

FASCICULE DE BREVET D'INVENTION

21 Numéro de dépôt : 1201900469
PCT/IL2018/050573

22 Date de dépôt : 24/05/2018

30 Priorité(s) :
US n° 62/510,921 du 25/05/2017

24 Délivré le : 02/12/2020

45 Publié le : 31.12.2020

73 Titulaire(s) :
EGGXYT LTD,
Keren HaYessod 36,
9100760 JERUSALEM (IL)

72 Inventeur(s) :
OFFEN, Daniel (IL)

74 Mandataire : S.C.P AKKUM, AKKUM & Associates,
Quartier Mballa II, Dragages,
B.P. 4966, YAOUNDE (CM).

54 Titre : Methods for gender determination of avian embryos in unhatched eggs and means thereof.

57 Abrégé :

The present invention relates to methods of fertilization and gender determination and identification in avian subjects. More specifically, the invention provides non-invasive methods using transgenic avian animals that comprise at least one reporter gene, specifically, RFP, integrated into at least one gender chromosome Z or W. The transgenic avian animals of the invention are used for gender determination and selection of embryos in unhatched avian eggs.

METHODS FOR GENDER DETERMINATION OF AVIAN EMBRYOS IN UNHATCHED EGGS AND MEANS THEREOF

FIELD OF THE INVENTION

5 The present invention relates to methods of gender determination and identification in avian subjects. More specifically, the invention provides non-invasive methods and transgenic avian animals for gender determination and selection of embryos in unhatched avian eggs.

BACKGROUND ART

10 References considered to be relevant as background to the presently disclosed subject matter are listed below:

- WO 2010/103111
- WO 2014/0296707
- US 6244214
- 15 • 06124456A2
- US2014069336A
- WO16005539
- WO 96/39505
- WO 97/49806
- 20 • Quansah, E., Long, J.A., Donovan, D.M., Becker, S.C., Telugu, B., Foster Frey, J.A., Urwin, N. (2014). Sperm-mediated transgenesis in chicken using a PiggyBac transposon system. Poultry Science Association Meeting Abstract. BARC Poster Day.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial
- 25 immunity. *Science*, 337(6096), 816-821.
- Cong, L., & Zhang, F. (2015). Genome engineering using CRISPR-Cas9 system. *Chromosomal Mutagenesis*, 197-217.
- US Patent 6, 244,214.
- WO 2014/0296707.
- 30 • WO 06124456A2;
- WO16005539.

- Véron N., Qu Z., Kipen PA., Hirst CE., Marcelle C. (2015). CRISPR mediated somatic cell genome engineering in the chicken. *Dev. Biol.* 407(1):68-74. doi: 10.1016/j.ydbio.2015.08.007. Epub 2015 Aug 13.
- CA2264450.
- 5 • Niu, Y., B. Shen, Y. Cui, Y. Chen, J. Wang et al., (2014). Generation of genemodified cynomolgus monkey via cas9/rna-mediated gene targeting in one-cell embryos. *Cell*, 156(4): 836–843.
- Hwang, W. Y., Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai et al., (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.*
10 31(3): 227–229.
- Nadège, V., Q. Zhengdong, P. A. S. Kipen, C. E. Hirst, M. Christophe et al., (2015). CRISPR mediated somatic cell genome engineering in the chicken. *Dev. Biol.* 407(1): 68–74.
- Bai, Y., L. He, P. Li, K. Xu, S. Shao et al., (2016). Efficient genome editing in
15 chicken DF-1 cells using the CRISPR/Cas9 system. *G3 (Bethesda)* pii: g3.116.027706.
- Doran T. et al., (2016). Sex selection in layer chickens. *ASAP Animal Production*, Adelaide.
- Tizard M. and Doran T. (2014). Precision genome engineering in the chicken: Th gap between science and market place. A presentation at IWRAB-II, in Brasilia.

20

Acknowledgement of the above references herein is not to be inferred as meaning that these are in any way relevant to the patentability of the presently disclosed subject matter.

25 **BACKGROUND OF THE INVENTION**

In the food industry, chicks are culled by billions on a daily basis via suffocation or grinding. The males are terminated since they are not useful for laying eggs or to be bread for meat and the weak or unhealthy females are being terminated as well. A method for *in-ovo*, or embryo sex-determination prior to hatching is thus highly desired
30 due to both ethical and economic considerations.

Specifically, visually identifying poultry fertile eggs is important to allow removal of unfertile eggs to save hatching costs (by prevention of hatching an unfertile egg), and to lower the bio-security risks involved in the continuation of the incubation of these contamination-prone unfertile eggs alongside the fertile eggs.

5

Visually identifying egg fertility at an early stage of the embryo, while inside the unhatched egg, involves outer light source candling and can be difficult, virtually impossible in early embryonic stages. An even greater challenge is to identify the sex of embryos, and currently there is no available method for the discrimination between
10 males and females in unhatched eggs that are found fertile. However, identification of fertility at an early embryonic stage and the sex determination of poultry are vital for aviculture, scientific research, and conservation. The determination of sex in young birds by morphological features is extremely challenging for most species. The gender may be determined by individual vent sexing which involves manually squeezing the
15 feces out of the chick, which opens up the chicks' anal vent slightly, allowing to see if the chick has a small "bump", which would indicate that the chick is a male. However, this method represents high risk of bird injury and mistakes in sex determination, together with cumbersome work conducted manually by trained personal.

20 Vent sexing or chick sexing is the method of distinguishing the sex of chicken and other hatchlings, usually by a trained person called a chick sexer or chicken sexer. Chicken sexing is practiced mostly by large commercial hatcheries, who have to know the difference between the sexes in order to separate them into sex groups, and in order to take them into different programs, which can include the growing of the desired group
25 and culling of the undesired group which does not meet the commercial needs. In example, a male hatched from an egg that comes from an egg layer commercial line of breed. That male will not have a good meat yield and will not lay eggs, and thus will be culled after sexing. Upon sexing, the relevant sex will continue its course to serve the purpose while the other sex or most of it will be culled within days of hatching being
30 irrelevant to egg production.

In farms that produce eggs, males are unwanted, and chicks of an unwanted sex are killed almost immediately to reduce costs to the breeder. Chicks are moved down a

- 4 -

conveyer belt, where chick sexers separate out the males and toss them into a chute where they are usually ground up alive in a meat grinder.

5 Identification and determination of the fertility of an egg and the sex of the embryos in eggs prior to their hatching, will enable the elimination of unfertile eggs, and the unwanted type of embryos while in their eggs, and thus will immensely reduce incubation costs (which includes the energy and efficiency costs alongside with air pollution and energy consumption). In addition, chicks' suffering will cease and pollution from culling will be prevented. An automated sexing device will additionally
10 result in reduced eggs production costs by eliminating the need for chick sexers, as well as reduce the size of the hatcheries needed since at early stage 50% of the eggs will be reduced deducted from the process, thus reducing the costs of hatching these eggs, and later on the need for any elaborate killing procedures.

15 In all commercial types of birds intended for breeding, laying, or meat production, there is a need to determine fertility and the sex of the embryo. There are great economic returns; in energy saving, biosecurity risk reduction, garbage disposal, sexing labor costs and sexing errors, culling costs and disposal, and animal welfare.

20 WO 2010/103111 describes an invasive method comprising a series of steps, among them introducing into the egg a labeled antibody, specifically designed to match a sex-specific antigen on the embryo.

WO 2014/0296707 describes luminance composition designed to serve as a biomarker
25 for quantifying or evaluating efficiency of vaccination being injected into the bird's egg. No sex determination is described or even hinted in this disclosure. In-ovo injection apparatus and detection methods was disclosed by US Patent 6, 244,214.

WO 06124456A2 discloses invasive methods of *in-ovo* sex determining of an avian
30 embryo by determining the presence of an estrogenic steroid compound in a sample of embryonic fluid (e.g., allantoic fluid or blood) from the avian egg. Determining the presence of the compound is done by measuring analytes in samples obtained from said avian egg by competitive immunoassay utilizing fluorescence microscopy.

Spectroscopic approaches were also described, among them US2014069336A which is based on screening the avian embryo feather color (pre-hatching) and determining the sex of the avian embryo, based on the feather color or WO16005539 which disclose a device obtaining a shell-specific spectral response to an incident light signal.

Further genetic approaches addressing the *in ovo* sexing include DNA sequencing of DNA samples obtained from fertilized eggs for detecting two specific genes located on the Z and W chromosomes of birds (WO 96/39505), or the use of oligonucleotide probes which hybridize to specific sequences of the female W chromosome (WO 97/49806). These methods are invasive and therefore do not provide a safe strategy.

The clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system is the state of the art gene editing system, allowing a simple construct design with high success rate (M. Jinek, 2012).

Niu et al. (2014) injected guide RNA (gRNA) and Cas9 RNA into monkey oocytes to modify three target genes, and Hwang et al. (2013) modified the *drd3* and *gsk3b* genes in zebrafish embryos to obtain a two-locus mutant. Cong and Zhang (2015) have modified the CRISPR system to edit any gene in living cells.

Veron and coworkers (2015), demonstrated that expression levels of somatic cells in chicken embryos were modified by electroporation of CRISPR gRNA plasmids directed against the PAX7 transcription factor (Nadège et al. 2015), Bai and coworkers edited the PPAR-g, ATP synthase epsilon subunit (ATP5E).

Quansah, E. et. al., disclosed sperm mediated transgenesis in chicken using a PiggyBac transposon system. In particular, they disclose that aGFP plasmid and Lipofectamine LTXTM 9LPX) combination had no effect on viability, mobility or fertility of chicken sperm.

CA2264450 discloses pluripotent cells comprising a nucleic acid sequence encoding a fluorescent protein marker selectively integrated into a heterologous sex chromosome in

the cell embryos and transgenic animals produced using the pluripotent cells and, the uses of such cells, embryos, and animals. This publication particularly relate to GFP in mammalian cells.

5 Doran T. et al., generally describes a strategy to differentiate between males and females pre-hatch, by adding a biological marker, the GFP reporter gene, to the sex chromosome, in ASAP Animal Production 2016, Adelaide. The integration of the GFP reporter gene into the sex chromosome of chicks using the CRISPR technology was also described in a presentation at IWRAB-II, in Brasilia, in 2014. However, as this general
10 publication discloses that the labeled chromosome is visualized *in ovo* by exposing the egg to UV light, due to the auto-fluorescence of the egg, no detectable signal is expected.

Thus, effective and non-invasive methods for sex identification during the egg stage,
15 prior to the hatching of the chick are currently not available. There is therefore a long-felt need for a method enabling accurate and safe sex identification of the embryos in unhatched eggs.

SUMMARY OF THE INVENTION

20 A first aspect of the invention relates to a method of gender determination of avian, or avian embryo in an unhatched egg, specifically, a fertilized unhatched egg. In some specific embodiments, the method may comprise the step of:

First, in step (a), providing or obtaining at least one transgenic avian subject or animal comprising at least one exogenous reporter gene integrated into at least one position or
25 location (also referred to herein as locus) in at least one of gender chromosome Z and W therefore. In a second step (b), obtaining at least one fertilized egg from the transgenic avian subject or of any cells thereof.

The next step (c) involves determining in the egg if at least one detectable signal is detected. In more specific embodiments, detection of at least one detectable signal
30 indicates the expression of said at least one reporter gene, thereby the presence of the W chromosome or Z chromosome in the avian embryo. In some specific embodiments, the integrated reporter gene may encode the Red fluorescent protein (RFP).

In a second aspect, the invention relates to an avian transgenic animal comprising, in at least one cell thereof, at least one exogenous reporter gene integrated into at least one position or location in at least one of gender chromosome Z and W. In some specific embodiments, the integrated reporter gene may encode the RFP.

5

In yet another aspect, the invention relates to a cell comprising at least one exogenous reporter gene integrated into at least one position or locus in at least one of gender chromosome Z and W.

10 In yet a further aspect, the invention provides a kit comprising:

(a) at least one Cas9 protein or any fragments or derivatives thereof, or at least one first nucleic acid sequence comprising at least one nucleic acid sequence encoding said at least one Cas9 protein; and at least one nucleic acid sequence encoding at least one guide RNA (gRNA) that targets at least one protospacer located within the gender
15 chromosome Z or W; and (b) at least one second nucleic acid sequence comprising at least one said reporter gene.

Still further, the invention provides a method for determining and detecting fertilization in an unhatched egg. More specifically, such method comprises comprise the step of:

20
25
30

First, in step (a), providing or obtaining at least one transgenic avian subject or animal comprising at least one exogenous reporter gene integrated into at least one position or location in both gender chromosome Z and W, in case of a female subject and into both gender chromosomes Z in a male. In a second step (b) obtaining at least one fertilized egg from the transgenic avian subject, or of any cells thereof. The next step (c) involves determining in the egg if at least one detectable signal is detected. In more specific embodiments, detection of at least one detectable signal indicates the expression of said at least one reporter gene, thereby the presence of the labeled maternal W chromosome or Z chromosome (in case of a female or the labeled paternal Z chromosome in the avian embryo.

These and further embodiments of the invention will become apparent by the hand of the following figures.

BRIEF DESCRIPTION OF THE DRAWINGS

In order to better understand the subject matter that is disclosed herein and to exemplify how it may be carried out in practice, embodiments will now be described, by way of non-limiting example only, with reference to the accompanying drawings, in which:

5

Figure 1A-1B. Complete set up for assessing fluorescence of eggs

Fig. 1A. shows a photograph of the device used by the methods of the invention for *in ovo* gender determination. The complete setup is composed of a laser source (h), laser source holder (g), a stand for the egg (e), the egg (f), a lens (d), a filter (c), a stand for the detector (b) and a detector (a). The different components are placed on a solid support (i).

Fig. 1B. shows schematic illustration of the device of the invention shown in Fig. 1A. The device comprises a laser source (h), laser source holder (g), a stand for the egg (e), the egg (f), a lens (d), a filter (c), a stand for the detector (b) and a detector (a). The different components are placed on a solid support (i).

15

Figure 2. Fluorescence intensity using a Blue laser (473 nm) with or without fluorescein

The fluorescence intensity [Pout] that was received on the detector when a complete egg was subjected to blue laser with or without 10 μ M or 1mM fluorescein (fl), and the intensity is represented as a function of the light source intensity [Pin].

20

Figure 3. Fluorescence intensity using a Green laser (532 nm) with or without Rhodamin

The fluorescence intensity [Pout] that was received on the detector when a complete egg was subjected to green laser with or without Rhodamin, and the intensity is represented as a function of the light source intensity [Pin].

25

Figure 4. Fluorescence intensity using a Green laser (532 nm) with or without dir

The fluorescence intensity [Pout] that was received on the detector, when a complete egg was subjected to green laser with or without dir, and the intensity is represented as a function of the light source intensity [Pin].

30

Figure 5. Fluorescence intensity using a Red laser (632.8 nm) with or without dir

The fluorescence intensity [Pout] that was received on the detector, when a complete egg was subjected to red laser with or without dir, and the intensity is represented as a function of the light source intensity [Pin].

5

Figure 6. Cells expressing GFP

Picture providing the fluorescent signal of HEK cells transfected with GFP-vector expressing GFP.

10 **Figure 7. Cells expressing RFP**

Picture providing the fluorescent signal HEK cells transfected with RFP-vector expressed RFP.

Figure 8. Control measurements

15 Fluorescence intensity using a Green laser and a red filter with different positions of the egg. The fluorescence intensity [Pout] that was received on the detector is represented as a function of the light source intensity [Pin].

Figure 9. RFP Measurements

20 Measurements of fluorescence intensity of eggs injected with different concentrations of RFP-expressing cells using a Green laser and the red filter. The fluorescence intensity [Pout] that was received on the detector is represented as a function of the light source intensity [Pin].

25 **Figure 10. RFP angle**

RFP fluorescence intensity measurements of excited eggs at the determined optimal position of the egg using the Green laser and the red filter. The fluorescence intensity [Pout] that was received on the detector is represented as a function of the light source intensity [Pin].

30

Figure 11A-11B. RFP-expressing cells with PBS or Glycerol

Figure shows fluorescence intensity of Eggs injected with different concentration of RFP-expressing cells with PBS or Glycerol using the Green laser.

- 10 -

Fig. 11A. shows the fluorescence intensity of Eggs containing different concentrations of RFP-expressing cells together in PBS. **Fig. 11B** shows the fluorescence intensity of Eggs containing different concentrations of RFP-expressing cells in Glycerol. The fluorescence intensity [Pout] that was received on the detector is represented as a function of the light source intensity [Pin].

Figure 12. GFP-expressing cells with PBS or Glycerol

Figure shows fluorescence intensity of Eggs injected with 30,000 GFP-expressing cells with PBS or Glycerol using the Blue laser. The fluorescence intensity [Pout] that was received on the detector is represented as a function of the light source intensity [Pin].

Figure 13. RFP vs. GFP

Figure shows ratio between the Fluorescence Protein intensity either from GFP-expressing cells or RFP-expressing cells and the auto fluorescence intensity. Parameter R describing the ratio between Fluorescence and auto fluorescence is represented as a function of the light source intensity [Pin].

Figure 14. Incorporation of RFP into female Z chromosome

Figure shows RFP transfected female chicken cell line with pDsRed with chicken ChZ Left & Right arms and CMV-hspCas9-H1-gRNA.

Figure 15A-15B. Schematic representation of the integrated sequence into chromosome Z of female chicken cell line

Bold sequences represent the flanking left and right arms of the Z chromosome and are denoted by SEQ ID NO: 60 and SEQ ID NO: 61 respectively; The sequence (1017bp) marked with Italics represents Left arm, as denoted by SEQ ID NO: 62; Underlined sequence (629bp) represents the CMV Enhancer, Promotor and MCS, as denoted by SEQ ID NO: 63; Bold Italics sequence (714bp) represents the dsRED2 reporter gene, as denoted by SEQ ID NO: 64; Sequence with no color-3332bp represents the dsRED plasmid content; Underlined Italics sequence (1026bp) represents the Right arm, as denoted by SEQ ID NO: 65. The left part of the sequence is presented by SEQ ID NO. 66 (all parts 5' to vector sequences, shown in Fig. 15A), and the right part of the

- 11 -

sequence is denoted by SEQ ID NO. 67 (all parts 3' to vector sequences, shown in Fig. 15B).

DETAILED DESCRIPTION OF THE INVENTION

5 Each day, billions of male chicks are being terminated via suffocation or grinding since they are not useful for laying eggs or to be bread for meat. The ability to determine the sex of the embryo before hatching is of high importance both ethically and financially. In the chicken- the genetic make-up of the sex chromosomes is ZZ for males and ZW for females. Meaning the W chromosome determines the gender of the female. This is
10 unlike humans, in which it is the Y from the father that determines the male gender.

The invention provides a non-invasive efficient method for gender determination, using a reporter gene integrated in a gender specific chromosomes of transgenic avian subjects. Expression of this reporter gene in an embryo of an unhatched egg clearly and
15 accurately identify the gender of said embryo.

Thus, a first aspect of the invention relates to a method of gender determination and optionally of selection of avian, or avian embryo in an unhatched egg, specifically, a fertilized unhatched egg. In some specific embodiments, the method may comprise the
20 step of:

First, in step (a), providing or obtaining at least one transgenic avian subject or animal comprising at least one exogenous reporter gene integrated into at least one position or location in at least one of gender chromosome Z and W. In a second step (b) obtaining at least one fertilized egg from the transgenic avian subject, specifically animal or of
25 any cells thereof.

The next step (c), involves determining in the egg if at least one detectable signal is detected. In more specific embodiments, detection of at least one detectable signal indicates the expression of the at least one reporter gene, thereby the presence of the W chromosome or Z chromosome in the avian embryo. Thus, in case the reporter gene has
30 been integrated into the Z chromosome of a female transgenic avian, identification of a detectable signal in the examined egg indicate that the embryo comprised therein has a maternal Z chromosome having a reporter gene integrated therein, and the embryo is thereby identified as male. In yet some other embodiments, in case the reporter gene has

been integrated into the W chromosome of a female transgenic avian, identification of a detectable signal in the examined egg indicate that the embryo comprised therein carries a maternal W chromosome and is therefore determined as female, thereby providing gender determination of the embryo.

5

It should be appreciated that the transgenic avian provided by the invention may be either a female or a male, as described in more detail herein after. In yet some further specific embodiments, the transgenic animal may be the gender having two different gender chromosomes, specifically, the heterogametic subject. In some particular
10 embodiments the heterogametic animal may be a female. In more specific embodiments, where the transgenic avian subject is a female, the egg identified by the method of the invention is laid by the transgenic female avian provided by the invention. In more specific embodiment, the transgenic female may be fertilized either by a transgenic male or in some other embodiments, by a wild type avian male. Still
15 further, fertilization may occur either by mating or by insemination of the transgenic avian female with sperms obtained from a transgenic or wild type avian male. In yet some other embodiments, where the transgenic avian is a male, egg identified by the method of the invention may be laid by either a wild type or transgenic female mated with the transgenic male provided by the invention, or inseminated by any cells thereof,
20 specifically sperm cells that comprise the exogenous reporter gene of the invention integrated into the gender chromosomes thereof.

The invention thus provides a method for detecting a gender of an avian embryo within an unhatched fertilized egg. It should be appreciated that the method of the invention
25 may be applicable for unhatched eggs of any embryonic stage of an avian embryo.

It should be noted that "**Embryonic development stage or step of avian embryo**", as used herein refers to the stage of **day 1** wherein the germinal disc is at the blastodermal stage and the segmentation cavity takes on the shape of a dark ring; the stage of **day 2** wherein the first groove appears at the center of the blastoderm and the vitelline
30 membrane appears; the stage of **day 3** wherein blood circulation starts, the head and trunk can be discerned, as well as the brain and the cardiac structures which begins to beat; the stage of **day 4** wherein the amniotic cavity is developing to surround the embryo and the allantoic vesicle appears; the stage of **day 5** wherein the embryo takes a

C shape and limbs are extending; the stage of **day 6** wherein fingers of the upper and lower limbs becomes distinct; the stage of **day 7** wherein the neck clearly separates the head from the body, the beak is formed and the brain progressively enters the cephalic region; the stage of **day 8** wherein eye pigmentation is readily visible, the wings and legs are differentiated and the external auditory canal is opening; the stage of **day 9** wherein claws appears and the first feather follicles are budding; the stage of **day 10** wherein the nostrils are present, eyelids grow and the egg-tooth appears; the stage of **day 11** wherein the palpebral aperture has an elliptic shape and the embryo has the aspect of a chick; the stage of **day 12** wherein feather follicles surround the external auditory meatus and cover the upper eyelid whereas the lower eyelid covers major part of the cornea; the stage of **day 13** wherein the allantois becomes the chorioallantoic membrane while claws and leg scales becomes apparent; the stage of **days 14 to 16** wherein the whole body grows rapidly, vitellus shrinking accelerates and the egg white progressively disappears; the stage of **day 17** wherein the renal system produces urates, the beak points to the air cell and the egg white is fully resorbed; the stage of **day 18** wherein the vitellus internalized and the amount of amniotic fluid is reduced; the stage of **day 19** wherein vitellus resorption accelerates and the beak is ready to pierce the inner shell membrane; the stage of **day 20** wherein the vitellus is fully resorbed, the umbilicus is closed, the chick pierces the inner shell membrane, breathes in the air cell and is ready to hatch; the stage of **day 21** wherein the chick pierces the shell in a circular way by means of its egg-tooth, extricates itself from the shell in 12 to 18 hours and lets its down dry off.

More specifically, the method of the invention may be applicable in determining the gender of an avian embryo *in-ovo*, inside the egg, at every stage of the embryonic developmental process. More specifically, from day 1, from day 2, from day 3, from day 4, from day 5, from day 6, from day 7, from day 8, from day 9, from day 10, from day 11, from day 12, from day 13, from day 14, from day 15, from day 16, from day 17, from day 18, from day 19, from day 20 and from day 21. More specifically, the method of the invention may be applicable for early detection of the embryo's gender, specifically, from day 1 to day 10. In some embodiments the method of the invention may enable determining the embryo's gender at day one of the embryonic development. In some further embodiments, the method of the invention may enable determining the

embryo's gender at day two. In some further embodiments, the method of the invention may enable determining the embryo's gender at day three. In further embodiments, the method of the invention may enable determining the embryo's gender at day four. In yet some further embodiments, the method of the invention may enable determining the embryo's gender at day five. In further embodiments, the method of the invention may enable determining the embryo's gender at day six. In further embodiments, the method of the invention may enable determining the embryo's gender at day seven. further embodiments, the method of the invention may enable determining the embryo's gender at day eight. In some further embodiments, the method of the invention may enable determining the embryo's gender at day nine. In yet some further embodiments, the method of the invention may enable determining the embryo's gender at day ten. In some specific the method of the invention may enable determining the embryo's gender between days 1 to 5 of the embryonic development.

As noted above, the method of the invention may be applicable for fertilized unhatched eggs. The term "**fertilized egg**" refers hereinafter to an egg laid by a hen wherein the hen has been mated by a rooster within two weeks, allowing deposit of male sperm into the female infundibulum and fertilization event to occur upon release of the ovum from the ovary. "**Unhatched egg**" as used herein, relates to an egg containing and embryo (also referred to herein as a fertile egg) within a structurally integral (not broken) shell.

The method of the invention is based on determination of a detectable signal formed by a reporter gene integrated into specific loci of the transgenic avian female or male laying the examined egg.

The "**Integration of foreign or exogenous DNA/gene into chromosome**" as used herein, refers hereinafter to a permanent modification of the nucleotide sequence of an organism chromosome. This modification is further transferred during cell division and if occurring in germinal cell lines, it will be transmitted also to offspring. In this case, the integrated reporter gene may be transferred to the embryo within the unhatched egg. The term "**exogenous**" as used herein, refers to originating from outside an organism that has been introduced into an organism for example by transformation or transfection with specifically manipulated vectors, viruses or any other vehicle. The integrated

exogenous gene according to certain embodiments, may be a reporter gene. The term "reporter gene" relates to gene which encodes a polypeptide, whose expression can be detected in a variety of known assays and wherein the level of the detected signal indicates the presence of said reported.

5

As noted above, the exogenous reporter gene may be integrated into the avian gender chromosomes Z or W. The avian "gender chromosome Z or W" as used herein refers to the chromosomal system that determines the sex of offspring in chicken wherein males are the homogametic sex (ZZ), while females are the heterogametic sex (ZW).
10 The presence of the W chromosome in the ovum determines the sex of the offspring while the Z chromosome is known to be larger and to possess more genes.

The method of the invention is based on the detection of a detectable signal that indicates and reflects the presence of the reporter gene and thereby the presence of a
15 specific gender chromosome. "Detectable signal" refers hereinafter to a change in that is perceptible either by observation or instrumentally. Without limitations, the signal can be detected directly. In some embodiments, detectable response is an optical signal.

It should be appreciated that in some specific embodiments, at least one transgenic
20 avian subject provided by the method of the invention, may comprise at least two different reporter genes, each reporter gene may be integrated into at least one position or location in one of gender chromosome Z or W. In case of at least two different reporter genes, in some embodiments, each of the gender chromosomes may be labeled differently. The evaluation of the detectable signal formed, may indicate the gender of
25 the examined embryo.

In yet some specific embodiments, the reporter gene comprised within the transgenic
avian of the invention may be at least one fluorescent reporter gene. Thus, in some
embodiments, the expressed polypeptide is a fluorescent protein and accordingly the
30 assay measures the levels of light emitted upon excitation, specifically, with the appropriate light source.

The term "**fluorescence**" refers to the emission of light by a substance that has been illuminated, absorbed light or other electromagnetic radiation. It is a form of luminescence i.e. an emission of light by a substance not resulting from heat.

A "**fluorescent protein**" refers to a protein that exhibits bright fluorescence when exposed to light. Fluorescent proteins have particular wavelength for the peak of the illumination excitation intensity and wavelength for the peak of fluorescence emission intensity. Excitation refers to the illumination of a fluorescent protein. In some embodiment, the excitation may be provided by a laser. In yet some embodiments, the signal may be detected upon the usage of an appropriate optical filter.

10 In certain embodiments, the fluorescent reporter gene may be red Fluorescent Protein (RFP), cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and auto-fluorescent proteins including blue fluorescent protein (BFP).

It should be noted that in some further specific embodiments, any fluorescent protein may be applicable in the present invention. More specifically, mutagenesis efforts in the original *Aequorea victoria* jellyfish green fluorescent protein have resulted in new fluorescent probes that range in color from blue to yellow, that may be applicable according to some embodiments, in the present invention. Longer wavelength fluorescent proteins, emitting in the orange and red spectral regions, have been developed from the marine anemone, *Discosoma striata*, and reef corals belonging to the class **Anthozoa**. Still other species have been mined to produce similar proteins having cyan, green, yellow, orange, and deep red fluorescence emission.

In some embodiments, the reporter gene that may be applicable in the methods, transgenic avian subjects, constructs, cells and kits of the invention may be the yellow fluorescent proteins. More specifically, the family of yellow fluorescent proteins was initiated after the crystal structure of green fluorescent protein revealed that threonine residue 203 (**Thr203**) was near the chromophore. Mutation of this residue to tyrosine was introduced to stabilize the excited state dipole moment of the chromophore and resulted in a 20-nanometer shift to longer wavelengths for both the excitation and emission spectra. Further refinements led to the development of the enhanced yellow fluorescent protein (**EYFP**), which is one of the brightest and most widely used fluorescent proteins.

The **Citrine** variant of yellow fluorescent protein is very bright relative to EYFP and may be also applicable in some embodiments of the invention. Another derivative,

named **Venus**, is the fastest maturing and one of the brightest yellow variants developed to date. The coral reef protein, **ZsYellow1**, originally cloned from a *Zoanthus* species native to the Indian and Pacific oceans, produces true yellow emission and is also applicable in the present invention.

5 In yet some further embodiments, the reporter gene that may be applicable in the methods, transgenic avian subjects, constructs, cells and kits of the invention may be the blue fluorescent proteins. The blue and cyan variants of green fluorescent protein resulted from direct modification of the tyrosine residue at position 66 (**Tyr66**) in the native fluorophore. Conversion of this amino acid to histidine results in blue emission
10 having a wavelength maxima at 450 nanometers, whereas conversion to tryptamine results in a major fluorescence peak around 480 nanometers along with a shoulder that peaks around 500 nanometers. More specifically, among the improved cyan fluorescent proteins that have been introduced, **AmCyan1** and an enhanced cyan variant termed **Cerulean** show the most promise. Derived from the reef coral, *Anemonia majano*, the
15 AmCyan1 fluorescent protein variant may be also applicable for the present invention.

In some specific embodiments, the reporter gene that may be applicable in the methods, transgenic avian subjects, constructs, cells and kits of the invention may be the red fluorescent proteins. The first coral-derived fluorescent protein to be extensively utilized was derived from *Discosoma striata* and is commonly referred to as **DsRed**.
20 Once fully matured, the fluorescence emission spectrum of DsRed features a peak at 583 nanometers whereas the excitation spectrum has a major peak at 558 nanometers and a minor peak around 500 nanometers. Still further, other variants that may be applicable in the present invention include but are not limited to **DsRed2**, **DsRed-Express** and **RedStar**. Several additional red fluorescent proteins showing a
25 considerable amount of promise have been isolated from the reef coral organisms. One of these proteins that may be applicable in the present invention, may be the **HcRed1**, which was isolated from *Heteractis crispa* and is now commercially available. HcRed1 was originally derived from a non-fluorescent chromoprotein that absorbs red light through mutagenesis to produce a weakly fluorescent obligate dimer having an
30 absorption maximum at 588 nanometers and an emission maximum of 618 nanometers. Still further, in some embodiments, the reporter gene that may be applicable in the methods, transgenic avian subjects, constructs, cells and kits of the invention may be any fluorescent protein, provided that such protein is not the GFP, and in some more

specific embodiments, not the Wild type GFP. However, in some embodiments, several mutants and variants of GFP that form the red-emitting spectral species with excitation and emission maxima at 555 and 585nm, may be also applicable in the present invention. In yet some further specific embodiments, such mutants or variants may comprise Serine 65 and/or Asparagine 68 residues. In yet some further embodiments, such mutants or variants of GFP may comprise mutation of at least one of F46L, V163A and I167V. In yet some further embodiments the variant may comprise any combination of Ser65, Asn68, F46L, V163A and I167V.

It should be understood that in yet some further embodiments, the reporter gene that may be applicable in the methods, transgenic avian subjects, constructs, cells and kits of the invention may be any of the fluorescent proteins disclosed in **Table 1** herein after. In yet some further embodiments the reporter gene that may be applicable in the methods, transgenic avian subjects, constructs, cells and kits of the invention may be any of the fluorescent proteins disclosed in **Table 1**, provided that said fluorescent protein is not the wild type GFP. In yet some other embodiments, the reporter gene that may be applicable in the methods, transgenic avian subjects, constructs, cells and kits of the invention may be any of the fluorescent proteins disclosed in **Table 1**, provided that such reporter gene is not any of the **Green Fluorescent Proteins**, disclosed by Table 1, more specifically, any one of EGFP, Emerald, Superfolder GFP, Azami Green, mWasabi, TagGFP, TurboGFP, AcGFP, ZsGreen, T-Sapphire.

Table 1: *fluorescent proteins:*

Protein	Excitation	Emission
(Acronym)	Maximum	Maximum
	(nm)	(nm)
GFP (wt)	395/475	509
Green Fluorescent Proteins		
EGFP	484	507
Emerald	487	509
Superfolder GFP	485	510
Azami Green	492	505
mWasabi	493	509
TagGFP	482	505

TurboGFP	482	502
AcGFP	480	505
ZsGreen	493	505
T-Sapphire	399	511
Blue Fluorescent Proteins		
EBFP	383	445
EBFP2	383	448
Azurite	384	450
mTagBFP	399	456
Cyan Fluorescent Proteins		
ECFP	439	476
mECFP	433	475
Cerulean	433	475
mTurquoise	434	474
CyPet	435	477
AmCyan1	458	489
Midori-Ishi Cyan	472	495
TagCFP	458	480
mTFP1 (Teal)	462	492
Yellow Fluorescent Proteins		
EYFP	514	527
Topaz	514	527
Venus	515	528
mCitrine	516	529
YPet	517	530
TagYFP	508	524
PhiYFP	525	537
ZsYellow1	529	539
mBanana	540	553
Orange Fluorescent Proteins		
Kusabira Orange	548	559
Kusabira Orange2	551	565
mOrange	548	562
mOrange2	549	565
dTomato	554	581

- 20 -

dTomato-Tandem	554	581
TagRFP	555	584
TagRFP-T	555	584
DsRed	558	583
DsRed2	563	582
DsRed-Express (T1)	555	584
DsRed-Monomer	556	586
mTangerine	568	585
Red Fluorescent Proteins		
mRuby	558	605
mApple	568	592
mStrawberry	574	596
AsRed2	576	592
mRFP1	584	607
JRed	584	610
mCherry	587	610
HcRed1	588	618
mRaspberry	598	625
dKeima-Tandem	440	620
HcRed-Tandem	590	637
mPlum	590	649
AQ143	595	655

As surprisingly shown by Examples 1 and 2, the auto fluorescent properties of the egg revealed by the present invention allow only the detection of embryos that carry cells expressing RFP.

5

Thus, in more specific embodiments, the reporter gene may be Red Fluorescent Protein (RFP). The term "**Red Fluorescent Protein**" or "RFP" refers hereinafter to a fluorescent protein that emits orange, red, and far-red fluorescence that has been isolated from anthozoans (corals) and anemones as well as any variants thereof. There are two main types of RFP proteins, DsRed and Kaede. DsRed-like RFPs are derived from *Discosoma striata* and include but are not limited to mCherry, zFP538, mKO,

10

mOrange, mRouge, E2-Crimson, mNeptune, TagRFP657, Keima, mKate, mStrawberry, mBanana, mHoneydew, niTangerine, mRaspberry, mPlum, mRFPmars and mRFPruby.

On the other end, Kaede is a natural fluorescent protein found in the stony coral *Trachyphyllia geoffroyi*, which irreversibly changes its emission wavelength from green
5 (518 nm) to red (582 nm) upon irradiation at ~400 nm. Kaede family members comprise for example EosFP, dendFP, mcavRFP, and rfloRFP, found respectively in the corals *Lobophyllia hemprichii*, *Dendronephthya*, *Monastrea cavernosa*, and *Ricordea florida*.

It should be appreciated that any of the RFPs described herein, of any source known in
10 the art, may be applicable for the methods, transgenic animals, cells, kits and devices of the invention.

In yet some specific embodiments, the RFP that may be used by the methods of the invention may be the RFP used in the pDsRed1-N1, that encodes a novel red fluorescent protein [RFP; Matz, M. V., et al. (1999) *Nature Biotech.* 17:969–973] that has been
15 optimized for high expression in mammalian cells (excitation maximum = 558 nm; emission maximum = 583 nm). RFP was isolated from an IndoPacific sea anemone-relative, *Discosoma* sp; DsRed1's coding sequence contains 144 silent base pair changes, which correspond to human codon-usage preferences for high expression in
20 mammalian cells.

In some embodiments, the RFP used by the methods and kits of the invention may be encoded by a nucleic acid sequence comprising the sequence as denoted by SEQ ID NO. 20. In yet some further embodiments, such RFP may comprise the amino acid sequence as denoted by SEQ ID NO. 21, or any homologs, variants, mutants or
25 derivatives thereof.

In yet some alternative specific embodiments, the RFP used by the invention may be the RFP encoded by the nucleic acid sequence as disclosed by GenBank: AF272711.1, having the amino acid sequence as disclosed by GenBank: AAG16224.1. In yet some further specific embodiments, the RFP used by the methods and kits of the invention
30 may be encoded by a nucleic acid sequence comprising the sequence as denoted by SEQ ID NO. 22. In yet some further embodiments, such RFP may comprise the amino acid sequence as denoted by SEQ ID NO. 23, or any homologs, mutants or derivatives thereof.

It should be noted that in some embodiments, the reporter gene used by the method of the invention may be any fluorescent protein, provided that said fluorescent protein is not green fluorescent protein (GFP). In other words, in some specific embodiments the reporter gene integrated into the gender chromosome of the transgenic animal provided
5 by the method of the invention may be any fluorescent protein with the proviso that said fluorescent protein is not GFP.

