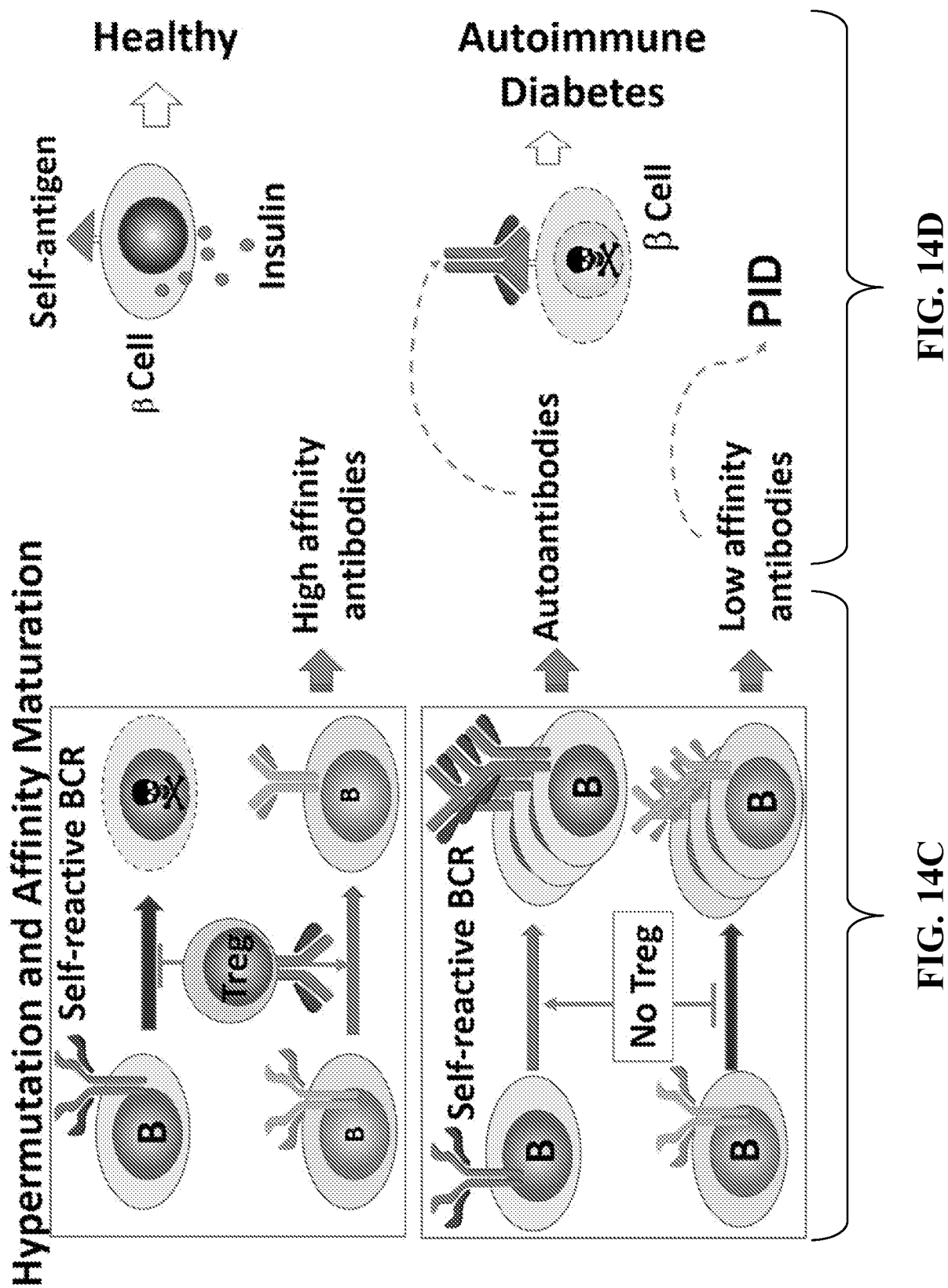


FIG. 14B



26/27

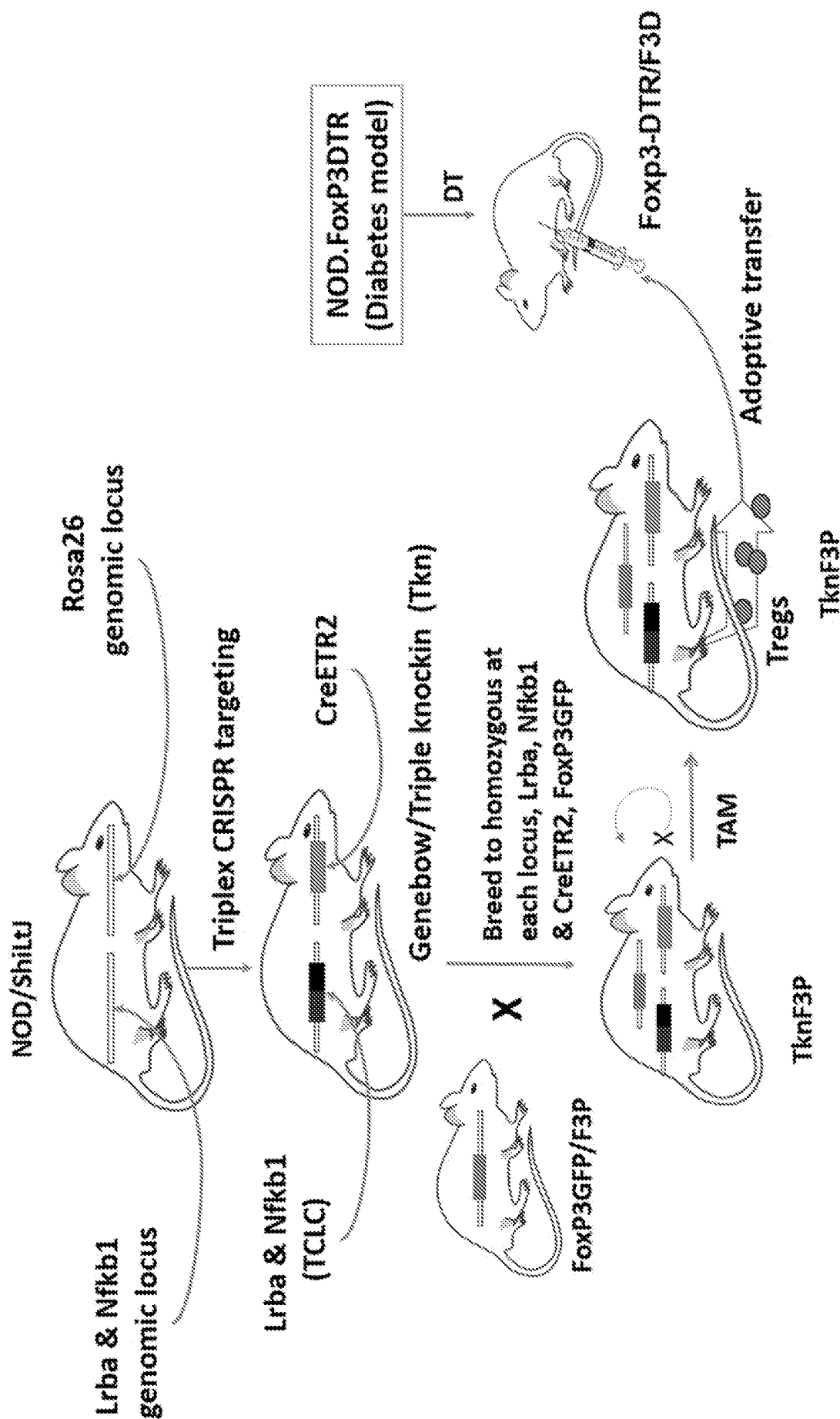


