Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2943794 C 2023/01/24

(11)(21) 2 943 794

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2015/03/26
(87) Date publication PCT/PCT Publication Date: 2015/10/01
(45) Date de délivrance/Issue Date: 2023/01/24
(85) Entrée phase nationale/National Entry: 2016/09/23
(86) N° demande PCT/PCT Application No.: US 2015/022660
(87) N° publication PCT/PCT Publication No.: 2015/148761
(30) Priorité/Priority: 2014/03/26 (US61/970,794)

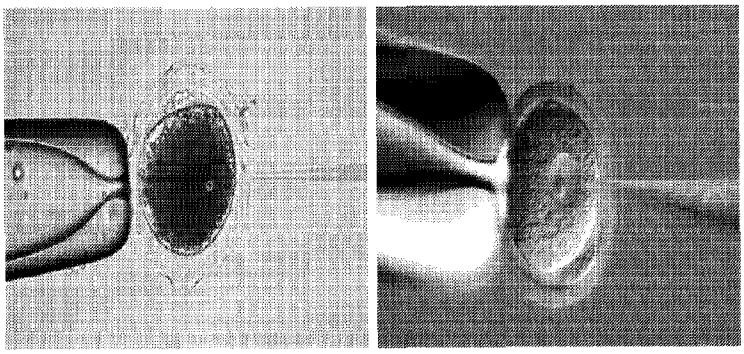
(51) Cl.Int./Int.Cl. *C12N 15/85* (2006.01), *A01K 67/027* (2006.01), *C12N 15/00* (2006.01), *C12N 15/09* (2006.01), *C12N 15/113* (2010.01), *C12N 15/12* (2006.01), *C12N 15/89* (2006.01), *C12N 15/90* (2006.01)

(72) Inventeurs/Inventors:
TELUGU, BHANU PRAKASH V.L., US;
PARK, KI-EUN, US

(73) Propriétaire/Owner:
UNIVERSITY OF MARYLAND, COLLEGE PARK, US

(74) Agent: BERESKIN & PARR LLP/S.E.N.C.R.L.,S.R.L.

(54) Titre : EDITION CIBLEE DU GENOME DANS DES ZYGOTES DE GRANDS ANIMAUX DOMESTIQUES
(54) Title: TARGETED GENOME EDITING IN ZYGOTES OF DOMESTIC LARGE ANIMALS



Pig zygote

Mouse zygote

(57) Abrégé/Abstract:
A method is provided of targeted genome editing of an animal using site specific homologous integration. A composition comprising a single stranded oligonucleotide or double stranded nucleic acid molecule comprising a nucleic acid molecule of interest and sequences flanking a target locus cleavage site is injected into the zygote of an animal along with a nuclease and a guide nucleic acid molecule that targets the nuclease to a target locus. The composition is injected into the zygote after fertilization and prior to formation of a nucleus. The nucleic acid molecule of interest is recombined into the genome with high efficiency. The process allows for integration of nucleic acid molecules into the genome of animals in which the pronuclei cannot be visually observed during injection.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 October 2015 (01.10.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/148761 A1

(51) International Patent Classification:
C12N 15/00 (2006.01) *C12N 15/85* (2006.01)

(21) International Application Number:
PCT/US2015/022660

(22) International Filing Date:
26 March 2015 (26.03.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/970,794 26 March 2014 (26.03.2014) US

(71) Applicant: **UNIVERSITY OF MARYLAND, COLLEGE PARK** [US/US]; Office of Technology Commercialization, 2130 Mitchell Building, College Park, Maryland 20742 (US).

(72) Inventors: **TELUGU, Bhanu Prakash V.L.**; University of Maryland, College Park, Office of Technology Commercialization, 2130 Mitchell Building, College Park, Maryland 20742 (US). **PARK, Ki-Eun**; University of Maryland, College Park, Office of Technology Commercialization, 2130 Mitchell Building, College Park, Maryland 20742 (US).

(74) Agents: **NEBEL, Heidi S.** et al.; McKee, Voorhees & Sease, P.L.C., 801 Grand Avenue, Suite 3200, Des Moines, Iowa 50309-2721 (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))*

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

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

(54) **Title**: TARGETED GENOME EDITING IN ZYGOTES OF DOMESTIC LARGE ANIMALS

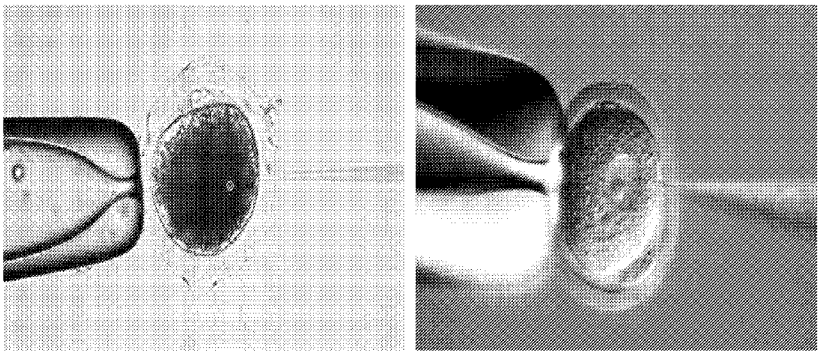


FIG. 1

(57) **Abstract**: A method is provided of targeted genome editing of an animal using site specific homologous integration. A composition comprising a single stranded oligonucleotide or double stranded nucleic acid molecule comprising a nucleic acid molecule of interest and sequences flanking a target locus cleavage site is injected into the zygote of an animal along with a nuclease and a guide nucleic acid molecule that targets the nuclease to a target locus. The composition is injected into the zygote after fertilization and prior to formation of a nucleus. The nucleic acid molecule of interest is recombined into the genome with high efficiency. The process allows for integration of nucleic acid molecules into the genome of animals in which the pronuclei cannot be visually observed during injection.

WO 2015/148761 A1

TITLE: TARGETED GENOME EDITING IN ZYGOTES OF DOMESTIC LARGE ANIMALS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority under 35 U.S.C. § 119 to provisional application Serial No. 61/970,794 filed March 26, 2014.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web. Said ASCII copy, created on March 22, 2015, is named U of Maryland and is 16,384 bytes in size.

BACKGROUND OF THE INVENTION

Modification of animal genomes has multiple uses. Where a gene is modified, by, for example, silencing the gene, prohibiting the expression of protein, changing expression product or expression levels, the function of the gene or region of the gene becomes apparent. Changes to the gene may also result in modification to a more desired phenotype. Treatment of disease is also possible, where defective genes can be modified. Effective gene targeting of animals and in particular large domestic animals in which the pronuclei are difficult to visually identify in a fertilized zygote has numerous challenges preventing reliable delivery of targeting constructs in order to effectively target a specific region in such animal genome.

SUMMARY OF THE INVENTION

The present method allows for injection of nucleic acid molecules into zygotes of large animals followed by incorporation of the nucleic acid molecule into the nucleus of the cell, such that one need not visually identify the pronuclei of the zygote. The method provides for injecting nucleic acid molecules into the zygote after fertilization of an egg by sperm and prior to formation of a nucleus. Methods in one embodiment employ

homologous recombination with the resulting nucleus that is formed having the nucleic acid molecule of interest targeted. A double stranded break may be induced at the target locus in an embodiment of the invention. A nuclease and either single stranded oligonucleotides or a double stranded targeting vector is provided with heterologous nucleic acid molecules of interest and sequences having homology to sequences flanking the site of the double stranded break, which allows for homologous recombination at the site. In an embodiment, the flanking sequences may be 50 base pairs (bp) or more when using single stranded oligonucleotides, and may be 500 bp or more when using a double stranded targeting vector. The method may be used for modifying expression of sequences in the animal by, for example, inhibiting expression, deletion, replacement of alleles, or introducing into the animal genome a sequence, which in one embodiment expresses a unique protein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a photomicrograph of porcine and murine zygotes.

Figure 2 shows CRISPR-mediated knockin of GFP transgene into the PRNP locus. A) is a schematic drawing of the targeting vector showing targeting arms homologous to endogenous PRNP locus. The term "upper" refers to the 500 bp of the targeted *PRNP* locus upstream of the cleavage site and "lower" refers to the 1000 bp of the downstream flanking sequence. UBC refers to (human ubiquitin C promoter), GFP refers to green fluorescent protein; bPA (bovine poly adenlyation transcription terminator sequence); PGK/EM7 – is a hybrid eukaryotic (phosphoglycero kinase) and prokaryotic (EM7, a synthetic bacterial promoter derived from the T7 promoter that enables the constitutive expression of the antibiotic resistance gene in E.coli
patents/US7244609) RNA polymerase II promoter sequences driving the expression of Neo/kan which are neomycin (or G418 for eukaryotic) and kanamycin resistance (for prokaryotic) selectable markers. The linearized targeting vector alongside, Cas9 mRNA and single guide (sgRNA)targeting *PRNP* are injected into the porcine zygotes. B) is a schematic showing an embodiment of the process. Cas9 induces double strand break in 3rd exon of *PRNP*. The broken DNA is then repaired with the targeting vector, resulted in the targeted allele.

Figure 3 A and B are photos showing representative embryos that have developed to the blastocyst stage show expression of GFP, with A) showing the embryo under green fluorescent microscope and B) showing bright field microscope view.

Figure 4 is a gel where use of primers, one within the targeting vector and another outside of the homologous region, shows specific band of right size confirming targeting to the intended locus.

Figure 5A and B show CRISPR mediated homology directed repair and editing of porcine genomes using a single stranded oligonucleotide incorporation into the zygote. A) is a schematic of Cas9 vector, sgRNA targeting *ZBED6* locus and a single stranded oligo containing a LoxP site and a EcoR1 restriction enzyme site injected into porcine zygotes. CMV and T7 refer to the promoters HA is HA tag, NLS is nuclear localization signal. B) is a gel showing the PCR amplicon contains the recombined EcoRI site, which can be digested to produce two fragments, confirming the recombination of a functional EcoR1 enzyme site.

Figure 6 shows sequencing of the recombined *ZBED6* allele (the entire sequence shown is SEQ ID NO: 13) showing the EcoR1 site (underlined) and the LoxP (SEQ ID NO: 11) site in the inset.

Figure 7 shows A) a schematic of the pronuclei with loxP targeting; and two gels, B) a gel showing ZBED6 with loxP targeting and C) PRNP with loxP targeting.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Transfer of genetic material into mammalian cells was first reported 50 years ago (Szybalska, E.et al. Genetics of human cell line. IV. DNA-mediated heritable transformation of a biochemical trait. *Proc Natl Acad Sci U S A* 48, 2026-2034 (1962)) when cells were made resistant to hypoxanthine, aminopterin and thymidine (HAT) selection medium by calcium phosphate mediated transfection. The efficiency of exogenous transfer was further improved by injection of DNA directly into tissue culture cells with the aid of micropipettes mounted on micromanipulators. (Graessmann, et al. Retransformation of a simian virus 40 revertant cell line, which is resistant to viral and DNA infections, by microinjection of viral DNA. *J Virol* 32, 989-994 (1979); Capecchi, et al. High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. *Cell* 22, 479-488 (1980)). This served as a prelude for injection of

transgenes into the pronucleus (PN) of mouse zygotes (Gordon, Jet al. Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl. Acad. Sci. USA* 77, 7380-7384 (1980)) and the production of the first transgenic mice. (Brinster, R.L. et al. Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* 27, 223-231 (1981)). Soon after the first demonstrated success in generating transgenic mice by PN injection, similar procedures were used to generate growth hormone transgenic pigs. (Hammer, R.E. et al. Production of transgenic rabbits, sheep and pigs by microinjection. *Nature* 315, 680-683 (1985)). The PN injection procedure remains a widely used technique for making transgenic domestic large animals, such as pigs, cattle, sheep, goats, camelids, dogs, cats, etc., however, only 1% of injected zygotes produce stable transgenic founders. (Niemann, H. Transgenic pigs expressing plant genes. *Proc Natl Acad Sci U S A* 101,7211-7212 (2004); Prather, et al. Genetically modified pigs for medicine and agriculture. *Biotechnology & Genetic Engineering Reviews* 25, 245-265 (2008)). In addition to low efficiencies in generating founder animals, the technique is challenging in pig and other large animal models, where the embryo is murky and the pronucleus is difficult to localize for injecting large DNA constructs (Figure 1).

Besides PN injection, other methods for generating transgenic animals have been attempted, which included sperm-mediated gene transfer, oocyte transduction, and transposons, with varying degree of success. (Park, et al. Role of stem cells in large animal genetic engineering in the TALENs-CRISPR era. *Reprod Fertil Dev* 26, 65-73 (2013)). Specifically in conjunction with the use of CRISPR/Cas system or other site specific nucleases, such DNA constructs can be targeted to a specific region of the gene, allowing for site-specific knockin of the candidate genes, modification of nucleotides or replacement of alleles. However, all the methods suffer from somewhat similar limitations, including (but not limited to): (1) random integration in the genome; (2) insertional mutagenesis; (3) positional silencing; (4) lack of control over the number of integrants and expression; (5) an inability to pre-screen for stable integrations before embryo transfer; and (6) the inability to delete endogenous genes (i.e. gene "knockout" or "KO", where the gene expression is inactivated or inoperative). "Gene targeting", whereby the intended gene sequences can be targeted for deletion, or exogenous DNA incorporated, offsets all the disadvantages described above.

Success in gene targeting can be attributed to the discovery that flanking homologous isogenic sequences combined with a positive and negative selection scheme facilitates the creation and identification of recombination events at the intended locus (Luciw, et al. Location and function of retroviral and SV40 sequences that enhance biochemical transformation after microinjection of DNA. *Cell* 33, 705-716 (1983); Kucherlapati, R.S. et al. Homologous recombination in monkey cells and human cell-free extracts. Cold Spring Harbor Symposia on Quantitative Biology 49, 191-197 (1984)). However, one of the major limitations of homologous recombination (HR)-based gene targeting for knocking out genes, introducing point mutations, or knocking-in genes is the poor efficiency of achieving correct recombination events, typically in the range of 1 in 10^6–10^7 cells. In addition, these modifications are often monoallelic in nature, requiring a second round of targeting for bi-allelic (homozygous) modifications. In pigs and other livestock species that lack genuine embryonic stem cells (ESC), somatic cells, typically fetal fibroblasts are used for gene targeting, and the modified cells are used as donor cells in nuclear transfer or cloning to generate the genetically modified (GM) animals. The gene targeting efficiencies are relatively poor when using somatic cells as precursors, which is further compounded by the relatively early senescence of the fibroblasts precluding characterization of the desired knock in (KI) or other genetic alterations in the course of the experiment. With the long gestation length and maturation to reproduction age of pigs and other large animals, the generation of homozygous knockin (or "KI", where a nucleic acid sequence is inserted at a particular locus) animals by nuclear transfer or cloning to produce recombinant pigs is both technically challenging and cost prohibitive.