Still further, in some specific and particular embodiments, when the reporter gene is
10 RFP, it should be noted that in some embodiments, the excitation wavelengths of RFPs are about 500-650 nm while the emission wavelengths may be about 550-650 nm.

In some other embodiments, detecting a detectable signal by the method of the invention may further comprise the step of subjecting said egg to an appropriate light
15 source. In yet some further embodiments, exposure of the egg to an appropriate light source comprising wavelength of between about 400 to about 650. In yet some further embodiments, light of any wavelength may be used with the proviso that said light source is not Ultra-Violet (UV) light (wavelength 10-400 nm).

In some specific embodiment, the step of subjecting said egg to a light source is
20 excitation of the fluorescent protein. In some particular embodiments, the excitation wavelength is between about 500 nm and about 650 nm. In some further embodiments, the excitation wavelength is about 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640 nm.

In some more specific embodiment, the excitation wavelength may range between about
25 515 to about 555. In yet some further embodiments, the wavelength may be 532 nm.

In some specific and non-limiting embodiment, the light source may be provided by a laser.

As used herein, the term "**laser**" refers to electromagnetic radiation of any frequency
30 that is amplified by stimulated emission of radiation. A laser also refers to a device that emits electromagnetic radiation through a process called stimulated emission. Laser light is usually spatially coherent, which means that the light either is emitted in a

narrow, low-divergence beam, or can be converted into one with the help of optical components such as lenses.

In some other embodiments, the laser may be a blue laser, or a red laser. In more specific embodiments, the light source is a green laser.

In certain embodiments, the laser may be employed with a filter.

As used herein, the term "**filter**" refers to "optical filter" which corresponds to a device that selectively transmits light of different wavelengths, usually implemented as plane glass or plastic devices in the optical path which are either dyed in the bulk or have interference coatings.

In some embodiments, the filter may be a green filter that transmits above 500 nm, or a dark green filter that transmits between 540 nm and 580 nm, or a red filter that transmits between 590 nm and 650 nm, or a red filter that transmits above 650 nm or above 660 nm.

In some particular embodiments, the light source may be a green laser with a wavelength of 532 nm. In some further embodiments, such laser may be provided with a red filter that transmits above 650 nm.

The chicken germ-line develops from a small population of primordial germ cells, migrating to the genital ridge from an extragonadal site, while undergoing phases of active migration, as well as passive circulation in the bloodstream. PGCs are located in the center of the epiblast of freshly laid, un-incubated egg, a developmental stage referred as stage X. During incubation PGCs migrate anteriorly and accumulate at the germinal crescent of stage 10HH embryo (approximately 33-38 hours of incubation), considered the main site for intravasation. Later, PGCs can be found in the circulation from stage 12HH to 17HH (approximately from 45 to 64 hours of incubation), reaching a peak concentration in stage 14HH (approximately 50-53 hours of incubation). PGCs leave the circulation at a site adjacent to the gonad anlage at the intermediate mesoderm, where they can be found as early as stage 15HH (55-55 hours of incubation). PGCs reach the genital ridge by active migration along the dorsal mesentery, and colonize both gonads, where they later differentiate into spermatogonia or oogonia.

Embryo is positioned on top of the yolk sac. The yolk is freely rotating in the egg, and the embryo always be facing the upper side of the yolk. During incubation, the blunt end of the egg (where the air sac is located) is facing up, and eggs are occasionally
5 rotated in 90°. As the embryo develops, formation of extra-embryonic tissues retracts the embryo away from the eggshell, and the embryo is less accessible.

Thus, in some specific embodiments, the egg is exposed to said light source. In yet some further embodiments, the egg is placed in a position facilitating exposure of the
10 embryo at any stage to the light source. In yet some further embodiments, a region containing the upper face of the egg yolk at stage X of the egg is excited with the light source.

In some further embodiments, said step of subjecting said egg to a light source is provided by a system, apparatus or a device that may comprise a laser source, a stand
15 for the egg, a lens, a filter, a stand for a detector and a detector.

As used herein, the term "**detector**" refers to any type of device that detects and/or measures light. In some embodiments, it should be noted that the detectable signal, specifically, the fluorescent signal may be detected using suitable fluorescent means. In
20 some embodiments, the detectable signal formed by the RFP reporter gene may be detected by light sensitive apparatus such as modified optical microscopes or Charge Coupled Device (CCD), a highly sensitive photon detector.

In yet some further embodiments, the methods of the invention may be performed using
25 any device, apparatus or system that provides at least one light source, at least one detector, at least one filter and at least one holding arm that places the egg in an appropriate position facilitating the exposure of cells expressing the reporter gene of the invention to said light source. In some particular and non-limiting embodiments, the method of the invention may use a device as illustrated in Figure 1. In yet some further
30 embodiments, the distance between the egg holder and lens and/or lens and detector may range between about 1 to 100 cm, specifically, 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 and 100 cm, specifically between about 5 to 25 cm, more specifically, about 20 cm.

It should be therefore appreciated that in some embodiments, the methods of the invention may further comprise the step of providing a system or device, for example, as described herein, that is adapted for determining the existence of a detectable signal in an examined egg. Specifically, an apparatus or device comprising a laser source, a stand for the egg, a lens, a filter, a stand for a detector and a detector. As indicated
5 above, in some embodiments for such device, a green laser source that display an excitation wavelength of about 532nm may be used. In yet some further embodiments, a filter, specifically a red filter, that transmits above 650nm, may be used.

In still further embodiments, the at least one transgenic avian subject or animal provided
10 by the method of the invention may be a female avian subject or animal. In more specific embodiments, where the at least one reporter gene is integrated into at least one position of female chromosome Z, detection of a detectable signal indicates that the embryo in the unhatched egg is a male. More specifically, an embryo containing the paternal unlabeled Z gender chromosome and further contains the maternal labeled Z
15 chromosome, contains two Z chromosomes and therefore is a male.

In yet some further embodiments, at least one transgenic avian subject or animal provided by the methods of the invention may be a female avian subject or animal. In some specific embodiments, where the at least one reporter gene is integrated into at least one position of female chromosome W, detection of a detectable signal, indicates
20 that the embryo in the unhatched egg is female. More specifically, the embryo contains a parental unlabeled z chromosome and a labeled maternal W chromosome is an heterogametic embryo containing the WZ chromosomes that is a female.

In yet some further embodiments, the transgenic animal provided by the methods of the invention may be a male subject having the reporter gene integrated into the Z
25 chromosomes thereof. In such case, a detectable signal determined in an egg fertilized by such transgenic male or any sperms thereof, indicates that the embryo carries a paternal Z chromosome comprising the transgenic reporter gene, and is therefore male. In still further embodiments, detection of a detectable signal in an egg laid by a transgenic female avian fertilized by a transgenic male avian, both carrying the reporter
30 gene of the invention integrated into the Z chromosomes thereof, may indicate in case of an intense signal that the embryo carries two copies of a reporter gene integrated into the female and male Z chromosomes thereof. In case of a less intense signal, for

- 26 -

example, a signal with reduced intensity of about 50%. The egg may be determined as a female.

As indicated herein before, the method of the invention involves the provision of transgenic avian animals. The preparation of transgenic avian animals, requires the use
5 of genetic engineering approach that may use specific gene editing compound or component. Non-limiting examples for gene editing components may comprise in some embodiments, the use of nucleases.

In some specific embodiments, the nucleases applicable in the methods of the invention may be RNA guided DNA binding protein nucleases.

10 As used herein, an **RNA guided DNA binding protein nuclease** is a nuclease which is guided to its cleavage site (or alternatively, a site for any other alternative activity), by an RNA molecule. This RNA molecule is referred as a guide RNA.

Thus, in yet more specific embodiments, the at least one reporter gene may be
15 integrated into the gender chromosome of the transgenic avian subject or animal provided by the method of the invention using at least one programmable engineered nuclease (PEN). The term "**programmable engineered nucleases (PEN)**" as used herein, refers to synthetic enzymes that cut specific DNA sequences, derived from natural occurring nucleases involved in DNA repair of double strand DNA lesions and
20 enabling direct genome editing.

In some more specific embodiments, PEN used by the methods of the invention may be any one of a clustered regularly interspaced short palindromic repeat (CRISPR) Class 2 or Class 1 system.

More specifically, the Clustered Regularly Interspaced Short Palindromic Repeats
25 (CRISPR) Type II system is a bacterial immune system that has been modified for genome engineering.

It should be appreciated however that other genome engineering approaches, like zinc finger nucleases (ZFNs) or transcription-activator-like effector nucleases (TALENs) that relay upon the use of customizable DNA-binding protein nucleases that require
30 design and generation of specific nuclease-pair for every genomic target may be also applicable herein.

CRISPR-Cas systems fall into two classes. Class 1 systems use a complex of multiple Cas proteins to degrade foreign nucleic acids. Class 2 systems use a single large Cas

protein for the same purpose. More specifically, Class 1 may be divided into types I, III, and IV and class 2 may be divided into types II, V, and VI.

As used herein, CRISPR arrays also known as SPIDRs (Spacer Interspersed Direct Repeats) constitute a family of recently described DNA loci that are usually specific to a particular bacterial species. The CRISPR array is a distinct class of interspersed short sequence repeats (SSRs) that were first recognized in *E. coli*. In subsequent years, similar CRISPR arrays were found in *Mycobacterium tuberculosis*, *Haloferax mediterranei*, *Methanocaldococcus jannaschii*, *Thermotoga maritima* and other bacteria and archaea. It should be understood that the invention contemplates the use of any of the known CRISPR systems, particularly and of the CRISPR systems disclosed herein. The CRISPR-Cas system has evolved in prokaryotes to protect against phage attack and undesired plasmid replication by targeting foreign DNA or RNA. The CRISPR-Cas system, targets DNA molecules based on short homologous DNA sequences, called spacers that exist between repeats. These spacers guide CRISPR-associated (Cas) proteins to matching (and/or complementary) sequences within the foreign DNA, called proto-spacers, which are subsequently cleaved. The spacers can be rationally designed to target any DNA sequence. Moreover, this recognition element may be designed separately to recognize and target any desired target. With respect to CRISPR systems, as will be recognized by those skilled in the art, the structure of a naturally occurring CRISPR locus includes a number of short repeating sequences generally referred to as "repeats". The repeats occur in clusters and are usually regularly spaced by unique intervening sequences referred to as "spacers." Typically, CRISPR repeats vary from about 24 to 47 base pair (bp) in length and are partially palindromic. The spacers are located between two repeats and typically each spacer has unique sequences that are from about 20 or less to 72 or more bp in length. Thus, in certain embodiments the CRISPR spacers used in the sequence encoding at least one gRNA of the methods and kits of the invention may comprise between 10 to 75 nucleotides (nt) each. More specifically, about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75 or more. In some specific embodiments the spacers comprise about 20 to 25 nucleotides, more specifically, about 20 nucleobases.

In addition to at least one repeat and at least one spacer, a CRISPR locus also includes a leader sequence and optionally, a sequence encoding at least one tracrRNA. The leader sequence typically is an AT-rich sequence of up to 550 bp directly adjoining the 5' end of the first repeat.

- 5 In some specific embodiment, the PEN used by the methods of the invention may be a CRISPR Class 2 system. In yet some further particular embodiments, such class 2 system may be a CRISPR type II system.

More specifically, three major types of CRISPR–Cas system are delineated: Type I, Type II and Type III.

- 10 **The type II CRISPR–Cas systems** include the 'HNH'-type system (Streptococcus-like; also known as the Nmeni subtype, for *Neisseria meningitidis serogroup A str. Z2491*, or CASS4), in which Cas9, a single, very large protein, seems to be sufficient for generating crRNA and cleaving the target DNA, in addition to the ubiquitous Cas1 and Cas2. Cas9 contains at least two nuclease domains, a RuvC-like nuclease domain near
15 the amino terminus and the HNH (or McrA-like) nuclease domain in the middle of the protein, but the function of these domains remains to be elucidated. However, as the HNH nuclease domain is abundant in restriction enzymes and possesses endonuclease activity responsible for target cleavage.

- Type II systems cleave the pre-crRNA through an unusual mechanism that involves
20 duplex formation between a tracrRNA and part of the repeat in the pre-crRNA; the first cleavage in the pre-crRNA processing pathway subsequently occurs in this repeat region. Still further, it should be noted that type II system comprise at least one of *cas9*, *cas1*, *cas2* *csn2*, and *cas4* genes. It should be appreciated that any type II CRISPR-Cas systems may be applicable in the present invention, specifically, any one of type II-A or
25 B.

Thus, in yet some further and alternative embodiments, the at least one *cas* gene used in the methods and kits of the invention may be at least one *cas* gene of type II CRISPR system (either typeII-A or typeII-B). In more particular embodiments, at least one *cas* gene of type II CRISPR system used by the methods and kits of the invention may be the *cas9* gene. It should be appreciated that such system may further comprise at least one of *cas1*, *cas2*, *csn2* and *cas4* genes.

Double-stranded DNA (dsDNA) cleavage by Cas9 is a hallmark of "*type II CRISPR-Cas*" immune systems. The CRISPR-associated protein Cas9 is an RNA-guided DNA

- 29 -

endonuclease that uses RNA:DNA complementarity to identify target sites for sequence-specific double stranded DNA (dsDNA) cleavage, creating the double strand breaks (DSBs) required for the HDR that results in the integration of the reporter gene into the specific target sequence, for example, a specific target within the avian gender chromosomes W and Z. The targeted DNA sequences are specified by the CRISPR array, which is a series of about 30 to 40 bp spacers separated by short palindromic repeats. The array is transcribed as a pre-crRNA and is processed into shorter crRNAs that associate with the Cas protein complex to target complementary DNA sequences known as proto-spacers. These proto-spacer targets must also have an additional neighboring sequence known as a proto-spacer adjacent motif (PAM) that is required for target recognition. After binding, a Cas protein complex serves as a DNA endonuclease to cut both strands at the target and subsequent DNA degradation occurs via exonuclease activity.

CRISPR type II system as used herein requires the inclusion of two essential components: a “guide” RNA (gRNA) and a non-specific CRISPR-associated endonuclease (Cas9). The gRNA is a short synthetic RNA composed of a “scaffold” sequence necessary for Cas9-binding and about 20 nucleotide long “spacer” or “targeting” sequence which defines the genomic target to be modified. Thus, one can change the genomic target of Cas9 by simply changing the targeting sequence present in the gRNA. Guide RNA (gRNA), as used herein refers to a synthetic fusion of the endogenous bacterial crRNA and tracrRNA, providing both targeting specificity and scaffolding/binding ability for Cas9 nuclease. Also referred to as “single guide RNA” or “sgRNA”. CRISPR was originally employed to “knock-out” target genes in various cell types and organisms, but modifications to the Cas9 enzyme have extended the application of CRISPR to “knock-in” target genes, selectively activate or repress target genes, purify specific regions of DNA, and even image DNA in live cells using fluorescence microscopy. Furthermore, the ease of generating gRNAs makes CRISPR one of the most scalable genome editing technologies and has been recently utilized for genome-wide screens.

The target within the genome to be edited, specifically, the specific target loci within the gender chromosomes Z or W, where the reporter gene of the invention is to be

- 30 -

integrated, should be present immediately upstream of a Protospacer Adjacent Motif (PAM).

The PAM sequence is absolutely necessary for target binding and the exact sequence is dependent upon the species of Cas9 (5' NGG 3' for *Streptococcus pyogenes* Cas9). In certain embodiments, Cas9 from *S. pyogenes* is used in the methods and kits of the invention. Nevertheless, it should be appreciated that any known Cas9 may be applicable. Non-limiting examples for Cas9 useful in the present disclosure include but are not limited to *Streptococcus pyogenes* (SP), also indicated herein as SpCas9, *Staphylococcus aureus* (SA), also indicated herein as SaCas9, *Neisseria meningitidis* (NM), also indicated herein as NmCas9, *Streptococcus thermophilus* (ST), also indicated herein as StCas9 and *Treponema denticola* (TD), also indicated herein as TdCas9. In some specific embodiments, the Cas9 of *Streptococcus pyogenes* M1 GAS, specifically, the Cas9 of protein id: AAK33936.1, may be applicable in the methods and kits of the invention. In some embodiments, the Cas9 protein may be encoded by the nucleic acid sequence as denoted by SEQ ID NO. 24. In further specific and non-limiting embodiments, the Cas9 protein may comprise the amino acid sequence as denoted by SEQ ID NO. 25, or any derivatives, mutants or variants thereof. Once expressed, the Cas9 protein and the gRNA, form a riboprotein complex through interactions between the gRNA "scaffold" domain and surface-exposed positively-charged grooves on Cas9. Cas9 undergoes a conformational change upon gRNA binding that shifts the molecule from an inactive, non-DNA binding conformation, into an active DNA-binding conformation. Importantly, the "spacer" sequence of the gRNA remains free to interact with target DNA. The Cas9-gRNA complex binds any genomic sequence with a PAM, but the extent to which the gRNA spacer matches the target DNA determines whether Cas9 will cut. Once the Cas9-gRNA complex binds a putative DNA target, a "seed" sequence at the 3' end of the gRNA targeting sequence begins to anneal to the target DNA. If the seed and target DNA sequences match, the gRNA continues to anneal to the target DNA in a 3' to 5' direction.

Cas9 will only cleave the target if sufficient homology exists between the gRNA spacer and target sequences. Still further, the Cas9 nuclease has two functional endonuclease domains: RuvC and HNH. Cas9 undergoes a second conformational change upon target binding that positions the nuclease domains to cleave opposite strands of the target

DNA. The end result of Cas9-mediated DNA cleavage is a double strand break (DSB) within the target DNA that occurs about 3 to 4 nucleotides upstream of the PAM sequence.

The resulting DSB may be then repaired by one of two general repair pathways, the efficient but error-prone Non-Homologous End Joining (NHEJ) pathway and the less efficient but high-fidelity Homology Directed Repair (HDR) pathway. In some embodiments, the insertion that results in the specific integration of the reporter gene of the invention to the specific target loci within the gender chromosomes W or Z, is a result of repair of DSBs caused by Cas9. In some specific embodiments, the reporter gene of the invention is integrated, or knocked-in the target loci by HDR.

The term "**Homology directed repair (HDR)**", as used herein refers to a mechanism in cells to repair double strand DNA lesions. The most common form of HDR is homologous recombination. The HDR repair mechanism can only be used by the cell when there is a homologue piece of DNA present in the nucleus, mostly in G2 and S phase of the cell cycle. When the homologue DNA piece is absent, another process called non-homologous end joining (NHEJ) can take place instead. Programmable engineered nucleases (PEN) strategies for genome editing, are based on cell activation of the HDR mechanism following specific double stranded DNA cleavage.

As discussed previously, Cas9 generates double strand breaks (DSBs) through the combined activity of two nuclease domains, RuvC and HNH. The exact amino acid residues within each nuclease domain that are critical for endonuclease activity are known (D10A for HNH and H840A for RuvC in *S. pyogenes* Cas9) and modified versions of the Cas9 enzyme containing only one active catalytic domain (called "Cas9 nickase") have been generated. Cas9 nickases still bind DNA based on gRNA specificity, but nickases are only capable of cutting one of the DNA strands, resulting in a "nick", or single strand break, instead of a DSB. DNA nicks are rapidly repaired by HDR (homology directed repair) using the intact complementary DNA strand as the template. Thus, two nickases targeting opposite strands are required to generate a DSB within the target DNA (often referred to as a "double nick" or "dual nickase" CRISPR system). This requirement dramatically increases target specificity, since it is unlikely that two off-target nicks will be generated within close enough proximity to cause a

- 32 -

DSB. It should be therefore understood, that the invention further encompasses the use of the dual nickase approach to create a double nick-induced DSB for increasing specificity and reducing off-target effects.

5 Thus, in certain embodiments, the at least one reporter gene may be integrated into the gender chromosome of the transgenic avian subject, specifically animal by homology directed repair (HDR) mediated by at least one CRISPR/CRISPR-associated endonuclease 9 (Cas9) system.

10 As indicated above, the CRISPR type II system comprises at least two elements, the nuclease, specifically, the Cas9 and the guide RNA. Therefore, for incorporating and integrating the reporter gene used by the invention, into the gender chromosomes of the transgenic avian subject, in addition to Cas9 or any nucleic acid sequence encoding Cas9, the invention further provides guide RNA or any nucleic acid sequence encoding
15 such gRNA that targets the Cas9 to the target site within a specific gender chromosome. In some further embodiments, the gRNA of the kit of the invention may comprise at least one CRISPR RNA (crRNA) and at least one trans-activating crRNA (tracrRNA).

In some alternative embodiments the kit of the invention may comprise nucleic acid
20 sequence encoding the at least one gRNA. Such nucleic acid sequence may comprise a CRISPR array comprising at least one spacer sequence that targets and is therefore identical to at least one protospacer in a target genomic DNA sequence. It should be noted that the nucleic acid sequence further comprises a sequence encoding at least one tracrRNA.

25

Still further, in some embodiments, the at least one reporter gene may be integrated into a gender chromosome of the transgenic avian subject, specifically animal by contacting or co-transfecting at least one cell of the avian subject, or animal or at least one cell introduced into the avian subject, or animal, with: (a) elements of the CRISPR Cas9
30 system, specifically, at least one Cas9 protein or at least one first nucleic acid sequence comprising at least one nucleic acid sequence encoding said at least one Cas9 protein; and at least one guide RNA (gRNA) that targets a protospacer within the gender chromosome Z or W, or at least one nucleic acid sequence encoding said at least one

- 33 -

gRNA; and (b) at least one second nucleic acid sequence comprising at least one reporter gene. It should be noted that in some embodiments, in case the elements of the CRISPR Cas9 system of (a) are provided in the form of nucleic acid sequence, they may be provided either in one nucleic acid molecule that comprises the different elements
5 (referred to herein as the "first nucleic acid sequence", or in two or more nucleic acid molecules, each comprises one of the elements indicated above, specifically, sequences encoding the gRNA and Cas9.

Thus, for the preparation of a transgenic avian animal used by the methods of the
10 invention, at least two nucleic acid molecules should be provided.

Still further, a "gRNA" or "targeting RNA" is an RNA that, when transcribed from the portion of the CRISPR system encoding it, comprises at least one segment of RNA sequence that is identical to (with the exception of replacing T for U in the case of RNA) or complementary to (and thus "targets") a DNA sequence in the target genomic
15 DNA, referred to herein as a protospacer. The CRISPR systems of the present disclosure may optionally encode more than one targeting RNA, and the targeting RNAs be directed to one or more target sequences in the genomic DNA. A "**proto-spacer**", as used herein, refers to the target sequence within the target chromosome. Such proto-spacers comprise nucleic acid sequence having sufficient complementarity
20 to a targeting RNA encoded by the CRISPR spacers comprised within the nucleic acid sequence encoding the gRNA of the methods and kits of the invention.

The methods of the invention, as well as the kits described herein after, provide nucleic acid sequences. As used herein, "**nucleic acids or nucleic acid molecules**" is
25 interchangeable with the term "**polynucleotide(s)**" and it generally refers to any polyribonucleotide or poly-deoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA or any combination thereof. "Nucleic acids" include, without limitation, single- and double-stranded nucleic acids. As used herein, the term "nucleic acid(s)" also includes DNAs or RNAs as described above that contain one or
30 more modified bases. As used herein, the term "**oligonucleotide**" is defined as a molecule comprised of two or more deoxyribonucleotides and/or ribonucleotides, and preferably more than three. Its exact size will depend upon many factors which in turn, depend upon the ultimate function and use of the oligonucleotide. The oligonucleotides

may be from about 8 to about 1,000 nucleotides long. More specifically, the oligonucleotide molecule/s used by the kit of the invention may comprise any one of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 5 900, 1000 or more bases in length.

Nucleic acid molecules can be composed of monomers that are naturally-occurring nucleotides (such as DNA and RNA), or analogs of naturally-occurring nucleotides (e.g., alpha-enantiomeric forms of naturally-occurring nucleotides), or modified 10 nucleotides or any combination thereof. Herein this term also encompasses a cDNA, i.e. complementary or copy DNA produced from an RNA template by the action of reverse transcriptase (RNA-dependent DNA polymerase).

In this connection an "**isolated polynucleotide**" is a nucleic acid molecule that is 15 separated from the genome of an organism. For example, a DNA molecule that encodes the reporter gene used by the methods and kits of the invention or any derivatives or homologs thereof, as well as the sequences encoding the CRISPR/Cas9 and gRNAs of the methods and kits of the invention, that has been separated from the genomic DNA of a cell is an isolated DNA molecule. Another example of an isolated nucleic acid 20 molecule is a chemically-synthesized nucleic acid molecule that is not integrated in the genome of an organism. A nucleic acid molecule that has been isolated from a particular species is smaller than the complete DNA molecule of a chromosome from that species. In some embodiments, the nucleic acid sequences used by the methods and kits (described herein after) of the invention, specifically, nucleic acid sequences comprising 25 sequences encoding the Cas9 and gRNA, or alternatively the reporter gene of the invention, may be provided constructed within a vector, cassette or any other vehicle. The invention thus further relates to recombinant DNA constructs comprising the polynucleotides of the invention, and optionally, further additional elements such as promoters, regulatory and control elements, translation, expression and other signals, 30 operably linked to the nucleic acid sequence of the invention. A non-limiting example for a vector provided by the invention may be any vector that comprises the nucleic acid sequences that encode the Cas9 and specific gRNA required for integration of the reporter gene into the target gender chromosome. In some specific embodiments, the

invention provides DNA constructs or vectors that comprise the nucleic acid sequences that encode Cas9 as denoted by SEQ ID NO. 25, and the gRNA7 as denoted by SEQ ID NO. 26, that targets the reporter gene to the gender chromosome Z. In yet some further embodiments, the invention further provides a vector or any DNA construct that
5 comprises the nucleic acid sequence that encodes the reporter gene of the invention, specifically, the RFP, flanked by left and right arms required for recombination and integration into the target locus within the gender chromosome, specifically, the gender chromosome Z. A non-limiting example for such vector may be a vector that comprises the nucleic acid sequences as disclosed in Figure 15. In yet some further embodiments,
10 the nucleic acid sequence that comprise the nucleic acid sequence as denoted by at least one of SEQ ID NO. 66 and/or SEQ ID NO. 67.

Still further, the invention further encompasses any host cell that comprises and/or express the vectors and DNA constructs of the invention, specifically, any of the vectors described by the invention.

15

As used herein, the terms "**recombinant DNA**", "**recombinant nucleic acid sequence**" or "**recombinant gene**" refer to a nucleic acid comprising an open reading frame encoding one of the CRISPR system of the invention, specifically, the CRISPR/Cas9 type II, along with the gRNA of the invention that target the Cas9 to the corresponding
20 protospacer in a specific locus within avian chromosomes Z and/or W. In yet some other embodiments, recombinant DNA as used herein further refers to a nucleic acid comprising an open reading frame encoding the reporter gene of the invention, specifically, transgene.

25 As referred to herein, by the term "**gene**" or "**transgene**" is meant a nucleic acid, either naturally occurring or synthetic, which encodes a protein product. The term "nucleic acid" is intended to mean natural and/or synthetic linear, circular and sequential arrays of nucleotides and nucleosides, e.g., cDNA, genomic DNA (gDNA), mRNA, and RNA, oligonucleotides, oligonucleosides, and derivatives thereof.

30 The phrase "**operatively-linked**" is intended to mean attached in a manner which allows for transgene transcription. As noted above, the transgene used by the method of the invention is prepared by providing nucleic acid sequence encoding the transgene that is to be incorporated into the gender chromosomes Z or W, specifically, the reporter gene,

and a gene editing system (e.g., the CRISP/Cas9 system). The term "**encoding**" is intended to mean that the subject nucleic acid may be transcribed and translated into either the desired polypeptide or the subject protein in an appropriate expression system, e.g., when the subject nucleic acid is linked to appropriate control sequences such as promoter and enhancer elements in a suitable vector (e.g., an expression vector) and when the vector is introduced into an appropriate system or cell.

It should be appreciated that in some embodiments, at least one of the first and the second nucleic acid sequences provided and used by the methods and kits of the invention may be constructed and comprised within a vector. "**Vectors**" or "**Vehicles**", as used herein, encompass vectors such as plasmids, phagemides, viruses, integratable DNA fragments, and other vehicles, which enable the integration of DNA fragments into the genome of the host, or alternatively, enable expression of genetic elements that are not integrated. Vectors are typically self-replicating DNA or RNA constructs containing the desired nucleic acid sequences, and operably linked genetic control elements that are recognized in a suitable host cell and effect the translation of the desired spacers. Generally, the genetic control elements can include a prokaryotic promoter system or a eukaryotic promoter expression control system. Such system typically includes a transcriptional promoter, transcription enhancers to elevate the level of RNA expression. Vectors usually contain an origin of replication that allows the vector to replicate independently of the host cell. In yet some alternative embodiments, the expression vectors used by the invention may comprise elements necessary for integration of the desired reporter gene of the invention into the avian gender specific chromosomes W and/or Z.

Accordingly, the term "**control and regulatory elements**" includes promoters, terminators and other expression control elements. Such regulatory elements are described in Goeddel; [Goeddel., et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990)]. For instance, any of a wide variety of expression control sequences that control the expression of a DNA sequence when operatively linked to it may be used in these vectors to express DNA sequences encoding any desired protein using the method of this invention.

A vector may additionally include appropriate restriction sites, antibiotic resistance or other markers for selection of vector-containing cells. Plasmids are the most commonly used form of vector but other forms of vectors which serve an equivalent function and which are, or become, known in the art are suitable for use herein. See, e.g., Pouwels et al., Cloning Vectors: a Laboratory Manual (1985 and supplements), Elsevier, N.Y.; and Rodriquez, et al. (eds.) Vectors: a Survey of Molecular Cloning Vectors and their Uses, Butterworth, Boston, Mass (1988), which are incorporated herein by reference.

To create the transgenic avian animal used by the methods of the invention, an avian cell comprising the reporter gene integrated into specific loci within the gender chromosomes Z or W thereof must be prepared. Such cell may be prepared in some embodiments, by contacting, introducing or co-transfecting the cell of the avian subject or any cell introduced to said subject, with the first and second nucleic acid sequences provided by the methods and kits of the invention or with any construct, vector, vehicle, comprising the same. In some embodiments, nucleic acid sequences encoding the gene editing elements, and nucleic acid sequences encoding the targeted reporter gene, specifically, RFP. "**Transfection**" as used herein is meant the process of inserting genetic material, such as DNA and double stranded RNA, into mammalian cells. The insertion of DNA into a cell enables the expression, or production, of proteins using the cells own machinery. Thus, co-transfection as used herein refers to simultaneous transfection of at least two different nucleic acid molecules or any vector comprising the same to each single cell. Still further, the nucleic acid sequences to be transfected can be transiently expressed for a short period of time, or become incorporated into the genomic DNA, where the change is passed on from cell to cell as it divides.

The invention therefore provides methods for an *in-ovo* gender determination of an avian embryo *in-ovo* based on expression of a reporter gene, specifically, Red Fluorescent Protein (RFP) integrated in the gender chromosome/s thereof (specifically, Z and/or W). "**Expression**" generally refers to the process by which gene-encoded information is converted into the structures present and operating in the cell. Therefore, according to the invention "expression" of a reporter gene, specifically, may refer to transcription into a polynucleotide, translation into a protein, or even posttranslational modification of the protein.

The insertion and specific integration of the reporter gene to the specific target locus within the gender chromosomes of the transgenic avian subject, may involve in some embodiments the provision of the CRISPR/Cas9 system that includes specific gRNA and the nucleic acid sequence of the reporter gene that should be integrated. To facilitate and enable integration, the reporter gene must be flanked in some embodiments, with sequences that may be homolog to the sequences flanking the targeted integration site and thereby enable recombination. Thus, in yet some further specific embodiments, the at least one reporter gene in the second nucleic acid sequence may be flanked at 5' and 3' thereof by homologous arms. It should be appreciated that in some embodiments, these arms are required and therefore facilitate HDR of the reporter gene at the integration site.

In more specific embodiments, the reporter gene in the second nucleic acid sequence used by the method of the invention, may be flanked with two arms that are homologous or show homology or identity of about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% to the whole, entire, or complete at least one nucleic acid sequence comprised within the target loci within the gender chromosomes Z or W, that serves as the integration site to facilitate specific integration via HDR. In certain embodiments, the target sequence is also referred to herein as at least one "proto-spacer" that is recognized by the "spacer" sequences that are part of the gRNA used by the invention, and provided by the first nucleic acid sequence.

25

The term "**Homologous arms**", as used herein refers to HDR templates introduced into specific vectors or viruses, used to create specific mutations or insertion of new elements into a gene, that possess a certain amount of homology surrounding the target sequence to be modified (depending on which PEN is used). In yet some further specific embodiments, where CRISPR is used as a PEN, the arms sequences (left, upstream and right, downstream) may comprise between about 10 to 5000 bp, specifically, between about 50 to 1000 bp, between 100 to 500, specifically, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000bp.

30

In yet some further embodiments, the targeting sequence within the gRNA encoded by the first nucleic acid sequence provided by the methods and kits of the invention, also referred to herein as the "spacer" sequence, exhibits homology or identity of about 70%,
5 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% to the entire, complete or whole at least one nucleic acid sequence comprised within the target loci within the gender chromosomes Z or W, referred to herein as the "proto-spacer".

10

In some embodiments, the at least one reporter gene in the second nucleic acid sequence may be operably linked to any one of a gender specific promoter, an embryonal specific promoter (for example α -Globin Promoter) and an inducible promoter (for example light-inducible promoters derived from the soybean SSU gene, or derived from parsley chalcone synthase CHS promoter, or an engineered version of EL222, a bacterial Light-Oxygen-Voltage protein that activates expression when illuminated with blue light). In
15 yet more specific embodiments, the reporter gene may be under the control of an embryonic promoter, thereby limiting the expression of the transgenic reporter gene to the embryonal stage, with no expression in the adult chick. In such embodiment, the
20 reporter transgene may be used and expressed only at the embryonal stage, for diagnostic purposes.

More specifically, "**Promoter**" as used herein, refers to a particular region of the DNA that has the ability to control the expression of the gene which is placed downstream.

25

Thus, "**Promoter specific for gender in chicks**" refers hereinafter to a promoter that will activate the expression of a gene, only in a specific chick gender (i.e. male or female). Still further, "**Promoter specific for development in chicks**" refers to a promoter that will activate the expression of a gene, only at specific stages of the chick development.

30

In some specific embodiments, the at least one reporter gene may be inserted and thereby integrated into at least one non-coding region of the target gender chromosome.

- 40 -

Such approach avoids the disruption of genes that may be required for development and maturation of the unhatched embryo.

"Non-coding region" as used herein, refers to components of an organism's DNA that do not encode protein sequences. Some noncoding DNA region is transcribed into functional non-coding RNA molecules, other functions of noncoding DNA regions include the transcriptional and translational regulation of protein-coding sequences, scaffold attachment regions, origins of DNA replication, centromeres and telomeres. The hypothesized non-functional portion (or DNA of unknown function) has often been referred to as "junk DNA".

In yet some specific embodiments, the at least one reporter gene used by the method of the invention for preparing the transgenic avian animal, may be integrated into at least one site at gender Z chromosome. In more specific embodiments, the specific loci in the Z chromosome may be any one of regions 9156874-9161874, as denoted by SEQ ID NO:15, 27764943-27769943, as denoted by SEQ ID NO:16, 42172748-42177748, as denoted by SEQ ID NO:17, 63363656-63368656, as denoted by SEQ ID NO:18 and 78777477-78782477, as denoted by SEQ ID NO:19 of Chromosome Z of female chicken.

In some specific embodiments, at least one reporter gene may be integrated into at least one site at gender Z chromosome locus 42172748-42177748, specifically, as denoted by SEQ ID NO. 17. In yet some further specific embodiments, a gRNA sequence suitable for integration into specific loci within the Z chromosome, may include but is not limited to gRNA7 directed to a protospacer located within Z chromosome locus chrZ_42174515_-1. In some specific embodiments, the gRNA may be the gRNA designated gRNA7. In more specific embodiments, such gRNA7 may comprise the nucleic acid sequence GTAATACAGAGCTAAACCAG, as also denoted by SEQ ID NO:26. In yet some further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, the sequence encoding the reporter gene, specifically, the RFP, provided by the invention, may be flanked by a left arm at the 5' thereof, and a right arm at the 3' thereof. In yet some further specific embodiments, the left arm may comprise the nucleic acid sequence as denoted by SEQ ID NO. 31, and right arm comprising the nucleic acid sequence as denoted by SEQ ID

NO. 32, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA7 of SEQ ID NO:26.

In yet some further specific embodiments, the at least one gRNA required to target the reporter gene to such specific location within the Z chromosome may comprises the nucleic acid sequence as denoted by any one of gRNA3: ACAGACCTATGATATGT, as denoted by SEQ ID NO. 11; gRNA4: CGATTATCACTCACAAG, as denoted by SEQ ID NO. 12; gRNA5: CTGGTTAGCATGGGGAC, as denoted by SEQ ID NO. 13; gRNA6: GTAAAGAGTCAGATACA, as denoted by SEQ ID NO. 14.

10 In yet some further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 41 and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 42 may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA3 of SEQ ID NO: 11. In further
15 embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 43, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 44, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA4 of SEQ ID NO:12. In still further embodiments, for integrating
20 the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 45, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 46, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA5 of SEQ ID NO:13. In some further embodiments, for integrating the reporter gene of the
25 invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 47, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 48, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA6 of SEQ ID NO:14. Further non-limiting examples for gRNA sequences suitable for integration into specific
30 loci within the Z chromosome, may include but are not limited to gRNA8 of Z chromosome locus chrZ_9157091_1, comprising the nucleic acid sequence ACAGACCTATGATATGTGAG, as also denoted by SEQ ID NO:27, gRNA9 of Z chromosome locus chrZ_27767602_-1, comprising the nucleic acid sequence

- 42 -

GAGCTTGTGAGTGATAATCG, as also denoted by SEQ ID NO:28, gRNA10 of Z chromosome locus chrZ_78779927_1, comprising the nucleic acid sequence GTAAAGAGTCAGATACACAG, as also denoted by SEQ ID NO: 29, and gRNA11 of Z chromosome locus chrZ_63364946_-1, comprising the nucleic acid sequence
5 CAGTGGGTACTGAAGCTGTG as also denoted by SEQ ID NO: 30.