FIG. 15

FIG. 16A

FIG. 16C

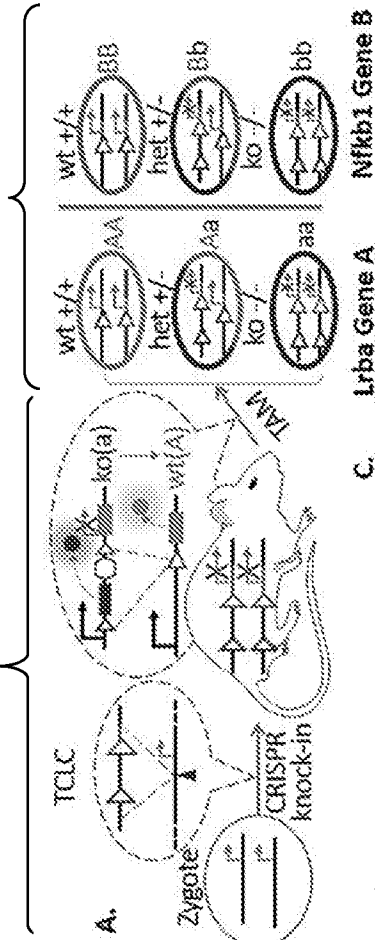


FIG. 16B

FIG. 16D

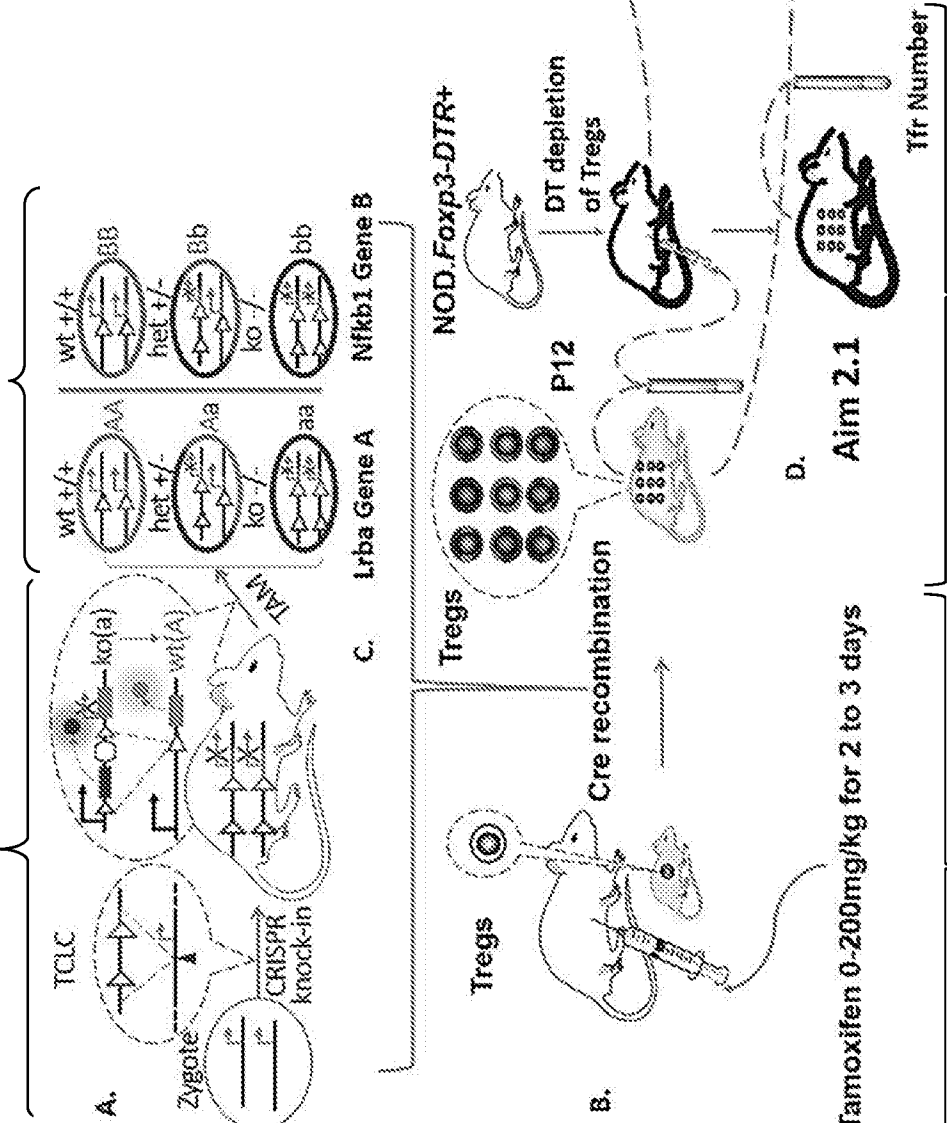
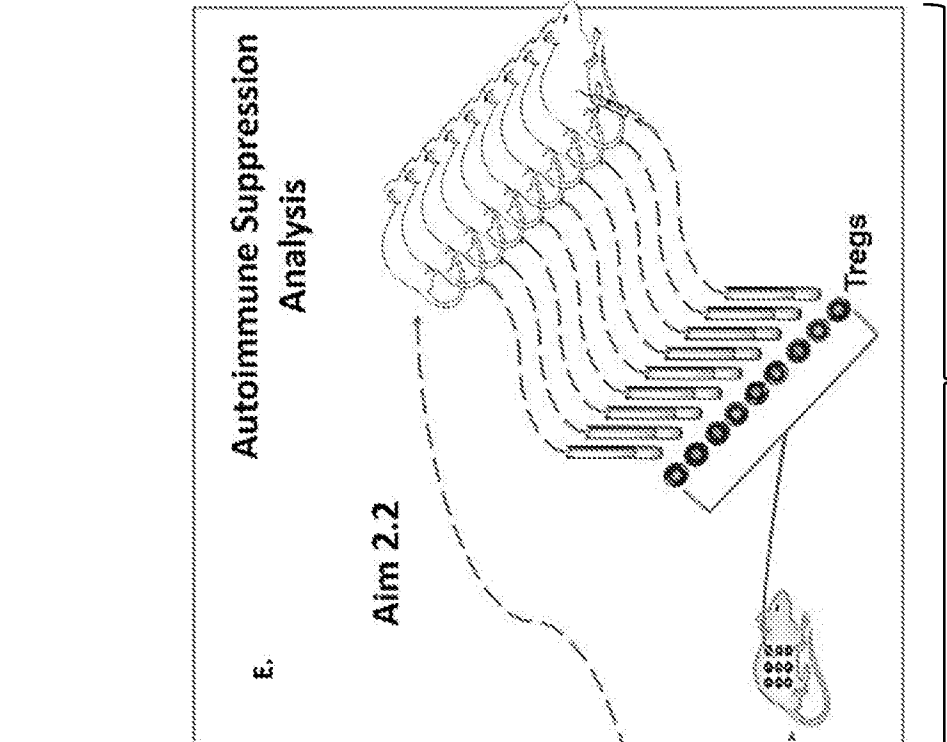