An improved method is shown here of targeting genome editing in zygotes that allows for genome targeting in animals including animals having a zygote with optical density such that the pronuclei are obscured as a result of lipid granules in the cytoplasm interfering with visualization. This makes it difficult to visually identify the pronuclei of the zygote with the naked human eye and inject nucleic acid molecules into the nuclei. A pronucleus is the haploid nucleus of a sex cell, here the male and female pronuclei present following fertilization of the oocyte (egg) by the spermatatozoa (sperm). The membranes of the pronuclei dissolve and the chromosomes align, then becoming part of a single diploid nucleus and cell division occurs. One skilled in the art appreciates that it is possible to determine the time frame prior to formation of the pronuclei into nucleus.

Using homologous recombination techniques, precise targeting of a nucleic acid molecule to a location of the genome is now possible using direct injection into such animal zygotic cells where the injection occurs after fertilization and prior to the pronuclei forming a nucleus. Using this method it has been found the nucleic acid molecule will be efficiently recombined into the nucleus. One embodiment provides for injection up to 18 to 24 hours after fertilization. In an embodiment the time frame is at up to five hours in one embodiment, and up to 12 hours in another embodiment. At embodiment provides injection occurs at three, four or five hours after fertilization. For about three to five hours after fertilization the cells are typically kept in a fertilization medium, and would be injected after removal from the medium. In a further embodiment injection occurs about up to 16 hours after fertilization, in an embodiment is three to four to five to six to seven to eight to nine to ten to twelve to thirteen to fourteen to fifteen hours after fertilization up to 16 hours, and in further embodiment is 8 to 12 hours after fertilization. This avoids the need to use fetal fibroblasts or a system of nuclear transfer or cloning. A nucleic acid sequence of interest is injected along with components which target the sequence to the desired target locus in the genome and provide for homologous recombination at the site. Such components in an embodiment include a nuclease for cleaving the target locus, along with a guide RNA that directs the nuclease to the targeted locus.

Any method which provides for targeting of the nucleic acid molecule of interest (NOI) to the target site of the target gene may be utilized in the method. The following is provided by way of example rather than limitation. A guide nucleic acid molecule is one that directs the nuclease to the specific cut site in the genome, whether via use of a binding domain, recognition domains, guide RNAs or other mechanisms. The guide nucleic acid molecule is introduced into the cell under conditions appropriate for operation of the guide nucleic acid molecule in directing cleavage to the target locus. A person of skill in the art will have available a number of methods that may be used, the most common utilizing a nuclease to cleave the target region of the gene, along with sequences which will recognize sequences at the target locus and direct cleavage to the locus. Any nuclease that can cleave the phosphodiester bond of a polynucleotide chain may be used in the methods described here. By way of example without limitation, available systems include those utilizing site specific nucleases (SSN) such as ZFNs (Zinc finger nuclease), (Whyte, J.J. et al. Gene targeting with zinc finger nucleases to produce cloned eGFP knockout pigs. *Mol*

Reprod Dev 78, 2 (2011); Whyte, et al. Cell Biology Symposium: Zinc finger nucleases to create custom-designed modifications in the swine (Sus scrofa) genome. *J Anim Sci* 90, 1111-1117 (2012)); TALENs (Transcription activator-like effector nucleases) (see, Carlson, D.F. et al. Efficient TALEN-mediated gene knockout in livestock. *Proc Natl Acad Sci U S A* 109, 17382-17387 (2012); Tan, W. et al. Efficient nonmeiotic allele introgression in livestock using custom endonucleases. *Proc Natl Acad Sci U S A* 110, 16526-16531 (2013); Lillico, S.G. et al. Live pigs produced from genome edited zygotes. *Scientific reports 3*, 2847 (2013)), and the CRISPR (Clustered regularly interspaced short palindromic repeats) -associated (Cas) nuclease system (Hai, T., Teng, F., Guo, R., Li, W. & Zhou, Q. One-step generation of knockout pigs by zygote injection of CRISPR/Cas system. *Cell Res* 24, 372-375 (2014)) that have permitted editing of animal genomes such as pig genomes with relative ease. The use of recombinases such as FLP/FRT as described in US Patent No. 6,720,475, or CRE/LOX as described in US Patent No. 5,658,772, can be utilized to integrate a polynucleotide sequence into a specific chromosomal site. Meganucleases have been used for targeting donor polynucleotides into a specific chromosomal location as described in Puchta *et al.*, *PNAS USA* 93 (1996) pp. 5055-5060. ZFNs work with proteins or domains of proteins binding to a binding domain having a stabilized structure as a result of use a zinc ion. TALENs utilize domains with repeats of amino acids which can specifically recognize a base pair in a DNA sequence. For a discussion of both systems see Voytas et al. US Patent No. 8,697,853. These systems utilize enzymes prepared for each target sequence.

The CRISPR/Cas nuclease system has evolved in archaea and bacteria as a RNA based adaptive immunity system to detect and cleave invading viruses and plasmids. (Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167-170 (2010); Wiedenheft, et al. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331-338 (2012)). Unlike ZFNs and TALENs, which require assembly of DNA binding domain (DBD) to direct the nuclease to the target site, the CRISPR/Cas system utilizes RNA as a guide. The CRISPR locus is a distinct class of interspersed short sequence repeats (SSRs) recognized in bacterial genes. The repeats are short elements occurring in clusters that are regularly spaced by unique intervening sequences with a substantially constant length. They were observed as an immunity

system in which nucleic acid molecules homologous to virus or plasmid sequences are integrated into the CRISPR loci. The foreign DNA or RNA is targeted and cleaved. The system has been adapted for targeted insertion of a nucleic acid molecule at a defined locus. In general, a CRISPR system is characterized by elements that promote the formation of a CRISPR complex at the site of a target sequence to which a guide sequence is designed to have complementarity, where hybridization between a target sequence and a guide sequence promotes the formation of a CRISPR complex. Full complementarity is not necessarily required, provided there is sufficient complementarity to cause hybridization and promote formation of a CRISPR complex. In the CRISPR system one enzyme, a CRISPR enzyme is used for targeting using short RNA molecules.

Two basic components are used with the system, a guide RNA and an endonuclease. The guide RNA is endogenous sequence specifying the target site and tracrRNA, needed to bind to the enzyme. The guide sequence provides target specificity and the tracrRNA provides scaffolding properties. These guide sequences are typically about 15 up to 20 to 25 base pairs (bp) that hybridize with the target site and direct binding of a CRISPR complex to a target sequence. A sequence encoding a CRISPR-associated enzyme may be provided on the same or different vectors. Non-limiting examples of Cas proteins include Cas1, Cas1B, Cas2. Cas3, Cas4, Cas5, Cas6. Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homologs thereof, or modified versions thereof. In one embodiment the enzyme is a type II CRISPR system enzyme and is Cas9 or variants or modifications thereof. These enzymes are known; for example, the amino acid sequence of *S. pyogenes* Cas9 protein may be found in the SwissProt database under accession number Q99ZW2. The enzyme or Cas9 protein can be used as a nickase or nuclease and cleave one or two strands of DNA. Cas9 has two functional domains, RuvC and HNH and when both are used both strands are cleaved. Cas9 nuclease forms a ribonuclease complex with target CRISPR RNAs (crRNAs) and transactivating RNAs (tracrRNA), and uses the chimeric RNA to target the genomic sequence and induce DSB. The CRISPR/Cas nuclease and other SSN can introduce a targeted double strand break (DSB) in the genomic DNA, which in the presence of a single stranded (SS) DNA oligonucleotide or a double stranded (DS)

targeting vector, result in homologous recombination (HR) based alteration of selected nucleotides or KI of transgenes respectively, into the target loci. In another embodiment a SS oligonucleotide having the nucleic acid molecule of interest may be used with Cas9 mRNA and sgRNA to target modification of a particular target gene region. In further embodiments the target gene is complementary to the gRNA sequence and will have a protospacer adjacent motif or PAM sequence. This aids in binding by Cas9. For a discussion of details of the CRISPR/Cas system see Cong et al., US Patent Nos. 8,932,814; 8,871,445 and 8,906,616.

Breaks in the genome can be repaired by the non-homologous end joining DNA repair pathway (NHEJ) or by homology directed repair pathway (HDR). NHEJ can disrupt the gene, by causing frame shifts or premature stop codons to occur. HDR is an embodiment that provides for insertion of a nucleic acid molecule that avoids such issues. With a double strand break a DNA repair template is provided in which sequences are provided that have homology to and hybridize with genome sequences flanking the cleavage site (homology arm). In one embodiment the DNA template or flanking sequences are transfected into the cell with the CRISPR/Cas vector.

Even though HDR-based gene targeting events are extremely rare, the efficiencies can be improved by several orders of magnitude (>1000-fold) by introducing a DSB at the target locus (Moehle, E.A. et al. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A* 104, 3055-3060 (2007)). Following DSB, either a SS oligo, or a DS vector with homology to the ends flanking the DSB, can produce animals with targeted genomic alterations or transgene integrations (Cui, Let al. The permissive effect of zinc deficiency on uroguanylin and inducible nitric oxide synthase gene upregulation in rat intestine induced by interleukin 1alpha is rapidly reversed by zinc repletion. *The Journal of Nutrition* 133, 51-56 (2003); Meyer, M et al. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc Natl Acad Sci U S A* 107, 15022-15026 (2010)).

In the inventors laboratory, CRISPR/Cas mediated HDR and editing of nucleotides and gene KI by injections directly into the porcine zygotes bypassing the need for nuclear transfer/cloning to generate genetically modified (GM) animals. The length of each homology arm will vary depending on the size of the modification to the genome. Here, the inventors have discovered far smaller homology arms may be utilized than previously

employed when not using the present method. Previously the sequence homologous to the target gene flanking sequences were required to be of a size in the region of 6000 bp. However, here with use of a double stranded break and site specific nuclease systems, it is demonstrated the methods need less than 6000bp homologous regions, and operate with only at least about 40 bp, at least about 50 bp and amounts in-between, or more upstream and about 40 bp, at least about 50 bp and amounts in-between, or more downstream sequences for use with single stranded oligo and about at least 300 bp, at least 500 bp up to 1000 bp and amounts in-between or more with a double stranded targeting vector. This provides for a much less laborious process in preparing such sequences, reducing the time for preparation of a target vector to one to two weeks or less instead of six months. These homologous arms are provided with the SS oligonucleotide or DS vector.

The present methods provide for expansion of use in gene editing to provide greater efficiencies, by 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, up to 100% increase and amounts in-between in the number of recombination events obtained. The fold increase in events recovered is up to1000 fold or more. This results in high efficiency and usefulness with large animal gene editing and use in many settings.

In referring to a target gene is meant to refer to any nucleic acid molecule within the animal genome desired to be modified as described or where it is desired to insert a nucleic acid molecule. The target polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., a regulatory polynucleotide or a junk DNA).

The process is useful with knockins and knockouts. One can upregulate gene transcription through insertion of a transcriptional activator, for example, or repress expression using transcriptional repressors. Without intending to be limiting, among the variety of utilities of gene editing include modifying (e.g., deleting, inserting, translocating, inactivating, activating, mutating) a target polynucleotide in a multiplicity of cell types. In some embodiments the polypeptide expressed may be modified. There are a broad spectrum of applications in, e.g., gene therapy, drug screening, disease diagnosis, prognosis, increasing or decreasing growth and body composition, and improving quality or composition of animal products. The methods are useful in any situation where modification of a disease gene is desired. Further examples include improved feed use and meat composition, enhanced reproductive performance, changing component content of

animal products such as milk (such as lactose content), and, reduce or eliminate phenotypes such as boar taint phenotype. The process is useful with so-called knockins where a sequence is inserted in the genome or knockouts where gene expression is reduced or eliminated or interrupted. This allows for understanding and control of the gene and its' downstream impact.

The target nucleotide may include a number of disease-associated genes and polynucleotides as well as signaling biochemical pathway-associated genes and polynucleotides. Examples of target polynucleotides include a sequence associated with a signaling biochemical pathway, e.g., a signaling biochemical pathway-associated gene or polynucleotide. Examples of target polynucleotides include a disease gene or polynucleotide. A "disease" gene or polynucleotide includes any gene or polynucleotide associated with impacting disease in and animal and can include a gene or polynucleotide which is yielding transcription or translation products at an abnormal level or in an abnormal form in cells derived from a disease-affected tissues compared with tissues or cells of a non-disease control. A disease gene is any gene associated with an increase or decrease in the risk of having or developing a disease or recovery from disease. It may be a gene that becomes expressed at an abnormally high level; it may be a gene that becomes expressed at an abnormally low level, where the altered expression correlates with the occurrence and/or progression of the disease. A disease gene also refers to a gene possessing mutation(s) or genetic variation that is directly responsible or is in linkage disequilibrium with a gene(s) that is responsible for the etiology of a disease. The transcribed or translated products may be known or unknown, and may be at a normal or abnormal level. A vast array of animal diseases may be treated, prevented or studied by use of the methods here with genes or proteins encoded associated with the disease. A few examples include the *PRNP* gene providing resistance to transmissible spongiform encephalopathies; upregulation *of SOCS1, SOD2, RBP4, HLA-B, HLA-G, PPP2R1A* and *TAP1* gene or downregulation of *IL18, TF, C4BPA, C1QA, C1QB* and *TYROBP* genes of immune response genes in protection from PRRSV; KI decoys to combat zoonotic flu disease, eliminating negative phenotypes such as boar taint, and introgress agriculturally beneficial traits.