In further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 33, and right arm comprising the nucleic acid sequence as
10 denoted by SEQ ID NO. 34, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA8 of SEQ ID NO:27. In still further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 35, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 36,
15 may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA9 of SEQ ID NO:28. In further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 37, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 38, may be used to integrate the
20 reporter gene of the invention to the specific loci directed by gRNA10 of SEQ ID NO:29. In yet a further embodiment, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 39, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 40, may be used to integrate the reporter gene of
25 the invention to the specific loci directed by gRNA11 of SEQ ID NO:30.

In yet some alternative embodiments, the at least one reporter gene may be integrated into at least one site at gender W chromosome. In more specific embodiments, the specific locus in the W chromosome may be location 1022859-1024215. In some
30 specific embodiments, the target locus may comprise the nucleic acid sequence as denoted by SEQ ID NO. 3.

- 43 -

In more specific embodiments, the at least one gRNA required to target the reporter gene to such specific location within the W chromosome may comprises the nucleic acid sequence as denoted by any one of SEQ ID NO. 1 and 2, these gRNAs are designated herein as gRNA1 and gRNA2, respectively.

5

In yet some more specific embodiments, the gRNA used by the method of the invention to prepare the transgenic avian female may comprise the nucleic acid sequence as denoted by SEQ ID NO. 1 (gRNA1). In such case, the at least one reporter gene comprised within said second nucleic acid sequence may be flanked at 5' and 3' thereof
10 by homologous arms comprising the amino acid sequence as denoted by SEQ ID NO. 4 and 5, that facilitate the integration thereof to said specific loci in W chromosome, respectively. It should be appreciated that these arms are also referred to herein as left and right arms, respectively.

15 In yet some alternative embodiments, the gRNA used for preparing the transgenic avian female of the invention may comprise the nucleic acid sequence as denoted by SEQ ID NO. 2 (gRNA2). In such case the at least one reporter gene comprised within the second nucleic acid sequence is flanked at 5' and 3' thereof by homologous arms comprising the amino acid sequence as denoted by SEQ ID NO. 6 and 7, respectively. It should be
20 appreciated that these arms are also referred to herein as left and right arms, respectively.

When genetic loci of zygote cells of an avian host, have been targeted and/or transfected with exogenous sequences, specifically, the reporter gene used by the invention, it may
25 be desirable to use such cells to generate transgenic animals. For such a procedure, following the introduction of the targeting construct into the embryonic stem (ES) cells, the cells may be plated onto a feeder layer in an appropriate medium, for example, DMEM supplemented with growth factors and cytokines, fetal bovine serum and antibiotics. The embryonic stem cells may have a single targeted locus (heterozygotic)
30 or both loci targeted (homozygotic). Cells containing the construct may be detected by employing a selective medium and after sufficient time for colonies to grow, colonies may be picked and analyzed for the occurrence of gene targeting. In some specific embodiments, PCR may be applied to verify the integration of the desired exogenous

- 44 -

sequences into the target loci, using primers within and outside the construct sequence. Colonies which show gene targeting may then be used for injection into avian embryos. The ES cells can then be trypsinized and the modified cells can be injected through an opening made in the side of the egg. After sealing the eggs, the eggs can be incubated
5 under appropriate conditions until hatching. Newly hatched avian can be tested for the presence of the target construct sequences, for example by examining a biological sample thereof, e.g., a blood sample. After the avian have reached maturity, they are bred and their progeny may be examined to determine whether the exogenous integrated sequences are transmitted through the germ line.

10

Chimeric avian are generated which are derived in part from the modified embryonic stem cells or zygote cells, capable of transmitting the genetic modifications through the germ line. Mating avian strains containing exogenous sequences, specifically, the reporter gene used by the invention, or portions thereof, with strains in which the avian
15 wild type loci, or portions thereof, is restored, should result in progenies displaying an *in-ovo* detectable gender.

Still further, transgenic avian can also be produced by other methods, some of which are discussed below. Among the avian cells suitable for transformation for generating
20 transgenic animals are primordial germ cells (PGC), sperm cells and zygote cells (including embryonic stem cells). Sperm cells can be transformed with DNA constructs by any suitable method, including electroporation, microparticle bombardment, lipofection and the like. The sperm can be used for artificial insemination of avian. Progeny of the inseminated avian can be examined for the exogenous sequence as
25 described above.

Alternatively, primordial germ cells may be isolated from avian eggs, transfected with the exogenous reporter gene of the invention by any appropriate method, and transferred or inserted into new embryos, where they can become incorporated into the developing
30 gonads. Hatched avian and their progeny can be examined for the exogenous reporter gene sequence as described by the invention.

In yet another approach, dispersed blastodermal cells isolated from eggs can be transfected by any appropriate means with the exogenous reporter gene sequence, or portions thereof, integrated to the gender specific chromosomes Z or W, followed by injection into the subgerminal cavity of intact eggs. Hatched avian subjects and their
5 progeny may be examined for the exogenous reporter gene as described above.

Chicken primordial germ cells (PGCs) are the precursors for ova and spermatozoa. Thus, in some aspects thereof, the invention provides the production of transgenic chickens via a germline transmission system using PGCs co-transfected with the
10 reporter gene construct and with the CRISPR/Cas9 gRNA construct that directs the integration of the reporter gene into the gender specific chromosomes W and/or Z. PGCs are sorted and transferred into the bloodstream of 2.5-day recipient embryos for germline transmission.

15 PGCs are the precursors of gametes and show unique migration activity during early embryogenesis in birds, when compared to mammals. In avian embryos, PGCs first arise from the epiblast and migrate to the hypoblast of the germinal crescent at stage 4 (HH), approximately 18–19 h after incubation. Between stages 10 and 12 (HH), PGCs move from the germinal crescent into the bloodstream and migrate through the
20 circulatory system until they reach the genital ridges and colonize the developing gonads. This contrasts with mammals in which PGCs migrate through the embryonic tissues to reach the developing gonads. It is this unique feature of avian PGC migration through the blood stream that has facilitated the major advance in the genetic modification of chickens.

25

Thus, in some specific embodiments, the "**Preparation of transgenic avian animal**" refers to a multi-step method involving genetic engineering techniques for production of chicken with genomic modifications wherein a) Primordial Germ Cells (PGCs) are isolated from the blood of two days-old chick embryos; b) a transgene construct is
30 incorporated into cultured PGCs, for example, by using lentiviral system, Piggybac transposon vectors, TALENS or CRISPR/Cas9 techniques; (c) transgenic PGCs are identified and injected into the circulatory system of embryos and migrate to the developing gonads; d) recipient embryos are incubated at 37°C until hatching (d)

- 46 -

hatched males are reared to sexual maturity and crossed with wild-type hens (e) offspring are screened to identify those derived from the transgenic PGCs.

Thus, in a second aspect, the invention relates to an avian transgenic animal or subject
5 comprising, in at least one cell thereof, at least one exogenous reporter gene integrated into at least one position or location (also referred to herein as locus) in at least one of gender chromosome Z and W.

The term "**avian**" relates to any species derived from birds characterized by feathers,
10 toothless beaked jaws, the laying of hard-shelled eggs, a high metabolic rate, a four-chambered heart, and a lightweight but strong skeleton. Avian species includes, without limitation, chicken, quail, turkey, duck, Gallinacea sp, goose, pheasant and other fowl. The term "hen" includes all females of the avian species. A "**transgenic avian**" generally refers to an avian that has had a heterologous DNA sequence, or one or more
15 additional DNA sequences normally endogenous to the avian (collectively referred to herein as "transgenes") chromosomally integrated into the germ cells of the avian. As a result of such transfer and integration, the transferred sequence may be transmitted through germ cells to the offspring of a transgenic avian. The transgenic avian (including its progeny) also have the transgene integrated into the gender chromosomes
20 of somatic cells.

In some specific embodiments, the at least one transgenic animal of the invention may comprise at least two different reporter genes. In such case, each reporter gene may be integrated into at least one position or location in one of gender chromosome Z and/or
25 W. In yet some further embodiments, each reporter gene may be integrated into at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten or more, positions or locations in one of gender chromosome Z and/or W.

30 In yet some further embodiments, the reporter gene comprised within the transgenic animal of the invention, may be at least one fluorescent reporter gene. It should be appreciated that any reporter gene and specifically, any fluorescent reporter gene disclosed by the invention, specifically in Table 1 herein before, in connection with

other aspects of the invention may be also applicable in connection with the transgenic avian subject provide by the invention.

In more specific embodiments, such fluorescent reporter gene may comprise or may be
5 at least one Red Fluorescent Protein (RFP). It should be noted however, that in some
embodiments, the reporter gene expressed as the transgene by the transgenic avian
subject of the invention, may be any fluorescent protein, e.g., RFP, YFP, BFP and the
like, provided that said fluorescent protein is not green fluorescent protein (GFP). In
other words, in some specific embodiments the reporter gene integrated into the gender
10 chromosome of the transgenic animal provided by the invention may be any fluorescent
protein with the proviso that said fluorescent protein is not GFP.

In certain embodiments, the at least one transgenic avian animal provided by the
invention, may be female. In more specific embodiments, the at least one reporter gene
15 in such transgenic avian female may be integrated into at least one position of the
female chromosome Z.

In yet some alternative embodiments, the at least one transgenic avian animal may be
female, having at least one reporter gene integrated into at least one position of the
female chromosome W.

20 In yet some further alternative embodiments, the transgenic animal provided by the
invention may be a transgenic male subject that carry the transgene reporter gene
integrated into at least one of its two gender Z chromosomes. In some further
embodiments, the transgenic animal provided by the invention may be a transgenic male
subject that carry the transgene reporter gene integrated into one of its two gender Z
25 chromosomes. In further alternative embodiments, the transgenic animal provided by
the invention may be a transgenic male subject that carry the transgene reporter gene
integrated into two of its two gender Z chromosomes.

In some specific embodiments, the at least one reporter gene may be integrated into the
30 gender chromosome of the transgenic animal of the invention using at least one PEN.
As noted above in connection with the methods of the invention, any genetic
engineering system or any PEN may be used, specifically, any RNA guided system, for
example, one of the CRISPR, TALEN, ZFN and the like.

More specifically, such PEN may be in certain embodiments, a CRISPR system may be used. In more specific embodiments, a CRISPR Class 2 system may be used in creating the transgenic animal of the invention. In yet some further specific embodiments, a
5 CRISPR type II system may be used.

In yet more specific embodiments, the at least one reporter gene may be integrated into the gender chromosome of the transgenic avian animal of the invention by HDR mediated by at least one CRISPR/Cas9 system.

10

In more specific embodiments, the at least one reporter gene may be integrated into a gender chromosome of the transgenic avian animal of the invention by contacting or co-transfecting at least one cell of this avian animal/subject, or at least one cell that is to be introduced into said avian animal with at least two nucleic acid sequences. More
15 specifically, such cell may be contacted or co-transfected with (a) at least one component of a gene editing system, specifically, at least one Cas9 protein or at least one first nucleic acid sequence comprising at least one nucleic acid sequence encoding said at least one Cas9 protein; and at least one gRNA targeting at least one protospacer within at least one gender chromosome, Z or W, or at least one nucleic acid sequence
20 encoding said at least one gRNA, thereby providing a CRISPR mediated integration; and (b) at least one second nucleic acid sequence comprising at least one reporter gene or any fragments thereof.

In more specific embodiments, the at least one reporter gene in the second nucleic acid
25 sequence may be flanked at 5' and 3' thereof by homologous arms. These arms exhibit homology to the integration target site within the target gender chromosome, thereby facilitating HDR at the integration site.

In yet more specific embodiments, the at least one reporter gene in the second nucleic
30 acid sequence may be operably linked to any one of a gender specific promoter, an embryonal specific promoter and an inducible promoter. Such promoter may in some embodiments, limit the expression of the reporter gene of the invention to the specific

desired gender (in case of gender specific promoter), the specific embryonic stage (embryonic specific promoter) or specific conditions (inducible conditions).

In yet some further specific embodiments, the at least one reporter gene comprised
5 within the transgenic avian animal of the invention may be integrated into at least one non-coding region of one of its gender chromosomes.

In yet some further alternative embodiments, the transgenic avian animal of the invention may comprise at least one reporter gene integrated into at least one site at
10 gender Z chromosome. In some particular and non-limiting embodiments, such avian transgenic animal may be female that carry the transgenic reporter gene integrated into the Z chromosome. In more specific embodiments, the specific loci in the Z chromosome may be any one of regions 9156874-9161874, as denoted by SEQ ID NO:15, 27764943-27769943, as denoted by SEQ ID NO:16, 42172748-42177748, as
15 denoted by SEQ ID NO:17, 63363656-63368656, as denoted by SEQ ID NO:18 and 78777477-78782477, as denoted by SEQ ID NO:19 of Chromosome Z of female chicken.

In some specific embodiments, at least one reporter gene may be integrated into at least one site at gender Z chromosome locus 42172748-42177748, specifically, as denoted by
20 SEQ ID NO. 17. In yet some further specific embodiments, a gRNA sequence suitable for integration into specific loci within the Z chromosome, may include but is not limited to gRNA7 directed to a protospacer located within Z chromosome locus chrZ_42174515_-1. In some specific embodiments, the gRNA may be the gRNA designated gRNA7. In more specific embodiments, such gRNA7 may comprise the
25 nucleic acid sequence GTAATACAGAGCTAAACCAG, as also denoted by SEQ ID NO:26. In yet some further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, the sequence encoding the reporter gene, specifically, the RFP, provided by the invention, may be flanked by a left arm at the 5' thereof, and a right arm at the 3' thereof. In yet some further specific
30 embodiments, the left arm may comprise the nucleic acid sequence as denoted by SEQ ID NO. 31, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 32, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA7 of SEQ ID NO:26.

- 50 -

In yet some more specific embodiments, the at least one gRNA required to target the reporter gene to such specific location within the Z chromosome in the transgenic avian subject, may comprises the nucleic acid sequence as denoted by any one of gRNA3: ACAGACCTATGATATGT, as denoted by SEQ ID NO. 11; gRNA4: 5 CGATTATCACTCACAAG, as denoted by SEQ ID NO. 12; gRNA5: CTGGTTAGCATGGGGAC, as denoted by SEQ ID NO. 13 ; gRNA6: GTAAAGAGTCAGATACA, as denoted by SEQ ID NO. 14.

In yet some further embodiments, for integrating the reporter gene of the invention into 10 the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 41, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 42, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA3 of SEQ ID NO:11. In further 15 embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 43, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 44, may be used to integrate the reporter gene in the transgenic avian subject of the invention to the specific loci directed by gRNA4 of SEQ ID NO:12. In still further 20 embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome in the transgenic avian subject, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 45, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 46, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA5 of SEQ ID NO:13. In some further embodiments, for integrating the reporter gene of the invention into the 25 specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 47, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 48, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA6 of SEQ ID NO:14.

30 Further non-limiting examples for gRNA sequences suitable for integration into specific loci within the Z chromosome of the transgenic avian subject of the invention, may include but are not limited to gRNA8 of Z chromosome locus chrZ_9157091_1, comprising the nucleic acid sequence ACAGACCTATGATATGTGAG, as also

- 51 -

denoted by SEQ ID NO:27, gRNA9 of Z chromosome locus chrZ_27767602_-1, comprising the nucleic acid sequence GAGCTTGTGAGTGATAATCG, as also denoted by SEQ ID NO:28, gRNA10 of Z chromosome locus chrZ_78779927_1, comprising the nucleic acid sequence GTAAAGAGTCAGATACACAG, as also
5 denoted by SEQ ID NO: 29, and gRNA11 of Z chromosome locus chrZ_63364946_-1, comprising the nucleic acid sequence CAGTGGGTACTGAAGCTGTG as also denoted by SEQ ID NO: 30.

In further embodiments, for integrating the reporter gene of the invention into the
10 specific locus within the Z chromosome of the transgenic avian subject of the invention, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 33, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 34, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA8 of SEQ ID NO:27. In still further embodiments, for integrating the reporter gene of the
15 invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 35, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 36, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA9 of SEQ ID NO:28. In further embodiments, for integrating the reporter gene of the invention into the
20 specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 37, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 38, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA10 of SEQ ID NO:29. In yet a further embodiment, for integrating the reporter gene of the invention into the specific locus within the Z
25 chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 39, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 40, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA11 of SEQ ID NO:30.

30 In yet some further alternative embodiments, the at least one reporter gene may be integrated into at least one site at gender W chromosome of the transgenic avian subject of the invention. In some particular embodiments, the integration site may be located at locus 1022859-1024215 at the W chromosome, specifically, galGal5_dna range of

chromosome W:1022859-1024215. In yet some further specific embodiments, such loci comprises the nucleic acid sequence as denoted by SEQ ID NO. 3.

For specific integration of the reporter gene of the invention at any position within the loci described above, specific gRNAs may be required. Therefore, in some particular and non-limiting embodiments, appropriate gRNAs used for the preparation of the transgenic avian animal of the invention may comprise the nucleic acid sequence as denoted by any one of SEQ ID NO. 1 and 2. In some specific embodiments, these gRNAs are referred to herein as gRNA1 and gRNA2, respectively.

In some particular embodiments, the transgenic avian animal provided by the invention has been prepared using a gRNA1 that comprises the nucleic acid sequence as denoted by SEQ ID NO. 1. To enable integration of the reporter gene of the invention in such specific location, the reporter gene that should be integrated, must carry in certain embodiments, particular arms facilitating incorporation thereof in the target integration site directed by the gRNA used. Thus, in some specific embodiments, the at least one reporter gene may be comprised within the second nucleic acid sequence, where this reporter gene is flanked at 5' and 3' thereof by homologous arms comprising the amino acid sequence as denoted by SEQ ID NO. 4 and 5, respectively.

In yet some alternative embodiments, the transgenic avian animal provided by the invention may be prepared using a gRNA2 that comprises the nucleic acid sequence as denoted by SEQ ID NO. 2. In such case, to enable integration of the reporter gene of the invention at the specific site recognized by said gRNA2, the at least one reporter gene comprised within the second nucleic acid sequence may be according to specific embodiments, flanked at 5' and 3' thereof by homologous arms comprising the amino acid sequence as denoted by SEQ ID NO. 6 and 7, respectively.

It should be further appreciated that the invention further encompasses any of the transgenic avian subjects as described herein for use in any of the methods of the invention, specifically, methods for *in ovo* gender determination as discussed herein.

In yet some additional embodiments, the invention further encompasses the use of any egg laid by any of the transgenic avian subjects provided by the invention, specifically, as disclosed above, for use in any of the methods of the invention. In yet another aspect,

the invention relates to a cell comprising at least one exogenous reporter gene integrated into at least one position or location in at least one of gender chromosome Z and W.

In some specific embodiments, the cell provided by the invention may be an avian cell.

5 In some particular embodiments, the avian cell provided by the invention may be a primordial germ cell (PGC).

The term "**germ cells**" refers to an embryonic cell that upon uniting with another germ cells develops into a gamete. "**Primordial germ cells (PGCs)**", as used herein relates to germline stem cells that serve as progenitors of the gametes and give rise to pluripotent embryonic stem cells. The cells in the gastrulating embryo that are signaled to become PGCs during embryogenesis, migrate into the genital ridges which becomes the gonads, and differentiate into mature gametes.

In some embodiments, the reporter gene integrated in at least one gender chromosome of the cell of the invention may be a fluorescent protein, specifically, RFP, YFP, BFP, and the like or any of the fluorescent proteins disclosed by Table 1. In yet some further 15 embodiments, the reporter gene integrated in the cell of the invention may be any fluorescent protein, provided that said fluorescent protein is not green fluorescent protein (GFP). In other words, in some specific embodiments the reporter gene integrated into the gender chromosome of the transgenic cell provided by the invention 20 may be any fluorescent protein with the proviso that said fluorescent protein is not GFP.

In yet more specific embodiments, the at least one reporter gene may be integrated into the gender chromosome of the transgenic avian subject or animal provided by the cell of the invention using at least one programmable engineered nuclease (PEN). The term 25 "**programmable engineered nucleases (PEN)**" as used herein, refers to synthetic enzymes that cut specific DNA sequences, derived from natural occurring nucleases involved in DNA repair of double strand DNA lesions and enabling direct genome editing. In more specific embodiments, PEN may be a clustered regularly interspaced short palindromic repeat (CRISPR) class 2, specifically, type II system.

30 In yet some further embodiments, the cell provided by the invention may comprise at least one reporter gene integrated into a gender chromosome of the cell. In more specific embodiments, such specific integration of the reporter gene may be enabled by contacting or co-transfecting the cell with: (a) at least one gene editing system,

comprising at least one Cas9 protein or at least one first nucleic acid sequence comprising at least one nucleic acid sequence encoding said at least one Cas9 protein; and at least one gRNA that targets at least one protospacer within at least one gender chromosome Z and/or W, or at least one nucleic acid sequence encoding said at least one gRNA; and (b) at least one second nucleic acid sequence comprising at least one said reporter gene.

In certain embodiments, the at least one reporter gene in the second nucleic acid sequence co-transfected to the cell of the invention, may be flanked at 5' and 3' thereof by homologous arms for HDR at the integration site.

In yet some specific embodiments, the cell provided by the invention may be prepared by integrating the at least one reporter gene of the invention into the Z chromosome of the cell. In certain embodiments, for preparing the cell of the invention, the at least one reporter gene may be integrated into at least one site at gender Z chromosome. In more specific embodiments, the specific loci in the Z chromosome may be any one of regions 9156874-9161874, as denoted by SEQ ID NO:15, 27764943-27769943, as denoted by SEQ ID NO:16, 42172748-42177748, as denoted by SEQ ID NO:17, 63363656-63368656, as denoted by SEQ ID NO:18 and 78777477-78782477, as denoted by SEQ ID NO:19 of Chromosome Z of female chicken.

In yet some further alternative embodiments, the cell provided by the invention may be prepared by using gRNAs that facilitate integration into at least one site at gender Z chromosome locus 42172748-42177748, specifically, SEQ ID NO. 17. In yet some further specific embodiments, a gRNA sequence suitable for integration into specific loci within the Z chromosome, may include but is not limited to gRNA7 of Z chromosome locus chrZ_42174515_-1, comprising the nucleic acid sequence GTAATACAGAGCTAAACCAG, as also denoted by SEQ ID NO:26, In yet some further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 31, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 32, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA7 of SEQ ID NO:26.

In more specific embodiments, the at least one gRNA required to target the reporter gene to such specific location within the Z chromosome of the cell of the invention may comprises the nucleic acid sequence as denoted by any one of gRNA3:
5 ACAGACCTATGATATGT, as denoted by SEQ ID NO. 11; gRNA4:
CGATTATCACTCACAAG, as denoted by SEQ ID NO. 12; gRNA5:
CTGGTTAGCATGGGGAC, as denoted by SEQ ID NO. 13; gRNA6:
GTAAAGAGTCAGATACA, as denoted by SEQ ID NO. 14.

In yet some further embodiments, for integrating the reporter gene of the invention into
10 the specific locus within the Z chromosome of the cell of the invention, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 41, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 42, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA3 of SEQ ID NO:11. In further embodiments, for integrating the reporter gene of the
15 invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 43, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 44, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA4 of SEQ ID NO:12. In still further embodiments, for integrating the reporter gene of the invention into the
20 specific locus within the Z chromosome of the cell of the invention, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 45, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 46, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA5 of SEQ ID NO:13. In some further embodiments, for integrating the reporter gene of the invention into the
25 specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 47, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 48, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA6 of SEQ ID NO:14.

Further non-limiting examples for gRNA sequences suitable for integration into specific
30 loci within the Z chromosome, may include but are not limited to gRNA8 of Z chromosome locus chrZ_9157091_1, comprising the nucleic acid sequence ACAGACCTATGATATGTGAG, as also denoted by SEQ ID NO:27, gRNA9 of Z chromosome locus chrZ_27767602_-1, comprising the nucleic acid sequence

- 56 -

GAGCTTGTGAGTGATAATCG, as also denoted by SEQ ID NO:28, gRNA10 of Z chromosome locus chrZ_78779927_1, comprising the nucleic acid sequence GTAAAGAGTCAGATACACAG, as also denoted by SEQ ID NO: 29, and gRNA11 of Z chromosome locus chrZ_63364946_-1, comprising the nucleic acid sequence
5 CAGTGGGTACTGAAGCTGTG as also denoted by SEQ ID NO: 30.

In further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome of the cell of the invention, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 33, and right arm comprising the
10 nucleic acid sequence as denoted by SEQ ID NO. 34, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA8 of SEQ ID NO:27. In still further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome of the cell of the invention, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 35, and right arm comprising the
15 nucleic acid sequence as denoted by SEQ ID NO. 36, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA9 of SEQ ID NO:28. In further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 37, and right arm comprising the nucleic acid sequence as
20 denoted by SEQ ID NO. 38, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA10 of SEQ ID NO:29. In yet a further embodiment, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 39, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 40,
25 may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA11 of SEQ ID NO:30.

In some particular embodiments, to target the integration of the reporter gene to chromosome W in the cell provided by the invention, specific gRNAs should be used.
30 In further particular embodiments, the gRNA may comprise the nucleic acid sequence as denoted by SEQ ID NO. 1 referred to herein as gRNA1. In such case, the at least one reporter gene comprised within the second nucleic acid sequence, may be flanked at 5'

- 57 -

and 3' thereof by homologous arms comprising the amino acid sequence as denoted by SEQ ID NO. 4 and 5, respectively.

In yet some further alternative embodiments, the cell provided by the invention may be prepared by using gRNA referred to herein as gRNA2. In certain embodiments, gRNA2 may comprise the nucleic acid sequence as denoted by SEQ ID NO. 2. In such specific embodiments, the at least one reporter gene comprised within the second nucleic acid sequence may be flanked at 5' and 3' thereof by homologous arms comprising the amino acid sequence as denoted by SEQ ID NO. 6 and 7, respectively.

10

Still further, the invention provides a method for detecting fertilization of an unhatched egg. In some specific embodiments, such method may comprise the step of (a), providing or obtaining at least one transgenic avian subject or animal comprising at least one exogenous reporter gene integrated into at least one position or location in both gender chromosome Z and W, in case the provided transgenic animal is a female subject and into both gender chromosomes Z in case the provided transgenic animal is a male. In a second step (b) obtaining at least one fertilized egg from the transgenic avian subject, specifically animal or of any cells thereof. The next step (c) involves determining in the egg if at least one detectable signal is detected. In more specific embodiments, detection of at least one detectable signal indicates the expression of said at least one reporter gene, thereby the presence of the labeled maternal W chromosome or Z chromosome (in case of a female or the labeled paternal Z chromosome in the avian embryo).

20

In yet some further embodiments, the transgenic animal provided by the method of the invention may be a male subject having the reporter gene integrated into both Z chromosomes thereof. In such case, a detectable signal determined in an egg fertilized by such transgenic male or any sperms thereof, indicates that the embryo carries a paternal Z chromosome comprising the transgenic reporter gene, and the egg is therefore fertilized. In still further embodiments, the transgenic animal provided by the method of the invention may be a female subject having the reporter gene integrated into both, Z and W chromosomes thereof. In such case, a detectable signal determined in an egg carried by such transgenic female, indicates that the embryo carries a maternal

30

Z or W chromosome comprising the transgenic reporter gene, and the egg is therefore fertilized. It should be appreciated that in further embodiments, for detecting a fertilized egg, the reporter gene, specifically, the RFP gene may be inserted into the two copies of any chromosomes, to create a homozygous transgenic avian subject, specifically, a male subject. A fertilized egg will be detected as carrying the RFP reporter gene.

In yet a further aspect, the invention provides a kit comprising:

(a) at least one component of at least one gene editing system, specifically, the CRISPR/Cas9 system that may comprise at least one Cas9 protein or at least one first nucleic acid sequence comprising at least one nucleic acid sequence encoding said at least one Cas9 protein; and at least one gRNA that targets at least one protospacer within at least one gender chromosome Z and/or W, or at least one nucleic acid sequence encoding said at least one gRNA; and (b) at least one second nucleic acid sequence comprising at least one said reporter gene.

In some further embodiments, the gRNA used by the method, as well by kit of the invention (described herein after) may comprise at least one CRISPR RNA (crRNA) and at least one trans-activating crRNA (tracrRNA).

In some alternative embodiments the kit of the invention may comprise nucleic acid sequence encoding the at least one gRNA. Such nucleic acid sequence may comprise a CRISPR array comprising at least one spacer sequence that targets and is therefore homolog or identical to at least one protospacer in a target genomic DNA sequence. It should be noted that the nucleic acid sequence may further comprise a sequence encoding at least one tracrRNA.

In some embodiments the CRISPR array according to the present disclosure comprises at least one spacer and at least one repeat. In yet another embodiment, the invention further encompasses the option of providing a pre-crRNA that can be processed to several final gRNA products that may target identical or different targets.

In yet some more specific embodiments, the crRNA comprised within the gRNA of the invention may be a single-stranded ribonucleic acid (ssRNA) sequence complementary to a target genomic DNA sequence. In some specific embodiments, the target genomic

- 59 -

DNA sequence may be located immediately upstream of a protospacer adjacent motif (PAM) sequence and further.

As indicated herein, the gRNA of the kits and methods of the invention may be
5 complementary, at least in part, to the target genomic DNA, in this case, the target locus
within the gender chromosomes Z or W, referred to herein as "protospacer". In certain
embodiments, "**Complementarity**" refers to a relationship between two structures each
following the lock-and-key principle. In nature complementarity is the base principle of
DNA replication and transcription as it is a property shared between two DNA or RNA
10 sequences, such that when they are aligned antiparallel to each other, the nucleotide
bases at each position in the sequences will be complementary (e.g., A and T or U, C
and G).

As indicated above, the genomic DNA sequence targeted by the gRNA of the kit of the
invention is located immediately upstream to a PAM sequence. In some embodiments,
15 such PAM sequence may be of the nucleic acid sequence NGG.

In certain embodiments, the PAM sequence referred to by the invention may comprise
N, that is any nucleotide, specifically, any one of Adenine (A), Guanine (G), Cytosine
(C) or Thymine (T). In yet some further embodiments the PAM sequence according to
20 the invention is composed of A, G, C, or T and two Guanines.

According to one embodiment, the polynucleotide encoding the gRNA of the invention
may comprise at least one spacer and optionally, at least one repeat. In yet some further
embodiments, the DNA encoding the gRNA of the invention may comprise at least 1, 2,
25 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28,
29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51,
52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74,
75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97,
98, 99, 100 or more, specifically, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 or
30 more spacers. In some embodiments, each spacer is located between two repeats. It
should be further understood that the spacers of the nucleic acid sequence encoding the
gRNA of the invention may be either identical or different spacers. In more
embodiments, these spacers may target either an identical or different target genomic

DNA. In yet some other embodiments, such spacer may target at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 5 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 or more target genomic DNA sequence. These target sequences may be derived from a single locus or alternatively, from several target loci.

As used herein, the term "**spacer**" refers to a non-repetitive spacer sequence that is designed to target a specific sequence and is located between multiple short direct repeats (i.e., CRISPR repeats) of CRISPR arrays. In some specific embodiments, spacers may comprise between about 15 to about 30 nucleotides, specifically, about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more nucleotides. More specifically, about 20-25 nucleotides.

15 The guide or targeting RNA encoded by the CRISPR system of the invention may comprise a CRISPR RNA (crRNA) and a trans activating RNA (tracrRNA). The sequence of the targeting RNA encoded by the CRISPR spacers is not particularly limited, other than by the requirement for it to be directed to (i.e., having a segment that is the same as or complementarity to) a target sequence in avian genomic DNA, specifically, in a target locus within the gender chromosomes Z or W, that is also referred to herein as a "**proto-spacer**". Such proto-spacers comprise nucleic acid sequence having sufficient complementarity to a targeting RNA encoded by the CRISPR spacers comprised within the nucleic acid sequence encoding the gRNA of the 20 methods and kits of the invention.

In some embodiments, a crRNA comprises or consists of 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 or 40 nt of the spacer (targeting) sequence followed by 19-36 nt of repeat sequence. In specific and non-limiting embodiments, the 30 targeting spacer may comprise or consist of a segment that targets any one of the genomic DNA sequence for which representative spacer sequences are indicated herein.

It should be noted that in some specific embodiments, the spacers of the CRISPR system of the invention may encode a targeting guide RNA (gRNA). A "gRNA" or "targeting RNA" is an RNA that, when transcribed from the portion of the CRISPR system encoding it, comprises at least one segment of RNA sequence that is identical to (with the exception of replacing T for U in the case of RNA) or complementary to (and thus "targets") a DNA sequence in the target genomic DNA. The CRISPR systems of the present disclosure may optionally encode more than one targeting RNA, and the targeting RNAs be directed to one or more target sequences in the genomic DNA.

10 In some embodiments, the Cas9 protein provided by the kit of the invention may be encoded by the nucleic acid sequence as denoted by SEQ ID NO. 24. In further specific and non-limiting embodiments, the Cas9 protein may comprise the amino acid sequence as denoted by SEQ ID NO. 25, or any derivatives, mutants or variants thereof.

It should be noted that "Amino acid sequence" or "peptide sequence" is the order in which amino acid residues connected by peptide bonds, lie in the chain in peptides and proteins. The sequence is generally reported from the N-terminal end containing free amino group to the C-terminal end containing amide. Amino acid sequence is often called peptide, protein sequence if it represents the primary structure of a protein, however one must discern between the terms "Amino acid sequence" or "peptide sequence" and "protein", since a protein is defined as an amino acid sequence folded into a specific three-dimensional configuration and that had typically undergone post-translational modifications, such as phosphorylation, acetylation, glycosylation, manosylation, amidation, carboxylation, sulfhydryl bond formation, cleavage and the like.

25 By "fragments or peptides" it is meant a fraction of said Cas9 or RFP (described herein after) molecules. A "fragment" of a molecule, such as any of the amino acid sequences of the present invention, is meant to refer to any amino acid subset of the Cas9 or RFP molecules. This may also include "variants" or "derivatives" thereof. A "peptide" is meant to refer to a particular amino acid subset having functional activity. By "functional" is meant having the same biological function, for example, having the ability to perform gene editing as the Cas 9 or as producing a detectable signal as the RFP.

It should be appreciated that the invention encompasses any variant or derivative of the RFP or Cas9 molecules of the invention and any polypeptides that are substantially identical or homologue. The term "*derivative*" is used to define amino acid sequences (polypeptide), with any insertions, deletions, substitutions and modifications to the amino acid sequences (polypeptide) that do not alter the activity of the original polypeptides. In this connection, a derivative or fragment of the Cas9 or RFP molecules of the invention may be any derivative or fragment of the Cas9 or RFP molecules, specifically as denoted by SEQ ID NO. 25, 21, 23, that do not reduce or alter the activity of the Cas9 or RFP molecules. By the term "derivative" it is also referred to homologues, variants and analogues thereof. Proteins orthologs or homologues having a sequence homology or identity to the proteins of interest in accordance with the invention, specifically that may share at least 50%, at least 60% and specifically 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, or higher, specifically as compared to the entire sequence of the proteins of interest in accordance with the invention, for example, any of the proteins that comprise the amino acid sequence as denoted by SEQ ID NO. 25, 21, 23. Specifically, homologs that comprise or consists of an amino acid sequence that is identical in at least 50%, at least 60% and specifically 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher to SEQ ID NO. 21, 23, 25 specifically, the entire sequence as denoted by SEQ ID NO. 21, 23, 25.

In some embodiments, derivatives refer to polypeptides, which differ from the polypeptides specifically defined in the present invention by insertions, deletions or substitutions of amino acid residues. It should be appreciated that by the terms "**insertion/s**", "**deletion/s**" or "**substitution/s**", as used herein it is meant any addition, deletion or replacement, respectively, of amino acid residues to the polypeptides disclosed by the invention, of between 1 to 50 amino acid residues, between 20 to 1 amino acid residues, and specifically, between 1 to 10 amino acid residues. More particularly, insertion/s, deletion/s or substitution/s may be of any one of 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids. It should be noted that the insertion/s, deletion/s or substitution/s encompassed by the invention may occur in any position of the modified peptide, as well as in any of the N' or C' termini thereof.

With respect to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein

sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologues, and alleles of the invention. For example, substitutions may be made wherein an aliphatic amino acid (G, A, I, L, or V) is substituted with another member of the group, or substitution such as the substitution of one polar residue for another, such as arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine. Each of the following eight groups contains other exemplary amino acids that are conservative substitutions for one another:

- 1) Alanine (A), Glycine (G);
- 2) Aspartic acid (D), Glutamic acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
- 7) Serine (S), Threonine (T); and
- 8) Cysteine (C), Methionine (M).

Thus, in some embodiments, the invention encompasses Cas9 or RFP molecules or any derivatives thereof, specifically a derivative that comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more conservative substitutions to the amino acid sequences as denoted by any one of SEQ ID NO. 21, 23, 25. More specifically, amino acid “substitutions” are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, i.e., conservative amino acid replacements. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar “*hydrophobic*” amino acids are selected from the group consisting of Valine (V), Isoleucine (I), Leucine (L), Methionine (M), Phenylalanine (F), Tryptophan (W), Cysteine (C), Alanine (A), Tyrosine (Y), Histidine (H), Threonine (T), Serine (S), Proline (P), Glycine (G), Arginine (R) and Lysine (K); “*polar*” amino acids are selected from the group consisting of Arginine (R), Lysine (K), Aspartic acid (D), Glutamic acid (E), Asparagine (N), Glutamine (Q); “*positively charged*” amino

- 64 -

acids are selected from the group consisting of Arginine (R), Lysine (K) and Histidine (H) and wherein “*acidic*” amino acids are selected from the group consisting of Aspartic acid (D), Asparagine (N), Glutamic acid (E) and Glutamine (Q).