FIG. 16E

FIG. 16F



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/028165**A. CLASSIFICATION OF SUBJECT MATTER****A01K 67/027(2006.01)i, C12N 15/85(2006.01)i**

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)

A01K 67/027; G06F 15/00; C12N 15/90; C12Q 1/68; G01N 33/50; C12N 15/85

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) & Keywords: non-human animal model, genotype, label, Lrba, NfκB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 2530163 A1 (VETERINARMEDIZINISCHE UNIVERSITÄT WIEN) 05 December 2012 See claims 1-4 and 11.	1-21
A	US 7991564 B2 (HUANG, JING et al.) 02 August 2011 See claim 1.	1-21
A	WO 01-91548 A2 (PHARMACIA & UPJOHN COMPANY) 06 December 2001 See claims 1-3.	1-21
A	US 2005-0123971 A1 (DI, XIAOJUN et al.) 09 June 2005 See claims 1-2.	1-21
A	WO 03-059054 A1 (BIO GENOMICS, INC. et al.) 24 July 2003 See claims 1 and 4.	1-21



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

27 July 2016 (27.07.2016)

Date of mailing of the international search report

27 July 2016 (27.07.2016)

Name and mailing address of the ISA/KR

International Application Division

Korean Intellectual Property Office

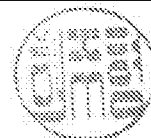
189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

HEO, Joo Hyung

Telephone No. +82-42-481-8150



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2016/028165

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 2530163 A1	05/12/2012	None	
US 7991564 B2	02/08/2011	US 2007-0128647 A1 US 2010-0144542 A1 US 2011-251798 A1 US 7634363 B2 US 8498825 B2	07/06/2007 10/06/2010 13/10/2011 15/12/2009 30/07/2013
WO 01-91548 A2	06/12/2001	AU 6526801 A EP 1353549 A2 US 2003-167488 A1 US 6984771 B2 WO 01-91548 A3	11/12/2001 22/10/2003 04/09/2003 10/01/2006 06/12/2001
US 2005-0123971 A1	09/06/2005	US 2004-0138821 A1 US 2005-0287575 A1	15/07/2004 29/12/2005
WO 03-059054 A1	24/07/2003	AU 2002-225486 A1	30/07/2003