A still further example of potential uses provides for introduction into the animal cell of interfering nucleic acid molecules. For example, double-stranded RNA molecules

(dsRNA) may be employed. In this process, in summary, RNA which is double stranded, in part, or completely, is produced based upon the sequence of the target nucleic acid molecule. Specifics of the means of producing the dsRNA may vary as one skilled in the art appreciates, and include, by way of example without intending to be limiting, the approach of Graham et al., US Patent No. 6,573,099 where two copies of a sequence corresponding to a target sequence is used, or that of Fire et al., US Patent 6,326,193, where the first strand is an RNA sequence corresponding to the target nucleic acid, and the second is one which is complementary to the target sequence. These strands hybridize with each other to form the inhibiting dsRNA. The strand which corresponds to the target nucleic acid molecule can correspond to all or a portion thereof, as long as a dsRNA is formed. Where a strand is used which is the complement (antisense) of the target nucleic acid is used, it can be complementary to all or a portion of the target nucleic acid molecule, so long as the dsRNA formed interferes with the target nucleic acid molecule. The dsRNA triggers a response in which the RNAse III Dicer enzyme process dsRNA into small interfering RNAs (siRNA) of approximately 21 - 23 nucleotides, which are formed into a RNA-induced silencing complex RISC which destroys homologous mRNAs. (See, Hammond, S. M., et al., *Nature* (2000) 404:293-296). Generally, sequences of up to 10 nucleotides 20 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 100 nucleotides, 200 nucleotides, 300, 500, 550, 500, 550, or greater and any amount in-between may be used.

One skilled in the art appreciates there are many uses available and which will become available for the methods described here.

The methods may be used in any animal. They are most useful with animals having pronuclei following fertilization that is at least partially visually obscured, making injection into the pronuclei difficult. Such pronuclei cannot be visually observed by the unaided human eye, that is, with the human eye and where visualized with a microscope, unaided by contrast, or other process to allow visualization. Such animals include "large domestic animals" that is, pigs, cattle, horses, dogs, cats, and other ruminant animals such as sheep, goats, oxen, musk ox, llamas, alpacas, guanicos, deer, bison, antelopes, camels, and giraffes, oxen, musk ox, llamas, alpacas, guanicos, deer, bison, antelopes, camels, and giraffes. Livestock animals are among those included in the animals for which the

processes can be used. Besides, in other animal models in which the cytoplasm is clear such as primates, rabbits, minks and other models, the procedure is still an attractive option, as it allows for not having to guess when the pronuclei are being formed, and instead the material can be injected into the cytoplasm.

The methods are particularly useful with swine. Pig is an economically important agricultural animal. Additionally, pigs are coveted for their biomedical applications. Similar to humans and mouse, the pigs are mono-gastric, and as such are playing a dominant role in investigations of nutrient uptake, trafficking and metabolism. (Patterson, et al. The pig as an experimental model for elucidating the mechanisms governing dietary influence on mineral absorption. *Experimental biology and medicine* 233, 651-664 (2008)). Advances in the field of animal genome editing have included sequencing of pig genome. (Groenen, M.A. et al. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491, 393-398 (2012)). Taken together, depending on the biological question that needs to be addressed, a suitable pig model is available for investigation. However, until now there has been a lack of incentive for the use of pig as "preferred models", due to the GM technologies that lag behind the mouse models. The present methods address these shortcomings.

In recent years, there is an increasing consensus that the mouse, although still a powerful genetic model species, has limitations and cannot fulfill the full spectrum of biomedical demands to address preventative medicine (obesity, infertility, cardiovascular disorders), identification of new or improved diagnostics, and models of farm derived zoonotic diseases. As an alternative, large domesticated animals such as pig are gaining favor, because they are more similar anatomically, physiologically, and immunologically to humans, while maintaining the advantages of being a litter bearing species and a relatively long life span permitting long term investigations. Examples where the mouse model has not met expectations due to differences in anatomy or physiology include, cystic fibrosis, ocular and cardiac diseases. (Rogers, C.S. et al. Disruption of the CFTR gene produces a model of cystic fibrosis in newborn pigs. *Science* 321, 1837-1841 (2008); Welsh, M et al. Development of a porcine model of cystic fibrosis. *Transactions of the American Clinical and Climatological Association* 120, 149-162 (2009); Zhou, L. et al. Differentiation of induced pluripotent stem cells of swine into rod photoreceptors and their integration into the retina. *Stem Cells* 29, 972-980 (2011); Whyte, J.J. et al. Vascular

endothelium-specific overexpression of human catalase in cloned pigs. *Transgenic Res* 20, 989-1001 (2011)). As almost all domestic pigs are crossbreeds, the resulting phenotype from these animal models is more reliable and applicable to the human diseases than inbred mouse strains. Likewise, the domestic pigs are better suited for zoonotic or infectious disease research, where they serve as reservoir or carriers of the disease, or are natural hosts to the pathogen. In addition to serving as models of human disease, the domestic pig is coveted for studies where the relatively long life span, close similarity in body size and physiology to humans offer an advantage. Briefly, pigs are preferred models for nutritional studies.

In referring to injection is meant any convenient method of inserting a device into the cell and passage of the nucleic acid molecules into the cell. By way of example without limitation, this can be accomplished with an injection pipette which may include a syringe holding the nucleic acid molecules. The pipette is inserted through the zona pellucida. Contrary to present techniques, one need not be able to visually observe the pronuceli in order to contact the injection device with the pronuclei. Instead, one need only introduce the nucleic acid molecules into the fertilized oocyte, without concern if the pronuclei are contacted, and introduction into the cytoplasm is sufficient. The injection occurs after fertilization and prior to the pronuclei forming a nucleus. Prior to fertilization, the female DNA will be at metaphase II stage and not have completed meiosis. The sperm stimulates meiosis and eventual formation of a nucleus. Here injection occurs before the nucleus forms. An embodiment provides the injection occurs up to 16 hours and, in an embodiment, 12 to 16 hours after fertilization, in another embodiment three, four, five or 6 to 12 hours after fertilization and in a further embodiment 16 hours after fertilization. As noted herein, one skilled in the art can determine when, in a particular animal, the nucleus will be formed and inject prior to that time. In pigs and cows for example, the time frame is 12 to 16 hours after fertilization. Without wishing to be bound by any theory it is believed this provides sufficient time to dissipate the molecules in the cytoplasm and for the nuclease recombination components to integrate the nucleic acid molecule into the eventual nucleus. Not only does this allow for introduction of nucleic acid molecules into large animals in which injection of the pronuclei was difficult, but also provides high event recovery, here, up to 100%, and avoids problems with production of a mosaic which can occur if recombination occurs after DNA replication.

As used herein, the terms nucleic acid or polynucleotide refer to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form unless indicated otherwise. As such, the terms include RNA and DNA, which can be a gene or a portion thereof, a cDNA, a synthetic polydeoxyribonucleic acid sequence, or the like, and can be single-stranded or double-stranded, as well as a DNA/RNA hybrid. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g. degenerate codon substitutions) and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer et al. (1991) *Nucleic Acid Res.* 19:5081; Ohtsuka et al. (1985) J. *Biol. Chem.* 260:2605-2608; Rossolini et al. (1994) *Mol. Cell. Probes* 8:91-98). The term nucleic acid is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

A "polypeptide" refers generally to peptides and proteins. In certain embodiments the polypeptide may be at least two, three, four, five, six, seven, eight, nine or ten or more amino acids or more or any amount in-between. A peptide is generally considered to be more than fifty amino acids. The terms "fragment," "derivative" and "homologue" when referring to the polypeptides according to the present invention, means a polypeptide which retains essentially the same biological function or activity as said polypeptide. Such fragments, derivatives and homologues can be chosen based on the ability to retain one or more of the biological activities of the polypeptide. The polypeptides may be recombinant polypeptides, natural polypeptides or synthetic polypeptides.

Thus when referring to a nucleic acid molecule of interest (NOI) is meant a nucleic acid molecule which is desired to be introduced into the animal cell nucleus. An NOI, then, by way of example without limitation, may be the target gene and may be modified; RNA; interfering RNA; a nucleic acid molecule that can have various impact on the target gene or another gene, as discussed herein; a nucleic acid molecule newly inserted into the genome and that may produce an additional polypeptide within the genome, or a combination of any of these. Any nucleic acid molecule desired to be introduced into the animal cell genome can be the NOI.

"Codon optimization" can be used to optimize sequences for expression in an animal and is defined as modifying a nucleic acid sequence for enhanced expression in the cells of the animal of interest, e.g. swine, by replacing at least one, more than one, or a significant number, of codons of the native sequence with codons that are more frequently or most frequently used in the genes of that animal. Various species exhibit particular bias for certain codons of a particular amino acid. Cas9 can be one of the sequences codon optimized for improved expression.

In one aspect, polynucleotides comprising nucleic acid fragments of codon-optimized coding regions which may produce RNA, encode polypeptides, or fragments, variants, or derivatives thereof, with the codon usage adapted for optimized expression in the cells of a given animal. These polynucleotides are prepared by incorporating codons preferred for use in the genes of the host of interest into the DNA sequence.

A "heterologous" nucleic acid molecule is any which is not naturally found next to the adjacent nucleic acid molecule. A heterologous polynucleotide or a heterologous nucleic acid or an exogenous DNA segment refers to a polynucleotide, nucleic acid or DNA segment that originates from a source foreign to the particular host cell, or, if from the same source, is modified from its original form in composition and/or genomic locus by human intervention. A heterologous gene in a host cell includes a gene that is endogenous to the particular host cell, but has been modified or introduced into the host. Thus, the terms refer to a nucleic acid molecule which is foreign or heterologous to the cell, or homologous to the cell but in a position within the host cell nucleic acid in which the element is not ordinarily found.

A nucleic acid may then be introduced into an animal host cell through the use of a vector, plasmid or construct and the like. A "vector" is any means for the transfer of a nucleic acid into a host cell. Vectors can be single stranded, double stranded or partially double stranded, may have free ends or no free ends, may be DNA, RNA or both. A variety of polynucleotides are known to be useful as vectors. A plasmid is a circular double stranded DNA loop. Referring to one or more expression vectors is meant to refer to one or more vectors comprising necessary regulatory elements for proper expression of the operably linked nucleic acid molecules. A vector may be a replicon to which another DNA segment may be attached so as to bring about the replication of the attached segment. A replicon is any genetic element (e.g., plasmid, phage, cosmid, chromosome, virus) that

functions as an autonomous unit of DNA or RNA replication *in vivo*, i.e., capable of replication under its own control. The term "vector" includes both viral and nonviral means for introducing the nucleic acid into a cell *in vitro*, *ex vivo* or *in vivo*. Viral vectors include but are not limited to adeno-associated viruses, lentiviruses, alphavirus, retrovirus, pox, baculovirus, vaccinia, herpes simplex, Epstein-Barr, rabies virus, and vesicular stomatitis virus. Non-viral vectors include, but are not limited to plasmids, liposomes, electrically charged lipids (cytofectins), DNA- or RNA protein complexes, and biopolymers. In addition to a nucleic acid, a vector may also contain one or more regulatory regions, and/or selectable markers useful in selecting, measuring, and monitoring nucleic acid transfer results (transfer to which tissues, duration of expression, etc.). Transformed cells can be selected, for example, by resistance to antibiotics conferred by genes contained on the plasmids, such as the *amp, kan, gpt, neo* and *hyg* genes. The techniques employed to insert such a sequence into the viral vector and make ether alterations in the viral DNA, e.g., to insert linker sequences and the like, are known to one of skill in the art. (See, e.g., Sambrook et al., 2001. Molecular Cloning: A Laboratory Manual, 3rd Edition. Cold Spring Harbor Laboratory Press, Plainview, NY). A "cassette" refers to a segment of DNA that can be inserted into a vector at specific restriction sites. The segment of DNA encodes a polypeptide of interest or produces RNA, and the cassette and restriction sites are designed to ensure insertion of the cassette in the proper reading frame for transcription and translation.

The nucleic acid molecule may be operably linked to a suitable promoter at the 5' end and a termination signal and poly(A) signal at the 3' end. As used herein, the term "operably linked" means that the nucleic acid molecule containing an expression control sequence, e.g., transcription promoter and termination sequences, are situated in a vector or cell such that expression of the polypeptide or RNA produced by the nucleic acid molecule is regulated by the expression control sequence. Methods for cloning and operably linking such sequences are well known in the art. Promoters may direct constitutive expression or tissue preferred expression. Tissue-preferred (sometimes called tissue-specific) promoters can be used to target enhanced transcription and/or expression within a particular cell or tissue. Such promoters express at a higher level in the particular cell region or tissue than in other parts of the cell or tissue and may express primarily in the cell region or tissue. Examples include promoters that secrete to the cell wall, retain expression in the

endoplasmic reticulum, or target vacuoles or other cell organelles. Other may direct expression primarily to muscle, neuron, bone, skin, blood or specific organs or cell types. Such promoters may also direct expression in a temporal manner, expressing at a particular stage of development or cycle of the cell. The promoter(s) utilized in one example may be polymerase (pol) I, pol II or pol III promoters. Examples of pol I promoters include the chicken RNA pol I promoter. Examples of pol II promoters include but are not limited to the cytomegalovirus immediate-early (CMV) promoter, the Rous sarcoma virus long terminal repeat (RSV-LTR) promoter, and the simian virus 40 (SV40) immediate-early promoter. Examples of pol III promoters includes U6 and H1 promoters. Inducible promoters may be used such as the metallothionein promoter. Other examples of promoters include, T7 phage promoter, T3 phage promoter, beta-galactosidase promoter, and the Sp6 phage promoter. An example of a DNA having a termination and poly(A) signal is the SV40 late poly(A) region. The use of these commercially available expression vectors and systems are well known in the art. The vector may contain multiple copies of a nucleic acid molecule of interest or a combination of nucleic acid molecules; also multiple vectors may be introduced simultaneously or sequentially into the cell.

Other components may be included in the vector or in vectors also introduced into the cells, such as polyadenylation sequences, enhancers, signal peptides, inducible elements, introns, translation control sequences or the like. As noted above, selectable markers allowing survival of cells with the vector or other identification of cells having the vector may be used.