5 Variants of the polypeptides of the invention may have at least 80% sequence similarity or identity, often at least 85% sequence similarity or identity, 90% sequence similarity or identity, or at least 95%, 96%, 97%, 98%, or 99% sequence similarity or identity at the amino acid level, with the protein of interest, such as the various polypeptides of the invention.

10 In some embodiments, the at least one reporter gene in the second nucleic acid sequence comprised within the kit of the invention, may be flanked at 5' and 3' thereof by homologous arms for HDR at the integration site.

In yet some further specific embodiments, the at least one reporter gene in the second
15 nucleic acid sequence of the kit of the invention may be operably linked to any one of a gender specific promoter, an embryonic specific promoter and an inducible promoter.

In yet some further alternative embodiments, the at least one reporter gene may be
20 integrated into at least one site at gender Z chromosome. In more specific embodiments, the specific loci in the Z chromosome may be any one of regions 9156874-9161874, as denoted by SEQ ID NO:15, 27764943-27769943, as denoted by SEQ ID NO:16, 42172748-42177748, as denoted by SEQ ID NO:17, 63363656-63368656, as denoted by SEQ ID NO:18 and 78777477-78782477, as denoted by SEQ ID NO:19 of Chromosome Z of female chicken.

25 In some specific embodiments, at least one reporter gene may be integrated into at least one site at gender Z chromosome locus 42172748-42177748. Non-limiting examples of the first nucleic acid sequence of the kit of the invention may comprise a gRNA, being the at least one of gRNA7 of Z chromosome locus chrZ_42174515_-1, comprising the
30 nucleic acid sequence GTAATACAGAGCTAAACCAG, as also denoted by SEQ ID NO:26. More specifically, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 31, and right arm comprising the nucleic acid sequence as

denoted by SEQ ID NO. 32, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA7 of SEQ ID NO:26.

Thus, in more specific embodiments, the first nucleic acid sequence of the kit of the invention may comprise a gRNA, being the at least one of gRNA3: 5 ACAGACCTATGATATGT, as denoted by SEQ ID NO. 11; gRNA4: CGATTATCACTCACAAG, as denoted by SEQ ID NO. 12; gRNA5: CTGGTTAGCATGGGGAC, as denoted by SEQ ID NO. 13; gRNA6: GTAAAGAGTCAGATACA, as denoted by SEQ ID NO. 14.

10

In further embodiments, the at least one reporter gene comprised within the second nucleic acid sequence of the kit of the invention, may be flanked at 5' and 3' thereof by homologous arms comprising the amino acid sequence as denoted by any one of SEQ ID NO. 41 to 48. More specifically, for integrating the reporter gene of the invention 15 into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 41, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 42, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA3 of SEQ ID NO:11. In further 20 embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 43, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 44, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA4 of SEQ ID NO:12. In still further embodiments, for integrating 25 the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 45, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 46, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA5 of SEQ ID NO:13. In some further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left 30 arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 47, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 48, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA6 of SEQ ID NO:14.

- 66 -

Further non-limiting examples of the first nucleic acid sequence of the kit of the invention may comprise a gRNA, being the at least one of gRNA8 of Z chromosome locus chrZ_9157091_1, comprising the nucleic acid sequence ACAGACCTATGATATGTGAG, as also denoted by SEQ ID NO:27, gRNA9 of Z chromosome locus chrZ_27767602_-1, comprising the nucleic acid sequence GAGCTTGTGAGTGATAATCG, as also denoted by SEQ ID NO:28, gRNA10 of Z chromosome locus chrZ_78779927_1, comprising the nucleic acid sequence GTAAAGAGTCAGATACACAG, as also denoted by SEQ ID NO: 29, and gRNA11 of Z chromosome locus chrZ_63364946_-1, comprising the nucleic acid sequence CAGTGGGTACTGAAGCTGTG as also denoted by SEQ ID NO: 30.

In further embodiments, the at least one reporter gene comprised within the second nucleic acid sequence of the kit of the invention, may be flanked at 5' and 3' thereof by homologous arms comprising the amino acid sequence as denoted by any one of SEQ ID NO. 31 to 40. More specifically, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 33, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 34, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA8 of SEQ ID NO:27. In still further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 35, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 36, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA9 of SEQ ID NO:28. In further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 37, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 38, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA10 of SEQ ID NO:29. In yet a further embodiment, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 39, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 40, may be used to integrate the

reporter gene of the invention to the specific loci directed by gRNA11 of SEQ ID NO:30.

5 In certain embodiments, the at least one reporter gene provided by the kits of the invention may be integrated into at least one non-coding region of the gender chromosome, specifically, to chromosome W. In such case, the first nucleic acid sequence of the kit of the invention may encode at least one gRNA comprising the nucleic acid sequence as denoted by any one of SEQ ID NO. 1 and 2, also referred to herein as gRNA1 and gRNA2, respectively.

10

In some specific embodiments, the first nucleic acid sequence of the kit of the invention may comprise a gRNA, being gRNA1. In some embodiments, such gRNA1 may comprise the nucleic acid sequence as denoted by SEQ ID NO. 1. In such case, the reporter gene comprised within said second nucleic acid sequence of the kit of the invention, may be flanked at 5' and 3' thereof by homologous arms comprising the amino acid sequence as denoted by SEQ ID NO. 4 and 5, respectively.

15

In yet some further alternative embodiments, the kit of the invention may comprise in the first nucleic acid sequence thereof, a sequence encoding gRNA2. In some specific 20 embodiments, such sequence encodes the nucleic acid sequence as denoted by SEQ ID NO. 2. In yet some further embodiments, the least one reporter gene comprised within the second nucleic acid sequence of the kit of the invention, may be flanked at 5' and 3' thereof by homologous arms comprising the amino acid sequence as denoted by SEQ ID NO. 6 and 7, respectively.

25

In some embodiments, the reporter gene comprised within the second nucleic acid sequence of the kit of the invention may be at least one fluorescent reporter gene. Still further, in some embodiments, the reporter gene provided by the kit of the invention may be any fluorescent protein, specifically, RFP, YFP and any of the fluorescent 30 proteins disclosed by the invention or any of the fluorescent proteins disclosed by Table 1. In some embodiments, the reporter gene used by the kits of the invention may be any fluorescent protein, provided that said fluorescent protein is not green fluorescent protein (GFP). In other words, in some specific embodiments the reporter gene

- 68 -

integrated into the gender chromosome of the transgenic animal provided by the kit of the invention may be any fluorescent protein with the proviso that said fluorescent protein is not GFP. In yet some further embodiments, the reporter gene may be RFP. In more specific embodiments, such RFP may comprise the amino acid sequence as denoted by SEQ ID NO. 21, or any homologs, variants, mutants or derivatives thereof. In yet some alternative specific embodiments, the RFP used by the invention may be the RFP that may comprise the amino acid sequence as denoted by SEQ ID NO. 23, or any homologs, mutants or derivatives thereof.

10 In yet some further embodiments, the kit of the invention may be suitable for use in the preparation of a transgenic avian animal comprising at least one exogenous reporter gene integrated into at least one position or location in at least one of gender chromosome Z and W.

15 In some embodiments, the method of the invention may use any of the kits of the invention as described herein.

Still further, it must be appreciated that the kits of the invention may further comprise any reagent, buffer, media or material required for the preparation of the transgenic avian animals of the invention. The kit of the invention may further comprise instructions as well as containers for the different components thereof.

In some embodiments, the kits of the invention may be used in the preparation of any of the transgenic avian subjects of the invention or of any cell thereof in accordance with the invention. In yet some further embodiments, the transgenic avian subjects and cells prepared using the kits of the invention, may be used by any of the methods of the invention, using any of the systems and devices as disclosed herein. Still further, in some embodiments, the kit of the invention may further comprise any apparatus, system, and/or device adapted for determining the presence of the reporter gene within the *in ovo* embryo examined by the methods of the invention. In some specific embodiments, such apparatus or device may be any of the devices disclosed by the invention, specifically, as shown in Figure 1.

In yet another aspect thereof, the invention provides a system, apparatus or a device that is specifically adapted for detecting a reporter gene within an *in ovo* embryo. In some embodiments, the device of the invention may comprise a laser source, a stand for the egg, a lens, a filter, a stand for a detector and a detector. The device of the invention
5 may optionally, in some embodiments, comprise a solid support that holds all elements of the system of the invention, for example, as shown in Figure 1. As used herein, the term "**detector**" refers to any type of device that detects and/or measures light.

In some embodiments, it should be noted that the detectable signal, specifically, the fluorescent signal may be detected by the device of the invention using suitable
10 fluorescent means. In some embodiments, the detectable signal formed by the RFP reporter gene may be detected by light sensitive apparatus such as modified optical microscopes or Charge Coupled Device (CCD), a highly sensitive photon detector.

In yet some further embodiments, the device, apparatus or system of the invention may
15 comprise at least one light source, at least one detector, at least one filter and at least one holding arm that places the egg in an appropriate position facilitating the exposure of cells expressing the reporter gene of the invention to said light source. In some embodiments, such egg may be a fertilized unhatched avian egg. In yet some further embodiments, such egg may be laid by any of the transgenic avian subjects of the
20 invention. In some particular and non-limiting embodiments, such device of the invention may be as illustrated in Figure 1. In yet some further embodiments, the distance between the egg holder and lens and/or lens and detector may range between about 1 to 100 cm, specifically, 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 and 100 cm, specifically between about 5 to 25 cm, more specifically,
25 about 20 cm.

It should be therefore appreciated that in some embodiments, the system or device of the invention may be adapted for performing any of the methods disclosed by the invention.

In some specific embodiments, the device of the invention may comprise an appropriate
30 light source that enables detection of the reporter gene in the examined egg. In yet some further embodiments, the light source may comprise wavelength of between about 400 to about 650. In yet some further embodiments, light of any wavelength may be used

with the proviso that said light source is not Ultra-Violet (UV) light (wavelength 10-400 nm).

In some specific embodiment, the appropriate light source is suitable for excitation of the fluorescent protein. In some particular embodiments, the excitation wavelength may
5 be between about 500 nm and about 650 nm. In some further embodiments, the excitation wavelength is about 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640 nm.

In some more specific embodiment, the excitation wavelength may range between about 515 to about 555. In yet some further embodiments, the wavelength may be 532 nm. In
10 some specific and non-limiting embodiment, the light source may be provided by a laser. In some other embodiments, the laser of the device of the invention may be a blue laser, or a red laser. In more specific embodiments, the light source is a green laser.

In certain embodiments, the laser may be employed with a filter. In some embodiments, the filter may be a green filter that transmits above 500 nm, or a dark green filter that
15 transmits between 540 nm and 580 nm, or a red filter that transmits between 590 nm and 650 nm, or a red filter that transmits above 650 nm or above 660 nm.

In some particular embodiments, the light source may be a green laser with a wavelength of 532 nm. In some further embodiments, such laser may be provided with
20 a red filter that transmits above 650 nm.

It should be thus further appreciated that the invention further encompasses the device as described herein for use in any of the methods of the invention, specifically, methods for *in ovo* gender determination as discussed herein, using any of the transgenic avian
25 subjects of the invention.

It should be appreciated that in certain embodiments, the oligonucleotide/s or polynucleotide/s used by the kit/s and method/s of the invention are isolated and/or purified molecules. As used herein, "**isolated**" or "**purified**" when used in reference to a nucleic acid means that a naturally occurring sequence has been removed from its
30 normal cellular (e.g., chromosomal) environment or is synthesized in a non-natural environment (e.g., artificially synthesized). Thus, an "isolated" or "purified" sequence may be in a cell-free solution or placed in a different cellular environment. The term "purified" does not imply that the sequence is the only nucleotide present, but that it is

essentially free (about 90-95% pure) of non-nucleotide material naturally associated with it, and thus is distinguished from isolated chromosomes.

5 All scientific and technical terms used herein have meanings commonly used in the art unless otherwise specified. The definitions provided herein are to facilitate understanding of certain terms used frequently herein and are not meant to limit the scope of the present disclosure.

10 Before specific aspects and embodiments of the invention are described in detail, it is to be understood that this invention is not limited to particular methods, and experimental conditions described, as such methods and conditions may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

15 As used in this specification and the appended claims, the singular forms "a", "an", and "the" include plural references unless the context clearly dictates otherwise. Thus for example, references to "a method" includes one or more methods, and/or steps of the type described herein and/or which will become apparent to those persons skilled in the art upon reading this disclosure and so forth.

20

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are now described.

25

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps. More specifically, the terms "comprises", "comprising", "includes", "including", "having" and their conjugates mean "including but not limited to". This term

30

encompasses the terms "consisting of" and "consisting essentially of". The phrase "consisting essentially of" means that the composition or method may include additional ingredients and/or steps, but only if the additional ingredients and/or steps do not materially alter the basic and novel characteristics of the claimed composition or method.

The term "about" as used herein indicates values that may deviate up to 1%, more specifically 5%, more specifically 10%, more specifically 15%, and in some cases up to 20% higher or lower than the value referred to, the deviation range including integer values, and, if applicable, non-integer values as well, constituting a continuous range. As used herein the term "about" refers to $\pm 10\%$.

It should be noted that various embodiments of this invention may be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible sub ranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed sub ranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range. Whenever a numerical range is indicated herein, it is meant to include any cited numeral (fractional or integral) within the indicated range. The phrases "ranging/ranges between" a first indicate number and a second indicate number and "ranging/ranges from" a first indicate number "to" a second indicate number are used herein interchangeably and are meant to include the first and second indicated numbers and all the fractional and integral numerals there between.

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the methods and compositions of the invention, and are not intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by

- 73 -

weight, molecular weight is average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

5 The examples are representative of techniques employed by the inventors in carrying out aspects of the present invention. It should be appreciated that while these techniques are exemplary of preferred embodiments for the practice of the invention, those of skill in the art, in light of the present disclosure, will recognize that numerous modifications can be made without departing from the spirit and intended scope of the invention.

10 It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable sub combination or as suitable in any other described embodiment of the
15 invention. Certain features described in the context of various embodiments are not to be considered essential features of those embodiments, unless the embodiment is inoperative without those elements.

20 Various embodiments and aspects of the present invention as delineated hereinabove and as claimed in the claims section below find experimental support in the following examples.

25 Disclosed and described, it is to be understood that this invention is not limited to the particular examples, methods steps, and compositions disclosed herein as such methods steps and compositions may vary somewhat. It is also to be understood that the terminology used herein is used for the purpose of describing particular embodiments only and not intended to be limiting since the scope of the present invention will be limited only by the appended claims and equivalents thereof.

30 It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise.

EXAMPLES

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the claimed invention in any way.

Standard molecular biology protocols known in the art not specifically described herein are generally followed essentially as in Sambrook et al., *Molecular cloning: A laboratory manual*, Cold Springs Harbor Laboratory, New-York (1989,1992), and in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Maryland (1988).

Reagents**Animals:**

Commercial White Leghorn chickens are obtained from Hendrix ISA, and Minnesota. Marker Line chickens are from the Pacific Agri-Food Research Centre, Agassiz, British Columbia, Canada.

Animal experiments were done in strict accordance to IACUC approved protocols and under supervision of the Crystal Bioscience IACUC committee ensuring that no animal suffers from illness nor dies during the course of the experiments.

Vectors:

PrecisionX™ Cas9 SmartNuclease Vector System is ordered from System Bioscience Inc., catalogue number CAS8/9xx series.

pDsRed1-N1 from Clontech#6921-1.

GFP plasmid- FUGW-H1-GFP-neomycin (Addgene), Catalog number 37632.

Cell lines: Female cells are of *Gallus gallus*, lymphoblast female cell line Con A - C1 – VICK, ordered from ATCC® Number: CRL-12135™. For the chicken cell culture experiment chicken fibroblast female cell line UMNSAH/DF-1 (ATCC® CRL-

- 75 -

12203™) from ATCC were used. Male cells are of *Gallus gallus*, epithelial male cell line LMH/2A ordered from ATCC® Number: CRL-2118™.

Primordial Germ Cell (PGC) line.

human embryonic kidney 293 (HEK-293) cell line (ATTC).

5

Reagents and materials required for preparing the avian transgenic subjects

PGC Isolation

1. Microelectrode pipette puller (Shutter Instrument Co.).
2. Micro grinder (NARISHIGE).
- 10 3. Small-diameter (25 µm) glass micropipette.
4. Mouth-controlled pipette (Sigma-Aldrich).
5. 1x Hank's balanced salt solution (HBSS) without CaCl₂ or MgCl₂ (Hyclone, Logan, UT).
6. 1x Dulbecco's phosphate-buffered saline (PBS), without Ca₂ or Mg₂ (Hyclone).
- 15 7. Trypsin/ethylenediaminetetraacetic acid (EDTA) (Gibco).
8. Chicken embryos in Hamburger and Hamilton (HH) stages 14–17.
9. Magnetic-activated cell sorting (MACS) buffer. 1x PBS supplemented with 0.5% bovine serum albumin (BSA) and 2 mM EDTA.
10. MiniMACS column (Miltenyi Biotec).
- 20 11. Stage-specific embryonic antigen 1 (SSEA-1) antibody (Santa Cruz Biotechnology, Santa Cruz, CA).
12. Anti-mouse IgM MicroBeads (Miltenyi Biotec).

PGC Culture

- 25 1. PBS (Hyclone).
2. HBSS (Hyclone).
3. PGC culture medium (50 mL). 3.75 mL fetal bovine serum (FBS, Hyclone), 1.25 mL chicken serum (Sigma-Aldrich, St. Louis, MO), 500 µL GlutaMAX-I Supplement (100x, Invitrogen), 500 µL nucleosides (100x, Millipore), 500 µL antibiotic/antimycotic (ABAM; 100x, Invitrogen), 500 µL insulin-transferrin-selenium supplement (100x, Gibco), 500 µL nonessential amino acids (NEAA; 100x, Invitrogen), 50 µL β-mercaptoethanol (1000x, Sigma-Aldrich), 10 µL of 50 ng/µL human basic fibroblast
- 30

- 76 -

growth factor (bFGF, Sigma-Aldrich), and add Knockout Dulbecco's modified Eagle's medium (DMEM; Invitrogen) up to 50 mL. Store at 4 °C.

4. Accutase (Millipore).

5 Germline Chimera Production and Transgenesis

Gene Transfer into Chicken PGC

1. Lipofectamine 2000 (Invitrogen).
2. Opti-MEM medium (Invitrogen).
3. Transfection medium (PGC culture media without ABAM; 50 mL): 3.75 mL FBS (Hyclone), 1.25 mL chicken serum (Sigma-Aldrich), 500 µL GlutaMAX-I supplement (100x, Invitrogen), 500 µL nucleosides (100x, Millipore), 500 µL insulin-transferrin-selenium supplement (100x, Gibco), 500 µL NEAA (100x, Invitrogen), 50 µL β-mercaptoethanol (1000 x, Sigma-Aldrich), and 10 µL of 50 ng/µL bFGF (Sigma-Aldrich), and add Knockout DMEM (Invitrogen) up to 50 mL. Store at 4 °C.
4. PBS, without Ca² or Mg².
5. HBSS, without Ca² or Mg².
6. Prepare plasmid DNA at 1 µg/µL.

In Vitro Proliferation and Selection of Transgenic PGCs

1. HBSS, without Ca² or Mg² (Hyclone).
2. Hemocytometer.
3. Trypan blue, 0.4% (Invitrogen).
4. Inverted fluorescence microscope (Leica Microsystems).
5. Geneticin® Selective Antibiotic (Gibco).

PGC Transplantation into Recipient Eggs

1. Small-diameter (25 µm) glass micropipette (made with microelectrode pipette puller (Shutter Instrument Co.) and a micro grinder (NARISHIGE)).
2. Mouth-controlled pipette.
3. HBSS, without CaCl₂ or MgCl₂ (Hyclone).
4. Recipient chicken embryos at HH stages 14–17.
5. PBS, without Ca² or Mg² (Hyclone).
6. Pipette washing solution. 0.1% hydrogen peroxide (Sigma-Aldrich) in autoclaved distilled water.
7. Forceps.

- 77 -

8. Stereomicroscope with illuminator.
 9. Hot-melt glue sticks and glue gun.
 10. Parafilm.
 11. Donor PGC prepared at 3×10^3 cells/ μ L in HBSS.
- 5 Testcross
1. Wild-type Korean Ogye (KO) female chicken.
 2. 1 mL syringe.
 3. PBS, without Ca² or Mg² (Hyclone).
- 10 Donor-Derived Progeny and Transgenic Chicken Validation
1. DNeasy Blood & Tissue kit (Qiagen).
 2. Donor PGC-specific primer set.
 3. Transgene-specific primer set.
 4. PCR machine.
- 15 5. PCR reagents: PCR buffer, dNTP, Taq polymerase.
6. Agarose.
 7. TAE buffer.
 8. Electrophoresis machine.

20 *Experimental procedures*

Fluorescence assay

The complete setup of the experiment comprises a laser source, a holder or stand enabling positioning of the egg in an appropriate position allowing the exposure of cells expressing the reporting gene within the egg, to said laser source, a lens, a filter, a stand for the detector and a detector as illustrated in **Figure 1**. The distance between the sample and the lens and between the lens and the detector is about 20cm. The focal plane of the lens is about 10 cm and the filter is deposited there.

The optimal conditions were detected by placing the filter before the detector and by covering all the surrounding area of the detector to ensure that a scattering light will not bypass the filter.

Three types of lasers were used with initial power I_0 , and filters as indicate in the following **Table 2**:

Table 2: light source used:

Laser type (Wavelength)	Power [W]	Filters
Blue laser (473±1nm)	$I_o = 17.3mW$	above 500 nm
Green laser (532nm)	$I_o = 3.9mW$	540 nm-580nm
Red laser (632.8nm)	$I_o = 3.3mW$	above 660 nm

Injection of transfected GFP or RFP-expressing cells into eggs

5 The HEK cells were maintained in DMEM medium, supplemented with 1% L-glutamine, 10% fetal bovine serum (FBS), and 1% PenStrep (penicillin + streptomycin) (Biological industries, Israel). The cells were grown in a humidified incubator at 37°C with 5% CO₂. Medium was renewed 2-3 times a week, and cell count was kept between 1 x 10⁵ and 3 x 10⁵ cells/mL. Cells were subcultured and resuspended in fresh medium
10 as follow:

1. Supernatant was removed from the flask for the cells to adhere to the flask surface.
2. The dislodged cells were resuspended in 5 mL of fresh medium.

Followed the 2nd subculture, cell were transfected with 2 µg DNA of plasmid using PolyJet™ In Vitro DNA Transfection Reagent (SignaGen Laboratories, MD, USA)
15 according to manufacturer protocol. Two Plasmids were transfected into the cells separately; a GFP plasmid and a RFP plasmid, as specified above.

Following, transfected cells (RFP) were prepared at a concentration of 10⁶ cells/ml. A total volume of 100 µl of cells was then injected to fresh chicken egg at the center, according to **Tables 3 and 4**.

Table 3: RFP injection

Egg number	Egg condition	Saline (ul)	Cell line (ul) of 1M cells/ul	Cell line follow transfection of GFP or RFP
1.	Untouched	0	0	-----
2.	Small hole at top	100	0	0.1ml saline
11.	Small hole at top	97	3	,3000 cells + RFP
12.	Small hole at top	97	3	,3000 cells + RFP
13.	Small hole at top	90	10	,10000 cells + RFP
14.	Small hole at top	90	10	,10000 cells + RFP
15.	Small hole at top	70	30	,30000 cells + RFP
16.	Small hole at top	70	30	,30000 cells + RFP
17.	Small hole at top	0	100	,100000 cells + RFP
18.	Small hole at top	0	100	,100000 cells + RFP

In the second experiment cell were grow and transfected as described earlier. Cells were then transferred to either PBS or 50% glycerol (to control disperse) and injected into fresh chicken egg at the center. The different injected eggs are disclosed by **Table 4**.

Table 4: RFP or GFP injection

Egg number	Egg condition	PBS (ul) or 50% glycerol	Cell line (ul) of 1M cells/ul	Cell line follow of transfection of GFP or RFP
1.	Untouched	0		Control-NON transfected
2.	Small hole at top	100	0	0.2 ml PBS
3.	Small hole at top	100	0	0.2 ml 50% Glycerol
4.	Small hole at top	99	1	1000 cells in PBS
5.	Small hole at top	99	1	1000 cells in 50% Gly
6.	Small hole at top	97	3	3000 cells in PBS
7.	Small hole at top	97	3	3000 cells in 50% Gly
8.	Small hole at top	90	10	10000 cells in PBS
9.	Small hole at top	90	10	10000 cells in 50% Gly
10.	Small hole at top	70	30	30000 cells in PBS
11.	Small hole at top	70	30	30000 cells in 50% Gly
12.	Small hole at top	99	1	1000 cells in PBS
13.	Small hole at top	99	1	1000 cells in 50% Gly
14.	Small hole at top	97	3	3000 cells in PBS

- 81 -

15.	Small hole at top	97	3	3000 cells in 50% Gly
16.	Small hole at top	90	10	10000 cells in PBS
17.	Small hole at top	90	10	10000 cells in 50% Gly
18.	Small hole at top	70	30	30000 cells in PBS
19.	Small hole at top	70	30	30000 cells in 50% Gly

Male chicken cell line

The Male cells, specifically, *Gallus gallus*, chicken (liver; chemically induced; transfected with chicken estrogen receptor gene), were grown in a waymouth's MB 752/1 medium supplemented with 1% L-glutamine, 10% fetal bovine serum (FBS), and 1% PenStrep (penicillin + streptomycin) (Biological industries, Israel). Culture flask (T25) was coated with 0.1% gelatin). Medium was renewed every 3 days. The cells were grown in a humidified incubator at 37°C with 5% CO₂.

10 Female chicken cell line

Chicken fibroblast female cell line **UMNSAH/DF-1** (ATCC® CRL-12203™) from ATCC were used. More details on the cell is presented in Table 5:

Table 5: Female chicken cell line

Organism	<i>Gallus gallus</i> , chicken
Cell type	fibroblast spontaneously transformed
Product Format	frozen
Morphology	fibroblast
Culture Properties	adherent
Age	10 days gestation
Gender	Female

The female cells were grown in DMEM medium supplemented with 1% L-glutamine, 10% fetal bovine serum (FBS), and 1% PenStrep (penicillin + streptomycin) (Biological industries, Israel). The cells were grown in a humidified incubator at 37°C with 5% CO₂. Medium was renewed 2-3 times a week.

5 PGCs cell culture

Embryonic gonads are retrieved from White Leghorn (WL) embryos at 6.5 days of incubation and then the retrieved gonadal cells are treated with magnetic activated cell sorter (MACS) for separation of PGCs in total embryonic gonadal cells. PGCs are grown in KO-DMEM (Life Technologies), of which 40% is preconditioned on buffalo rat liver cells (BRL, ATCC), and supplemented with 7.5% fetal bovine serum (Hyclone), 2.5% irradiated chicken serum, 1X non-essential amino acids, 2mM glutamine, 1mM sodium pyruvate, 0.1mMβ-mercaptoethanol (all from Life Technologies), 4ng/ml recombinant human fibroblast growth factor, 6ng/ml recombinant mouse stem cell factor (both from R&D Systems) and grow on an irradiated feeder layer of BRL cells. The cells are passaged 3 times per week onto fresh feeder layers.

Restriction-free (RF) cloning

The insertion of gRNAs into Cas9-SmartNucleaseTM vector is performed by applying the Restriction Free method [Peleg Y et al., J. Struct. Biol. 172(1):34-44 (2010) 2010]. Primers are ordered from Sigma-Genosys (Rehovot, Israel) and subsequent RF reactions were carried out using Phusion polymerase (Thermo Scientific, Hudson, NH, USA). Plasmid purification is carried out using the MEGAspin kit and DNA-spin plasmid DNA purification kit, respectively (Intron Biotechnology, Daejeon, South Korea).

Transfection into chromosomes W or Z of chickens

25 Cell are co-transfected with 2 µg DNA of each plasmid (see details below) by using PolyJetTM. In Vitro DNA Transfection Reagent (SignaGen Laboratories, MD, USA) according to manufacturer protocol. Co-transfection of two plasmid is performed as follow:

1. pDsRed with chicken Chromosome Z Left and Right arms or Chromosome W Left and Right arms and CMV-hspCas9-H1-gRNA (assay).
2. pCMVGluc with chicken Chromosome Z Left & Right arms or Chromosome W Left & Right arms and CMV-hspCas9-H1-gRNA (control).

Cells are then plated with Neomycin-resistant and seeded in a 48-well plate at a density of 10^5 cells per well. After 3 days, 40 μ g/ml Neomycin is added to select for cells with a stable integration of the reporter genes.

5 DNA analysis for integration into chromosome Z or W

For verification of chromosomal integration, the following strategy was used. Cells were isolated and DNA of half of them was extracted and run on an agar gel (1%). The genomic DNA (leaving the plasmid DNA on the gel) was cut and cleaned. A PCR assay was run with the following primers. The products were separated on a gel and then
10 sequenced. Since the insert was too long to be detected by the PCR assay, sliding window PCR assay was designed that embark on initial PCR product.

a. Assay for cell line with the insert (includes part of the Left arm, CMV Enhancer, Promotor and MCS and part of the dsRED2):

Left primer-TTGGTGTGTGCTAATAGGCAGT as denoted by SEQ ID NO 49; right
15 primer-TAGTCCTCGTTGTGGGAGGT as denoted by SEQ ID NO 50, the product size: 1489bp.

b. Assay for cell line with the insert (includes part of the flanking (to the arms) Z chromosome, Left arm, and part of the CMV Enhancer, Promotor):

Left primer- GCATTGATCTGTCCAGTTGC as denoted by SEQ ID NO 51, right
20 primer- TACTGCCAAAACCGCATCAC as denoted by SEQ ID NO 52, product size: 1462bp.

Preparation of transgenic chickens

Concentrated vehicle (that may be either lentivirus at a titer of about 10^7 MOI) or
25 plasmid DNA) is injected to 25 embryos in new laid eggs. Injections are carried out weekly three injections. The injected embryos hatch 3 weeks after injection. These are G0 birds. Immediately after hatch, the DNA is extracted from CAM samples of the hatched chicks and detection of the presence/absence of vector DNA is carried out by semi-quantitative PCR. Blood sample G0 chicks at 2-3 weeks of age and repeat PCR
30 screen. G0 birds are raised to sexual maturity, 16-20 weeks for males, 20-24 weeks for females. Cockerels are tested for semen production from approximately 16 weeks.

Hens are inseminated, fertile eggs collected daily. The G1 chicks are hatch 3 weeks later and each individual chick wing banded and a chick chorioallantoic membrane

- 84 -

(CAM) sample taken from the shell. Extract DNA from CAM samples and carry out PCR screen for presence of transgene, predicted to be single copy level. Repeat screen to confirm and sex chicks on DNA from blood sample 2-3 weeks later.

At a few weeks of age a blood sample is taken from G1 birds to prepare genomic DNA for PCR analysis. G1 birds are used for breeding G2.

PGC Isolation

1. Incubate fresh chicken eggs (EGK stage X), for example, Korean Ogye (KO), at 37 °C for 50–54 h (HH stage 14–17) for blood PGC isolation. Place incubated eggs horizontally on the egg plate and gently wipe eggshell using sanitized cotton with 70% ethanol.
2. Using sharpened forceps, cautiously crack the eggshell (<1 cm diameter) for the isolation of whole blood cells.
3. Collect about 2–3 µL whole blood cells from the embryonic dorsal aorta using a 25 µm thinly ground glass needle and mouth pipette. Mix with 500 µL PBS. For microinjection of PGCs, a mouth pipette and a glass micropipette of 20–25 µm diameter may be used. The glass micropipette is made with a microelectrode pipette puller (Shutter Instrument Co.) and ground at 25 degrees using a micro grinder (NARISHIGE).
4. Transfer whole-blood cells to a 1.5 mL sterile tube, centrifuge (250 x g, 5 min), and remove the supernatant.
5. For gonadal PGC isolation, incubate fresh chicken eggs at 37 °C for 5.5 days (HH 20–26). Extract embryonic gonads from embryos at HH stages 20–26 with sharpened forceps, and incubate with 500 µL of 0.05% trypsin/EDTA at 37 °C incubator for 5 min.
6. Add 50 µL FBS for inactivation, and centrifuge (250 x g, 5 min) and remove the supernatant.
7. Resuspend whole blood cells or dissociated gonadal cells in 1 mL PBS and apply anti-SSEA-1 antibody (Santa Cruz Biotechnology, SC-21702) at 1:200 dilution, and then incubate the mixture for 15 min at room temperature (RT).
8. Wash cells to remove unbound primary antibody by adding 5 mL MACS buffer (0.5% BSA and 2 mM EDTA in PBS, pH 7.2) per 10⁷ total cells and centrifuge (250 x g, 5 min).

9. Aspirate supernatant completely and re-suspend cell pellet in 80 μ L MACS buffer per 10^7 total cells.
10. Add 20 μ L rat anti-mouse IgM MicroBeads (Miltenyi Biotec. 130-047-301) per 10^7 total cells. For higher cell numbers, scale up buffer volume accordingly.
- 5 11. Incubate cells with antibody at 2–8 °C for 20 min.
12. Wash cells by adding 2 mL MACS buffer per 10^7 total cells and centrifuge (250 x g, 5 min).
13. Aspirate the supernatant completely and re-suspend up to 10^8 total cells in 500 μ L MACS buffer.
- 10 14. Place column in the magnetic field of a suitable MACS separator.
15. Prepare column by rinsing with 500 μ L MACS buffer.
16. Apply cell suspension onto the column, and wash the column with 500 μ L buffer three times. Add new buffer when the column reservoir is empty.
17. Remove column from the separator and place it on a suitable collection tube.
- 15 18. Add 1 mL MACS buffer to the column and immediately flush out the magnetically labeled cells by firmly pushing the plunger into the column.
19. Transfer the isolated cells containing PGCs to the pre-warmed PGC culture medium (1 mL PGC medium per 1×10^5 purified PGCs).

20 **PGC Culture**

1. Transfer the isolated cells containing PGCs to the pre-warmed PGC culture medium (1 mL PGC medium per 1×10^5 purified PGCs) and incubate at 37°C.
2. After 7–14 days of growth, most of the primary cultured PGCs form colonies and lose aggregates of cell colonies.
- 25 3. The suspended PGC colonies can be withdrawn gently with medium after gentle pipetting and centrifuging (200 x g, 5 min).
4. Disaggregate the cell pellet with Accutase (1 mL Accutase solution per $\sim 5 \times 10^5$ PGCs).
5. Centrifuge (200 x g, 5 min) and re-suspend the cell pellet in a PGC culture medium.
- 30 6. Seed the suspended PGCs in a 12-well plate (1×10^5 cultured PGCs in 1 mL PGC culture medium per well of a 12-well plate).
7. PGCs can be routinely subcultured every 3–4 days.

Gene Transfer into Chicken PGCs, Germline Chimera Production, and Transgenesis

Gene Transfer into Chicken PGCs

1. Mix 4 µg helper plasmid containing CAGG-PBase (pCyL43B) and 6 µg the piggyBac transposon (pCyL50) containing the reporter gene (RFP) and optionally, selectable marker (e.g., neomycin resistance gene) with 100 µL Opti-MEM (Gibco) and incubate for 5 min at RT. For CRISPR/Cas transfection, mix 2.5 µg CMVRFP expression plasmid vector, 2.5 µg Cas9 and guide RNA expression plasmid vectors with 100 µL Opti-MEM (Gibco) and incubate for 5 min at RT.
2. Mix 10 µL Lipofectamine 2000 reagent (Thermo Fisher-Invitrogen) with 100 µL Opti-MEM and incubate for 5 min at RT.
3. Mix the plasmids with Opti-MEM and Lipofectamine 2000 reagent with Opti-MEM and incubate for 20 min at RT.
4. During incubation, harvest cultured PGCs and centrifuge (200 x g, 5 min). Discard the supernatants.
5. Add 1 mL Accutase (Millipore) to the harvested PGCs and incubate for 10 min at 37 °C.
6. Determine the number of PGCs using a hemocytometer (Marienfeld) and seed 5 x 10⁵ PGCs in a 12-well plate with 1 mL PGC culture medium without antibiotics.
7. Apply DNA-Lipofectamine complex to PGCs and incubate for 1 day at 37 °C in a CO₂ incubator.
8. Harvest the transfected PGCs and centrifuge (200 x g, 5 min). Remove the supernatants.
9. Wash the PGCs with 1 mL HBSS three times and suspend them with the PGC culture medium with antibiotics.

In Vitro Proliferation and Selection of Transgenic PGCs

1. Select the transfected PGCs in the PGC culture medium containing 100 µg/mL G418 (for the neomycin resistance gene) the day after transfection.
2. Monitor reporter gene expression using fluorescence microscopy.
3. Subculture the PGCs every 3–4 days. A complete selection period requires up to 3 weeks.
4. If there is no drug-selection marker in the transfected plasmid vectors, PGCs that express fluorescent protein can be sorted by FACS.

PGC Transplantation into Recipient Eggs

1. Incubate recipient eggs up to HH stages 14–17 at 37 °C in air with 60–70% relative humidity.
- 5 2. Harvest the transfected PGCs and centrifuge (200 x g, 5 min) and remove the supernatants.
3. Add 1 mL Accutase (Millipore) to the harvested PGCs and incubate for 10 min at 37 °C.
4. Centrifuge (200 x g, 5 min). Discard the supernatants.
- 10 5. Suspend the PGCs in HBSS (Hyclone).
6. Make a small window on the pointed end of the recipient egg and microinject a 2 µL aliquot containing more than 3,000 PGCs with a micropipette into the dorsal aorta of the recipient embryo.
7. Seal the egg window of the recipient embryo with parafilm using the glue gun, and
- 15 incubate the egg with the pointed end down until hatching at 37 °C in air with 60–70% relative humidity.

Monitoring Genetically Modified PGCs in Embryonic Gonads

1. Incubate recipient eggs up to HH stages 28–30 at 37 °C in air with 60–70% relative humidity.
- 20 2. Dissect the gonad at embryonic day 6.
3. Monitor fluorescent protein expression in the embryonic gonads using fluorescence microscopy.