A nucleic acid molecule is introduced into a cell when it is inserted in the cell. A cell has been "transfected" by exogenous or heterologous DNA or RNA when such DNA or RNA has been introduced inside the cell. When referring to integration of a nucleic acid molecule into a cell is meant that the molecule has recombined and become part of the genome.

The presence of the nucleic acid molecule of interest may be determined by any convenient technique, such as identifying the presence of a marker gene; detecting the presence of the inserted sequence via PCR or the like; detecting expression product from animal cells, tissue or fluids; Northern or Western blot analysis; or any other readily available method.

The examples presented are provided by way of illustration and not meant to limit the scope of the invention.

EXAMPLES

CRISPR mediated gene knock-ins:

In the inventors' laboratory, major prion protein (*PRNP*) has been used as a so-called "safe harbor" locus for targeted KI of transgenes. In mouse, transgenes are typically knocked into the *Rosa26* locus, which is ubiquitously expressed in all target tissues, and has been shown not to be vital for the survival of offspring if deleted. (SEQ ID NO: 1 is the PRNP protein Gene ID: 494014*)* Therefore, *Rosa26* has been the preferred site for inserting transgenes to prevent both: a) inadvertent silencing of the transgene; and b) insertional mutagenesis caused by random insertion of a transgene into an off-target site in the genome. In mouse, besides *Rosa26*, *PRNP* is another ubiquitously expressed gene, and is not required for viability, thereby qualifying as a safe harbor locus. Such a safe harbor gene is one that meets the criteria of having no known adverse effect on the cell by introduction of the vector or nucleic acid molecule into the site and has transcriptional competence across cell types. Besides mice, *PRNP-/-* cattle are also healthy. (Richt, J.A. et al. Production of cattle lacking prion protein. *Nat Biotechnol* 25, 132-138 (2007)). Any safe harbor gene may be utilized as desired for knockin of a nucleic acid molecule.

In pigs, and other large animals, the *Rosa26* locus has not been characterized adequately. Hence, *PRNP* was selected as a safe harbor locus for targeted KO and KI of transgenes and establishing recombinant pigs. The results are discussed below.

A) Assembly of targeting vector:

A gene targeting vector was generated consisting of 500 bp upper arm and 1000 bp lower arm homologous to the Cas9 cut site *(*SEQ ID NO: 2 and 3 respectively) bordering exon-3 of *PRNP* gene. Using primers bearing unique restriction sequences AscI and XhoI, 1000 bp of sequence of upper arm was cloned into the corresponding sequences in the GFP (green fluorescent protein; see SEQ ID NO: 4 and GenBank U73901) expressing piggyBac vector. (SEQ ID NO: 5, 6). Likewise, using primers consisting of BsiWI and MluI, 1000 bp of sequence downstream of the cut site were amplified and cloned into the

same GFP piggyBac vector. (SEQ ID NO: 7, 8 . The final targeting vector is generated by digesting the vector with AscI and HpaI and the linearized vector without prokaryotic sequences is injected into the embryos.

B) CRISPR mediated targeted gene knock-in in porcine embryos:

We investigated whether the gene targeting can be achieved by direct injections into the embryos thereby avoiding the need for targeting in an intermediate cell type and generating GM animals by cloning experiments. As noted in Figure 2a and 2B, by coinjection of targeting vector and CRISPRs (for generating DSB), we have demonstrated successful KI of GFP transgene into *PRNP* locus in embryos. In Figure 2, UBC refers to (human ubiquitin C promoter), GFP refers to green fluorescent protein; bPA (bovine poly adenlyation transcription terminator sequence); PGK/EM7 – is a hybrid eukaryotic (phosphoglycero kinase) and prokaryotic (EM7, a synthetic bacterial promoter derived from the T7 promoter that enables the constitutive expression of the antibiotic resistance gene in E.coli patents/US7244609) RNA polymerase II promoter sequences driving the expression of Neo/kan which are neomycin (or G418 for eukaryotic) and kanamycin resistance (for prokaryotic) selectable markers .In the figure "Upper" refers to the upper 500 bp homologous sequences flanking the Cas9 cut site, and the "lower" refers to 1000 bp downstream sequences flanking the cleavage site.

Briefly, maturing oocytes from sows were purchased from ART Inc. (Madison, WI) and shipped to the lab overnight in their commercial maturation medium #1. Twenty-four hours after being placed in the maturation medium #1 (provided by ART), 50 to 75 cumulus-oocyte complexes (COCs) were placed in 500 µl of tissue culture medium 199 (TCM 199) containing 0.14% PVA, 10 ng/ml epidermal growth factor, 0.57 mM cysteine, 0.5 IU/ml porcine FSH, and 0.5 IU/ml ovine LH and cultured for an additional 20 hours at 38.5°C and 5% CO2 in air, 100% humidity. (Abeydeera, Let al. Maturation in vitro of pig oocytes in protein-free culture media: fertilization and subsequent embryo development in vitro. *Biol Reprod* 58, 1316-1320 (1998)). COCs were vortexed in 0.1% hyaluronidase in HEPES-buffered medium containing 0.01% PVA for 4 minutes to remove the cumulus cells following maturation. Groups of 30-35 mature, denuded oocytes were placed in 100 µl of a modified Tris-buffered medium (mTBM) and fertilized according to established protocol (Abeydeera, L.R. & Day, B.N. Fertilization and subsequent development in vitro

of pig ooytes inseminated in a modified tris-buffered medium with frozen-thawed ejaculated spermatozoa. *Biol Reprod* 57, 729-734 (1997)) using fresh extended boar semen. 1-2 ml of extended semen was mixed with Dulbecco's Phosphate Buffered Saline (DPBS) containing 1 mg/ml BSA to a final volume of 10 ml and centrifuged at 1000xg, 25°C for four minutes; spermatozoa were washed in DPBS a total of three times. After the final wash, spermatozoa were resuspended in mTBM medium and added to oocytes at a final concentration of $5x10^5$ spermatozoa /ml, and co-incubated for 5 hours at 38.5°C and 5% CO_2. Five hours following fertilization, the presumptive zygotes were injected with 100 ng/ul of Cas9 mRNA alongside 50 ng/ul sgRNA targeting PRNP, and 2 ng/ul of linearized targeting vector using Eppendorf® Transjector. Approximately, 50 nl of the mixture is injected into the porcine zygotes, and cultured in porcine *in vitro* culture PZM3 medium (Yoshioka, K et al. Birth of piglets derived from porcine zygotes cultured in a chemically defined medium. *Biol Reprod* 66, 112-119(2002)) for 6 additional days. The injected embryos, which have now grown to blastocyst stage were screened for the expression of inserted GFP cassette. In Figure 3A and 3B, expression of GFP was apparent in the expanded blastocysts at day 6 of *in vitro* culture.

Targeted integration of GFP was also confirmed by PCR amplification of blastocyst DNA using primers both within and outside the targeting vector (Figure 4). (SEQ ID NO: 9, 10). A final confirmation was obtained by sequencing of the PCR amplicon. We have performed transfers of these GFP KI embryos to recipient sows and the animals are pregnant at day 60 of gestation, and will likely give rise to the birth of live animals.

Single stranded oligo mediated genome editing in embryos:

With the sequencing of animal genomes it is now possible to identify quantitative trait loci (QTL) and in some cases quantitative trait nucleotides (QTN) that can influence the phenotype in question. It is imperative that these QTL and QTNs be tested. The ability to specifically alter nucleotides to accurately reflect superior or disease genotypes will be a powerful approach for improving animal agriculture and for generating disease models respectively. We have investigated if it will be possible to alter nucleotides to obtain desired phenotypes. This approach has the most direct applicability to animal agriculture

where the negative traits such as boar taint, horned phenotype, disease susceptibility, and other parameters be altered with pinpoint accuracy to benefit the animals.

We have tested if specific loci can be targeted for introducing nucleotides by using single stranded oligo bearing the desired nucleotide modifications. In this case, we have investigated if we can target the PRNP loci to modify one of the existing nucleotides to generate a novel EcoR1 restriction enzyme (Figure 5A). The promoters used were CMV and T7, HA refers to _HA tag, NLS is the nuclear localization signal which may be used to facilitate nuclear localization. In addition, for ease of analysis, a 34 nucleotide LoxP site (SEQ ID NO: 11 has been engineered on the oligo (Figure 5A). As shown in Figure 5B, we were able to alter a candidate *ZBED6* locus by injecting the LoxP and EcoR1 restriction enzyme bearing oligo, along with Cas9 mRNA and sgRNA targeting *ZBED6. (SEQ ID NO: 12).* ZBED6 is a zinc finger protein that regulates expression of the insulin-like growth factor 2 gene in pigs. See Markljung et al., "ZBED6, a novel transcription factor derived from a domesticated DNA transposon regulates IGF2 expression and muscle growth. *PLoS Biol.* 7(12) e1000256. As mentioned above, the engineered *ZBED6* locus could be visualized by a 34 nucleotide increase in the length of the PCR product, which was then verified by EcoR1 digestion. A final confirmation was evident from the sequencing of the targeted loci where the restriction site is shown as underlined (Figure 6). We were able to repeat the results with *PRNP* (See Figure 7. Cas9 was introduced at 100 ng/ul; sgRNA at 50 ng/ul and LoxP oligo at 2ng/ul)

We have established a very efficient way for generating transgenic and/or edited animals, by eliminating intermediate steps and performing gene targeting directly in the embryos. The advantages with our approach are two-fold, one in generating targeting vectors and second in delivery of constructs. In terms of the targeting vector, a principle advantage is the less stringent requirement for lengthy homology arms in the targeting construct. In our experiments, gene targeting could be achieved with as small 300, up to 500, up to 1000 bp of homology on each targeting arm, a DNA fragment that is readily obtained by PCR of genomic DNA or can be synthesized as a G-block and obtained from IDT Inc. (Figure 2A). This is in stark contrast to at least 6 kb of homology required for gene targeting in porcine fetal fibroblasts (PFF). In terms of delivery of the DNA constructs, we have identified a window of time post fertilization, where DNA and the CRISPR/Cas can be injected into the cytoplasm of the primitive pig zygote that allows the

DNA to become packaged in the pronucleus when the nuclear envelopes are formed, thereby allowing for homologous recombination mediated gene targeting in the embryos. This strategy eliminates another major bottleneck in generating transgenic animals in pigs where locating pronucleus in the zygote for injecting transgenes has been a limiting factor. Additionally, the efficiencies of gene targeting were 80 to 100% with all of the blastocysts screened, showing accurate targeting of a KI construct in embryos. A final advantage is the ability to edit only few nucleotides precisely in the embryos, which have the ability to alter phenotype for agricultural applications, and for generating biomedical models. In summary, we have developed technologies to usher large animal biotechnology into the genomics era.

LIST OF SEQUENCES

SEQ ID NO: 1 the PRNP protein

NCBI Reference Sequence: NC_010459.4

ACGTACGCGGCCAAAAGAGTCTCAACTCCCTCCCAGAGACTCAGATTTCCGACCAGCTTGGCAGATCCCG
GGCGCCGGAGCGCCAGAGCGCGCGCGCGCCGCCGCCGCCTCCCTTCCCCGCCCGCGCGCCTCGCCACC
CCTCGGCGCCAGACACTGACAGCCCCGGAGCTGCGAGCGTCTTCTGTTCCAGCGGTGGCAGGTAAACAGC
CGGGTCGCCCCGAGAACTGGGGGTGCCAAGGTCGGGAGTCAGAACCCCCCGCCTTGGAGCTTAGGCGCAA
GGGTGAGCGGCCCACTTGGGGCGCTAGGGAAACCTTGACAGAGAGGGGGGCGGGGGACTTCCGGCGGGCG
GGGGGCGCGACAGGCCTGCCCCCGCTTGTTCGTCGGCTTAACATTGGCCCCGCTCACATTTGCTTTCCAG
GTGGGGTGAGGGCCCTTTACTCGGAATGGGGCTTGGTGGGGGAAGGGGGTGCCTGGGGCTGGCCGCGACC
CCTGGGCAGGGAGGCGGATGGTGGCGGGGAGTCTAGCCCTCTCCCACGCGCGTCCGCGGCTGGCGGAGGA
TGGGAGAGGCGCGCCGGCACCGGGGCCGCAGGGCCGGGAACTCAAGGAGGGTCGCCCACGGCCTCCGCTG
AGGCGCTCGGAGGAGAACCAGGGAGAGGGGCCGGCGTGCGGCTCTCCCGAGGCCCCAGGGGCACGTGGGG
TGGGGATGATGGCCCCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAGCCCCAAGGGGCCCGTTGGGGG
GGGGTGATGGCCCAGGCCCCGGCGGGAGGGGCCCTTGCCGCCGCCACCATGGAGCTCCTCCGGAGAGACG
GCCCACCTGCCATCCACGCGGCCCGGGGGTGCTGGGGAGGCCGCAGCGGGGTGGGCCTCCTGCATCCCTG
GGGATGGAGACCGGGCACGTGGTCCGACCTTCGGGCCTCGGCGGGCGGAGTCTCGGGCGGCCCCACCCTC
ACTGTCGGCCGAGACGTCCAGGTCCCTGGGGGCCGAGTGTCCGGAGAGCCCAGGTGCTTTGCTTTTCACT
GTCTCCACTGCCCCTCCCCGAAGGGAACTGGGCGGTCAGAGGCTCCTCTCAAGGGGATCAGGGCCGGACA
GCTCTCCTCCTGATCCACTTGTGTGGGCTGCCCTTTCAGGCTTGATTGTAAACAACGAGGAGAAGTCCAT
TTTTACTGTCCCTTTTCATTTTTTTCCCTTTTCTAAATTTGTAGCAACCATGGCAAATCAGTTTTTAAAA
TCATAACCCACAGCCATCAATCCACCCTTACCATCTCAAAGCCACTGCTTTCTGTTTTTCAGTTTTCTGT
CTCCAGATTCGTACATAACGCAAAGAAATTTCAACTGCCCTGATTATGATTATCCTCCTCTAATGGCCGA
GTTATTTTCTTCTGCTTAAAGCGTCGCAGTTTAATAAGCAGTTCCCCGAATGCTGAACATTTGAAATGTT