Testcross

1. Collect semen twice in 1 week from sexually mature recipients and wild-type (WT)
- 25 chickens.
2. Introduce 50 µL semen from mature male recipients and WT roosters to WT laying hens and mature female recipients, respectively.
3. Collect eggs from WT laying hens and mature female recipients the day after artificial insemination, and incubate the egg with the pointed end down until hatching at
- 30 37 °C in air with 60–70% relative humidity.

EXAMPLE 1***Fluorescence analysis using different colored-lasers and filters for visual gender identification in poultry*****Optical fluorescent system**

5 In order to demonstrate the feasibility of visually identify gender of in-ovo poultry, the use of green fluorescence as compared to blue fluorescence or red fluorescence reporter genes was evaluated.

The autofluorescence of complete eggs or of eggs separated into egg white, egg yolk and shell was determined using a blue laser, a green laser or a red laser. Comparison of
10 an empty egg (shell) to a complete egg indicates no significant difference in autofluorescence level (data not shown).

The scattering and autofluorescence intensity provided by the different components and parts of the egg is described in **Table 6**.

15 Importantly, it was noted that the background noise (without laser) was about 4nW, while in the presence of the laser (532nm), it was about 12nW due to light scattering that bypass the filter (low pass filter that passes above 660nm).

Table 6. Scattering and autofluorescence intensity of the parts of the egg using 3 different fluorescent lasers

Laser	Scattering Intensity			Autofluorescence Intensity		
	Empty container	Container + egg white	Container + egg yellow	Empty container	Container + egg white	Container +egg yellow
Blue laser (473±1nm)	2mW	1.02mW	1.2mW	16.5nW	26.12±6.25nW	10.33±1.5nW
Green laser (532nm)	370µW	484µW	415µW	13.6nW	14.5nW	16nW
Red laser (632.8nm)	300µW	160µW	156µW	13nW	8.33±1.86nW	10.82±3.42nW

20

The use of a blue laser (473 nm) together with a green (+500 nm) or red filter (590-650 nm) with or without addition of different concentrations of fluorescein (1µM, 10mM) into a complete egg was assayed and is summarized in **Figure 2**. It appears that there is

a large green autofluorescence, however fluorescence intensity could be detected upon using high fluorescein concentration.

5 The use of a green laser (532 nm) together with a dark green (540-580 nm) or red filter (+660 nm) with or without addition of Rhodamine B (10 μ M) into a complete egg was assayed and is summarized in **Figure 3**. It appears that above 40mW, the fluorescent emission of the dye is considerably detectable with the green filter (almost 2 times larger). However, with the red filter, no difference could be observed.

10 The use of a green laser (532 nm) together with a red filter A (590-650 nm) or a red filter B (+660 nm) with or without addition of dir (10 μ M) into a complete egg was assayed and is summarized in **Figure 4**. It appears that above 80mW, the fluorescent emission of the dye is detectable using Red filter A while by using the Red filter B (+660nm), only a modest difference can be observed.

15 The use of a red laser (632.8 nm) together with a red filter (+660 nm) or a green filter (540-580 nm) with or without addition of dir (10 μ M) into a complete egg was assayed and is summarized in **Figure 5**. Almost no intensity difference was shown with or without fluorescent dye.

All of the experiments were repeated using only the egg shell. Similar results were
20 observed (data not shown).

These observations demonstrate the feasibility of detecting RFP through the egg shell.

EXAMPLE 2

Fluorescence of RFP or GFP transfected cells into eggs

25 To illustrate the feasibility of *in-ovo* detection of RFP expressing cells in an embryo, the inventors next examined if cells transfected with a fluorescent protein, specifically, RFP or GFP, can be detected upon injection thereof into a complete egg.

More specifically, HEK cells were transfected either with GFP or RFP vectors as
30 described in in Experimental procedures. The fluorescence of both GFP and RFP transfected cells was examined following excitation. **Figure 6** (GFP) and **Figure 7** (RFP) clearly demonstrate that both GFP and RFP are expressed by the transfected cells and fluorescent signals can be observed.

HEK transfected cells were then injected into fresh eggs as described in Experimental procedures and the fluorescence intensity of eggs was assayed.

5 First, the fluorescence intensity of eggs with or without the Red Fluorescent Protein RFP-expressing cells was characterized by excitation of the eggs with the green laser (532nm) and a Red filter of +650nm. **Figure 8** shows that the fluorescence intensity of egg without RFP measured with five **different positions of the egg** is quite similar. **Figure 9** shows the measured fluorescence intensity using different concentrations of RFP-expressing cells when the egg was excited at the center. The correlation between
10 the RFP concentration and the intensity on the detector was clearly observed.

Several tests were performed in order to check if in the presence of RFP expressing cells, the position of the egg, i.e. the exact place of excitation provided by the laser, was influencing the fluorescence intensity. It appears that the intensity on the detector greatly changed with the position of the egg during excitation. This was expected since,
15 following injection of RFP-expressing cells into the egg, RFP is not uniformly distributed. **Figure 10** presents the maximum intensity observed at a specific position of the egg during excitation, which corresponds to the position where the concentration of RFP-expressing cells is the highest. This figure demonstrate that a minimal amount of about 3000 cells can be detected.

20 In embryo, it is known in the art that gender cells, PGCs are positioned in the center of the epiblast of freshly laid egg at developmental stage X of the egg. It should be noted that the embryo always facing the upper side of the egg yolk sack.

The results exposed above suggest that upon excitation of the region situated few millimeters behind the egg shall and nearby the yolk, maximal RFP emission may be
25 detected.

The experiments were repeated one week following injection of the eggs with RFP expressing cells. The results were similar demonstrating that the RFP dye did not diffuse inside the egg.

30

The influence of PBS and glycerol on the fluorescence intensity of eggs using the green laser (532 nm) and the red filter of +650 nm was examined with or without the presence of RFP-expressing cells. It appears that neither glycerol nor PBS affect fluorescence

intensity measurements of RFP expressing cells into eggs as illustrated in **Figure 11A and 11B**.

Following, the influence of PBS and glycerol on the fluorescence intensity of eggs using the blue laser (483 nm) and the filter of +500 nm was examined with or without the presence of GFP-expressing cells. It appears that GFP-expressing cells could not be detected in any given concentration (1000, 3000 10,000 and 30,000 cells), as illustrated by **Figure 12** for example with 30,000 GFP expressing cells. The same results were obtained using three different filters: +500 nm, 530-550 nm and 540-580 nm.

10

To further evaluate the surprising advantage of using RFP as a reporter gene, the fluorescence of GFP and RFP-expressing cells were finally compared as illustrated in **Figure 13**. The parameter R was defined to indicate the ratio between the fluorescence protein (FP) intensity (I_{FP}) and the autofluorescence intensity ($I_{\text{autofluorescence}}$):

$$R = \frac{I_{FP}}{I_{\text{autofluorescence}}}$$

15

The R parameter was calculated for different concentrations of both GFP and RFP-expressing cells. As indicated by the Figure, it appears that GFP-expressing cells are not detectable in any of the given concentration (1000 and 30,000 cells as shown in **Figure 13**).

On the other hand, RFP-expressing cells provided significant R parameter values even at low concentration (1000 RFP-expressing cells). At high concentration of RFP-expressing cells (30000 RFP-expressing cells), the R value reached 3, and thus clearly demonstrating that the RFP reporter gene can be easily detected in excited eggs.

The main findings are summarized below:

1. There is a strong autofluorescence at the blue while there is no autofluorescence in the green and red.
2. The main autofluorescence originates from the eggshell.
3. The system can detect green and red fluorescence dyes with relatively low dyes concentration and with low excitation power.

30

- 92 -

4. Blue fluorescence is hardly detected (only in mM dyes concentration) due to the autofluorescence.
5. RFP is much easily and significantly detectable inside the eggs.
6. GFP type is not detectable in this system.
- 5 7. The exact location of the RFP-expressing cells inside the egg is significant to define the region of excitation.
8. The RFP dye did not diffuse into the eggs during 14 days from injection.
9. Nor PBS or Glycerol affects RFP detection.

These results clearly demonstrate that an embryo expressing a RFP gene integrated in a
 10 gender chromosome, can be detected in early developmental stage, where about 1000 cells express RFP.

EXAMPLE 3

Design of guide RNAs vector

15 In order to incorporate the RFP reporter gene into the gender chromosomes W or Z, the CRISPR/Cas9 mediated HDR method is selected. Relevant gRNA sites are then sought from both gender chromosomes.

Female Z chromosome integration

The regions 9156874-9161874, as denoted by SEQ ID NO:15, 27764943-27769943, as
 20 denoted by SEQ ID NO:16, 42172748-42177748, as denoted by SEQ ID NO:17, 63363656-63368656, as denoted by SEQ ID NO:18 and 78777477-78782477, as denoted by SEQ ID NO:19 of Chromosome Z of female chicken are analyzed for guide RNA design. More specifically, for directing the integration of the RFP into Z chromosome, a gRNA designated gRNA7 of Z chromosome locus chrZ_42174515_-1,
 25 comprising the nucleic acid sequence GTAATACAGAGCTAAACCAG, as also denoted by SEQ ID NO:26, was next prepared.

The procedure was performed according to the “PrecisionV Cas9 SmartNuclease Vector System” user manual. More specifically, two primers were designed for the cloning:

30 ChZgRNA_Forward: TGTATGAGACCACGTAATACAGAGCTAAACCAG, as denoted by SEQ ID NO. 53, ChZgRNA_Reverse: AAACCTGGTTAGCTCTGTATTACGTGGTCTCA, as denoted by SEQ ID NO. 54.

- 93 -

The primers were annealed to generate a duplex in a reaction mix that contains 5 μ M of each primer. The mixture was incubated at 95°C for 5 minutes and then removed to cool down to RT.

Later the duplex was cloned into the provided linearized Cas9 vector by ligation reaction that was set as follows: 1 μ l of the linearized vector, 3 μ l of the annealed oligo mix, 1 μ l of ligation buffer and 0.25 μ l of the Fast ligase. The mixture was incubated for 5-7 minutes at 25°C.

The mixture was transformed into DH5 α competent cells and the cells were cultured on an LB plate containing 50 μ g/ml Kanamycin and incubated overnight at 37°C.

10 The next day, two colonies were picked randomly and were grown in LB/Kanamycin medium overnight at 37°C with shaking.

The next day, plasmid DNAs were prepared using Qiagen mini-prep kit and were sequenced using the provided sequencing primer.

In yet some further embodiments, four additional guide RNAs are selected, synthesized and cloned separately into the Cas9 SmartNuclease vector containing the wild type Cas9 nuclease (Horizon) by Restriction free cloning protocol: gRNA3: ACAGACCTATGATATGT, as denoted by SEQ ID NO. 11; gRNA4: CGATTATCACTCACAAG, as denoted by SEQ ID NO. 12; gRNA5: CTGGTTAGCATGGGGAC, as denoted by SEQ ID NO. 13; gRNA6: 20 GTAAAGAGTCAGATACA, as denoted by SEQ ID NO. 14.

Further non-limiting examples for gRNA sequences suitable for integration into specific loci within the Z chromosome, may include but are not limited to gRNA8 of Z chromosome locus chrZ_9157091_1, comprising the nucleic acid sequence ACAGACCTATGATATGTGAG, as also denoted by SEQ ID NO:27, gRNA9 of Z chromosome locus chrZ_27767602_-1, comprising the nucleic acid sequence 25 GAGCTTGTGAGTGATAATCG, as also denoted by SEQ ID NO:28, gRNA10 of Z chromosome locus chrZ_78779927_1, comprising the nucleic acid sequence GTAAAGAGTCAGATACACAG, as also denoted by SEQ ID NO: 29, and gRNA11 of Z chromosome locus chrZ_63364946_-1, comprising the nucleic acid sequence 30 CAGTGGGTACTGAAGCTGTG as also denoted by SEQ ID NO: 30.

These gRNAs have few predicted off-target sites, none of which were in known coding sequences.

Female W chromosome integration

The region 1022859-1024215 of Chromosome W of female chicken, comprising the nucleic acid sequence as denoted by SEQ ID NO. 3, is analyzed for guide RNA design. Two guide RNAs are selected, synthesized and cloned separately into the Cas9 SmartNuclease vector containing the wild type Cas9 nuclease (Horizon) by Restriction free cloning protocol: gRNA1: GCACTAGGAACCAGCAGCAG, as denoted by SEQ ID NO. 1 and gRNA2: GTAGCCCCAAGAGGGCTAGG, as denoted by SEQ ID NO. 2.

The predicted parameters of these two gRNAs are presented in Table 7:

10

Table 7

gRNA parameters	gRNA1	gRNA2
sgRNA designer	0.506	0.63
sscore	0.8677	0.5323
sgRNA scorer	94.8	99.9

EXAMPLE 415 ***Design of RFP targeting vector***

Flanking sequences homologous of the appropriate flanking sequences indicated above of female W chromosome or of the female Z chromosome loci, are introduced into the RFP-expressing vector, specifically, the pDsRed1-N1 plasmid.

20 For, integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 31, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 32, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA7 of SEQ ID NO:26.

25

More specifically, to integrate the reporter gene of the invention to the specific loci directed by gRNA7, specifically, GTAATACAGAGCTAAACCAG, as denoted by

- 95 -

SEQ ID NO:26, cloning of the 'Left arm' was performed upstream the CMV enhancer in pDsRed1-N1 plasmid, using the RF methodology.

The following primers were used:

Forward primer:

- 5 ACGGGGTCATTAGTTCATAGCCCATATATGGTGAACGGATAGGCAGCAAGC,
as denoted by SEQ ID NO. 55 (contains the 5' Left arm sequence and 30 nucleotides of the site of integration in the plasmid).

Reverse primer:

- GGCGGGCCATTTACCGTAAGTTATGTAACGTGAGGAAGGGTCTGTTACTGG
10 A, as denoted by SEQ ID NO. 56.

The Right arm was cloned downstream the Neo resistance gene, using the following primers.

Forward primer:

- 15 TTCTATCGCCTTCTTGACGAGTTCTTCTGAGTCATGTTCGGTGGAGGAGAAA,
as denoted by SEQ ID NO. 57 (3' Neo resistance sequence marked in green)

Reverse primer:

- GTGATGGCAGGTTGGGCGTCGCTTGGTTCGGGCATGGGATGTTAAAGAGAAG
CT, as denoted by SEQ ID NO. 58 (In orange vector sequences 3' to the Neo gene).

- 20 Cloning was verified by sequencing to ensure proper integration.

Note that there are 193 nucleotides deletion from the CMV enhancer due to repetition of 16 nucleotides (CGGTAAATGGCCCGCC, as denoted by SEQ ID NO. 59) in the enhancer region flanking the site of integration.

- 25 In yet some further embodiments, the RFP reporter gene may be cloned for using either the gRNA3, as denoted by SEQ ID NO. 11, gRNA4 : as denoted by SEQ ID NO. 12, gRNA5, as denoted by SEQ ID NO. 13, gRNA6, as denoted by SEQ ID NO. 14.

- For cloning using the gRNA3, "Left arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 41, and the "Right arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 42 are provided. For cloning using the gRNA4, "Left arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 43, and the "Right arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 44 are provided. For cloning using the gRNA5, "Left arm" comprising the nucleic acid sequence as
- 30

denoted by SEQ ID NO. 45, and the "Right arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 46 are provided. For cloning using the gRNA6, "Left arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 47, and the "Right arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 48 are provided.

5 In further embodiments, for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 33, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 34, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA8 of SEQ ID NO:27. In still further embodiments,
10 for integrating the reporter gene of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 35, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 36, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA9 of SEQ ID NO:28. In further embodiments, for integrating the reporter gene
15 of the invention into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 37, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 38, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA10 of SEQ ID NO:29. In yet a further embodiment, for integrating the reporter gene of the invention
20 into the specific locus within the Z chromosome, left arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 39, and right arm comprising the nucleic acid sequence as denoted by SEQ ID NO. 40, may be used to integrate the reporter gene of the invention to the specific loci directed by gRNA11 of SEQ ID NO:30.

25 For the female W chromosome, the reporter gene, specifically RFP may be cloned for using either the Guide 1 (gRNA1), as denoted by SEQ ID NO. 1 or Guide 2 (gRNA2): as denoted by SEQ ID NO. 2. For cloning using the gRNA1, "Left arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 4, and the "Right arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 5 are provided. For cloning using the
30 gRNA2, "Left arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 6, and the "Right arm" comprising the nucleic acid sequence as denoted by SEQ ID NO. 7 are provided.

Still further, a "left arm" for the region upstream to the CMV-promoter comprises the nucleic acid sequence as denoted by SEQ ID NO. 8, and a "right arm" for the region downstream the Neomycin-resistance, may comprise the nucleic acid sequence as denoted by SEQ ID NO. 9, or SEQ ID NO.10 for the region downstream the polyA site.

5

EXAMPLE 5

Integration of the RFP reporter gene into the Z chromosome of female chicken cell line

Female chicken cell line was co-transfected with two plasmids i.e. pDsRed containing the chicken Chromosome Z left and right arms and the plasmid containing the CMV-hspCas9-1-gRNA as detailed in the Experimental procedure section.

10

The employed gRNA sequence was the gRNA 7 as denoted by **SEQ ID NO: 26**. **Table 8** provides more details on the used guide RNA.

15 **Table 8. Characteristics of the used gRNA 7**

protospacer gRNA	strand	cut site	chromosome	guideId
TTAAACAATC AAACCA, SEQ ID NO. 26	-1	17 515	chrZ	NS0000001009 NS0000001008chrZ00000017515-1

20

Transfected cells were then selected with red fluorescent. As can be seen in **Figure 14** the transfection was positive. The fluorescent cells were collected, the genomic DNA was extracted and the sliding PCR assay as described in the Experimental procedures was performed. PCR fragments were then sent for Sanger sequencing and the integrated sequences as illustrated in **Figure 15** were obtained, which demonstrate successful integration of the RFP gene into the Z chromosome.

In conclusion, these results demonstrate that

25

- successful growing procedure of female chicken cell line was established
- successful transfection of the RFP reporter gene and CRISPR Cas9 system into female chicken cell line was established
- a strong fluorescence signal was generated in red excitation

- successful point integration of the RFP reporter gene into the Z chromosome was established

EXAMPLE 6

5 *Germline transmission of CRISPR-treated cells*

The use of primordial germ cells (PGCs) to produce transgenic chickens has many advantages in animal biotechnology and biomodeling. Because chicken embryos are oviparous, PGCs at an early stage are readily accessible and can be manipulated in vitro for practical applications, including the restoration of genetic material, genome editing, and transgenic research. Significant efforts have been made to establish culture systems for chicken PGCs from different embryonic origins, and it has been demonstrated that only chicken PGCs can be maintained indefinitely in vitro without losing their properties.

15 More specifically, in chickens, PGCs first separate in the epiblast in Eyal-Giladi and Kochav (EGK) stage X embryos and then move down the hypoblast of the area pellucida, and subsequently to the germinal crescent and enter the blood stream. After migration through the circulatory system, they arrive at the genital ridges.

These unique characteristics in germline development make it possible to use them to produce transgenic chickens via the injection of genetically manipulated PGCs into the blood vessels of recipient eggs.

The two above described vectors, specifically, the gRNA/Cas9 and the reporter-gene vectors are co-transfected to PGCs as detailed in experimental procedures. After stable clones are identified, the cells are expanded and confirmed for the RFP integration by PCR. Confirmed clones are injected into recipient chicken embryos at Stage 14–16 (H&H). The injected embryos are transferred to surrogate shells and incubated until hatch at 37°C. The sex of the chicks is determined after hatch by PCR for the W- or Z chromosomes.

30

Female and Male chimeras are grown to sexual maturity and bred to wild type male and female chickens. Hatched chicks are evaluated for the expression of RFP, and the germline progeny are confirmed by PCR to carry targeted RFP.

CLAIMS

1. A method of gender determination of avian fertilized unhatched egg, the method comprising the step of:
- 5 (a) providing at least one transgenic avian subject comprising at least one exogenous reporter gene integrated into at least one position or location in at least one of gender chromosome Z and W, wherein said reporter gene encodes a protein having an excitation wavelength of 500-650 nm and an emission wavelength of 550-650 nm;
- 10 (b) obtaining at least one fertilized egg from said transgenic avian subject, or of any cells thereof;
- (c) determining in said egg if at least one detectable signal is detected, wherein detection of said at least one detectable signal indicates the expression of said at least one reporter gene, thereby the presence of said Z chromosome or W chromosome in an avian embryo comprised within said fertilized unhatched egg.
- 15
2. The method according to claim 1, wherein said reporter gene is at least one fluorescent reporter gene.
3. The method according to claim 1 or 2, wherein said reporter gene is Red
- 20 Fluorescent Protein (RFP).
4. The method according to any one of claim 1 to 3, wherein said determining if a detectable signal is detected comprises the step of subjecting said egg to a light source.
- 25 5. The method according to any one of claims 1 to 4, wherein said at least one transgenic avian subject is a female avian subject, and wherein said at least one reporter gene is integrated into at least one position of: (a) female chromosome Z, thereby detection

of a detectable signal indicates that said embryo in said unhatched egg is a male; or (b) female chromosome W, thereby detection of a detectable signal indicates that said embryo in said unhatched egg is a female.

5 6. The method according to any one of claims 1 to 5, wherein said at least one reporter gene is integrated into said gender chromosome of said transgenic avian animal using at least one programmable engineered nuclease (PEN).

7. The method according to claim 6, wherein said PEN is a clustered regularly interspaced short palindromic repeat (CRISPR) type II system.

8. The method according to any one of claims 1 to 7, wherein said at least one reporter gene is integrated into a gender chromosome of said transgenic avian subject by contacting or co-transfecting at least one cell of said avian subject or at least one cell introduced into said avian animal, with:

- 15 (a) at least one Cas9 protein or at least one first nucleic acid sequence comprising at least one nucleic acid sequence encoding said at least one Cas9 protein; and guide RNA (gRNA) that targets at least one protospacer within at least one gender chromosome Z and/or W, or at least one nucleic acid sequence
20 encoding said at least one gRNA; and
- (b) at least one second nucleic acid sequence comprising at least one said reporter gene.

-101-

9. The method according to any one of claims 1 to 8, wherein said at least one reporter gene is integrated into at least one site at: gender Z chromosome locus 42172748-42177748 or gender W chromosome locus 1022859-1024215.

5 10. An avian transgenic subject comprising at least one exogenous reporter gene integrated into at least one locus in at least one of gender chromosome Z and W, wherein the said reporter gene encodes a protein having an excitation wavelength of 500-650 nm and an emission wavelength of 550-650 nm.

10 11. The avian transgenic subject according to claim 10, wherein said reporter gene is at least one fluorescent reporter gene.

12. The avian transgenic subject according to claim 10 or 11, wherein said reporter gene is Red Fluorescent Protein (RFP).

15

13. The avian transgenic subject according to any one of claims 10 to 12, wherein said at least one transgenic avian subject is female, and wherein said at least one reporter gene is integrated into at least one position of: female chromosome Z or female chromosome W.

20 14. The transgenic avian subject according to any one of claims 10 to 13, wherein said at least one reporter gene is integrated into at least one site at: gender Z chromosome locus 42172748-42177748 or gender W chromosome locus 1022859-102421.

15. A kit comprising:

(a) at least one Cas9 protein or at least one first nucleic acid sequence comprising at least one nucleic acid sequence encoding said at least one Cas9 protein; and at least one gRNA that targets at least one protospacer within at least one gender chromosome Z and/or W, or nucleic acid sequence encoding said at least one gRNA; and

(b) at least one second nucleic acid sequence comprising at least one said reporter gene, wherein a said reporter gene encode a protein having an excitation wavelength of 500-650 nm and an emission wavelength of 550-650 nm.