TCGTTTTTTCTTGCAAAAAAACCTTCCAGGTATAACAGTATTAAAGAAAGAATAGGAATAGGAGTTCCCT
GGTGGCCTGGTGGTTAAGGATCCCATATTGTCGCTGCTGTGGCTAGAGTTCCATCCCTGGCCCGGAAACT
TCTGTATACCTCAGGCAAGCCAACAAAAAGAAAAAAAGAAGGGAAGAAAAGAAAAAAATAGTAACAGAAA
ATTGAATTAAAATGTCAAACCCCTGGAAAAATTAACAACTATCTCAGATTTGAGGAAGAAAAAAAAAATC
AACAGTTTTACCTGCAGTAAACACTGAGGCGCTCTTACAATGAAAAAGAACTAACCGGAAAGGGAAAGAA
AGCAGAGCTAGGATGTGATTTGTATATGATTTGTATCTGACACAAAATTTTCATAGTTATGAAAGGAAAA
TATGAAAATTATAATAAATGATGAATCAATTATAGAATAAAATGTAAATTAAAGCACTCTGGATTATCAT
TTTACAATTACTAACAAATACAACAACAGTAATAGTAACAGCAACCACTGTTGGAAGATTGCCTAGAAAT
TTTCACATTCAGTTATTTTGAGAGGTGGCAGGACTTTGGGGTTAGAAGAATGGCTCTGCACCTAATTTCA
TAGCATGGAATGGGTTACCTAATTTCCTCATCCTCCTTTTGTGCATTCATAAAATAGAGGAAATTATACC
TACTTCAGGAAATTGCTAAGATTAACCATCTATGTAAAACTGACCTTTGGTATGTAATCCTTTTTCTATT
CTTGGGCACACTGCACTGGGGACATTGTGAATTTACTGTAAAGCAATTTGGTAATGCATTAGGCCATGAC
CCAGAATTTACCTGCAATGGAAGAAACAACAAGAAAAAAAAGGGGGGGAAAAAAAAGCCATATGCAGTCA
CAGAATTCAGAATGATCTAAAATTAGACCCAACGGAAAAACAACCTAAATGTATAACAGCAGGGCAGCAG
CTGAGGAAATCATGGCTCTTTAACTGAAAGAAACATTATGTAACTATCAAAAGTCGGTGGTACATGAGAA
AAAACTGATAAATCAGTGTAGATTACATTACCAAACTGTATCTACTTACTCAGTGAATGCATATGTGGAA
AATCTGAAAGGAAAAGCATACAAAGGAATTGAATAGAATTAAAATGAAGTTAGAAATTATGGGTCAGTTT
AATTTTCTTTTTTATTTATATACTGCTTTTCTAAGTAACGAAATTTTTAAGATAATTATTCATCGTTCTC
AATTTTTGTAATGAAATGGAGCCATAGTTTTCTGTATGTGGGAGAGGGAGAGCTCAGTGTCATGTTAGCT
GACTTCATTTCTTCAGATATTGTCTCTTTTTTTTTCTCGTGTAGATGGATTAGGGGGTTAGTAATGTATA
TATTAGGATCCTTCTTTAAGGAATAGNNTCTAAAAATTATTATCAATTCTAGNTTTTTCTTTTTTTGGTG
TTTCTAAGGACCCTGAATATATTTGAAAACTGAACAGTTTCAGCTAAGCCGAAGCATTCTGTCTTCCCCA
GAACACAAATCCAGCCTCGCTGAGCCATTACAGATGTACATACTGCAGAGTCCCTTTGATCTCGTGATGG
CTCAGAGTCCCTATCTGTCACAGAAAAGAAGGGCCTAGAAAGGCTATGATGACAGTCCTTGCTCTGAGGC
AACTTAGGGTTATATGTGTAACCCACACCTCTTAAATCATCCTTTCTTGTAAAATCATTTTGATTTTGCA
GGCAAAGGGCCACGGATTTCTTTAGAATCACTCTGAGTTATACAAGGAATCAACCATTTAAAAAATACAA
TAAAAAAGCAATTACAAATATTTCTGTACCAGATTAACACTGAAGGTGACTATCAGCCAACAAAAGGTTG
ACAGTTTTTCTAAGCTGTGGTTTTAGTTTACCTATTCCATTCTCCCTTTTCAGTTCTTATTTCCTATTCC
AAGACAACTGTATATCAGCTGTGAACTGCTGCATGTGAAACATTCAAAACATGCCTCTACATAAAGTGGT
GGCTGCTGCAGAAGAATCTGAGTGGATAGCATATATAGCAGTTCTTACAGTGGGAGTATTTTTTTCTCAA
TTTGTTCATCCTTTTTCTTCTAACTATATTCTGCTTATTGCTGTTCCAGTATTGTGTGATCACATCAAGG
GAGGGGTACCCTGTTATGCTTTATAAGTGTTAAGTTTTGTGCTCCTGGGACCTAGCTCTGAAGCCTGGCT
AGGAGTGCAGTTCTCTGGGAAGCTTTCCTGTCTTTCCTGAGCTAAGGTGGAACTCAGCAGTCATATTTTC
ATCTTTTGAAGTCATTGTCCCTTAAAGAATCTAATGCAAGCTATATAAAATTTTTACAAACCATTCTGGG
GTGGGGGGTTTTCACTGGCACTCTGAGACTTAGCCCTCGACTGAATAGGGCAAGGGGCCCCCAACCTGGC
CATCCTCAGAAATAGGTGAAGAGCTTTTAAAAATGTAACATTGCTAGAACCTATCCCAGGCTGACAGCCA
GGTGTTTTCTAGTGATGGGATTTAGGAATCTATATTGTAGCAGGATTTAACAGTATCTGATAACAGTATA
TTCTCTCTTCTTTCCTCTTCCTTTTTAAAGTTGAGTTTGTTATTTGTACACATTCCTTACCCCTTATGGT
TTTATTGAACTCAAAAAAATTGAGCTGGAGTTTTGTCATTTTATTGAGATTATAAATTCACTTAGGGGTA

ACTGACATCTTTAGAATACTAAATTGTTCCAACCAAGACTTTCCCATTTTTATGGGTCATTTTCTTTTTT
TTTCTTTTTTTTTTTTTTAATTTTTTTTTTTTTTTTTTTTTTTTTTGGCCTTTTTTTGTCTTTTGTTGTTG
TTGCTATTTCTTGGACCGCTCCCGCGGCATATGGAGGTTCCCAGGCTAGGGGTCGAATCGGAGCTGTAGC
CACCGGCCTACGCCAGAGCCACAGCAACGGTGGATCCGAGCCGCGTCTGCAACCTACACCACAGCTCACG
GCAACGCCGGATCGTTAACCCACTGAGCAAGGGCAGGGACTGAACCCGCAACCTCATGGTTCCTAGTCGG
ATTCGTTAACCACTGCGCCACGACGGGAACTCCTTTATGGGTCATTTTCTATCTTCTTTATTAATGTTTA
AATTTTTTCCCCAGTCTTAATTTCTCTTAAGTCAGTCCTAGTTACATTACAGTTCTGGTTGAAATTATGG
CTATTTTATTAATTTTTTTCCTAGTTGATTATTGCTGCTGTAGAGAAATGCTATTGATTTTTATAGGCCA
ATCTTGTATCCCACAACCTTGCCAAACTCTAGTACTGAGTCTAGTACTGAAGTCCCCTTGTGGCGCAGGG
GGTTAAGGATCAGGCATTGCAAGCAGCTGTGGTGCAGGCTGCAATTGCAGAGTGGTTTCAATCCCTGGCC
CAGAAACTTCCACATGCCACAAGTACAGCCAAAACAAACAAACAAACAAACAAAAACTTGTGTCAGGTCT
GTTGAATTTTTTATGTATATGATCATCCAAATTCCAAATAGTGACAGCTGGAGCTCTTTCCTTGCAGTTC
TTGCACCAGTTACTTCGTTGCTTACAGCGTTGGAAGGACCTCCAGTCCCATGACAGTTGGCATCCTTGTC
TTGCTTCTGATTTTAAAGGGACTGCATTCCAAAGTTATCTATTAAGTACTGTTTTGTGGTATAATCTTTA
TCAAGTTAAGGAAATTCCCTCCTATTCCTGGTCTACTAAAAGTATTTTTTTTAGTGATAAATAGATATTG
ACCTTTATCAAATACTTAATGTTTTAAAATGTATCAGAGGATTAGGCCACAGAATGACTGTTCAGGAGTA
TCTTATATGAAATGTTATAACACAGCCAGGCATTCAGAAACCCAGTCTCCATTCCCAGCAATCTCACTAA
TTAGCATTTTAATCGCAGATATGTCACTTGACTTAGCTGAATCTTGGTGGCCTAATTTGTAGAGTGCGAG
ATTGGGATTAAATAATACAAGGTCTTTCTTATTTTATCACAAACAGAGATTTTCCATGGTCTCATTAAGT
TTTGAGTCTTCCTGGAGTTGCGGCTGTGGCTCAGTGGTAGTGAACCTGACGGGTATCCATGAGGACATGG
TTCAATCCCTGGCCTTCTCAGTGGGTTAAGAATCCAGCCTTGCTGTGAGCTGTGATGTAGGTCACAGAAG
TGGCTTGGATCCTGCGTTGCTGTGGCTGTGGTATAGGCCAGCAGCTGTAGTTCCTATTTGACTCCTAGCC
AGGGAATTACTATATGCTGCAGGTGGGGCCCAAGTTTTGAGTCTTCCGTATCCTTAGAAATGAGTCTCTC
CTTCAAAATGTAGTTGTTATTTTTTTATGCTTTTCATAAGTAGATGTTGCTTTTGTCATTCTCAACTGGT
TTTAAAACGATCCTGTCTTAATATAGTCTATGATTATCCAGCTTTACCTAGTATTCAGTGTTTAAGTACT
AATGGCAACATATGCCTGCTCCCCTTACACACTGTTGCCTGTATCTCCAGACCCTGTTCATTCCTCAGCT
CCAGATTCCCCAAGCTAACTGCCTACTTGTCTCTCTTGTCCCACATGCAGCTGGAGGATCGCATTGCCAA
AGCCACACCCCTCATTTGTCTCCCTCCAAATTTGCTCCACCTTCTGTGTTCTGTCTACTCAATTGCCCGA
TCTGGAAACTTGGGTACCACCTCGATTCTTCCATACTCTTACTTTCCACATCAAACTGCAAGCCTTCCAG
AAGGTGCTAATCCACCTTGTTCAGGCCCCTGAGGTGTTGGAACAGAATTGTGAGCAGTATGGAGGCGAAG
AGTTGAGCCCCATCAGCAAATGGATCCTGGTTAGGTTCATTAACCGCTAAGCCATGAAGGGAACTCCCTA
ATTGGCATATTTTGAGCAGTTTGTGCTGATAATGTTCAAATAAATTTTTATTTACACCAAGACCTTTCAA
AAGGTGTTTCTTGGTGACCTAGCAGGTTAAGGATCCAGGGTTGTAACTGCTATGGCTCGGGTCAGTGCTG
TAGCATGGGTGAGCTTGATCCCTGGCTCTGGGACTTCAGCATGCCATGCGTGCAGCCAAAAAAAAAAAAA
AATTTTTTTAATAAATGATAGTAGAAACAAATGAAATATTAACTACTTTATTTCTTTAATCATTCTTCCT
CTGAATATTATACAATTTAATCTTTAACTCTGAAGGTTATGAAAAATAATTTTATATTTGATGCCACTTT
TATTTGAAGATACCCTGGTGCTGGGGATAAAACTAGTATCAGGGTTATAGCATGTGCCGTTTTTGGTTTT
ATGGCACATGGTGTTTTGGAGTAGTACATGGTTTGACATGAGGTCCAGTAGAGTCTGTTTAATGGTCATT
CTATTCTGTGTCTTTAATCACTGCCCTTGGAACCCAGCACCCAGGGTAAATGACATCTTCAGTCTGAAAG

GATGTTTAATTTGATGAATGTACATTGTTTATTCTTAAGTTATTCTAGCCCTTTTATGTCTTAATATGTG
TCTTTTGCACCAAAAATTAGGAGGAAGTAACTCACTCATGAACTACCGAATGCCAAACTAAAACTTATTC
AGTGACTGAGTGACTGTAAATCTTGTGAACACAGACTTTCTTTTTTTTTTTATTAAAAATGTTCTTTTTA
AAAAAATTTTTTTTGGCCAACCACTCATGGCATGTTTGATCAGGCCAGGGATCAAACCTGAGGCAGAGGC
GTAACCAGAGCCACAGCAGTGACAATGCTGAATCCTTAACCTGCTGAGCCACCGGGGAACTCCTGAATGC
AGCCTTTCTTACCCCATTTTGAGCAGCCTCATTTGTCTGCCTATAGAAAGAAAGATGAATTTATAGATGG
GCTTGCAAAATCTTGCATGTTTTCTGTTTCCAAAAACTGTTGAGAATTCCTTTGAGGAAATTATTGTAGG
TATTTATATATTCAGAGGATATTAATTTCTTTACTTAATAGATACTTATTGTGTACCTGCCCTATACCAG
GCACTGCCCAAGCCCCTGAGGATATAGCAGCAAACAAAAGAGACTCATTCCCTGCTTACATTCCCACTCA
AGGAAGAAGGACACAAGTCAGCTATTTAAAAATTATTTCATGATCTCAACACCTACCTGGGGGCTTCACT
TCCAGAGGGGCCTGCCAGGTCCCTCAAGGTAGTCCACACAGCAGTTAGGGCCTTGGCTTCACTCCCAGCC
CCAGCCCCCACAAGTGACCGCTCAGGTACAAAGGAATGAGTCTGTGCGTAGAGGGTTCCCCTTGGCACAG
GGGCCCTAGGCCCCATAGGTGCTGATAGCCCTCATTATAGCCCCATTAGCAGAGCTGCCAGGGTCCCTAC
AAGAAGGTGCCTGGTTACAAGAATTATTACTTTACTTATTTATGCACCATGTGGAAGTTTCCAGGCTAGA
AGTCGAATCAGAGATGTAGCTGCTGGCCTACGCCACAGCCATACCAGATCTGAGCTGCATCTGCGACCTC
ACCACAGCACACGGCAAGGCCAGGGATCGAACCCGCGTCCTCATGGATACTAATCCATGGCTGGTTTCTG
CTGAGTCATGATGGGAACTCCCAGAATCACAAAAGCTCAAAGCTGTGGCAAACCCACCAGGTGTTTATTG
GTTTTTCCAGTTTAGATACAATGTATCAAGCAGAGGTTATTTTTACCATAAGCATGTTGCTGGCATTCCA
CCTTTATCTTTTCTAAGAAACAGAGCCAGAAAATTATCTGAAGGTCAAATTTGTCCTTAGAGAAGGAGAA
AGAGTTGAGTTAACCCTTCACCTACAGTTGTTTTTGTTGTAAGTGTTGCACAGGAGACAAATGGAGTATA
AAGAACATTCACAGCTGATGCCGCTACTATGTTCATTATGCTGCAGACATTAAGTGATTTCAATATAAAC
AGGACACTGACACCCTCTTTATTTTGTATTTTGCAGATAAGTAATCATGGTGAAAAGCCATATAGGTGGC
TGGATCCTCGTTCTCTTTGTGGCCGCATGGAGTGACATAGGGCTCTGCAAGAAGCGACCAAAGCCTGGCG
GAGGATGGAACACTGGGGGGAGCCGATACCCAGGGCAGGGTAGTCCTGGAGGCAACCGCTATCCACCCCA
GGGAGGGGGTGGCTGGGGACAGCCCCACGGAGGTGGCTGGGGACAGCCCCACGGAGGCGGCTGGGGACAG
CCCCACGGTGGCGGCTGGGGACAGCCCCATGGTGGCGGAGGCTGGGGTCAAGGTGGTGGCTCCCACGGTC
AGTGGAACAAGCCCAGTAAGCCGAAAACCAACATGAAGCATGTGGCAGGCGCCGCTGCAGCTGGGGCAGT
GGTAGGGGGCCTCGGCGGTTACATGCTGGGGAGTGCCATGAGCAGACCCCTGATACACTTTGGCAGTGAC
TATGAGGACCGTTACTATCGTGAAAACATGTACCGTTACCCCAACCAAGTGTACTACAGGCCAGTGGATC
AGTACAGCAACCAGAACAGTTTTGTGCATGACTGCGTCAACATCACCGTCAAGCAGCACACAGTGACCAC
GACCACCAAGGGGGAGAACTTCACCGAGACGGACGTCAAGATGATAGAGCGCGTGGTGGAACAGATGTGC
ATCACCCAGTACCAGAAAGAGTACGAGGCGTACGCCCAAAGAGGGGCCAGTGTGATCCTCTTCTCCTCCC
CTCCTGTGATCCTCCTCATCTCTTTCCTCCTTTTCCTCATAGTGGGCTGAGGGTGGCCTTTCTGTCGGCA
TCATCTTCTTAATCTTTATCAGGTTGGGGGAGGGAATATCTACCTGCAGCCCTTTAGTGGTGGTGTCTCA
TTCTTGCTTCTCTCTTCGGCACCCATAGGCTAACATCCATGGCGCTTGTAGCACTGGAAAAGGAGAGTAG
ACCTGAGATGTGATGTATTTAAGCCCCATTTGATTGAATCCTTCATAGGCCAGTGCTAGTGCTGGACTGG
TAAGAGCGTAACAGCAAATAATCATTGGTTGATCTGGGCTCATTTTTTGTCTGGTGCAACAGATTGAGGC
TAAAACAATTCTCAAAACACACTTCAAGTACCTTTACCTAAATACCTCCAGCTCCTTCTCCAGCTAGAGC
TCAGTACACAAATGCCCCGCCATAGTAGTGATTTTGTAGCAACTTTCCCATTTAAGAAAACCTGACTACA

TTTTCCTGTTCAAATAGCATTTCTACTGAGTTGGGGAGGAGGCCACATAATACTCATTCAAAAAAATGAA
ACTGGAAATCCTTAGCTCCTGGGCCCAGGGTCAGCCCAGTGGAAAGCATGTGTCCTGTGTCTGCAGAGAA
CTAAGGATATTTTGCAATTTGCAGTACAGGTTACACAGCAGCTATTGCATCAAGAATGGATGTCTGTGCA
ACACTAGACTTCTGGGCAGAGGGCATTTTCACAGGCAATGAACATAACTCACATAATATGAAAGGCTCTG
AAACTTAAAAAATTCCCACCTGTGTGAGGAACCCTCAGAGGCAGCCTTCTGTTATGGATGTTTAAAGCAC
CTTCATGGGGTAGTTCTTTCTTTAGTAATACAAACTATAGATAATTAAGGTAGTAGGACATGAAACAATC
TTCTGGACATTGAGAACAAATCTCTTTTGTTTGTTTATCTGGGAACTGGAGTGATTTTGCCATTTCTTGG
ATGAAGCCAGGAGATTTTAACATAGAGGAAGCTGCAGCTATAAAAACATCATATTTATTCATTTGATTGA
GTCTTTCATGGGCCAGTGCCGGTGTTGGGCTAGCAAGCATATGATACCAAATATAGAGGGTTATGAAGAA
AATGATTAGTGTACAAAAAAGAGAAATGCTTACATTTCTTTATTGCTGTGTCATAATTGTCAAAAATCAG
AATTAGGTCCTCAATTTCTATAATTGGCTTTTGAATCAAAGAATAGGAAGGCATTCCCCCCCAAAAAAGT
TAAAGATGATAGAAATATGATCCATTCATATTAGGAAAAGAAATTCTGGTACTGTTATTTAAATAAGGCA
AAATTATTCCCTGGATTGTTTGGTGTTGTCACCTAGCAGATATACATTACTCTTCTGCATTGTTATTGGC
TTGCACTTTGTGGGTATCCTATGTAAAAAAAATATATGTATATATATATATTGCATATGACAAACTTGGA
GATTTTGGTTAGAGCTGTTAACATCTGAATTATCAAATGCATTACTTGTTTTTGTAAGGTACTAAATATT
TAATAATACTTAAAGGAAACCCTTTTGTGTGGTCCTTCGGCTTACAATGTGCACTGAATAGTTTTGTATA
AGGATCCAGAGTGGCATTTGAAATTCGCATGTGCTTTATATTTTCTATATTTGTAACTTTGCATGTACTT
GTTTTATTGTATTAAAAGTTTATAAATGTTTATTATCTGACTAAAATTAAAACAGGAGCTACAATGAG

SEQ ID NO: 2 500bp of upper homologous arm

CTTGGCACAGGGGCCCTAGGCCCCATAGGTGCTGATAGCCCTCATTATAGCCCCATTAGC
AGAGCTGCCAGGGTCCCTACAAGAAGGTGCCTGGTTACAAGAATTATTACTTTACTTATT
TATGCACCATGTGGAAGTTTCCAGGCTAGAAGTCGAATCAGAGATGTAGCTGCTGGCCTA
CGCCACAGCCATACCAGATCTGAGCTGCATCTGCGACCTCACCACAGCACACGGCAAGGC
CAGGGATCGAACCCGCGTCCTCATGGATACTAATCCATGGCTGGTTTCTGCTGAGTCATG
ATGGGAACTCCCAGAATCACAAAAGCTCAAAGCTGTGGCAAACCCACCAGGTGTTTATTG
GTTTTTCCAGTTTAGATACAATGTATCAAGCAGAGGTTATTTTTACCATAAGCATGTTGC
TGGCATTCCACCTTTATCTTTTCTAAGAAACAGAGCCAGAAAATTATCTGAAGGTCAAAT
TTGTCCTTAGAGAAGGAGAAAGAGTTGAGTTAAC

SEQ ID NO: 3 1000 bp lower homologous arm

CGGTTACATGCTGGGGAGTGCCATGAGCAGACCCCTGATACACTTTGGCAGTGACTATGA
GGACCGTTACTATCGTGAAAACATGTACCGTTACCCCAACCAAGTGTACTACAGGCCAGT
GGATCAGTACAGCAACCAGAACAGTTTTGTGCATGACTGCGTCAACATCACCGTCAAGCA
GCACACAGTGACCACGACCACCAAGGGGGAGAACTTCACCGAGACGGACGTCAAGATGAT
AGAGCGCGTGGTGGAACAGATGTGCATCACCCAGTACCAGAAAGAGTACGAGGCGTACGC
CCAAAGAGGGGCCAGTGTGATCCTCTTCTCCTCCCCTCCTGTGATCCTCCTCATCTCTTT
CCTCCTTTTCCTCATAGTGGGCTGAGGGTGGCCTTTCTGTCGGCATCATCTTCTTAATCT

TTATCAGGTTGGGGGAGGGAATATCTACCTGCAGCCCTTTAGTGGTGGTGTCTCATTCTT
GCTTCTCTCTTCGGCACCCATAGGCTAACATCCATGGCGCTTGTAGCACTGGAAAAGGAG
AGTAGACCTGAGATGTGATGTATTTAAGCCCCATTTGATTGAATCCTTCATAGGCCAGTG
CTAGTGCTGGACTGGTAAGAGCGTAACAGCAAATAATCATTGGTTGATCTGGGCTCATTT
TTTGTCTGGTGCAACAGATTGAGGCTAAAACAATTCTCAAAACACACTTCAAGTACCTTT
ACCTAAATACCTCCAGCTCCTTCTCCAGCTAGAGCTCAGTACACAAATGCCCCGCCATAG
TAGTGATTTTGTAGCAACTTTCCCATTTAAGAAAACCTGACTACATTTTCCTGTTCAAAT
AGCATTTCTACTGAGTTGGGGAGGAGGCCACATAATACTCATTCAAAAAAATGAAACTGG
AAATCCTTAGCTCCTGGGCCCAGGGTCAGCCCAGTGGAAAGCATGTGTCCTGTGTCTGCA
GAGAACTAAGGATATTTTGCAATTTGCAGTACAGGTTACACAGCAGCTATTGCATCAAG

SEQ ID NO: 4 Green fluorescent protein encoding sequence

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
GGCGACGTGAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAC
GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACC
CTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAG
CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC
TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTG
GTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCAC
AAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAAC
GGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC
GACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCAC
TACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC
CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCACGGCATGGACGAGCTGTACAAGTAA

SEQ ID NO: 5, 6, 7, 8 primers used to obtain the upper and lower arms

PRNP upper arm MluI for: CTAG*ACGCGT*CTTGGCACAGGGGCCCTAG

PRNP upper arm BsiW1 rev: GACTA*CGTACG*CCTACCACTGCCCCAGCTG

PRNP lower arm XhoI for: GACT*CTCGAG*CGGTTACATGCTGGGGAGT

PRNP lower arm AscI rev: GTCA*GGCGCGCC*TTGATGCAATAGCTGCTGT

Targeting vector is generated by digesting with AscI (introduced by the primer) and HpaI enzyme on the endogenous sequence.

SEQ ID NO: 9 and 10 primers inside and outside vector to confirm GFP

Inside the GFP expression vector: CCAGCTGGGGCTCGACTAGA

Outside the homology arm: CCATTTTGAGCAGCCTCATTTG

SEQ ID NO: 11 LoxP nucleotide sequence
ATAACTTCGTATAATGTATGCTATACGAACGGTA

SEQ ID NO: 12: single stranded oligo used to target ZBED
GGTGGCAGAA GGAGTGGATA AAGAGGCAAA ATTGCCTGCC AAAAAGAAAA GAAAGAAGGGTTTGC (EcoR1 site follows in italics) *GAATTC* (LoxP site follows in lower case) taccgttcgtatagcatacattatacgaagttat AGGGGAAAA GGCGACGAAA GAAACTGATC CTTGCAAAAA AGTTTAGTAA GGATTTGGGA TCTGGGAGGC CTGTTGCAAG

SEQ ID NO: 13 is the *ZBED6* sequences shown in Figure 6 with the EcoRI and LoxP insert Blastocyst 1 and 5
AGAAGGGTTTGCGAATTCTACCGTTCGTATAGCATACATTATACGAAGTTATAGGGGAAAAGGCGAC

SEQ ID NO: 14 and 15 flanking sequence for ss oligo
Upper: GGTGGCAGAA GGAGTGGATA AAGAGGCAAA ATTGCCTGCCAAAAAGAAAA GAAAGAAGGGTTTGCGAATTC

Lower: AGGGGAAAA GGCGACGAAA GAAACTGATC CTTGCAAAAA AGTTTAGTAA GGATTTGGGA TCTGGGAGGC CTGTTGCAAG

SEQ I DNO:16
ZBED insert from Balstocyst 3 and 4 from Figure 6
AGAAGGGTTTGCGAATCCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGAC

SEQ ID NO:17
ZBED insert blastocyst 2 from Figure 6
AGAAGGGTTTGCGAATTCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGGC

What is claimed is:

1. An *in vitro* method of targeted gene modification of an animal genome of an animal having pronuclei that cannot be visually observed by an unaided human eye, the method comprising:

producing at least one zygote comprising said animal genome of said animal by fertilizing an oocyte of the animal with animal spermatozoa, wherein said animal spermatozoa is capable of fertilizing said oocyte;

injecting into said zygote a composition comprising:

i. a nuclease enzyme or a nucleic acid encoding a nuclease, wherein said nuclease introduces a double strand break at a target locus of said animal genome; and

ii. one or more nucleic acids and/or expression vectors comprising:

a) at least one guide nucleic acid molecule that targets said nuclease to said target locus;

b) a single stranded oligonucleotide or double stranded nucleic acid molecule comprising a heterologous nucleic acid molecule of interest (NOI);

c) a first sequence homologous to a first flanking sequence upstream of said target locus; and

d) a second sequence comprising a second flanking sequence downstream of said target locus,

wherein said first and second flanking sequences comprise at least about 40 to about 50 nucleotides when said NOI comprises a single stranded oligonucleotide, and wherein said first and second flanking sequences comprise at least about 300 to about 1000 nucleotides when said NOI comprises a double stranded nucleic acid molecule;

wherein said composition is injected into the cytoplasm of said at least one zygote before said zygote forms a single nucleus;

wherein said NOI recombines into said target locus of said animal genome; and

wherein said animal is a swine or a ruminant animal.