5

10

1/10

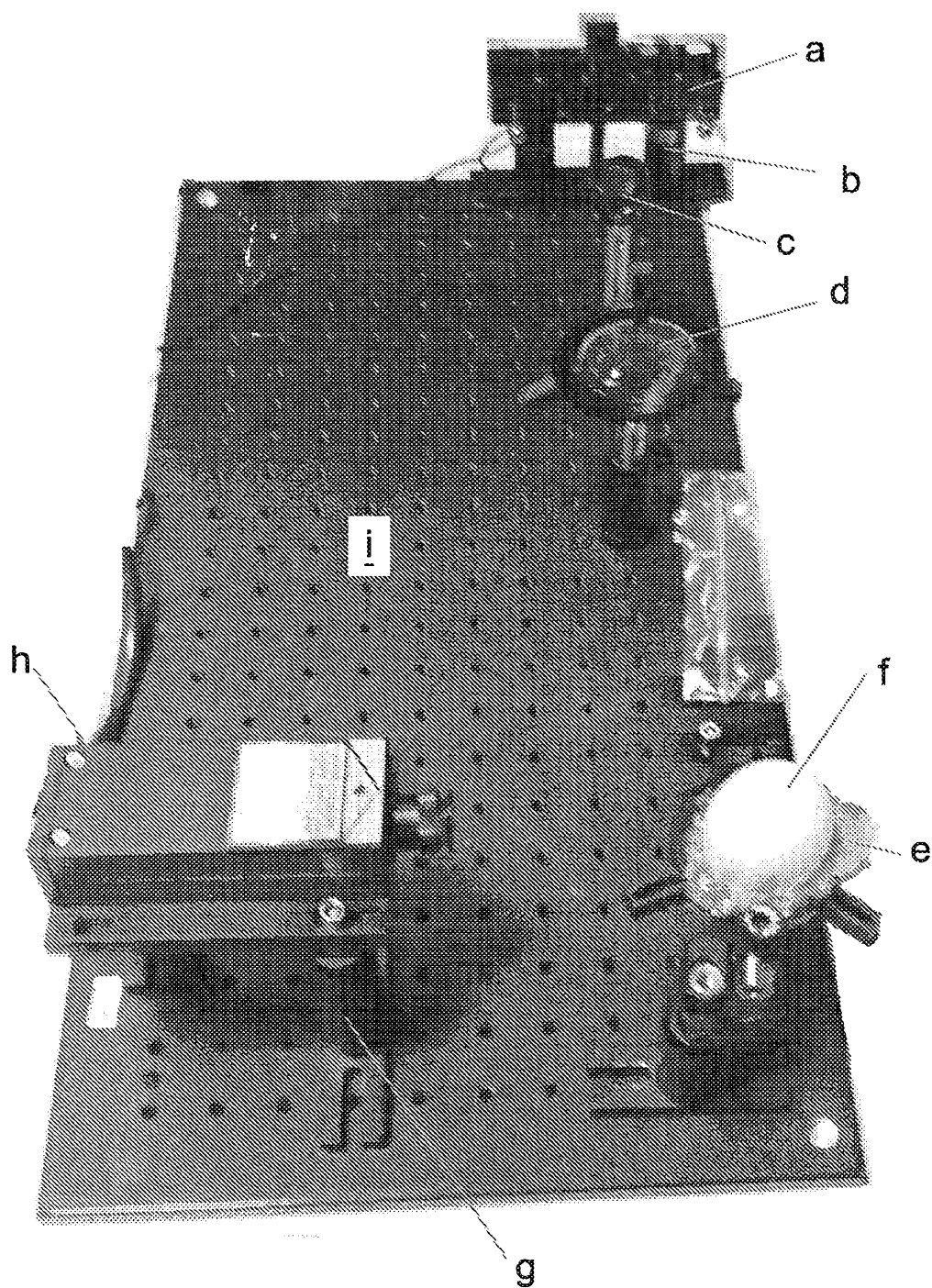


Fig. 1A

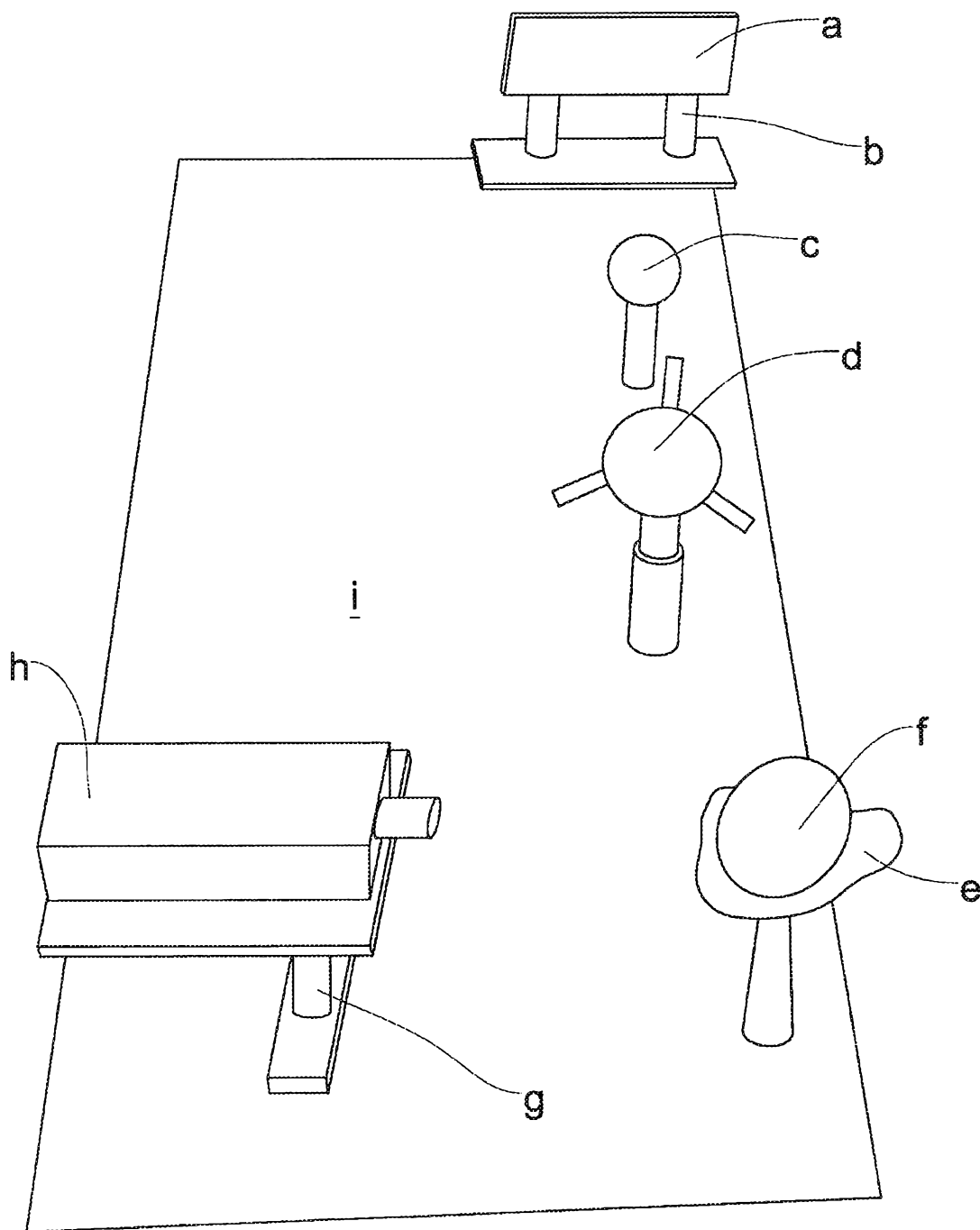


Fig. 1B

3/10

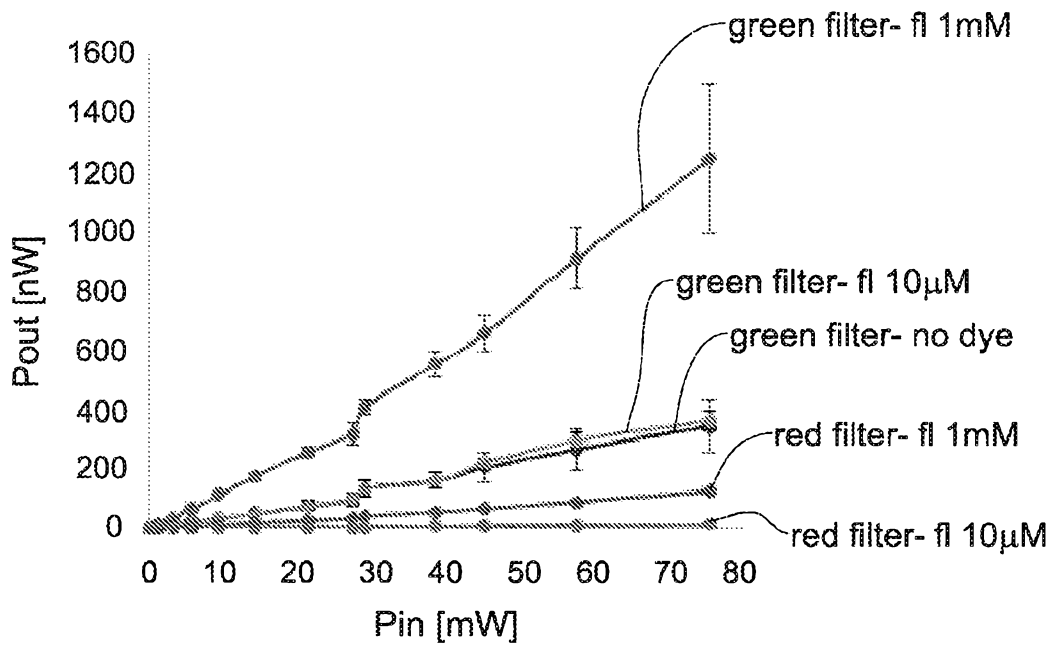


Fig. 2

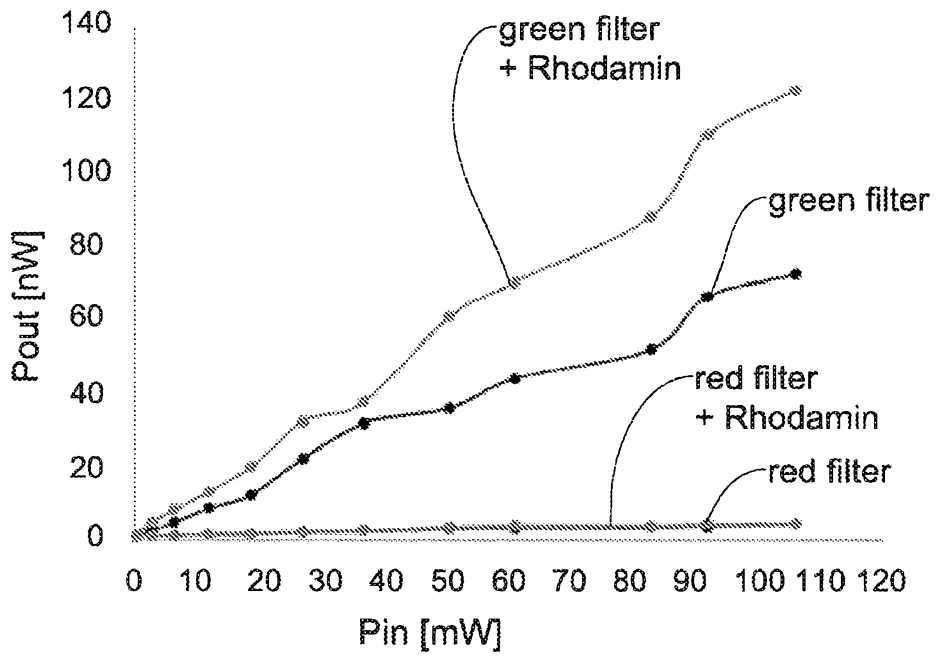


Fig. 3

4/10

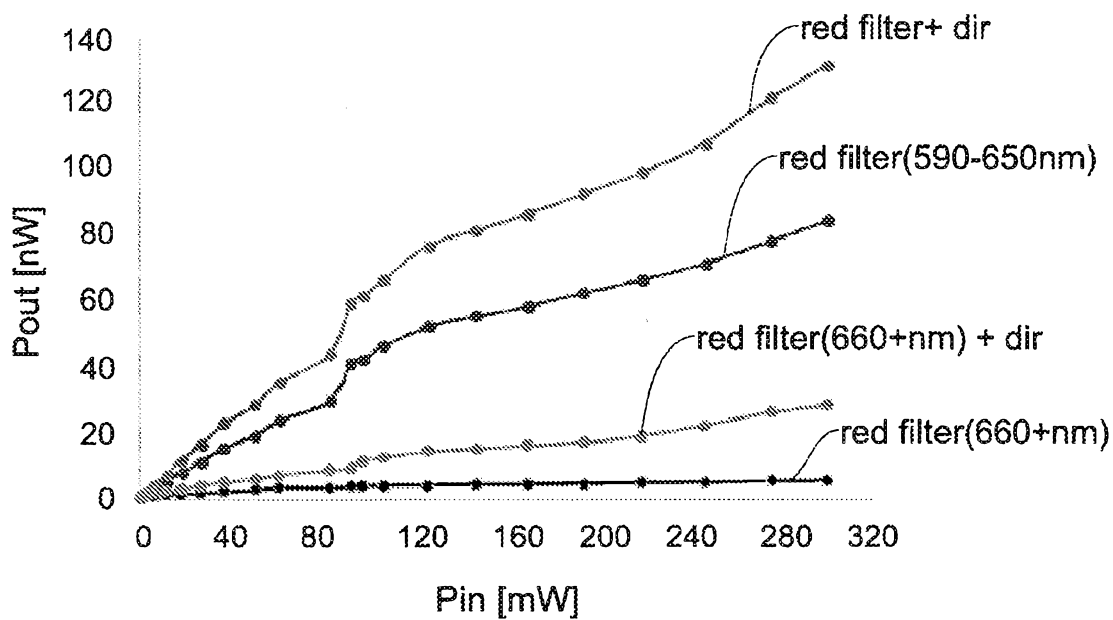


Fig. 4

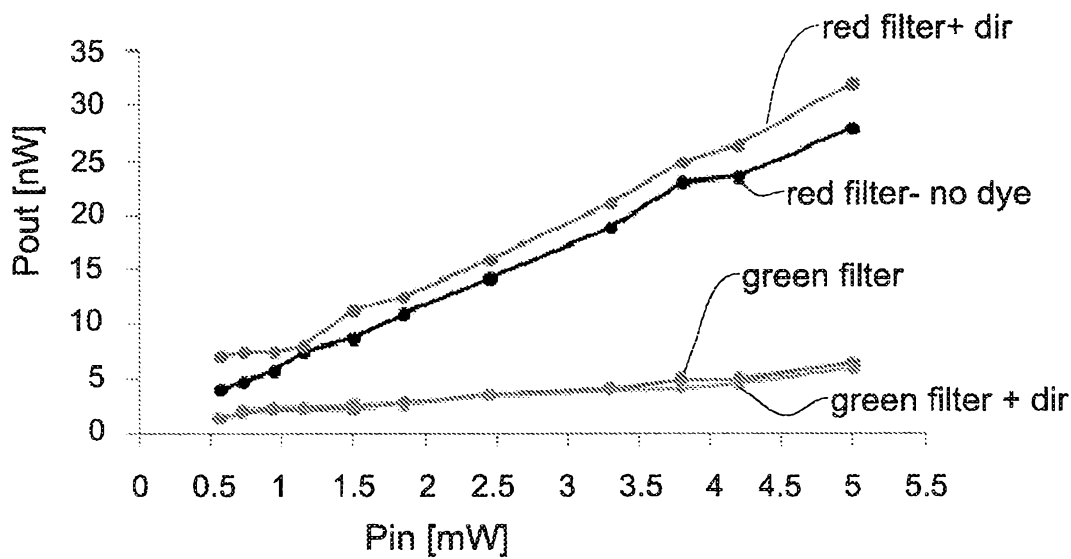


Fig. 5

5/10

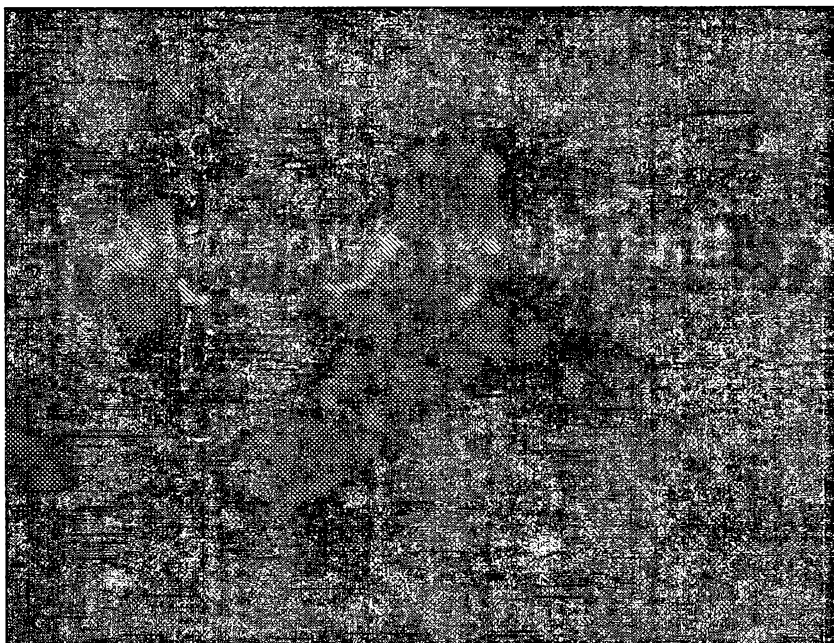


Fig. 6

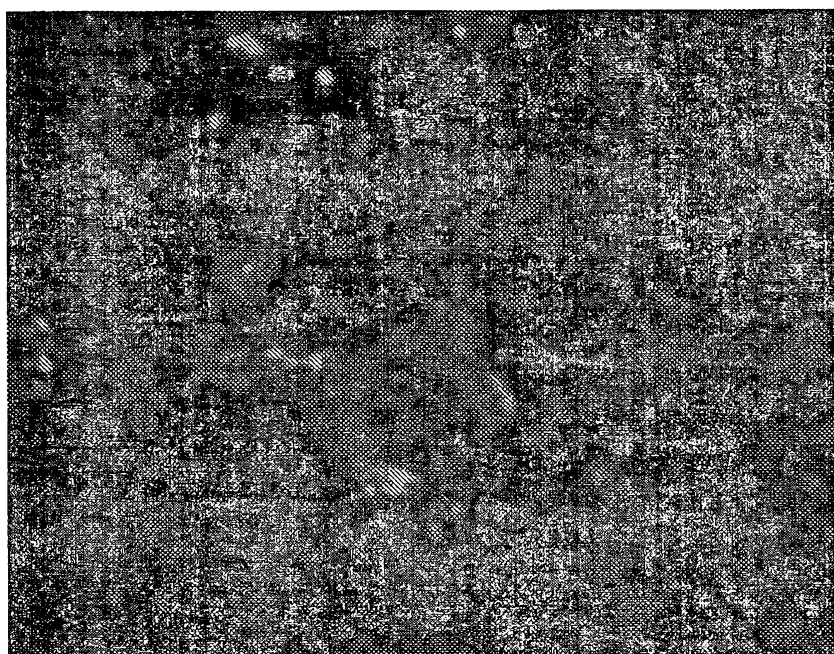


Fig. 7

6/10

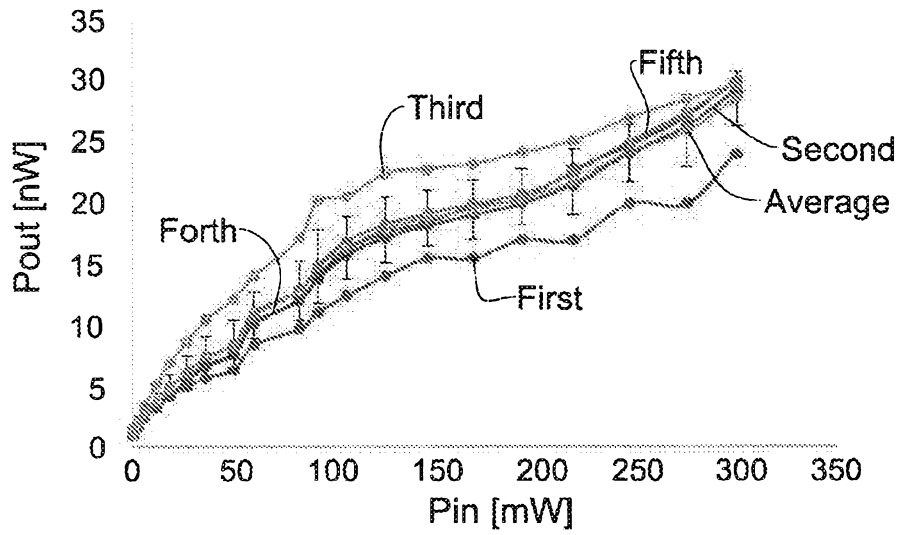


Fig. 8

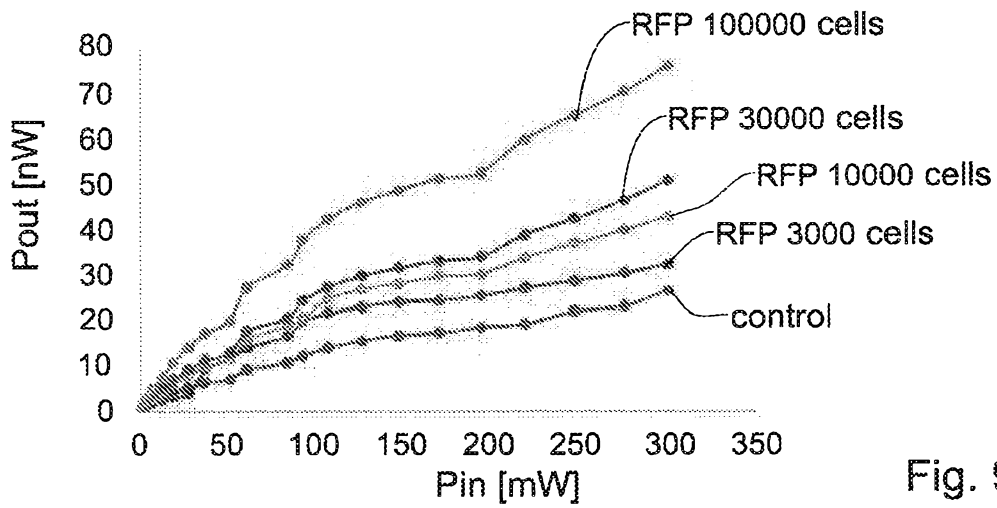


Fig. 9

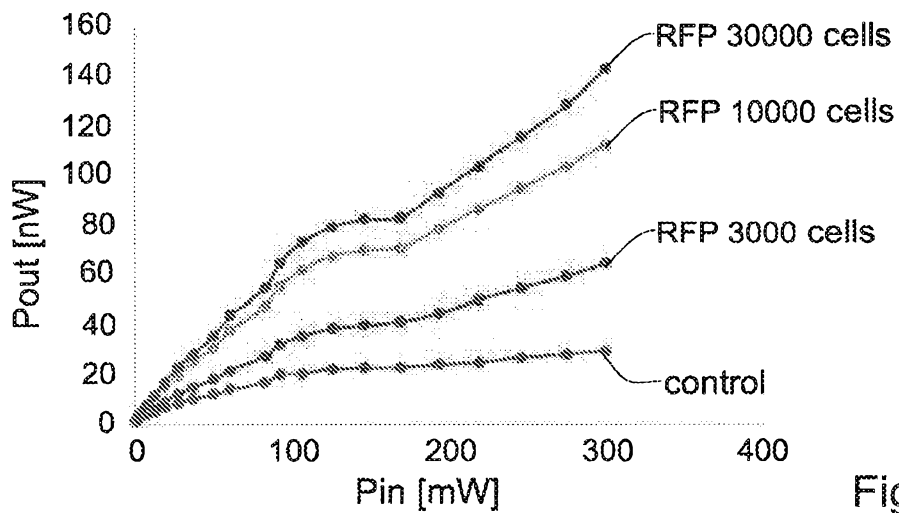


Fig. 10

7/10

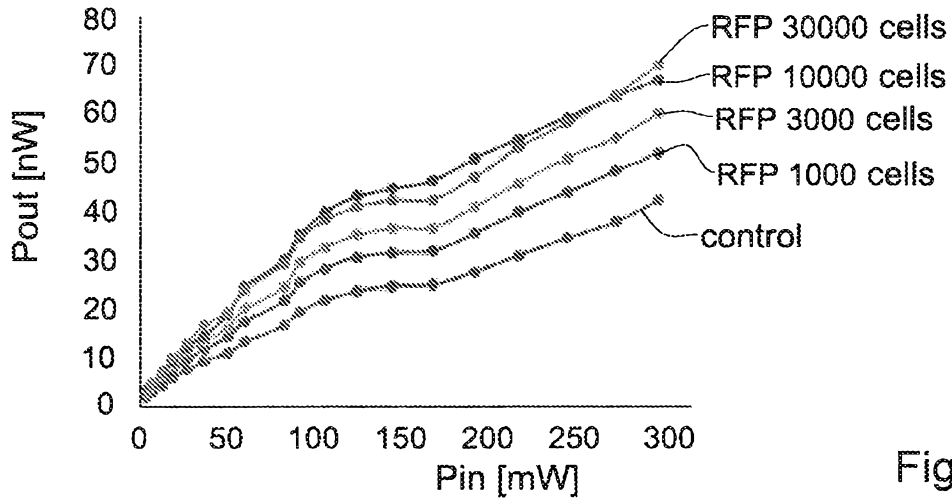


Fig. 11A

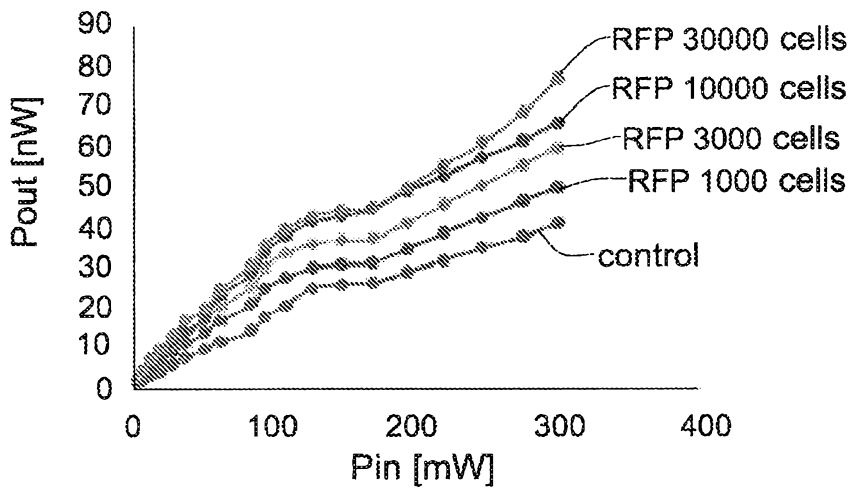


Fig. 11B

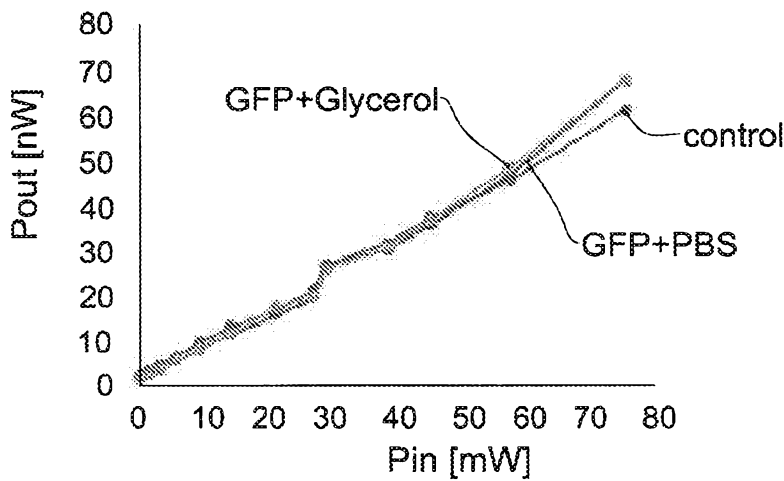


Fig. 12

8/10

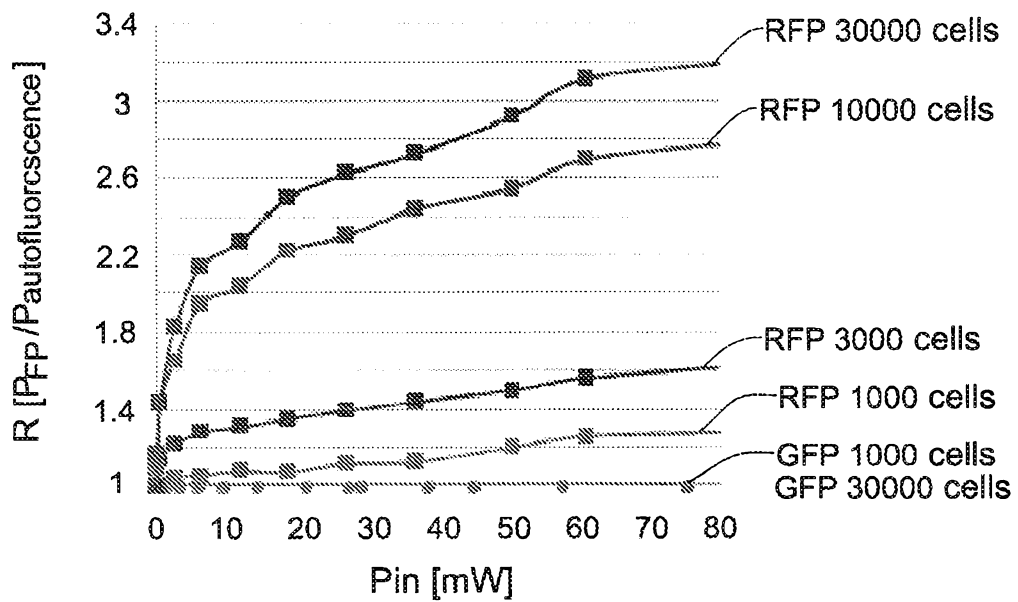


Fig. 13

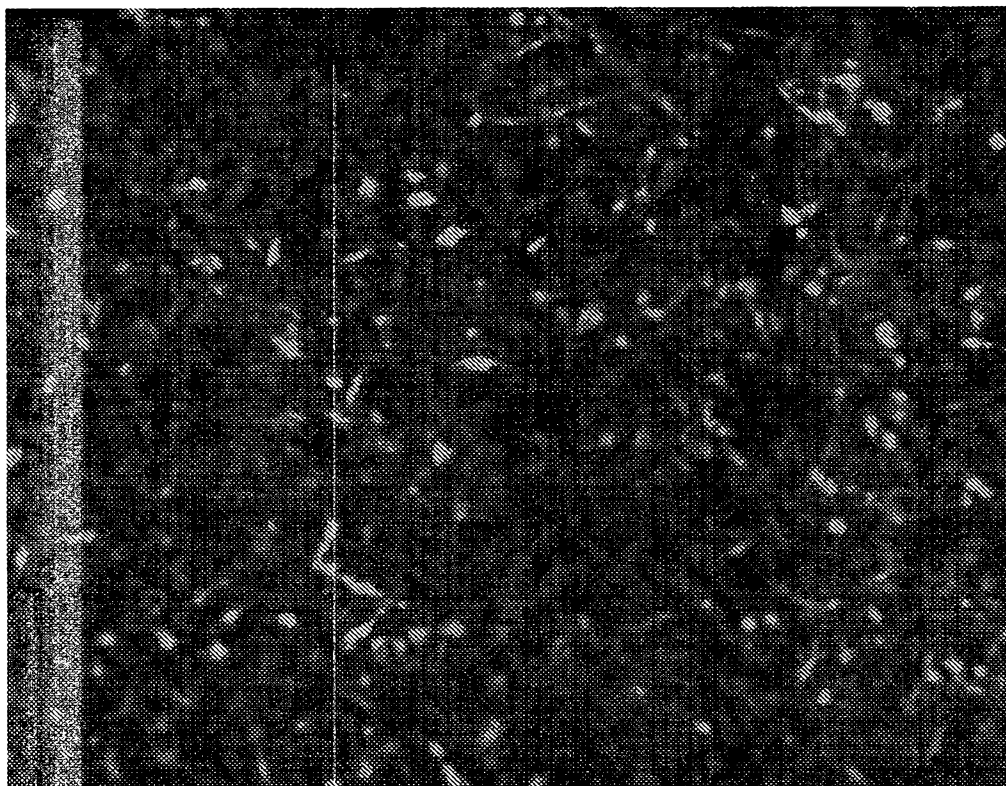


Fig. 14

TGGAGATCCCTTATTGGTGAAGCTTCCCTGCAAACCTAGAATCTTCTGATGCAAGTGAGCAATCTTCTCCCTTCCCTAGCAGCTCTCTGTTGCACC
 AGTCCTTCCATCCTTGCACCTTTGTTGTAATGCTTTTCCCATGGCATTGATCTGTCCAGTTGCTACAGTGTAGAAACAAGAAAGTGTTTTCTCAAACA
 TTATACATGCTTACTATCAGATCTTACTTAACATTGATAAGAAGTACACCTTTCATTACAACAGTGAACGGATAGGCAGCAAGCATTCAGAACAATGAA
 AAAGATCTGCGTTCATTTGTAAGCAAACAAAGGTGATGACGGAACGTTTATTTGAGATACAGGCACCATAGTACAAGATATCAAAGAGGAATTAATTC
 AACCTTCCCAACAACAGTTGGGAAAAAAAAAAGGATAAGCTGAGGAATGCAGCCTTCTCCGATTTTTACTCTGGTGCCGGAAGGAGGAGGAGCGGGG
 GAAAACCCAGCGGACTCAGACTGCATACGCACTCTCGGAAGTACCTCTGAGGCGGGTSGAGAGGAAGGGTGGATCTAGGCTCCCCCGCTCCACACACT
 CACGTGCTCGTGTACTGGTTCAGGCTGTGACTTGGCAACTCCTCTACATCTATCCATCTAGTTAGACTCAACAAGACCATCGTGAATCAGGAGTCTGAA
 GATGCTTAACTCTCTCAGAACGCTTAGTTCTCCTGGCCTGGTAAATGCTATGTTCCTCATACTGCCTCTCTGAAGAATGCTTCGAATGCTTCTAGTGAT
 GCTCTAAAGTTCTAAACAGAAAAATCTGGAGACAGTTCTGGTCTTTAGATAGAAAAATGCCAACATGCCAAAGGATGGTTACATCCTTCAAGCAACT
 TGTTCATGCTGTACAATAGACTCATGTAATAACTTAGCCGTAGTCAATCGTATCTTATTTACCTGTTTGGTTATTACATTTTCCCTGGTACTGCTTTAT
 ATTTAGTCAAGTTGTCCTTTTAAAGACAAATTTTTTGGTGTGTGCTAATAGGCAGTTACCAAATGTTCTAGAGGGAGGGAAATATAATCAGTGAGTATGTAG
 ATGTATATAGATGTATAACCAGTAAGTATACAGAGAAGTTAGGGTGGTCTTTTGGAGAGCATATAGCTGGAAAGATTATCACATGGGAAAAGGTAATCAG
 AAATAAATGGAAAAATGCTCACCAGTGTATAGGCTCACAAGAAAACCTCCCAAGGAGCAATCTCCAGTAACAGACCCCTTCTCAGTIGGAGTTCCG
CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCAT
AGTAAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCA
TATGCCAAGTACGCCCTTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCC
TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTT
GACTCACGGGGATTTCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT
CGTAAACAACCTCCGCCCATTTGACGCAAAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCG
TCAGATCCGCTAGCGCTACCGGACTCAGATCTCGAGCTCAAGCTTCGAATTTCTGCAGTCCGACGGTACCGCGGGCCCCGGGATCCACCG
GTCCGCCACCATGGCCTCTCCGAGAACGTCATCACCGAGTTCATGCGCTTCAAGGTGCGCATGGAGGGCACCGTGAAACGGCCACGAGT
TCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCCACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCT
TCGCTGGGACATCCTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCCGACTACAAGAAGCT
GTCTTTCCCGAGGGCTTCAAGTGGGAGCGCGTGTAACTTTCAGAGGACGGCGGGCTGGCGACCGTGACCCAGGACTCCTCCCTGCA
GGACGGCTGCTTCACTACAAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAAGAAGACCATGGGCTGG
GAGGCCCTCACCGAGCGCTGTACCCCGCGACGGCGTGTGAAGGGCGAGACCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCAC
TACCTGGTGGAGTTCAAGTCCATCTACATGGCCAAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGACGCCAAGCTGGACATCA
CCTCCCAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGCCACCCTGTTCTGTAGCGGCCGCGACTCTA
GATCATAATCAGCCATACCACAT

9/10

Fig. 15A



[SV40 poly(A) signal- fl ori- AmpR promoter- sv40 promoter- NeoR/KanR- HSV1Kpoly(A) signal-ori]

GAGTCATGTCGGTGGAGGAGAAAAGTGGCTGTAGAAAAGGTAACGGACAAGAAAATCCCTCAC TGGCTTAGCTGTAAAGGTCAGAGGAAAATCATCTTCTG
 TTCTCCATACACATCTCCTGTAATTAGGCACCTGTGCCCTTTAAAAAAGTAAAGTAATCAAAGAGTCATTTGGTGAAAATAGAAGCCAAACAGTTACAA
 AATTTAGTTAATAATTCAAAGACAAAACAACCTCTGGTTCCAAATATAAAACCCCTTTGTGTAAACACTGCCGGAAGGTGCCAAAGAGCCACCTACACAAAG
 AGATGCCACAGAGTAATTTGCGTACCCCAACAATAGGATCATATATGGGGCAAACCAAAGACCAACAAGACCCCTTGCAATAATCCCTTGATTGGCAG
 AAGCAAGTGCAGCGTTAGATTTAGTATGTGTGACAATTCTGAGAAATAGGAAGGAATTCAC TGAAACTGTGTGAAGATTTGTATGTTAAATACATATAA
 TGAGGGTTATTTAGCTCTGGGGTGTGCATGCTATGTGGAGAGATCCCATGTGCCAGCAC TGAATAGACTAATGTCAGCTTTTTAAACCATCATTG
 GTTTGAAGAGTTTATTATGATTTTCAGAAACAGAATTATGTATTCAGTGTCTGTACATTCTTACAAGCCTGCTGTTCTGTACACATGAAAAGCTACTT
 ATGGTGCAAATCAGTATAGAGAAGCAGTTTTGACATACAGTGGTCATGACTGGGTGATCTGCTGTCAGAAGGTAAGACACTGGTATACAGAATTGCAA
 CGTGAGATGAAGACCTCAAGTGCACCTCTGTATCATACCCTTCTCTTGCTATTATCTATT CAGTGTGTTCTAAATAATTCAGGCAGAAGGCTTGGT
 TTCACATCAGGTGTTACATTTAAGTATTCCTTGCCATTTTTGCTGTTTGTATGTGTAAC TTCAGATTCTCACATTGACTGTGTAGTGTAAATTTGCA
 AATAGATGTAACAGCTTCTCTTTAACATCCCATGCACAAACTTTTAATCAC TTTCAATCAATGTTGGAAATCTGCAGTACATAAAATTTATTTTGATTC
 ATTATTCAGAGAATCAGAGAAATAAGGACTTTTTTAGAGGTTTATGTTTATTTTAGAAAATATAAATAAATCTTGTATACATAGATTTCCAAC TAGTAA
 AATTTTAGCTATCGCTGTAGCATTCTATGAGGAGTTTTACTTTTTACTGTCTCTTATGAACTTAGTGGAAACAAAGCTAG

10/10

Fig. 15B

SEQUENCE LISTING

- <110> EGGXYT LTD.
- 5 <120> METHODS FOR GENDER DETERMINATION OF AVIAN EMBRYOS IN UNHATCHED EGGS AND MEANS THEREOF
- <130> 2562455
- 10 <150> US 62/510,921
<151> 2017-05-25
- <160> 67
- 15 <170> PatentIn version 3.5
- <210> 1
<211> 20
<212> DNA
- 20 <213> Artificial Sequence
- <220>
<223> Guide 1 for W chromosome
- 25 <400> 1
gcactaggaa ccagcagcag
20
- 30 <210> 2
<211> 20
<212> DNA
<213> Artificial Sequence
- 35 <220>
<223> Guide 2 for W chromosome
- <400> 2
gtagcccca gagggctagg
- 40 <210> 3
<211> 1460
<212> DNA
<213> Artificial Sequence
- <220>
<223> galGal5_dna range=chrW:1022859-1024215
- 50 <400> 3
tcgggctaac attagaccg ttgtgtctta atcctgccgt cgggtataca tataactaaaa
60

gaacctcctt tccctcctat aaatcggagc gagacaaggc attctttttt tattattatt
120

5 ttggttgccct gcccatcaga tgtaatgaaa gttgcacagg gtagggctaa gaaggtaaag
180

gagttaaatt cccccgggag gttagagatt ccctgtattg tggttgcctg cccattagac
240

10 gtaacaaaag ttgtgaaggg taggactaag gatccaggta aagaggttag tccccaggg
300

ttagagtccc tcctcagaag taaatgtaa acataaccac ggtaattgtc gttgagtgcg
360

15 cttgtttgat tgtaattgt ttgtgctggt attgtatgct gattgtttgt gttgttatta
420

20 tacaagagc tcacatgcat cgttaaaatg ggcatggat cacctagggg tgaggggaatt
480

ttgaaaaagt cacttttggg ctgtgtttta agtcattgga aagaaatagc tggatcccct
540

25 ggtgggggttg cacagaagga tgatttaagt acgatggatg cctttgaggc aggcactagg
600

aaccagcagc agtggcagat ttgtgatcat ggaaagaatt ggtgggagca tacggcacca
660

30 ggactcagat aagatagcta gctaagggtga tgaattgtag atggtgtcag ataacgaaag
720

aagaaaaaca aatgtaataa agtgttaaat gccaatatgt tttttataat gaggcaatat
780

cctgaatggc attagtgtgg gatcaatttg gccagatga ccctctgatt ttaatggtag
840

40 aaaatgataa gagagagagg gagaaaggaa gagagtgtaa aacgtgttgt tcagcatgta
900

gtattggact gagatgcttc aagcgaaaca agaatgagca ggaggaggat ttagagatgc
960

45 tagtagcccc aagagggcta gggggcctca aaatcatgga acccaactcc ctggagtaga
1020

agggtcaggt aatgagtcca aaagagggac tataggaacc atgccagtaa cagagaggac
1080

50 tcacagccag catagggctg tgctgtaagc cctggtgagg cagttggaaa tgacagttcg
1140

gtagtagtta aggtaccctt ttaaattaca gatttaaaca tttggaaagc agctgctggc
 1200

5 agctaccgtg atgacctaata acgggtagct aatgcttttg aatgatgat taaaactcag
 1260

gaactggata ggaaagatat ggaatttatt atgcacatgt tgtttgatag tacagaaaa
 1320

10 taaatgattc actagaccac acggacccaa gtggaggatc aggtaatggc aggggttttg
 1380

cttggatgta atgctggaaa ggacaagagt cctttcctgt cctggggagg atttagaggt
 1440

15 ggttagacct tggggaggag
 1460

20 <210> 4
 <211> 592
 <212> DNA
 <213> Artificial Sequence

25 <220>
 <223> For cloning using the gRNA1, W chromosome, "Left arm"

<400> 4
 tcgggctaac attagaccg ttgtgtctta atcctgccgt cgggtataca tataactaaaa
 30 60

gaacctcctt tcctcctat aaatcggagc gagacaaggc attctttttt tattattatt
 120

35 ttggttgctt gcccatcaga tgtaatgaaa gttgcacagg gtagggctaa gaaggtaaag
 180

gagttaaatt cccccgggag gttagagatt ccctgtattg tggttgctg cccattagac
 240

40 gtaacaaaag ttgtgaaggg taggactaag gatccaggta aagaggttag tccccaggg
 300

ttagagtccc tcctcagaag taaatgtaa acataaccac ggtaattgtc gttgagtgcg
 45 360

cttgtttgat tgtaattgt ttgtgctgtt attgtatgct gattgttgt gttgttatta
 420

50 tacaagagc tcacatgcat cgtaaataatg ggcattggat cacctagggg tgaggggaatt
 480

ttgaaaaagt cacttttggg ctgtgtttta agtcattgga aagaaatagc tggatcccct
 540

ggtgggggtg cacagaagga tgatttaagt acgatggatg cctttgaggc ag
592

5

<210> 5
<211> 848
<212> DNA
<213> Artificial Sequence

10

<220>
<223> For cloning using the gRNA1, W chromosome, "Right arm"

15

tggcagattt gtgatcatgg aaagaattgg tgggagcata cggcaccagg actcagataa
60

20

gatagctagc taaggtgatg aattgtagat ggtgtcagat aacgaaagaa gaaaaacaaa
120

25

tgtaataaag tgttaaatgc caatatgttt ttataatga ggcaatatcc tgaatggcat
180

30

tagtgtggga tcaatttggc ccagatgacc ctctgatttt aatggtagaa aatgataaga
240

35

gagagagga gaaaggaaga gagtgtaaaa cgtgttggtc agcatgtagt attggactga
300

40

gatgcttcaa gcgaaacaag aatgagcagg aggaggattt agagatgcta gtagcccaa
360

45

gagggctagg gggcctcaa atcatggaac ccaactcctt ggagtagaag ggtcaggtaa
420

50

tgagtccaaa agagggacta taggaacat gccagtaaca gagaggactc acagccagca
480

55

tagggctcgtg ctgtaagccc tgttgaggca gttggaaatg acagttcggg agtagttaag
540

60

gtaccctttt aaattacaga tttaaactt tggaaagcag ctgctggcag ctaccgtgat
600

65

gatcctaaac gggtagctaa tgcttttgaa atgatgatta aaactcagga actggatagg
660

70

aaagatatgg aatttattat gcacatggtg tttgatagta cagaaaaata aatgattcac
720

75

tagaccacac ggaccaagt ggaggatcag gtaatggcag gggttttgct tggatgtaat
780

gctggaaagg acaagagtcc tttcctgtcc tggggaggat ttagaggtgg ttagaccttg
840

5 gggaggag
848

<210> 6
<211> 962
10 <212> DNA
<213> Artificial Sequence

<220>
15 <223> For cloning using the gRNA2, W chromosome, "Left arm"

<400> 6
tcgggctaac attagaccgc ttgtgtctta atcctgccgt cgggtataca tatactaaaa
60

20 gaacctcctt tccctcctat aaatcggagc gagacaaggc attctttttt tattattatt
120

ttggttgctt gcccatcaga tgtaatgaaa gttgcacagg gtagggctaa gaaggtaaag
180

25 gagttaaatt cccccgggag gttagagatt cctgtattg tggttgcttg ccattagac
240

30 gtaacaaaag ttgtgaaggg taggactaag gatccaggta aagaggtag tccccaggg
300

ttagagtccc tcctcagaag taaatgtaa acataaccac ggtaattgct gttgagtgcg
360

35 cttgtttgat tgtaattgt ttgtgctgtt attgtatgct gattgtttgt gttgttatta
420

tacaaagagc tcacatgcat cgtaaaatg ggcattggat cacctagggg tgaggaatt
480

40 ttgaaaaagt cacttttggg ctgtgtttta agtcattgga aagaaatagc tggatcccct
540

45 ggtgggggtt cacagaagga tgatttaagt acgatggatg cctttgaggc aggcactagg
600

aaccagcagc agtggcagat ttgtgatcat ggaagaatt ggtgggagca tacggcacca
660

50 ggactcagat aagatagcta gctaaggatg tgaattgtag atggtgtcag ataacgaaag
720

aagaaaaaca aatgtaataa agtgtaaat gccaatatgt tttttataat gaggcaatat
780

cctgaatggc attagtgtgg gatcaatttg gccagatga ccctctgatt ttaatggtag
840

5 aaaatgataa gagagagagg gagaaaggaa gagagtgtaa aacgtgttgt tcagcatgta
900

gtattggact gagatgcttc aagcgaaca agaatgagca ggaggaggat ttagagatgc
960

10 ta
962

15 <210> 7
<211> 478
<212> DNA
<213> Artificial Sequence

20 <220>
<223> For cloning using the gRNA2, W chromosome, "Right arm"

<400> 7
gggcctcaaa atcatggaac ccaactccct ggagtagaag ggtaggtaa tgagtcctaaa
25 60

agagggacta taggaacat gccagtaaca gagaggactc acagccagca tagggctctg
120

30 ctgtaagccc tgttgaggca gttggaaatg acagttcggg agtagttaag gtaccctttt
180

aaattacaga tttaaacatt tggaaagcag ctgctggcag ctaccgtgat gatcctaaac
240

35 gggtagctaa tgcttttgaa atgatgatta aaactcagga actggatagg aaagatatgg
300

aatttattat gcacatgttg tttgatagta cagaaaaata aatgattcac tagaccacac
40 360

ggaccaagt ggaggatcag gtaatggcag gggttttgct tggatgtaat gctggaaagg
420

45 acaagagtcc tttcctgtcc tggggaggat ttagaggtgg ttagaccttg gggaggag
478

<210> 8
50 <211> 669
<212> DNA
<213> Artificial Sequence

<220>

<223> A "left arm" for the region upstream to the CMV-promoter

<400> 8
 5 aggcaaggct tgaccgacaa ttgcatgaag tcgggctaac attagaccg ttgtgtctta
 60
 atcctgccgt cgggtataca tataactaaaa gaacctcctt tccctcctat aaatcggagc
 120

10 gagacaaggc attctttttt tattattatt ttggttgctt gcccatcaga tgtaatgaaa
 180
 gttgcacagg gtagggctaa gaaggtaaag gagttaaatt cccccgggag gttagagatt
 240

15 ccctgtattg tggttgctg cccattagac gtaacaaaag ttgtgaaggg taggactaag
 300
 gatccaggta aagaggttag tccccaggg ttagagtccc tcctcagaag taaatgtaaa
 20 360
 acataaccac ggtaattgct gttgagtgcg cttgtttgat tgtaattgt ttgtgctggt
 420

25 attgtatgct gattgtttgt gttgttatta tacaagagc tcacatgcat cgtaaaaatg
 480
 ggcatggat cacctagggg tgaggaatt ttgaaaagt cacttttggg ctgtgtttta
 540

30 agtcattgga aagaaatagc tggatcccct ggtgggggtg cacagaagga tgatttaagt
 600
 acgatggatg cotttgaggc aggactagg aaccagcagc gatgtacggg ccagatatac
 35 660
 gcgttgaca
 669