Date Reçue/Date Received 2022-03-21

2. The method of claim 1, wherein the composition is injected into the cytoplasm of the at least one zygote 12 to 16 hours after the fertilization of the oocyte.

3. The method of claim 1, wherein expression and/or function of said target locus is modified after said NOI is recombined into said target locus of said animal genome.

4. The method of claim 1, wherein said NOI expresses a polypeptide.

5. The method of claim 1, wherein said animal is a ruminant animal.

6. The method of claim 5, wherein said ruminant animal is a bovine.

7. The method of claim 1, wherein said animal is a swine.

8. The method of claim 1, wherein said nuclease is a Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) - associated (Cas) nuclease.

9. A method of predicting an impact of integration of a nucleic acid molecule of interest (NOI) into a human genome, said method comprising the method of claim 1, wherein said animal is porcine, and said method further comprising:

producing a porcine animal from said zygote, wherein the porcine animal is recombinant for the NOI;

determining any genotypic or phenotypic change in said porcine animal as a result of integration of the NOI; and

predicting the impact of integration of the NOI into a human genome based on said change in said porcine animal.

10. The method of claim 1, wherein at least 80% to 100% of said at least one zygote comprise said double strand break at the target locus.

Date Reçue/Date Received 2022-03-21

11. Use of a composition for targeted gene modification of an animal genome of an animal having pronuclei that cannot be visually observed by an unaided human eye, wherein the composition comprises:

i. a nuclease or a nucleic acid encoding a nuclease, wherein said nuclease is for introducing a double strand break at a target locus of said animal genome;

ii. at least one guide nucleic acid molecule for targeting said nuclease to said target locus;

iii. a single stranded oligonucleotide or double stranded nucleic acid molecule comprising a heterologous nucleic acid molecule of interest (NOI);

iv. a first sequence homologous to a first flanking sequence upstream of said target locus and a second sequence comprising a second flanking sequence downstream of said target locus,

wherein said first and second flanking sequences comprise at least about 40 to about 50 nucleotides when said NOI comprises a single stranded oligonucleotide, and wherein said first and second flanking sequences comprise at least about 300 to about 1000 nucleotides when said NOI comprises a double stranded nucleic acid molecule;

wherein the composition is for injecting into the cytoplasm of a zygote before a single nucleus forms, wherein said zygote comprises said genome of said animal and has been produced by fertilizing an oocyte of the animal with animal spermatozoa;

wherein said NOI is for recombining into said target locus of said animal genome; and

wherein said animal is a swine or a ruminant animal.

12. The use of claim 11, wherein the composition is for injecting into the cytoplasm of the zygote 12 to 16 hours after the fertilization of the oocyte.

13. The use of claim 11, wherein expression and/or function of said target locus is modified after said NOI is recombined into said target locus of said animal genome.

14. The use of claim 11, wherein said NOI is suitable for expression of a polypeptide.

Date Reçue/Date Received 2022-03-21

15. The use of claim 11, wherein at least one phenotype of an animal produced from said zygote is modified after said NOI is recombined into said animal genome.

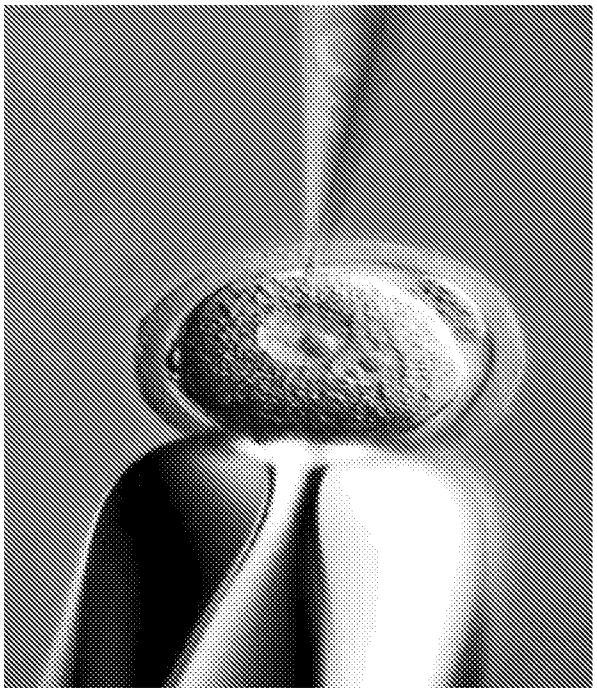
16. The use of claim 11, wherein said animal is a ruminant animal.

17. The use of claim 16, wherein said ruminant animal is a bovine.

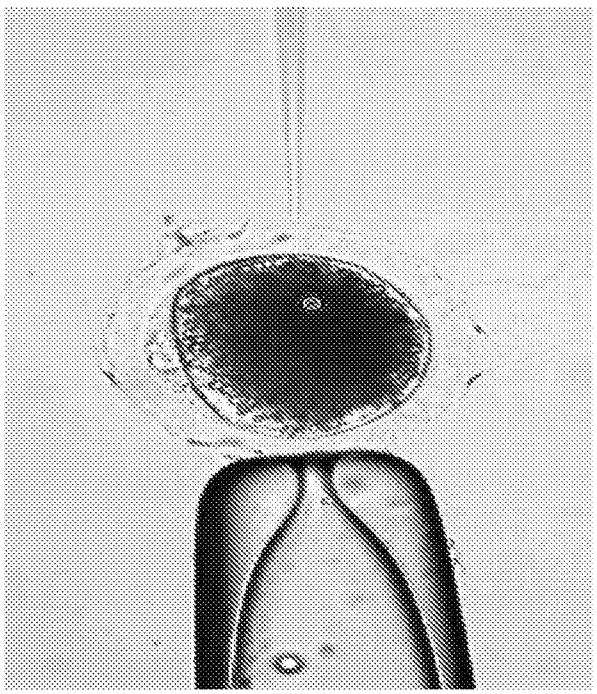
18. The use of claim 11, wherein said animal is a swine.

19. The use of claim 11, wherein said nuclease is a Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) - associated (Cas) enzyme.

Date Reçue/Date Received 2022-03-21



Mouse zygote



Pig zygote

FIG. 1

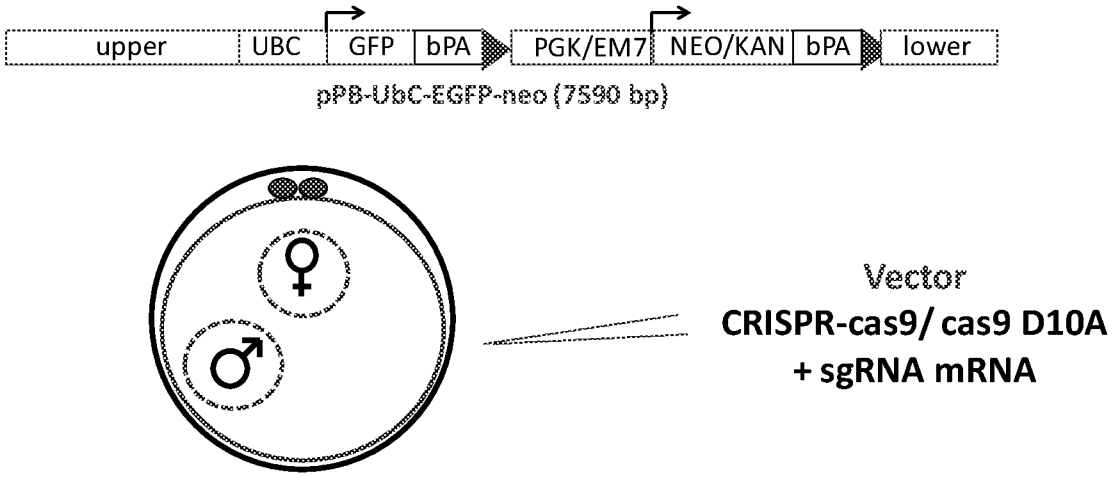

upper
UBC
GFP
bPA
PGK/EM7
NEO/KAN
bPA
lower
pPB-UbC-EGFP-neo (7590 bp)
Vector
CRISPR-cas9/ cas9 D10A
+ sgRNA mRNA

FIG. 2A

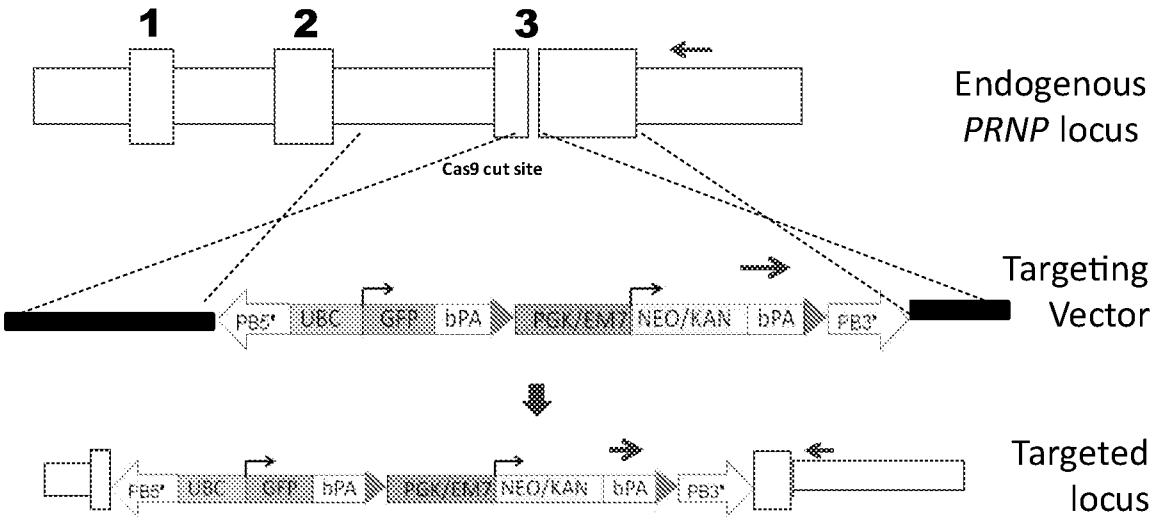

1
2
3
Cas9 cut site
Endogenous
PRNP locus
PB5'
UBC
GFP
bPA
PGK/EM7
NEO/KAN
bPA
PB3'
Targeting
Vector
Targeted
locus

FIG. 2B

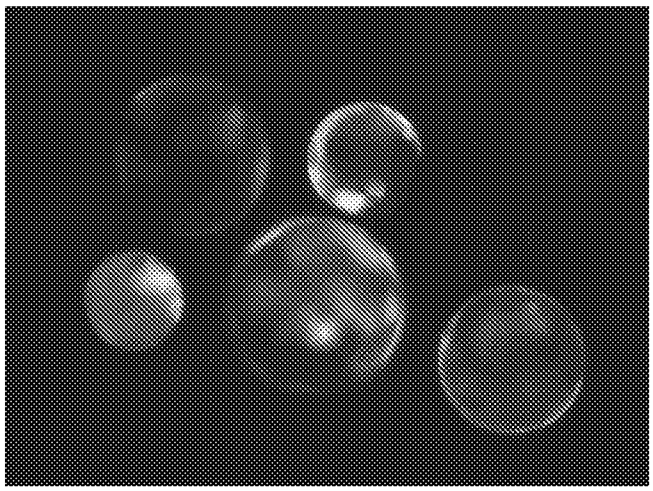


FIG. 3A

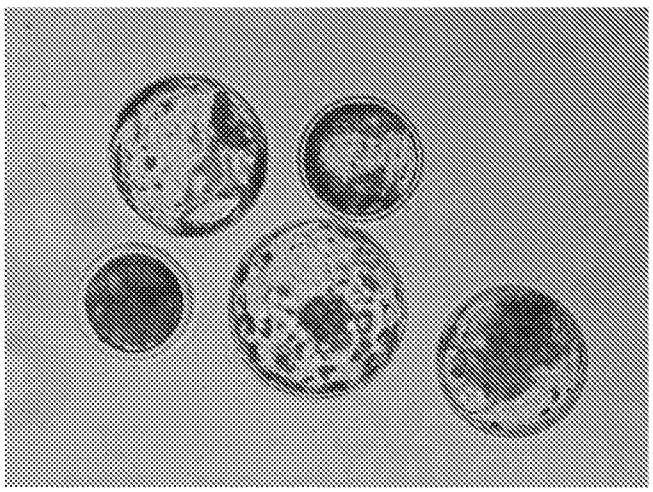


FIG. 3B

FIG. 4

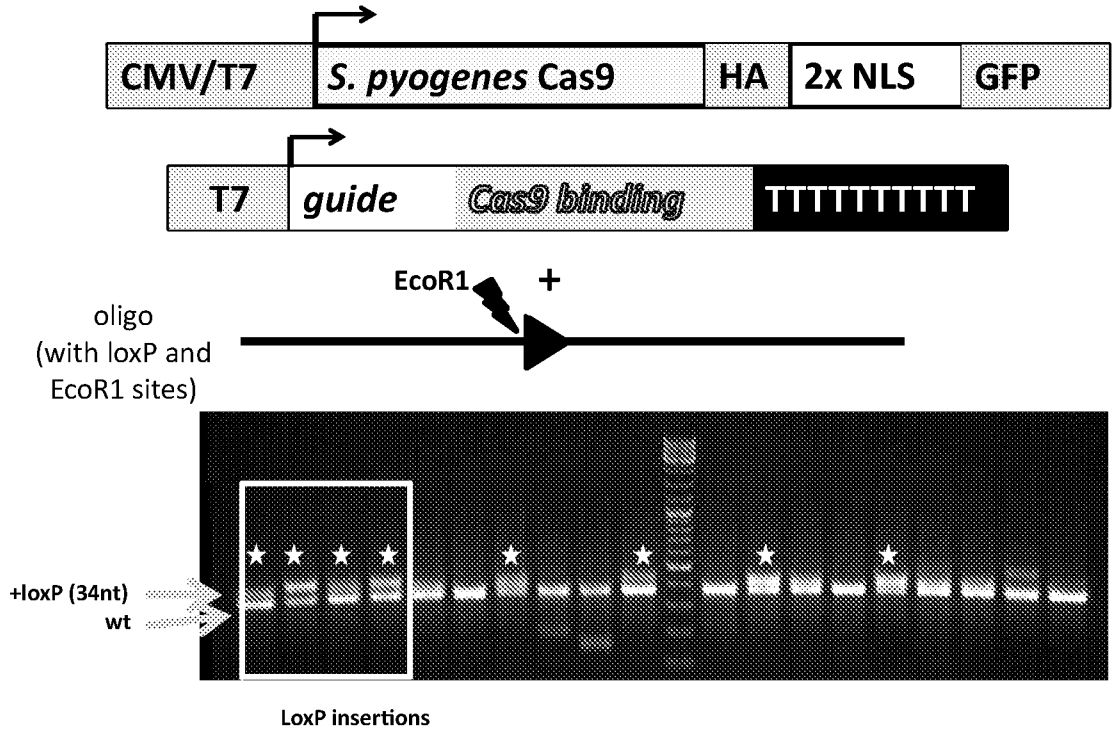

CMV/T7
S. pyogenes Cas9
HA
2x NLS
GFP
T7
guide
Cas9 binding
TTTTTTTTT
EcoR1
+
oligo
(with loxP and
EcoR1 sites)
+loxP (34nt)
wt
LoxP insertions

FIG. 5A

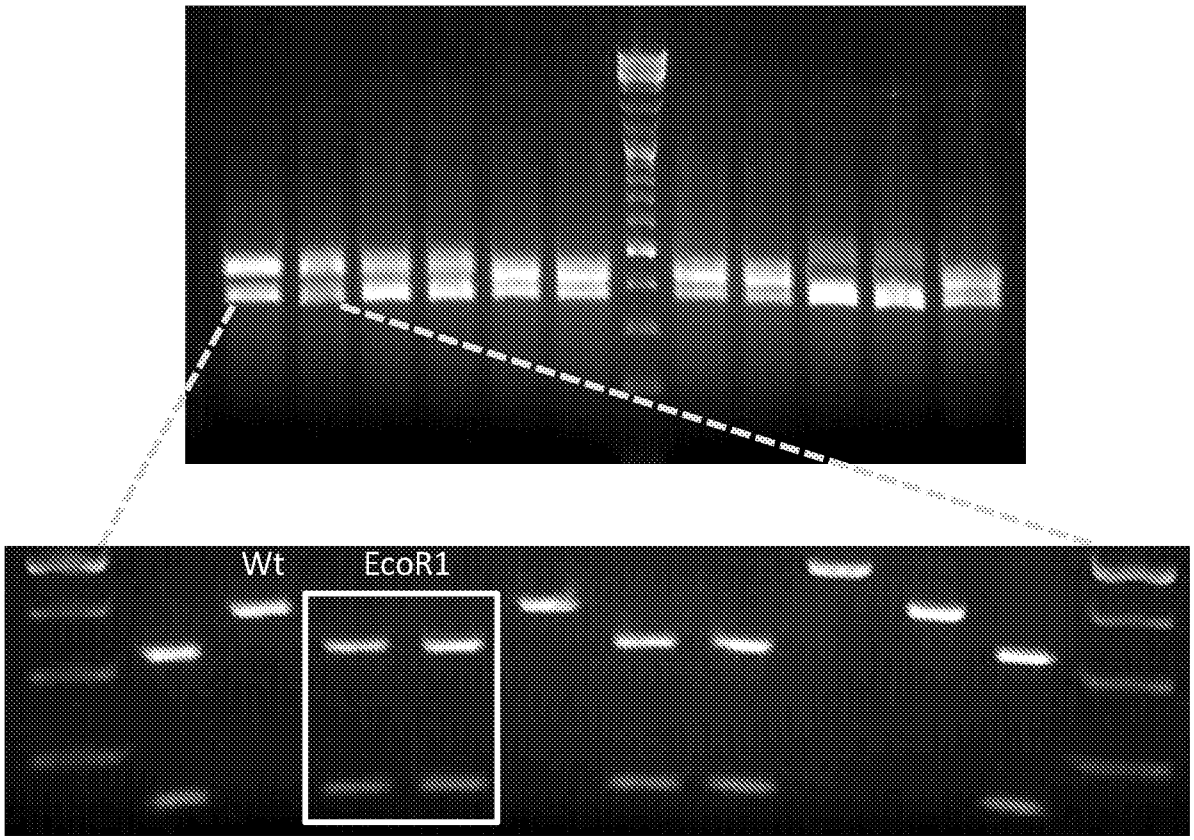

Wt
EcoR1

FIG. 5B

Crispr-Cas9&sgRNA+Loxp oligo (Zbed6)

Blastocyst 1

```
Zbed6 AGAAGGGTTTGCGAATTA----------------------------------AGGGGAAAAGGCGAC
1     AGAAGGGTTTGCGAATTCTACCGTTCGTATAGCATACATTATACGAAGTTATAGGGGAAAAGGCGAC
2     AGAAGGGTTTGCGAATTCTACCGTTCGTATAGCATACATTATACGAAGTTATAGGGGAAAAGGCGAC
3     AGAAGGGTTTGCGAATTCTACCGTTCGTATAGCATACATTATACGAAGTTATAGGGGAAAAGGCGAC
```

Blastocyst 2

```
Zbed6 AGAAGGGTTTGCGAATTA--------------------------------AGGGGAAAAGGCGAC
1     AGAAGGGTTTGCGAATTCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGGC
2     AGAAGGGTTTGCGAATTCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGGC
```

Blastocyst 3

```
Zbed6 AGAAGGGTTTGCGAATTA--------------------------------AAGGGGAAAAGGCGAC
1     AGAAGGGTTTGCGAATTCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGAC
2     AGAAGGGTTTGCGAATTCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGAC
3     AGAAGGGTTTGCGAATTCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGAC
```

Blastocyst 4

```
Zbed6 AGAAGGGTTTGCGAATTA--------------------------------AGGGGAAAAGGCGAC
1     AGAAGGGTTTGCGAATCCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGAC
2     AGAAGGGTTTGCGAATCCTACCGTTCGTATGCATACATTATACGAAGTTAAGGGGAAAAGGCGAC
```

Blastocyst 5

```
Zbed6 AGAAGGGTTTGCGAATTA----------------------------------AGGGGAAAAGGCGAC
1     AGAAGGGTTTGCGAATTCTACCGTTCGTATAGCATACATTATACGAAGTTATAGGGGAAAAGGCGAC
2     AGAAGGGTTTGCGAATTCTACCGTTCGTATAGCATACATTATACGAAGTTATAGGGGAAAAGGCGAC
```

FIG. 6

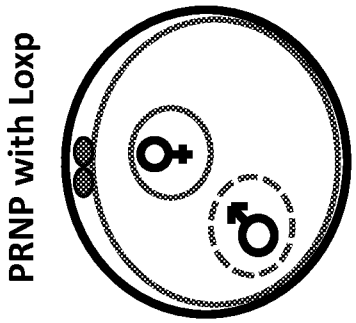
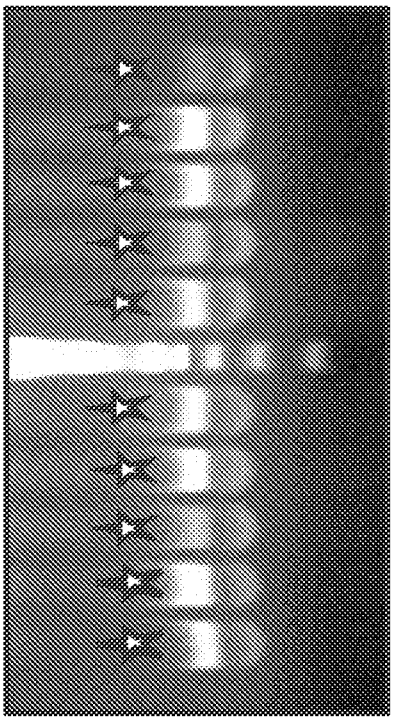
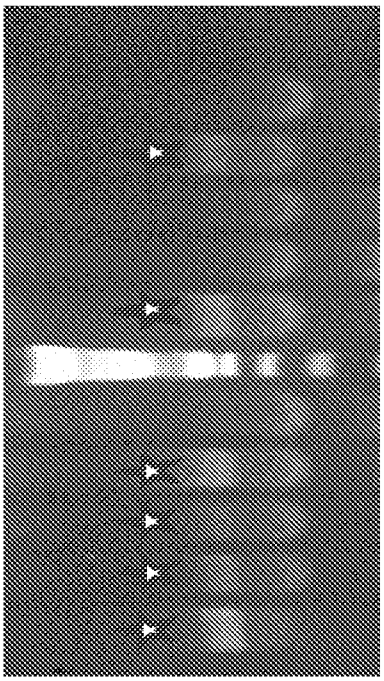


FIG. 7A



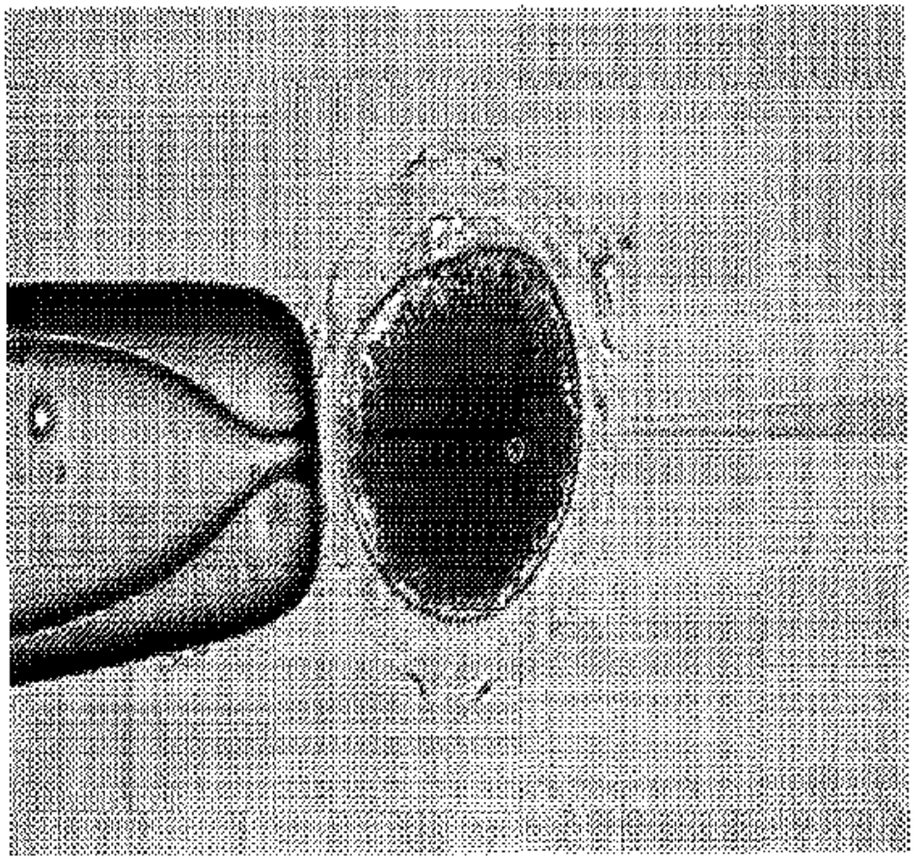
PRNP with loxp
(100% targeting)

FIG. 7C

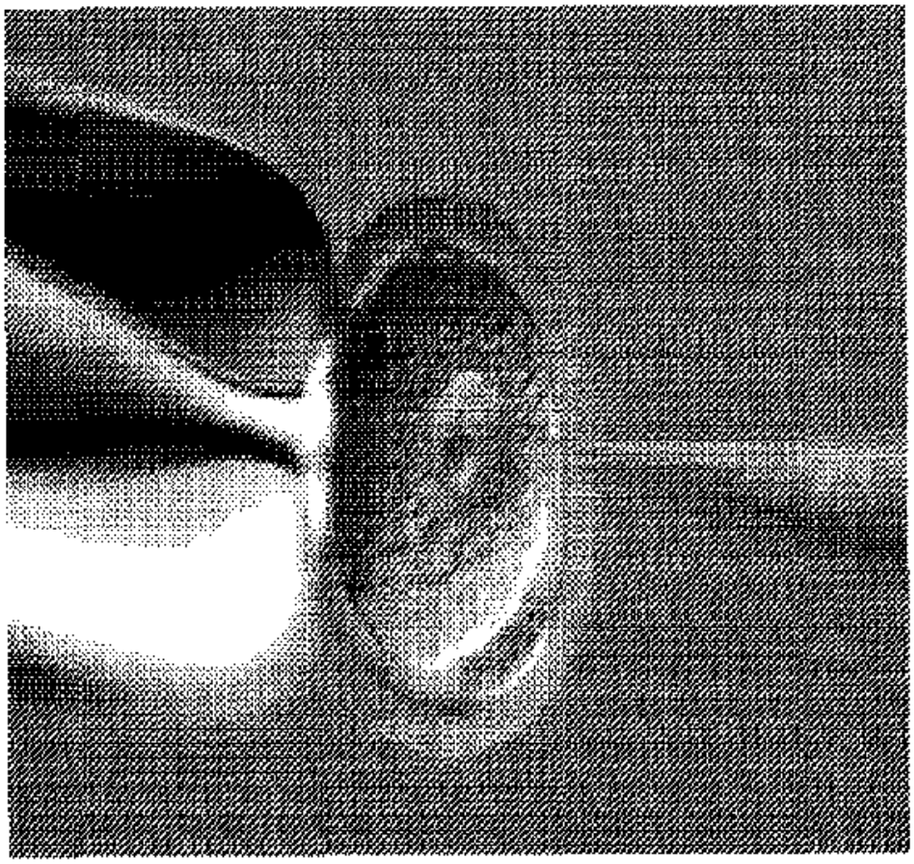


Zbed6 with loxp
(70% targeting)

FIG. 7B



Pig zygote



Mouse zygote