40
 <210> 9
 <211> 911
 <212> DNA
 <213> Artificial Sequence

45
 <220>
 <223> A "right arm" for the region downstream the Neomycin-resistance

<400> 9
 50 ttctatcgcc ttcttgacga gttcttctga cagtggcaga tttgtgatca tggaaagaat
 60
 tgggtgggagc atacggcacc aggactcaga taagatagct agctaaggtg atgaattgta
 120

gatggtgtca gataacgaaa gaagaaaaac aatgtaata aagtgttaa tgccaatatg
180

5 tttttataa tgaggcaata tcctgaatgg cattagtgtg ggatcaattt ggcccagatg
240

accctctgat tttaatggta gaaaatgata agagagagag ggagaaagga agagagtgtg
300

10 aaacgtggtg ttcagcatgt agtattggac tgagatgctt caagcgaaac aagaatgagc
360

15 aggaggagga tttagagatg ctagtagccc caagagggct agggggcctc aaaatcatgg
420

aaccceaactc cctggagtag aagggtcagg taatgagtcc aaaagagggg ctataggaac
480

20 catgccagta acagagagga ctcacagcca gcatagggtc gtgctgtaag ccctgttgag
540

gcagttggaa atgacagttc ggtagtagtt aaggtaccct tttaaattac agatttaaac
600

25 atttggaaag cagctgctgg cagctaccgt gatgatccta aacgggtagc taatgctttt
660

30 gaaatgatga ttaaaactca ggaactggat aggaaagata tggaatttat tatgcacatg
720

ttgtttgata gtacagaaaa ataaatgatt cactagacca cacggacca agtggaggat
780

35 caggtaatgg caggggtttt gcttggatgt aatgctggaa aggacaagag tcctttcctg
840

tcctgggggag gatttagagg tggttagacc ttggggagga gggttcgaaa tgaccgacca
900

40 agcgacgccc a
911

45 <210> 10
<211> 911
<212> DNA
<213> Artificial Sequence

50 <220>
<223> A "right arm" for the region downstream the polyA site

<400> 10

atctgtgtgt tggttttttg tgtgtctaga cagtggcaga tttgtgatca tggaaagaat
 60
 5 tgggtgggagc atacggcacc aggactcaga taagatagct agctaagggtg atgaattgta
 120
 gatgggtgtca gataacgaaa gaagaaaaac aaatgtaata aagtgttaaa tgccaatgatg
 180
 10 ttttttataa tgaggcaata tcctgaatgg cattagtgtg ggatcaattt ggcccagatg
 240
 accctctgat tttaatggta gaaaatgata agagagagag ggagaaagga agagagtgta
 300
 15 aaacgtgttg ttcagcatgt agtattggac tgagatgctt caagcgaaac aagaatgagc
 360
 20 aggaggagga tttagagatg ctagtagccc caagagggtc agggggcctc aaaatcatgg
 420
 aacccaactc cctggagtag aagggtcagg taatgagtcc aaaagagga ctataggaac
 480
 25 catgccagta acagagagga ctcacagcca gcatagggtc gtgctgtaag ccctgttgag
 540
 gcagttggaa atgacagttc ggtagtagtt aaggtagcct tttaaattac agatttaaac
 600
 30 atttgaaaag cagctgctgg cagctaccgt gatgatccta aacgggtagc taatgctttt
 660
 35 gaaatgatga ttaaaactca ggaactggat aggaaagata tggaatttat tatgcacatg
 720
 ttgtttgata gtacagaaaa ataaatgatt cactagacca cacggaccca agtggaggat
 780
 40 caggtaatgg caggggtttt gcttggatgt aatgctggaa aggacaagag tcctttcctg
 840
 tcctggggag gatttagagg tggttagacc ttggggagga gcaagtaaga tgcttttctg
 900
 45 tgctgcaata g
 911
 50 <210> 11
 <211> 17
 <212> DNA
 <213> Artificial Sequence

<220>
<223> Guide 3 for Z chromosome

5 <400> 11
acagacctat gatatgt
17

10 <210> 12
<211> 17
<212> DNA
<213> Artificial Sequence

15 <220>
<223> Guide 4 for Z chromosome

20 <400> 12
cgattatcac tcacaag
17

25 <210> 13
<211> 17
<212> DNA
<213> Artificial Sequence

30 <220>
<223> Guide 5 for Z chromosome

35 <400> 13
ctgggtagca tggggac
17

40 <210> 14
<211> 17
<212> DNA
<213> Artificial Sequence

45 <220>
<223> Guide 6 for Z chromosome

50 <400> 14
gtaaagagtc agataca
17

50 <210> 15
<211> 4980
<212> DNA
<213> Artificial Sequence

<220>
<223> ENSGALG00000025945_ENSGALG00000023622|chrZ:9156874-9161874

<400> 15
cctgcaggac atgtggacac tctgccaggt actgggggag gatccccctcc actgctcctc
60
5
gggagatggc aaacatctag agagggacac tgcagggatg ctgcatgaac acctccagcc
120
10
ttctgcaggc tgtgatgggg tcagggctga aggtcaacac aagcaatgac tgcttgctga
180
caatgtgggc attgccgacc acagacctat gatatgtgag tggtgagtgg ggccaggaga
240
15
gcaggacagg agtgtggact cggaggcgcg ggcagaggag gtagctcaga gcatgagata
300
tattgccag gtgtcagtgg acttctggca gtgcatactg cagagagctt tggttgctgg
360
20
ccatctggat gatgacaaaa tctcagccac agaccatgtg ggttggaaga gtcctcagga
420
gatcaccttt ccacctctg ctaaagcaca ttcctacag tatgtttcac atgacaccgt
480
25
gctggcagat ttttaatatc atcagagaat gacactccat atcctctctg ggcagccttt
540
30
ttcagtgctc tgtcaccac aaagtaaagt attgctcat gttctaattg aacttcctat
600
gttctagttt gtgaccattg aacctgttc tgtcactgga cagtgctgaa ggagcctggc
660
35
cctatccact tgactcccac actttacata tttacaagca tcgataagac cccctcagt
720
40
cttctccaga ctagatagcc ccaggtctcc caaccattcc ttttatggga gatgcttcag
780
gccctcatc acctctgtgg ccctctgctg gactgctgc agtagttcca ggttttttt
840
45
tttgaactgg ggagcccaga acttgacacg gtattccaga tgtggcctca ccagcgcaga
900
ggagagggga aggctcacct ccctcgacct ggtggccatg ttttttttaa cagacctcag
960
50
gataccactg gccttcttgg cccaagggca cactgctggc tcatggccca tgggttgget
1020

tccaggactc tcaggtcctt ctccgtagag ctctcttcaa gcaggacaac cctcagcctg
1080

5 tactgatgtg tgcggtgcgg ttattcctcc ccacgtgcaa gactctacat ctgccattgt
1140

taaatctcat aagcttcctc tctgcccac tctccactgt gtacgggct tagtgaatgg
1200

10 cagcacagac ttctggtgtg aggatagcat tcagaacaag ctgctgtacc ttcccagtca
1260

cagaggtgag ggtgactggc ttgtactttc ccagctttgc cttcttgccc ttttggaga
1320

15 ctggagcgac attggccttc tggaggaaac ggtcctcagg caccctcct gttctccatg
1380

acctttgaaa gatgatcgag agcaccttg cagtcacttc tgccagctcc ttcagcacgc
1440

20 atgagtgcac cccgccagtg cccatggatt tgtgttcatt gaattagcct aggtgacctc
1500

25 tgatcagatc cttcttgacc aaggggaaga cttcctttcc ccaggttctt atctccaggg
1560

cctgggattc ccaaggggca atcttagcag gaaaggatga agtaaaggag gtgtgcagtt
1620

30 acagtgcttt catggtgtcc tccattactc caatgaggtc acccactca ttcagcagca
1680

gacatacatt ctcccagtct tcctttttct actgccatac taaagtagtc cttcatgttg
1740

35 tccttgactg aaggtcaccc aacttgattt ccagtgagcc ttagcattcc ttgtagcatt
1800

40 catgcatgtc ctgacaacat tcctattttc ttcccaagtg gccagagacc ttttccacac
1860

tccagaggct ttgttcttcc attactgctt ttccatgagc ttcttgctca tctctgcagg
1920

45 tctcatgcca cctttgcccg atttcttagt cttagagatg caccgatctt gaatttggaa
1980

gaagtgggtgc ttgaatgaca accagctctc ttgggtcccc ttgctttcca aagctotaac
2040

50 ccatggaatg cctccatgta tgtctgtgaa gaggtcaaag tcagctctcc tgaagtogtg
2100

gttgcaatcc tacaaactgc ctgattttctc ccatgggaga tcccgaactc cactatctca
2160

5 cagtcaactgc actcaaggct tccccagcc ttcacatccc cagctacagg aagagaagag
2220

gctcgaggct acttgacgct gtctgaggaa ggctttctca gcttctctct cttcatcagg
2280

10 cggcctggtg taaacaccca caacagcgtc accatcata gcctgtccct tagtcctcac
2340

ccataaactt tctactcatt catcatgcat cccttggcag agctctgtac attccagctg
2400

15 ttccctcaca taaagagcaa ctccaccacc tcacaccgac gtctgtctgt cctaaaaagt
2460

acagagccat ccctgagagc attccagtca tgcgagctgc cccaccacgt ccctctgctg
2520

20 atcgcaactga gatcgaggcc ctacgacccat ccattgatct atgactcttc ctgtttactc
2580

25 cccgcaactgt gtgtgatggt acacaggcac ttcagagaag aaaagagagc acaggatgcc
2640

gatttccctt ggcaggacc cagggggccgc acgagccccc cacctctatc agcaggctctg
2700

30 agcaggactg gagctgcagg cagagcacag ctccagccctg agacacctcc atgctgcctc
2760

ctgcacagcc tcggccccac agccccggg aaggagagca cgggtggggt gccgggagcc
2820

35 acgctgcgct ctgacacact gcgtgtcccc tgctcccagc gcgagctgca gggcccgtgc
2880

40 aggaacgagc tgggccccaa cgtgaccctc ggccccggtc actgctgaac gacagcgttc
2940

agctctccag gctgcaaaga gcgagccgtg aatggagca acacgcggga ttcccgggtg
3000

45 tgaattccta aacggattcg aagggtgcca acaggcagcc actgcgtccg cagctgtgct
3060

tcggtacctg gagaccogcg agcccgtctg ttcccaccag cctccagcag ccggtgctcg
3120

50 cggatcaaaa gaccgcogtg aaggccgcag ggccgtccct gggacggagc gccgcccgtc
3180

tccgcccga gccgaggctc egccctcccc gccgagctga cgctggcagc ggcagagagc
3240

5 cggagcggag ccgaggcgtt ttccggcggc agcggccgca cggggctcgc cgagggccgc
3300

gcagagcgtt gtgcccggcg ggccgcgtcg ccgctccggc ggcaccgoga ttccggcagc
3360

10 gagcgcacg agcggccgct gcccggctgc aggcgcgggt cggggccaca ccgaggcggg
3420

gccgcagcct gccttcgcg ggaccgcagc tccggcctcc ccgaatttg cccgcagcc
3480

15 cccgggctgt caccgtgctg tcaccgtgcc gttcacgggt tacgggctgg ttccggccgg
3540

20 ttccggcccg ttccgtgact ggccgggacag agggattcgt ggatccgcgc ctccgggaag
3600

ggaaacggga cccgaataa ttaaaaacag cggccacgat ctgaggggtt acaaagtac
3660

25 taaaacgtac taaaacaaca cagctgctta acacctcta ttccaaaggc ccatcgacac
3720

ccctggcac cctcaaagta tccccttctg aggtaatggg taccaagaat gcacggagcc
3780

30 cctgggcccag tcacagcagg gtgcttttgc cagtctgtac cggtgaggcc tatttcggcc
3840

acaacacag tcaattcttg ggaaccccc ttcacccat gaatataagt tgattctgtc
3900

ccttgggtgac tcgagggcgt gacagtgcac tgtgcaccag ggtccaccag ggcctcgtat
3960

40 tcctgagggt ctgatgtgcc acaccacgaa cccacacggg tcaatacact ccatcatcct
4020

ttcccccat ggtgggggca gggtcctct atttgttacc tctgtcgttc ttagagcggg
4080

45 tgccatgtcc tgcagcaatg ggagcgatca cctctttggg gattaatctt ggtgggtaat
4140

ctgccttgta attatcttac ctgtgcttga aatgccct ccagggtgaaa gcaaaactgca
4200

50 gcctgcattc tgtggccaaa ggaatggaaa agaaggcgtt agcaatgcca gtgggtggcat
4260

accacttggc ttcttttgac tccagttcat actggagttc taacatgtcc agaacagcag
 4320

5 cgctcagtgg ggggtgtgact tcattctggc catggtagtc tactgtcagc ctccattctg
 4380

cactggcttt atgcactgga gcacagggtc tatggggagc ggcggaggga gctgggatgg
 4440

10 ttcagtctgc agatgaggag gctcagggga gaccttattg ctctctgtaa ctccatgaag
 4500

ggaggttgta gtgagctggg ggtcggcttc ttctctcatg tgactagtgg taggactaga
 4560

15 gggaatggcc tccagttgca gcaggggaga ttcaggctgg atgttaggaa atgggctgga
 4620

20 gtgggctgcc cggggagggtg gtggagtcac cgaccctgga ggtgctcaag gaatgtttag
 4680

acaccgtggt tagggacatg gtatagttag aactgttggt gatgggtgga cagttgggct
 4740

25 ggatgatctt gtaggtcttt ttaaagcttg gtgattcttt gattcgataa gtgcaaagtg
 4800

ttgcatttgg gtcagagcaa tctcacgggt gtgcacaggc tgggagaaga actcctggga
 4860

30 gaagaagaga gcagccctgc agagaaggac ttgacatgaa agacttaacg tcagccagac
 4920

35 tccctctgct ctgccctcat gagcctctcc tactctgtaa aggtttgggc ccccagcaca
 4980

<210> 16
 <211> 4980
 40 <212> DNA
 <213> Artificial Sequence

<220>
 <223> ENSGALG00000015036_ENSGALG00000000438|chrZ:27764943-27769943

45 <400> 16
 aggagtggct tctgaggctg gaaacaaggt gctcccctct tgtcaaacag ccaattccat
 60

50 ctagctccaa cccatcctcc actggccaaa gctgagccca tcgctgatgg tgatagcatc
 120

tgtaatagca tctcagaaag gctaaaagca ctgcgagca gctgagagaa aagagtgagg
 180

gaaaatggga cagaaacagc tctgcagccc tcagaggcaa tgcagaggag aggcaggagg
240

5 tgctccaggc atggagcatt agtttcttgc agtccaggag ttacaggtga tgggtgcaggc
300

tctgtgcagc ccagtgatga ctgtttcagg atccatgctg cagcctgtgc cagagccagg
360

10 ggataggctt ggctgottag tgcaggacgt ggagagccca tgctgagtgc aagctcccaa
420

15 aaggctagga cagaacaggt gtagtggagt agatcttgta ctgtggagca gtcttttcct
480

ggaggactaa cctgtgaaga ggaggactgc gctgaagcgg ttcatagaaga atgaaacctt
540

20 gtgggaggga cccacactg aaaaaggaga aaaccacgat gaggaaggag cagcagagag
600

ctccccagcc ctgtgtgcac tgccaggaga gaaggaagta gaaaacgtgg gaatgtgcag
660

25 gtgtaaagct gagtctggga aaaagggtggg ggtggaataa aggtatTTTT agtcctatac
720

30 ctttctcadc aacctacttt aattgtagat tggccataaa ttaaagtgtt tccccaggtc
780

aagtctgttt cttgcttgtg atgggtgatg atctccctgt cctcatccca atcctgaagc
840

35 tatttcatca tatattctct ccctgccctg tttcaggaca gggagtgcaa gcagtttggc
900

agccagacga ggttaaccca gcaccagagg tttagtcaca tgcaagtgag agctattgcc
960

40 tcaacttaaa tgggagaaaa catttttagca ggttgagtct cacottgttt atccagatta
1020

45 tacagccatg ccagcacatg ggctgcatgc agcctggccc aaccatcct ctagccaccc
1080

cctgccccat actgaaaatg cagcggctct gcctgcca gctgccaagc ctggctccag
1140

50 gcactcacc attgcttaca aacaacaaaa cgtcagcgtg tgagcgcgcc ctgtgagctg
1200

caactggcat ggggagggca cagtcagcac aaatacatga cctgcaaatt ttcacatgga
1260

atgtatggag ttgcaatcag atactcagct caaaaaattt aatcatggct ccttacctgt
1320

5 ctttttttta acctttctct tagctcagag aggtatacct cacttcacaa tcatgccgta
1380

ttcatgtcat tgctttctgg cagtacgtac gtttgtgaac agctgttttc aagggatgatg
1440

10 cacaggaaga gtaaaatttc attaaaaaat ctccaaggag cacatggaga gctcactcag
1500

15 aagcctcacc gctgccactg aacaagactg atgcattagt ttcatgaaat caaggctatg
1560

tgtcccactg gtttcattat tttgttactc tttctttaa acaaaaaat tgttttattg
1620

20 cttatataca tcaactatgt tgttatata tttccaaagg ccgctccaaa agtaatgctt
1680

cctgtttcat tctgttggcc caccgtgtca gaggcagatg gtggtacggc agtagagatg
1740

25 gaaccttccc accaatattc cattacatgt tgccatgtga tagacagcag cagatgggca
1800

30 ctctgacaga atggtatctg acatggaagc gtggatggag caaagttgtg tcaactgaatt
1860

cctccataag gaaatgagca cccactaaca ttcactgatg cttgctgaac atttctggag
1920

35 accaaacagt ggatgtgagc acagtgaggt ggtgggtggt gcatttcagc agtggtgaca
1980

40 atgacgttgg gtcacctctg ctggtgcagg ttttgatgag tgcagcctgt gaggtcttgt
2040

tcacatcaatg agaaaatgca caactaatgg tggatgactgc tgaaaaatag tgttttgtag
2100

45 ctgagaattt tctctgtcaa acagtgtcaa ttgtgctctt tgtatctggt gtagtttcca
2160

tgggaataaa taggagcat tacttgcaga ctgacctaca tatatgtggc cccaggcaag
2220

50 tcaacagctg tacgaaaagt ccatgcttcc tcattataaa cggaggaaaa aaagttgttt
2280

acagctgtaa tgggattata aaaggcaaat ggggatagta cagtggatgag aaccagatct
2340

atgaaggaaa agccctaaga aaaagagagt gcacagatac tottaacat ctaataactg
2400

5 tttcctccat cctacagctc agagttaaga cttacagagg actctagtag ttagtaagat
2460

gaatacgagc taatagtggc aaaaataatc ccagtgcctc aacactgacc tgggaaaaag
2520

10 gggcatgtat agaccttctg atattgtgat gctgtgtttg tacacttatt atctacattt
2580

tcagaaatta ggttaaactt cagaaaactg aagatctcca gggctttag cagaccctga
2640

15 ccaccagact ggtccccgat tatcactcac aagctctcaa accgattgtg gctgtttctc
2700

20 agaggcaact ctttgcgctc tagccctctc ataacatggg gcaacttccc ctgccccacc
2760

ttcccttctc gtatcttctg aaaagcttgt agccctcaat tgtcacgctc cagccatag
2820

25 aatcatocca ccacgtccct gtgttcacaa ttaccttcta gtcttctaata tgcaccatgg
2880

cttccaattc ctctgttta tcaccacagt tgcattcatt gatgtagagg cacttcagtc
2940

30 gggctatcag caatgttacc acctgtgagg agccctcttg agatcttcta ggacaattta
3000

35 gaggttttcc ccaactgttc ctaaaattac agcattccct gtccttctt agtgatatag
3060

aactccatcc ccttccacca tcaaacttag cttaaagctc ggtaaagat ccagccagtc
3120

40 tgcttcccaa aactcttgct cagttgcatt ctatctgatg cctacatgcc caacaccttc
3180

aaggcctgtc caagatcata gaacacaaag gcttgatcat gacaacatcc atgcagccaa
3240

45 tcattcacat gaccattcc tctctcctt tccaggtccc aacggccaac tgagagttgt
3300

50 gctctgagct cttcagtgtc ctccgaggg atgtaaagtc ccttttaatg ttctgcattt
3360

tctttgtcac agccttctga gaccctaact gaatgtggag gaggggagag taatcatccc
3420

gctctgatta tcagatacat tagcctcttc ttggcacctc tggcagacag aagacttccc
3480

5 tggagagatt acctggatga taactggggg cctcagtgcc tcacagcagc aagtccccag
3540

ttactaagac tctacacctc tttggtggcg ctagtctga tgtagttctt cagttgctta
3600

10 acatgcttgt tgtttcccaa gatttgactg ctatctgcac cctccctttc ttccagcccc
3660

15 agagccttat atcatgagct gagaacattc tgtgttcaaa atacaagaca catttatgca
3720

atctcgatt ctctacccat atttttgtga tagaaattag ggctacttat ttcccatagt
3780

20 gatttccatg ctgcctacca ggttctaagt caggtcaccc tgccagcttc ttccaaacca
3840

tgctcaggag tgcatacaat cacacgcaga cccaaggaaa aaatgtacca aacaatatta
3900

25 ccacgcatgc aaactcctca gtagctctgt cccatatgct acaattaggt ctgtataggc
3960

30 ttgttagatc ttgtctactt gattagtgtt ttccagcagg cctacaatcc aacctaaact
4020

gaacagacat taaagacatg tttcactaat acaaaacaac tccacctagt tcaacagtct
4080

35 gctcttttgt gtatggtacc tgaaactctt ctggttactg aacttatccc aacatttttc
4140

ctcctaagtc atactttttc tagacgtgat gtgaaagtac agtggttatac cctggtgctg
4200

40 cattaatttg cataaccaca gcacacaaag cataatcata acacacaaag gctcctatg
4260

aaaacatctg cttaaagtca taggtcactt atttattctc cctctagtta tgtaccagaa
4320

tgatgtatgt taataaataa tagattatta aatccatcaa aaattagga agagctctag
4380

50 cacaaaggca aagcaaagaa gctttgaaga agtgcaagta aaatgaatta tttccaaaca
4440

actagcttct cactgtccca ctaagaaca agactattct tatgcactgc tgccttaaa
4500

attatcagct gctttgtggt tatttcatgg agtaacacat gcaatatttt ttacaagaaa
4560

5 taagaaacca cagcagcaga agacgaatct gcttttattg aacgcagaaa tagatggaca
4620

ctgaaacaac agaagtttct accatgtttt gctgctgtga gagcaaatct gctgttgatc
4680

10 aactattagt ttgaaaagt atatgcaact ctgctatcaa ctccctcccc acatcacatc
4740

tcgtttcaag gagaaccttc ctgtccaggc aactagccag ggattttaat tttaacattt
4800

15 gctatctcag ataatctttt aatatttcta ggccgtgcac ataaccaca tagccctaac
4860

20 tccagatggt cccagcatca cttagcagta gttccaactg aagcataaat acagcactgg
4920

cacaacaaat gaagacaatt caagcacgcc tttacaagc atacagaaca tacagtttta
4980

25

<210> 17
<211> 4980
<212> DNA
30 <213> Artificial Sequence

<220>
<223> ENSGALG00000012629_ENSGALG00000012628|chrZ:42172748-42177748

35 <400> 17
acaaattaa tgtatthtct ctgatacta ttttgattat aagctattaa cttgatttgt
60

40 aaaagcaccg atattatggt taatgtatgt atggtaactg ttgagtcaca gctgaacct
120

ctgattgacc acctgcggaa aggtctgggt gagccctggg agcacaggca aaaacaattc
180

45 agttgtgcaa ctggaaggcc tctcttagac ctcatthtaag ggctgactgt cactggggaa
240

ggatctgatt ctggagatcc cttattggtg aagctthtccc ctgcaaacct agaatcttct
300

50 gatgcaagtg agcaatcttc thccctthccc tagcagctct ctgthgcacc agthcctthca
360

tccttgacc tttgttgtaa tgcttttccc atggcattga tctgtccagt tgctacagtg
420

5 tagaacaag aaagtgtttt ctcaaacat tatacatgct tactatcaga tcttacttaa
480

cattgataag aagtacacct ttcattacaa cagtgaacgg ataggcagca agcattcaga
540

10 acaatgaaaa agatctgcgt tctattgtaa agcaaaaaa ggtgatgacg gaacgtttat
600

ttgagataca ggcacatag tacaagatat caaagaggaa ttaattcaac cctcccaaca
660

15 acagttggga aaaaaaagg ataagctgag gaatgcagcc ttctccgatt tttactctgg
720

20 tgccggaagg aggaggagga gcgggggaaa acccagcggc ctcagactgc atacgcactc
780

tcggaagtac ctctgaggcg gggtgagagg aagggtggat ctaggctccc cccgctccac
840

25 aactcacgt gctcgtgtac tggttcaggc tgtgacttgg caactcctct acatctatcc
900

atctagttag actcaacaag accatcgtga ttcaggagtc tgaagatgct taactctctc
960

30 agaacgtcta gttctcctgg cctggtaaat gctatgtttc tcatactgcc tctctgaaga
1020

atgcttcgaa tgcttctagt gatgctctaa agttctaaac agaaaaatct ggagacagtt
35 1080

ctggctctta gatagaaaa atgccaacat gccaaaggat ggttacatcc ttcaagcaac
1140

40 cttgttgcat gctgtacaat agactcatgt aataacttag ccgtagtcat cgtatctctt
1200

atctacctgt tcgttattac attttctctgg tactgcttta tatttagtca gttgtccttt
1260

45 taagacaaat tttttggtgt gtgctaatag gcagttacca aatggtctag agggagggaa
1320

tataatcagt gagtatgtag atgtatatag atgtataacc agtaagtata cagagaagtt
50 1380

agggtggtct tttgagagca tatagctgga aagattatca catgggaaaa ggtaatcaga
1440

aataaatgga aaaatgctca ccagtgtcat aggctcacia gaaaactccc caaggagcaa
 1500

5 tctccagtaa cagacccttc ctcagtggca gtcagccctt aaatggcatc taggagaggt
 1560

gtagccaggc accaccctt ctgtcacata ggtgaattgc cttcagctgt gtcctgcag
 1620

10 ctgactcagt gttcacctca ggtgatcaat cagaggtcca gaccatgatt caacatttcc
 1680

catgcgagg gagatgaaac aagactgtca tctcatcttc tgaagctgat aggagaatgg
 1740

15 ttacaataag taggtcccct tccactgggt tagctctgta ttactcagtt taaaacttta
 1800

tgtagttaat ggagttccat agaatcacat cgctggaaaa tgtatgtgat agctataccc
 1860

20 aaaaagtgac acagcatcac agcctctacc aggtcctctg cctccagaca ggaccttgag
 1920

25 ctgtgtgatc cctttgggat gtttctgttt aagacagggg acacttacac attaggaac
 1980

tggagtcatg tcgggtggag agaaagtggc tgtagaaaag gtaacggaca agaaatcctt
 2040

30 cactggctta gctgtaaagg tcagaggaaa atcatcttct gttctccata cacatctcct
 2100

35 gtaattaggc acctgtgcc tttaaaaaag taaagtaatc aaagagtcac ttggtgaaaa
 2160

tagaagccaa acagttacaa aatttagtta atattcaaag acaaaacaac tctggttcca
 2220

40 aatataaaac cctttgtgta aacactgagg gaaggtgcca aagagccacc tacacaaaga
 2280

gatgccacag agtaatttgc gtaccccaac aataggatca tatatgggca aaccaaaga
 2340

45 ccaacaaaga ccccttgcaa taatcccttg attggcagaa gcaagtgcag cgttagattt
 2400

50 agtatgtgtg acaattctga gaataggaag gaatttcaact gaaactgtgt gaagatttgt
 2460

atgttaaata catataatga gggttattta gctctggggg gtgcatgcta tgtggagaga
 2520

tccccatgtg cccagcaactg caatagacta atgtcagctt tttaaaccat catttggttt
2580

5 gaagagttta ttatgatttt cagaaacaga attatgtatt cagtgtctgt acattcttac
2640

aagcctgctg ttctgtacac atgaaaaagc tacttatggg gcaaatcagt atagagaagc
2700

10 agttttgaca tacagtgggc atgactgggt gatctgctgt cagaaggtaa gacactggta
2760

tacagaattg caaacgtgag atgaagacct caagtgcaac tctgtcatca taccacttct
2820

15 cttgctatta tctattcagt gtttgttcta aatattcagg cagaaggctt ggtttcacat
2880

20 cagggtttac atttaagtat tctttgcca ttttttgctg tttgtatgtg taacttcaga
2940

ttctcacatt gactgtgtag tgtaaatttg caaatagatg taacagcttc tctttaacat
3000

25 cccatgccaa acttttaatc actttcaatc aatgttggaa atctgcagta cataaatttt
3060

attttgatcc attattcaga gaatcagaga aataaggact tttttagagg tttatgttta
3120

30 ttttagaaaa tataaataaa tcttgatac atagatttcc aactagtaaa atttttagct
3180

atcgctgtag catttctatg aggagtttta ctttttactg tctcttatga aacttagtgg
3240

35 aaacaaagct aggtttcata aaacgtctaa tctcatattc atctctctcc aaaataggt
3300

40 gaccttcaaa acaagaacag ttgcttatg taaacatttc ttacattgta ttctattaca
3360

ttatcatatt catgaagata ctttagaatt ttgttgctgt tagttctggt gtgactggct
3420

45 gcttttctca gcttaagagc tctgcagaag tgatgataat atatattctc tttgctgagt
3480

50 tacaggaatt tctgctttc ctctactgta cgttggagcc caggagatca gatccaacta
3540

ccagaatttt atatagctct ttgtgatcag tccttcocta gttgctgta ttctattga
3600

tgcaaacctt tttatggtgt attaaatgaa gaagatataa tgaggacagt ctaataatca
 3660

5 aagttttttt ctgcagccag tagtacaggt aggggtgacat gccttgcata aacagatcca
 3720

gatgaaaggt gtgggtgagg caataagtct tcaggccctg ggctaaggca ggacctgttt
 3780

10 tgtttcatga gatatgccta ggcctctcta tagacagggg gagatgtcag agaattctga
 3840

agcatcagca gaggctatct tggttttcag gccctgtgct aagtatgaga ggaagttgag
 3900

15 aacagataat gtgtgcaaca aagttagcaa catgcttttt caaacagag gaaacgatcc
 3960

20 ataaatactg gaacagaaca acctcacaga aatcactgac tgtccctttg ctggccatgg
 4020

catcctgaag gttgtgctgg tagcactcat tggggttctg tagctgataa ggcagtgcaa
 4080

25 ggacattggg tctgcacctg ctatgaacag gcactggggc ctccaggagc tgctgggtatt
 4140

agtcatgggc atcagaggtg aagaaggcac ttgtagttct cacagctctg gcactgaccg
 4200

30 atatttctgc tgtagttaat gggattatth gcatgtagtc aacagtttac aggctgattt
 4260

35 ggctctgctc cgcaactagc aggacagagg gacccacaga tccatgtcct ggaaggggat
 4320

aagggaaaag gatagagaga tgggcctaaa acaaaacaga agcaacgatc tgaggataaa
 4380

40 caaagtaatt tactaaataa gatagtggaa tgcattgataa cacactatag tacaacataa
 4440

tgcaatgtaa ttagaactgg agctaataaa tcaaatataa agagagtgtc tgaaagccaa
 4500

45 aggccttact ctaatgctgt gatgagatgg cacccacaaa agagacgagc cagatgattg
 4560

50 aaagggatgt atgggtgtct accaaggcct gttatctggt tggcctgagc agaaatgata
 4620

gcagaacagc aaacagtcct ctggggaatg tagtagttct tctcttctgg gacaggtagc
 4680

tggacctaga gcattaactc ttaaaccccc agtacactac atgatgttat gatgtggaat
 4740

5 accaataacc aaacatcata aaacatgac atagtcttat gtcgctgtgt tgtaactgga
 4800

agctgaggat ctccactcag tatttctgct gaatgaagca gatctttcag ttgaaggaga
 4860

10 ccaacttaat catccacttc caaccctta ctatggacag gttcacggtt cactagatca
 4920

agctgcccag ggccccatcc aacctagcct tgaatgcctc cagggatgag gcatccacca
 4980

15

<210> 18
 <211> 4980
 <212> DNA
 20 <213> Artificial Sequence

<220>
 <223> ENSGALG00000015617_ENSGALG00000027714|chrZ:63363656-63368656

25 <400> 18
 gaaaatctct gccactgatt gttttgcagc ttagttgtct agttattcgt gagatgttga
 60

30 ttcaaacctt tgctcagatc aggcacagaa gcagtgctga actttgtttt tccgtagcct
 120

tgctagtgtt gtaactgctg tcttagtgaa aacctgctta tttgcatccc cttcccaact
 180

35 ttgggaaagt tgctgtctgt gattccagaa ttaggactcc ctcggttgca acatgactca
 240

gattttcaga ccttggcttc agaggcacag tggcttttgg tgcagagaaa aatgaggctg
 300

40 agaagaccag ggccatgatg agtgaggaag atgcaacaga acaagataga gaagagggaa
 360

45 gagtttttta ggagctccac cagacatact gtagaagcag tgtagatgca cagtaccagt
 420

aatgtactcc agccaaccac tcggtgaaga cccatgaaaa cctaataagg tttagtaagg
 480

50 ccaagtgcag ggtgctgtac ttgggtcggg gcaatcccag gtatttatac aaactagggg
 540

aagatctcct tgagagcagc cctgcagaga aggacttggg ggtcctggtg gatgagaagc
 600

tggacgtgag ccagcagtgt gcacctgcag cccggaaggc caactgtggt ctgggctgca
660

5 ttaaagaagg ggtgaccagc agggagaggg aggtgattgt cccctctac tcagctcttg
720

tgaggcccca tctggagtac tgcattccagg cctggagccc ccagtacagg aaggacgtgg
780

10 agctcttggg gcgggtccag aggaagacca ctaagatgat cagagggctg gagcacctct
840

15 cctatgagga aaggttgagg gaactgtcct tgtttagctt agagaagaga aggcttcagg
900

gaaacctcag tgcagctttc cagtacttgg aaggagcaca taaacaggaa ggggaatggc
960

20 tgtttacgag ggtggatagt gacaggataa gggggaatgg ttttaaactg agacgggaga
1020

gtttttaggtt agatattagg aggaagtttt tcacacagag ggtggtgacg cactggaaca
1080

25 ggttgccccaa ggaggttgtg gatgccccat ccttgagggc attcaaggcc aggctggatg
1140

30 tggctctggg cagcctgctc tgggtggctgg caaccctcca catagcaggg gggttgaaac
1200

tgcgtgatca ttgtagtctt tttcaaccca ggccattcta tgattccatg aagactctgt
1260

35 aaggacccca atgcatcag ctgtccccac agcttcagta cccactgcca ccacattttt
1320

actgagatag gtagtttggg agccttcctc ccatcttggg taggggaggg ttaactgctg
1380

40 gcatttttgt tctttcctgc aatatgtggt ttctgcatgc attctctttt caccctaaat
1440

45 tttaatgtgg acggactttg aggatattct gcatctgccc aatatttctt acagcctcac
1500

atctactgat ttatgcacag tatcttatat ataatgtata ttatattatg tatattatgc
1560

50 actgaatcac atctactgtg tatgcacagt aacttctcct atggtactgt aagcccagaa
1620

atcccagata tgcactaca acgtgtctgc tatgtattgc ttctgtgaga cacagatgtg
1680

ataatcagag gctgtacaca gtagaaatgc atacatatcc tactgtgaaa tctctgaaat
1740

5 ctagccttaa tcttggaaca gaaatgaata tgggtgacatc ttgatctgat agaattgggt
1800

gccagtagca cagctgtaat catccatgat atgaatcaaa ccaagcacag gtaaacaggt
1860

10 gagagaaatc atgaacaatt acatgcaaac ggaccagcta aaatgtgttt gtttgttgtt
1920

15 tttttttttt caggtgattc ttgattacag taagatcaga agctgctaca ttagcagacc
1980

agccactgca ctcaaggctg tgattcacag cttgcagacc tgacagcagt tctgtggaag
2040

20 aggcaggtcc ctgtcaacca gtttaatcaa taaatcagtc tcgtgtacac aaataatggt
2100

atcctgcacc actgctgggtg ttacactatt tcacccaagt ttatcaccag caaactgagt
2160

25 cttatcgttc ctactgtgct ttgcttttct tgcttagcaa aacatatcag aaaggtttct
2220

30 agttaaaaaa aaaaaaaaaa aaaaaggatt atactgcagc catcaaagag ttctgttaaa
2280

gaagttgtgg ttcattttct tttcaagttc tgaagacaat ctgacaatag aggagatgag
2340

35 ggttttagtgg aggtaaatgg agctgtttct gaaagtaggg caccgtcgtg aggtaatctc
2400

accactagat gttgctgtaa gcatgatgag gggagaagct ggagcacaaa aggtgtaaac
2460

40 ccaatggttg atcatgcgca ctcgacaga ggaaatcttc aaaatgggaa agttagcttt
2520

45 gaactgattt ggtgggttcc cattaaaaat gggggactta ctgccttatt gtctttatgt
2580

accatttgg gggatagatg ccatacagaa acaccagagc ttacgggtgtg aattcattgc
2640

50 acatcttatt agatcaacaa aatggaataa aagaatacag aagattacta ctccattggg
2700

catgtggagc ttttacaggc ctggataaat tagatcttaa aaacaaacaa acaaacaaac
2760

aaaaaaactc cttgcaaaga taatgttatg taatattagt tgcaagaagt aagcaaacac
2820

5 acagaactgg gagcagaagc aaagcactaa gttattaaag caagttgcac attttgagtt
2880

gcattttgcc actggtttta taaacatggt tagcatgtct ggtcagaatt tgggcaccag
2940

10 gatgctttta agatgtctgt ctatggaacc tgtcagtgtc caagaataac ttctgttatt
3000

tggatgctgc accaaagaat tcagaggaag acgagccaag ccagacgtta tcatagtcac
15 3060

tagtaaagtg gttctaagcc taattaagac atgtcagaac tatgtgttgt gcaaccaa
3120

20 cctccaaaag agaaatcaga ggtgaacttg tgcaataaat atagaagaca cgtaa
atcct
3180

gaggcagtta gctaaccata tgaagccaat catacctgac tgctggacgc aggagactga
3240

25 accttacaga ccctggagaa tcactgtttg gcctagttag gcctgaatgg aatttacc
aa
3300

gattcatgac ttttaacatgg atcaggtgca aagaaaagaa ggctcagtta gttcctacag
30 3360

gctaccagat ctttttccac ctctgctcag cctggagctg tgggctctac ctgccctga
3420

35 agtgaagtgg catcaactgc aacacttttt gcagaggcaa aaatcatcaa gtcgtgcctc
3480

tgcattttta ggtgatggat tccaagagat agattccaaa cccaacactg agattcctct
3540

40 ggtgatgcag atgactgctt actgggatcc ttctctctct atagccta
aat
ccc
atgcagc
3600

acaaaaagtt gggcacctga atactatcta tcttctctca tctgacctct gctgctggct
45 3660

catcacgcat ggctcctgag cctgacagag ccatggggaa agacatggcc tgaccttttc
3720

50 agacagcagc acttctctca gtgtgctgca ctaacgacaa tatta
aaaaat
aacactgcag
3780

gaaaacatgc tgttttctag caggctggag gaaaaaaaag gtgttcaagg tgagactcta
3840

taagatcact ctttccaaac acagcagtta actcttttct cacttgaagc tgaaattcat
3900

5 cttcagtagg tgacggttac aagcagctct ccaccggtaa gtgaaagttt gctggggttt
3960

tttcaggcac ttcaagacac gaaggtaatt gctgcaacac cttcctagct cttcctgagt
4020

10 actcagccca aagctacact cttcagggct actctcctag cacggctctc tgtaattagc
4080

ctggctttct aacttatgaa ttaacgctgc agctactcaa cagggcagct ttgcagctat
15 4140

gaactttatg catacacttt ttattctggg tggtaacaaat tcattgctaa gaaatcttat
4200

20 aaactctatg gaacactttt aaaagcctac atgtgattaa aaacaaaaaa acaaaaaaca
4260

ctgatatcac atgctatatt ctctagaaaa gagtattttg gaaacaattg taattcacag
4320

25 ctcaggagaa agagcaattg ttttcctttg tgctggagct ggaattctat gcagtataaa
4380

tacagatgtc caaacaaact cattaatgg tagtcatagc aaatggagga aaaaaaagg
30 4440

gggggggggg ggggaactgaa gacactgaag aaattctgct gaaattcatg gagctgcaac
4500

35 aagaattgat attgattact tcaccatatt ccgctctaataactagcac aacctgact
4560

gttgtgaaaa caagcatgtg atttgacca ataagtgtt cttttagtaa ccaaaaccaa
4620

40 acaacaacaa accttctggc tgaccagatg gaggataaat tctaacctga tcattctgac
4680

acaaatgagg aattatgtag gttggaactg cattccaaag caaagcatgt ttttggtta
45 4740

gttgaacaca cagaattcat ttttcttgct gattcttgta actaagggca ggtatgaatt
4800

50 ggttccctta aaagaagtcc ccagaagcca tggctctcaa caagtcttct cagattagca
4860

aggatgaacg actgagaaat caaaaccaa attcagatct aagagccttc aacagatttg
4920

ctaaatgaaa ctagcaaata ctgtagaagt taacttgatt ttgaaataaa acaggatttc
4980

5

<210> 19
<211> 4980
<212> DNA
<213> Artificial Sequence

10

<220>
<223> ENSGALG00000008188_ENSGALG00000008197|chrZ:78777477-78782477

<400> 19

15 aaaggagttt cttgcctgaa acccctaag attttgtgta ttgttgttgc cctggtttca
60

gctgggacag gactgatatc cttcacagta tctgggatga tgctgtgttt tggttttagg
120

20

agaaaaccag tgctgatggc acaccgaggt tgtagttgcg actgagcagt gctgcacaga
180

25

gccaaggatg tttcagctgc ccacactggc cttcctctaa gggggtgtga gggcacatgg
240

agctggaggg aaccaaaggg gtattccaca ccatatgtca ctgagcaaag acattgctgc
300

30

ccaggcagag aatacagtgg tagaaatagc tcacataggt gcccatctac ccagaactg
360

ggccagcgaa gaacatcaga acaacaagca tgcgaatcag gctgctcaga ttgatgtggc
420

35

tcaggtagat ctcaactgtc aactgtcacc cataaaatgg atgagacagg agcaaggcct
480

40

gcctgtccaa actgaccttg gggccccggg cacagctgct cccgagcaag gagaggcaga
540

gaattgtgga ggaaaccctc ttccacctag ctccaagggt gcacgctcaa tgcttggatg
600

45

aagattcaag ttctcaatga agaactgcat tcaaagacat ctttgagggt gaccagact
660

gcctctggga tgaatccaga acaaacagaa tcttcccagc aacaccttg tttatcatac
720

50

acgttcagaa atgcagtgcc tgcttggtat ttttttaaac tcatcacaga ggaatctctg
780

agtggagcag aggagtgttt cctcttgctt ttttcttccc cctttctgta agaaatgcat
 840

5 atgccagttt cccctaaat gttttcaaac tccaaattgt ttctgctgg tgacattctc
 900

ctctccagca ctgcctacac cccagctgcg tctgatagga aaagcaccca gtaccatgct
 960

10 ggcattggcat ggtcagtgc acccactggg tagcatgggg acaaactgcc agagcagagg
 1020

gctgctgaaa cgccgagtgc caactgacac tgttcagctg acagcctcac gagagtgctg
 1080

15 cgaggggtaa gtgaggcaac aaaatacaag tactcaaaaa tagaatggaa tggaatggaa
 1140

20 taacagaggg aatggggaac aggctgcagt gagaaggaaa ccccgtgca ccgaccaacc
 1200

cgctcccagt agcctggctc ctaccttgcc ccaactccagc agctccacga ggtactcgaa
 1260

25 gtgccccatgt ccggccatcc ccagcatggc acagctgtca cctgggagca gcgggggtcc
 1320

ggcaggcatc gccggcacc acagcctggg ctgggctgag gatgggcttg gtgggcagct
 1380

30 ggggtgctgcc gtggcagctg agtgccgtcg tgcagggtgg ctcgggtgctc tctgcagccc
 1440

35 acggctcggg acccgtggtg ggcatactgc tgggcaaggg gacggctcgg atgtgcatcc
 1500

cagagcaggg tgcattgatgt gtgcgcagcg gcacacagca gcgcacttgt ccccgaggca
 1560

40 ccacccggtc tgatgaaatt ccagggttttg ttttaagatga aatgcatgt gggcatacat
 1620

ttgcacatga gaacacctgc taatgaaaca cacagagttg tatgtgcagt ctccctggct
 1680

45 ggtggcatgt ctgtgggggc agttaggcag gaagaaaagc tccaccatct gtttctgtgt
 1740

50 cccgggcttg gtgctgccac tagagatggg atgcacacag catgcctgcg tgctgccgag
 1800

gcagctcagt acctgctgct gggggagagc agccctcagg gagctgctgc aaagaccccg
 1860

tgggatggct ggtgaacagt gctggagaca tggcagatgt cgcaccctg ctccacacca
 1920

5 cctccgctg cagctgctg cacgatgcca tttcagcgtg ccacggcttg tttcttctc
 1980

ttgcagtctc aggttgccgc agcagcagct gcctgaggtc tggaggcaaa gggcagagca
 2040

10 cccagccccac tgctctccat gagtgagcgg gtagggggca cgcctgctt cgcttctggt
 2100

atcgggtggc ctgtggagca catcaccctt ggacagtga aacatgtcaa aggggtgttt
 2160

15 ataacagtat agcatatccc tttgaggctg aatttctga gcttatggat atggtagtaa
 2220

tgctatacca ggctctctct gcaggttgct cactgaaaca atataccctt tctttctcaa
 2280

gaaatgggac tcatgaatag ctaccaggcc tttctgctac tagatactgt agaattacat
 2340

25 atatcagatg gctcattaat cacatagctt gttattggga cagaggcagg ttttcagcat
 2400

ttccacact gcttttctgc aatggatta ggagtaaaga gtcagatata caggggacca
 2460

30 tcaactcactt cctaggccat gcttcgtagt ttaactccat gcaaggtatt ttcttgctct
 2520

gatctaactc tagatcacta aatggcactt gtgcaggaca cttttccact gtcattttgg
 2580

35 ggggtagaag tgtggtcaga gccagggagg cttgcagggc ccacaccagc tgtggcccaa
 2640

40 gcagctgctg tacaatggct gtatggcagt atcgatgtga acttctgat aataaacaga
 2700

actcagcagc aaataaaccc agattgttct gatcagtaac gagaggctgt gcaaaaagct
 2760

45 gagaaatgta cagcccttcc tgctcaccac tacagccctt catgcagcgg cctgctgcac
 2820

tagggcctgc ttggccagag gcagggctgc agcttgagat acgcagcacc cagtacccca
 2880

50 gcaggctgcc atgtgcttct gtcaggccat gaattgcaca agttggtata gttttataca
 2940

gtgtcatgca gaagcacaac tcttaccttg gcacactgat atggaaagca aagtgtaaag
3000

5 gaactgtatg tttcttggtc tgtgagttcc attggttctc ccaggacagc agtgacagct
3060

ctagattttc tggttgatct ctccatcta agtttttact tctgaatata tctgtgcagc
3120

10 aaagcctgat gctttttttt tccccctcc cctgggatcg tgaaatatta ttttcccatc
3180

tccataatac agatgatcca ctcaaactg catcctcagt cgcagtcctc agaagggcct
3240

15 tttcccacca tacacctaca cgctacacca gactttattg gtcaccatgc cttgagggat
3300

20 tctgcctaac acaggcacac aggaagggct gataccagta attcattgct gaggagagaa
3360

ggggcacttc ttttatagta gagccacaga atcaccagga ttggaaaaca cctccaagat
3420

25 catccaatcc aaccaccac ctaccaccag tatttcccac taaacctgac cccttggtag
3480

cacatctaaa agcatcttct gcaccttcac agcttcattg cctttctcag gacgtgctcc
3540

30 agggccttgg tgtctttctt gtattgagtg gcccaaaatt gaacacaata ctcgaggtgc
3600

35 agcttcacca aagctgagtg cagaggggtg atcacctcta ctaggccttg ttttaccct
3660

tccccctgca tacctagttt aaagcccttt cagcgagccc taccagctcc tgggctagga
3720

40 tccatttccc cctttaagag aggtgaaacc caccgcagc catcaggcca ggcactgagt
3780

taaactgctc catcagctaa ccaccccaat gccctgatgt cccttctgat agtcctcagg
3840

45 cttctctcag cgacttcac actgccagcc tgagctatcc gtaatggata acaaccagag
3900

50 ggctgaacca gactagggag tctcctggta atgtccctaa gtaatgaatt catgggggaa
3960

atggatctgt gcagatgcac tgccatctca cagggtcaa ctgtcacaga cagttgtcac
4020

aaaccagagc agcagagaga gcagcaaagg agatatcata aatagtggaa aagggtgcc
4080

5 gtgtttgtta ttgactgctt cactgtgtgt ttggaaatca caagaaaac aacaaaaaa
4140

ccccacagaa tacacatttc caagcatgca tgcttcaggt aaggactgcg gcccgttta
4200

10 attcccctgg aagcgtccag aggcctacac acaaatcagg ccacaacaaa caggttccgt
4260

gcaagcagtt tgtgagttgt atgaaaatag ctttcttcca taaagcgtgg cattgatggg
4320

15 aactgttact agcagatgaa aatctgccat gcacgacacc cttaatggtt ggagtggttc
4380

20 gactgagttt tgcattcttc attctgatgg atgctggcaa aaccaagatc acccaacacg
4440

taacactgaa tatagctcac caacagcaac gagccacacc agatactgag gcatctcaa
4500

25 ggctacgtct ggtggaaaga agagacagtc tcaccctgca gtgcaaaagc atcagggctt
4560

ctgcgtaagg tttattgatc ccaaagatag gcaagttaat attcatcagt ttaatagaca
4620

30 ggtactcaca ttactggaga tggcatgcca ggagagcctc tgccaagtgc tcaactgccag
4680

tggcagccgg ccagtgacaca gcttcttggt gctgtggcca tctatataac agaaggaagc
4740

35 aattaagcat cacctatcaa gataagcaag atcattcaaa gagaataagg gtacttggtt
4800

40 agagctgcag agcacttctt cacgaatgag tttcttccag ctacatgcag atgttacagg
4860

cagtttgctt gtgttctagg gataacacaa acccctctc atgggctagc tgctgctca
4920

45 ggctgtggtc tgtatTTTTG cacgtgcact gctagccttt acttcacatt tcctttact
4980

50 <210> 20
<211> 681
<212> DNA
<213> Artificial Sequence

<220>

<223> pDsRed1

<400> 20

5 atggtgcgct cctccaagaa cgtcatcaag gagttcatgc gcttcaaggt gcgcatggag
60

ggcaccgtga acggccacga gttcgagatc gagggcgagg gcgagggccg cccctacgag
120

10 ggccacaaca cgtgaagct gaaggtgacc aagggcggcc ccctgccctt cgctgggac
180

15 atcctgtccc ccagttcca gtacggctcc aaggtgtacg tgaagcacc cgccgacatc
240

cccgactaca agaagctgtc cttccccgag ggcttcaagt gggagcgcgt gatgaacttc
300

20 gaggacggcg gcgtggtgac cgtgaccacag gactcctccc tgcaggacgg ctgcttcatc
360

tacaaggtga agttcatcgg cgtgaacttc ccctccgacg gccccgtaat gcagaagaag
420

25 accatgggct gggaggcctc caccgagcgc ctgtaccccc gcgacggcgt gctgaagggc
480

30 gagatccaca aggccctgaa gctgaaggac ggcgccact acctggtgga gttcaagtcc
540

atctacatgg ccaagaagcc cgtgcagctg cccggctact actacgtgga ctccaagctg
600

35 gacatcacct cccacaacga ggactacacc atcgtggagc agtacgagcg caccgagggc
660

cgccaccacc tgttctgta g
681

40

<210> 21

<211> 226

<212> PRT

45 <213> Artificial Sequence

<220>

<223> pDsRed1

50 <400> 21

Met Val Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys
1 5 10 15

Val Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly
 20 25 30
 5
 Glu Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys
 35 40 45
 10 Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro
 50 55 60
 15 Gln Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile
 65 70 75 80
 20 Pro Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg
 85 90 95
 25 Val Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser
 100 105 110
 30 Ser Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val
 115 120 125
 35 Asn Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp
 130 135 140
 40 Glu Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly
 145 150 155 160
 45 Glu Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val
 165 170 175
 50 Glu Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly
 180 185 190
 Tyr Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp
 195 200 205
 Tyr Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu
 210 215 220
 Phe Leu

225

5 <210> 22
<211> 876
<212> DNA
<213> Discosoma sp

10 <220>
<221> misc_feature
<222> (1)..(876)
<223> red fluorescent protein (FP593)
GenBank: AF272711.1

15 <400> 22
agtttcagcc agtgacaggg tgagctgcca ggtattctaa caagatgagt tgtccaaga
60

20 atgtgatcaa ggagttcatg aggttcaagg ttcgtatgga aggaacggtc aatgggcacg
120
agtttgaaat aaaaggcgaa ggtgaaggga ggccttacga aggtcactgt tccgtaaagc
180

25 ttatggtaac caagggtgga cctttgccat ttgcttttga tattttgtca ccacaatttc
240

30 agtatggaag caaggtatat gtcaaacc ctgccgacat accagactat aaaaagctgt
300
catttcctga gggatttaa tgggaaaggg tcatgaactt tgaagacggt ggcgtgggta
360

35 ctgtatccca agattccagt ttgaaagacg gctgtttcat ctacgaggtc aagttcattg
420
gggtgaactt tccttctgat ggacctgta tgcagaggag gacacggggc tgggaagcca
480

40 gctctgagcg tttgtatcct cgtgatgggg tgctgaaagg agacatccat atggctctga
540

45 ggctggaagg aggcggccat tacctcgttg aattcaaaag tatttacatg gtaaagaagc
600
cttcagtgca gttgccaggc tactattatg ttgactccaa actggatatg acgagccaca
660

50 acgaagatta cacagtcgtt gagcagtatg aaaaaacca gggacgccac catccgttca
720
ttaagcctct gcagtgaact cggctcagtc atggattagc ggtaatggcc acaaaaggca
780

cgatgatcgt tttttaggaa tgcagccaaa aattgaaggt tatgacagta gaaatacaag
840

5 caacaggctt tgcttattaa acatgtaatt gaaaac
876

10 <210> 23
<211> 230
<212> PRT
<213> Discosoma sp

15 <220>
<221> MISC_FEATURE
<222> (1)..(230)
<223> red fluorescent protein D. GenBank: AAG16224.1

20 <400> 23

Met Ser Cys Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
1 5 10 15

25 Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Lys Gly Glu
20 25 30

30 Gly Glu Gly Arg Pro Tyr Glu Gly His Cys Ser Val Lys Leu Met Val
35 40 45

35 Thr Lys Gly Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ser Pro Gln
50 55 60

40 Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
65 70 75 80

45 Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
85 90 95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Ser Gln Asp Ser Ser
100 105 110

50 Leu Lys Asp Gly Cys Phe Ile Tyr Glu Val Lys Phe Ile Gly Val Asn
115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Arg Arg Thr Arg Gly Trp Glu

130 135 140

5 Ala Ser Ser Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Asp
145 150 155 160

10 Ile His Met Ala Leu Arg Leu Glu Gly Gly Gly His Tyr Leu Val Glu
165 170 175

15 Phe Lys Ser Ile Tyr Met Val Lys Lys Pro Ser Val Gln Leu Pro Gly
180 185 190

20 Tyr Tyr Tyr Val Asp Ser Lys Leu Asp Met Thr Ser His Asn Glu Asp
195 200 205

25 Tyr Thr Val Val Glu Gln Tyr Glu Lys Thr Gln Gly Arg His His Pro
210 215 220

30 Phe Ile Lys Pro Leu Gln
225 230

<210> 24
<211> 4107
<212> DNA
<213> Streptococcus pyogenes

35 <220>
<221> misc_feature
<222> (1)..(4107)
<223> M1 GAS
Cas9 protein id: AAK33936.1

40 <400> 24
atggataaga aataactcaat aggcttagat atcggcacia atagcgtcgg atgggcgggtg
60

45 atcactgatg aatataaggt tccgtctaaa aagttcaagg ttctgggaaa tacagaccgc
120

cacagtatca aaaaaaatct tataggggct cttttatttg acagtggaga gacagcggaa
180

50 gcgactcgtc tcaaacggac agctcgtaga aggtatacac gtcggaagaa tcgtatttgt
240

tatctacagg agatTTTTTc aaatgagatg gcgaaagtag atgatagttt ctttcatcga
300

cttgaagagt cttttttggt ggaagaagac aagaagcatg aacgtcatcc tatttttggg
360

5 aatatagtag atgaagttgc ttatcatgag aaatatccaa ctatctatca tctgcgaaaa
420

aaattggtag attctactga taaagcggat ttgcgcttaa tctatttggc cttagcgc
480

10 atgattaagt ttcgtggtca ttttttgatt gagggagatt taaatcctga taatagtgat
540

gtggacaaac tatttatcca gttggtacaa acctacaatc aattatttga agaaaaccct
15 600

attaacgcaa gtggagtaga tgctaaagcg attctttctg cacgattgag taaatcaaga
660

20 cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggctt atttgggaat
720

ctcattgctt tgcattggg tttgaccct aattttaaat caaatttga tttggcagaa
780

25 gatgctaaat tacagctttc aaaagatact tacgatgatg atttagataa tttattggcg
840

caaattggag atcaaatatgc tgatttgtt ttggcagcta agaatttatc agatgctatt
30 900

ttactttcag atatcctaag agtaaatact gaaataacta aggctcccct atcagcttca
960

35 atgattaaac gctacgatga acatcatcaa gacttgactc ttttaaagc tttagttcga
1020

caacaacttc cagaaaagta taaagaaatc ttttttgatc aatcaaaaaa cggatagca
1080

40 ggttatattg atgggggagc tagccaagaa gaattttata aatttatcaa accaatttta
1140

gaaaaaatgg atggtactga ggaattattg gtgaaactaa atcgtgaaga tttgctgctc
45 1200

aagcaacgga cctttgacaa cggctctatt ccccatcaaa ttcacttggg tgagctgcat
1260

50 gctattttga gaagacaaga agacttttat ccatttttaa aagacaatcg tgagaagatt
1320

gaaaaaatct tgacttttcg aattccttat tatgttggtc cattggcgcg tggcaatagt
1380

cgttttgcat ggatgactcg gaagtctgaa gaaacaatta ccccatggaa ttttgaagaa
 1440

5 gttgtcgata aagggtgcttc agctcaatca tttattgaac gcatgacaaa ctttgataaa
 1500

aatcttccaa atgaaaaagt actaccaaaa catagtttgc tttatgagta ttttacggtt
 1560

10 tataacgaat tgacaaaggt caaatatggt actgaaggaa tgcgaaaacc agcatttctt
 1620

15 tcaggtgaac agaagaaagc cattgttgat ttactcttca aaacaaatcg aaaagtaacc
 1680

gttaagcaat taaaagaaga ttatttcaaa aaaatagaat gttttgatag tgttgaaatt
 1740

20 tcaggagttg aagatagatt taatgcttca ttaggtacct accatgattt gctaaaaatt
 1800

attaaagata aagatttttt ggataatgaa gaaaatgaag atatcttaga ggatattggt
 1860

25 ttaacattga ccttatttga agataggag atgattgagg aaagacttaa aacatatgct
 1920

cacctctttg atgataaggt gatgaaacag cttaaacgtc gccggttatac tggttgggga
 1980

cgtttgtctc gaaaattgat taatggtatt agggataagc aatctggcaa aacaatatta
 2040

35 gattttttga aatcagatgg ttttgccaat cgcaatttta tgcagctgat ccatgatgat
 2100

agtttgacat ttaaagaaga cattcaaaaa gcacaagtgt ctggacaagg cgatagttta
 2160

40 catgaacata ttgcaaattt agctggtagc cctgctatta aaaaaggat tttacagact
 2220

gtaaaagttg ttgatgaatt ggtcaaagta atggggcggc ataagccaga aaatatcggt
 2280

attgaaatgg cacgtgaaaa tcagacaact caaaagggcc agaaaaattc gcgagagcgt
 2340

50 atgaaacgaa tcgaagaagg tatcaagaa ttaggaagtc agattcttaa agagcatcct
 2400

gttgaaaata ctcaattgca aatgaaaag ctctatctct attatctcca aatggaaga
 2460

gacatgtatg tggaccaaga attagatatt aatcgtttaa gtgattatga tgtcgatcac
2520

5 attgttccac aaagtttctt taaagacgat tcaatagaca ataaggtctt aacgcgttct
2580

gataaaaatc gtggtaaadc ggataacggt ccaagtgaag aagtagtcaa aaagatgaaa
2640

10 aactattgga gacaacttct aaacgccaaag ttaatcactc aacgtaagtt tgataattta
2700

acgaaagctg aacgtggagg tttgagtgaa cttgataaag ctggttttat caaacgccaa
15 2760

ttggttgaaa ctgcgcaaat cactaagcat gtggcacaaa ttttgatag tcgcatgaat
2820

20 actaaatcgc atgaaaatga taaacttatt cgagagggta aagtgattac cttaaaatct
2880

aaattagttt ctgacttccg aaaagatttc caattctata aagtacgtga gattaacaat
2940

25 taccatcatg cccatgatgc gtatctaaat gccgtcgttg gaactgcttt gattaagaaa
3000

tatccaaaac ttgaatcgga gtttgtctat ggtgattata aagtttatga tgttcgtaaa
30 3060

atgattgcta agtctgagca agaaataggc aaagcaaccg caaaatattt cttttactct
3120

35 aatatcatga acttcttcaa aacagaaatt acacttgcaa atggagagat tcgcaaaccg
3180

cctctaatac gaaactaatg gaaactgga gaaattgtct gggataaagg gcgagatttt
3240

40 gccacagtgc gcaaagtatt gtccatgccc caagtcaata ttgtcaagaa aacagaagta
3300

cagacaggcg gattctcaa ggagtcaatt ttaccaaaaa gaaattcgga caagcttatt
45 3360

gctcgtaaaa aagactggga tccaaaaaaa tatgggtggt ttgatagtcc aacggtagct
3420

50 tattcagtcc tagtgggtgc taaggtggaa aaagggaaat cgaagaagtt aaaatccggt
3480

aaagagttac tagggatcac aattatggaa agaagttcct ttgaaaaaaa tccgattgac
3540

tttttagaag ctaaaggata taaggaagtt aaaaaagact taatcattaa actacctaaa
 3600

5 tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaatta
 3660

caaaaaggaa atgagctggc tctgccaagc aaatatgtga attttttata tttagctagt
 3720

10 cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag
 3780

15 cagcataagc attatttaga tgagattatt gagcaaatca gtgaattttc taagcgtggt
 3840

attttagcag atgccaattt agataaagtt cttagtgc ataacaaa tagagacaaa
 3900

20 ccaatacgtg aacaagcaga aaatattatt cattttatc cggtgacgaa tcttgagct
 3960

cccgctgctt ttaaatattt tgatacaaca attgatcgta aacgatatac gtctacaaaa
 4020

25 gaagtttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt
 4080

30 gatttgagtc agctaggagg tgactga
 4107

<210> 25
 <211> 1368
 35 <212> PRT
 <213> Streptococcus pyogenes

<220>
 40 <221> MISC_FEATURE
 <222> (1)..(1368)
 <223> M1 GAS
 Cas9 protein id: AAK33936.1

45 <400> 25

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
 1 5 10 15

50 Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
 20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 35 40 45

5 Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 50 55 60

10 Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80

15 Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
 100 105 110

20 His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 115 120 125

25 His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 130 135 140

30 Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 145 150 155 160

35 Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
 180 185 190

40 Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
 195 200 205

45 Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
 210 215 220

50 Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
 225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
 260 265 270
 5

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
 275 280 285
 10

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300
 15

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
 305 310 315 320
 20

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 325 330 335
 25

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
 340 345 350
 30

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365
 35

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380
 40

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400
 45

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
 405 410 415
 50

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430
 55

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445
 60

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460
 65

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
 465 470 475 480

5 Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 485 490 495

10 Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510

15 Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525

20 Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
 530 535 540

25 Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560

30 Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575

35 Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590

40 Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605

45 Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620

50 Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640

55 His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 645 650 655

60 Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
 660 665 670

65 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700
 5

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
 705 710 715 720

10

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
 725 730 735

15

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
 740 745 750

20

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
 755 760 765

25

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
 770 775 780

30

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
 785 790 795 800

35

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
 805 810 815

40

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
 820 825 830

45

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
 835 840 845

50

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
 850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
 865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
 885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
 900 905 910

5 Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
 915 920 925

10 Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
 930 935 940

15 Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
 945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975

20 Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
 980 985 990

25 Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
 995 1000 1005

30 Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
 1010 1015 1020

35 Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

40 Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

45 Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
 1070 1075 1080

50 Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
 1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
 1115 1120 1125
 5

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
 1130 1135 1140
 10

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
 1145 1150 1155
 15

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
 1160 1165 1170
 20

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
 1175 1180 1185
 25

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
 1190 1195 1200
 30

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
 1205 1210 1215
 35

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
 1220 1225 1230
 40

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
 1235 1240 1245
 45

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
 1250 1255 1260
 50

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
 1265 1270 1275
 55

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
 1280 1285 1290
 60

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
 1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
 1310 1315 1320

5 Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335

10 Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

15 Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

20 <210> 26
 <211> 20
 <212> DNA
 <213> Artificial Sequence

25 <220>
 <223> gRNA7 for Z chromosome locus chrZ_42174515_-1
 <400> 26
 gtaatacaga gctaaaccag
 20

30 <210> 27
 <211> 20
 <212> DNA
 <213> Artificial Sequence

35 <220>
 <223> gRNA8 for Z chromosome locus chrZ_9157091_1
 <400> 27
 acagacctat gatatgtgag
 40 20

45 <210> 28
 <211> 20
 <212> DNA
 <213> Artificial Sequence

50 <220>
 <223> gRNA9 for Z chromosome locus chrZ_27767602_-1
 <400> 28
 gagcttgtga gtgataatcg
 20

<210> 29
 <211> 20
 <212> DNA
 5 <213> Artificial Sequence

 <220>
 <223> gRNA10 for Z chromosome locus chrZ_78779927_1

 10 <400> 29
 gtaaagagtc agatacacag
 20

 15 <210> 30
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 20 <220>
 <223> gRNA11 for Z chromosome locus chrZ_63364946_-1

 <400> 30
 cagtgggtac tgaagctgtg
 25 20

 <210> 31
 <211> 1012
 30 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Left arm for gRNA7 integration to Chromosome Z
 35 ZNF367_HABP4_ENSGALG00000012629_ENSGALG00000012628_chrZ_42174515_
 -1

 <400> 31
 gtgaacggat aggcagcaag cattcagaac aatgaaaaag atctgcgttc tattgtaaag
 40 60

 caaacaagg tgatgacgga acgtttatth gagatacagg caccatagta caagatatca
 120

 45 aagaggaatt aattcaaccc tccaacaac agttgggaaa aaaaaaggat aagctgagga
 180

 atgcagcctt ctccgatttt tactctggtg cgggaaggag gaggaggagc gggggaaaac
 240

 50 ccagcggact cagactgcat acgcactctc ggaagtacct ctgagggcggg gtgagagga
 300

ggggtggatct aggotccccc cgctccacac actcacgtgc tegtgtactg gttcaggctg
 360
 5 tgacttggca actcctctac atctatccat ctagttagac tcaacaagac catcgtgatt
 420
 caggagtctg aagatgctta actctctcag aacgtctagt tctcctggcc tggtaaattgc
 480
 10 tatgtttctc atactgcctc tctgaagaat gcttcgaatg cttctagtga tgctctaaag
 540
 ttctaaacag aaaaatctgg agacagttct ggtctttaga tagaaaaaat gccaacatgc
 600
 15 caaaggatgg ttacatcctt caagcaacct tgttgcacgc tgtacaatag actcatgtaa
 660
 20 taacttagcc gtagtcatcg tatctcttat ttacctgttc gttattacat tttcctggta
 720
 ctgctttata tttagtcagt tgtcctttta agacaaattt tttgggtgtg gctaataaggc
 780
 25 agttaccaa tgttctagag ggaggaata taatcagtga gtatgtagat gtatatagat
 840
 gtataaccag taagtataca gagaagttag ggtggctctt tgagagcata tagctggaaa
 900
 30 gattatcaca tgggaaaagg taatcagaaa taaatggaaa aatgctcacc agtgtcatag
 960
 35 gctcacaaga aaactcccca aggagcaatc tccagtaaca gacccttct ca
 1012
 <210> 32
 <211> 1023
 40 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Right arm for gRNA7 integration to Chromosome Z
 45 ZNF367_HABP4_ENSGALG00000012629_ENSGALG00000012628_chrZ_42174515_
 -1
 <400> 32
 50 gtcacatgctg tggaggagaa agtgggctgta gaaaaggtaa cggacaagaa atccctcact
 60
 ggcttagctg taaaggctcag aggaaaatca tcttctgttc tccatacaca tctcctgtaa
 120

ttaggcacct gtgcccttta aaaaagtaaa gtaatcaaag agtcatttgg tgaaaataga
 180

5 agccaaacag ttacaaaatt tagttaatat tcaaagacaa aacaactctg gttccaaata
 240

taaaaccctt tgtgtaaaca ctgcggaag gtgccaaaga gccacctaca caaagagatg
 300

10 ccacagagta atttgcgtac cccaacaata ggatcatata tgggcaaacc aaaagaccaa
 360

caaagacccc ttgcaataat cccttgattg gcagaagcaa gtgcagcgtt agatttagta
 420

15 tgtgtgacaa ttctgagaat aggaaggaat ttcactgaaa ctgtgtgaag atttgtatgt
 480

20 taaatacata taatgagggt tatttagctc tgggggtgtgc atgctatgtg gagagatccc
 540

catgtgcccc gcaactgcaat agactaatgt cagcttttta aaccatcatt tggtttgaag
 600

25 agtttattat gattttcaga aacagaatta tgtattcagt gtctgtacat tcttacaagc
 660

ctgctgttct gtacacatga aaaagctact tatggtgcaa atcagtatag agaagcagtt
 720

30 ttgacataca gtggatcatga ctgggtgatc tgctgtcaga aggtaagaca ctggatataca
 780

35 gaattgcaaa cgtgagatga agacctcaag tgcaactctg tcatcatacc acttctcttg
 840

ctattatcta ttcagtgttt gttctaaata ttcaggcaga aggcttggtt tcacatcagg
 900

40 tgttacattt aagtattctt tgcccatttt ttgctgtttg tatgtgtaac ttcagattct
 960

cacattgact gtgtagtgta aatttgcaaa tagatgtaac agcttctctt taacatccca
 1020

45 tgc
 1023

50 <210> 33
 <211> 1014
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Left arm for gRNA8 integration to Chromosome Z
ENSGALG00000025945_AVD_ENSGALG00000025945_ENSGALG00000023622_chrZ
_9157091_1

5

<400> 33

ctctgcatgg tcacaacggc cccaacagcc tccatcccg atgtcaccaa tcagttcact
60

10

attactggtg agcaaccagt ccagcactgc atccccctgg tggggctgtc tgttacgtgg
120

ctcaggaagt tatectcagt gcattccagg tgctcctgga tcgctgtag ctcaccgtgc
180

15

tacttttcca gcaggtgtca gggtggtaga agtccccag caggacgaga gcctgcgatt
240

20

gtgacagctt ctgtagctgg agggagaatg cttcatcaac aggctccgct tgatcagtgt
300

ccctgtactg ataccacaag gcttcttttg ctgctcccat cttaactctc accccatcgt
360

25

gggtgttatac gtgcataaca tcatgcagtg cactaagagt taaagagtta atgctccagt
420

tccgtcgatt gtcaataatt ccaggtgtac ctttctcaga agagaagaac taaaaatccc
480

30

ttaagatctt actgttccat tttcattctc cattcagaag gaaaggtaaa actgcagtgt
540

35

gtgctcccta gcattagagt aatttctccc ttagctcggt gccaccattt tgtttttagg
600

cccacctct taccttttcc ctacaccct cctgtgtctc tgagacgtgg gtccatgggt
660

40

ccctcaacca tagtagacac aaactcagat caaaccaagc cagagctttg acaacccta
720

agctttcagc ttactcaggg ttgctctgca gggactgctc ttcacactct attccttct
780

45

tgatatagag ggcaatgtct cctctctctc ttccttacct gtcctctctc atctgctgt
840

50

tgctgtcagt ggccacactt cagccatggg aatcatccca cctgctttt gtgaaggaaa
900

ctacattatg gttctcaggt agcacagtgg cttccatcac ctctgtttg cttccaagc
960



tgtgtgtggt ggtgtagagg cacttcagct gggcagctga agaactggac gcac
1014

5 <210> 34
<211> 1032
<212> DNA
<213> Artificial Sequence

10 <220>
<223> Right arm for gRNA8 integration to Chromosome Z
ENSGALG00000025945_AVD_ENSGALG00000025945_ENSGALG00000023622_chrZ
_9157091_1

15 <400> 34
cacctttcca cctcctgcta aagcacattc cctacagtat gtttcacatg acaccgtgct
60
ggcagatttt taatatcatc agagaatgac actccatata ctctctgggc agcctttttc
20 120
agtgctctgt cacccacaaa gtaaagtatt tgctcatggt ctaatggaac ttctatggt
180

25 ctagtttgtg accattgaac cttgttctgt cactggacag tgctgaagga gcctggcct
240
atccacttga ctcccacact ttacatattt acaagcatcg ataagacccc ctcagtctt
300

30 ctccagacta gatagcccca ggtctcccaa ccattccttt tatgggagat gcttcaggcc
360
cctcatcacc tctgtggccc tctgctggac tgctgcagc agttccaggt tttttttttt
35 420
gaactgggga gcccagaact tgacacggta ttccagatgt ggcctcacca gcgcagagga
480

40 gaggggaagg ctcacctccc tcgacctggt ggccatggtt tttttaacag acctcaggat
540
accactggcc ttcttgccc aagggcacac tgctggctca tggcccatgg gttggcttcc
600

45 aggactctca ggtccttctc cgtagagctc ctctcaagca ggacaaccct cagcctgtac
660
tgatgtgtgc ggtgcgggta ttctcccca cgtgcaagac tctacatctg ccattggtaa
50 720
atctcataag cttcctctct gcccaactct ccaactgtgta cgggccttag tgaatggcag
780

cacagacttc tgggtgtgagg atagcattca gaacaagctg ctgtaccttc ccagtcacag
 840

5 aggtgagggg gactggcttg tactttccca gctttgcctt cttgcccttt tggaagactg
 900

gagcgacatt ggccttctgg aggaaacggg cctcaggcac cctcctgtt ctccatgacc
 960

10 tttgaaagat gatcgagagc accttggcag tcacttctgc cagctccttc agcaacgcatg
 1020

agtgcacccc gc
 1032

15

<210> 35
 <211> 1025
 <212> DNA
 20 <213> Artificial Sequence

<220>
 <223> Left arm for gRNA9 integration to Chromosome Z
 RIC1_ERMP1_ENSGALG00000015036_ENSGALG00000000438_chrZ_27767602_-1

25

<400> 35
 gcagtacgta cgtttgtgaa cagctgtttt caagggtgat gcacaggaag agtaaaattt
 60

30 cattaaaaaa tctccaagga gcacatggag agctcactca gaagcctcac cgctgccact
 120

gaacaagact gatgcattag tttcatgaaa tcaaggctat gtgtcccact ggtttcatta
 180

35 ttttgttact ctttctttaa aacaaaaaaaa ttgttttatt gcttatatac atcaactatg
 240

ttgttatata ttttccaaag gccgctccaa aagtaatgcc tcctgtttca ttctgttggc
 40 300

ccaccgtgtc agaggcagat ggtggtacgg cagtagagat ggaaccttcc caccaatatt
 360

45 ccattacatg ttgcoatgtg atagacagca gcagatgggc actctgacag aatggtatct
 420

gacatggaag cgtggatgga gcaaagttgt gtcactgaat tcctccataa ggaaatgagc
 480

50 acccactaac attcactgat gcttgctgaa catttctgga gaccaaacag tggatgtgag
 540

cacagtgagg tgggtgggtgg tgcatttcag cagtgggtgac aatgacgttg ggtcacctct
600

5 gctgggtgcag gttttgatga gtgcagcctg tgaggtcttg ttcacatg gagaaaatgc
660

acaactaatg gtgggtgactg ctgaaaaata gtgttttgta gctgagaatt ttctctgtca
720

10 aacagtgcta attgtgctct ttgtatctgt tgtagtttcc atgggaataa ataggaggca
780

ttacttgca actgacctac atatatgtgg ccccaggcaa gtcaacagct gtacgaaaag
840

15 tccatgcttc ctcatataa acggaggaaa aaaagttggt tacagctgta atgggattat
900

20 aaaaggcaaa tggggatagt acagtgggtga gaaccagatc tatgaaggaa aagccctaag
960

aaaaagagag tgcacagata ctcttaacca tctaataact gtttctcca tcttacagct
1020

25 cagag
1025

<210> 36
30 <211> 1040
<212> DNA
<213> Artificial Sequence

<220>
35 <223> Right arm for gRNA9 integration to Chromosome Z
RIC1_ERMP1_ENSGALG00000015036_ENSGALG00000000438_chrZ_27767602_-1

<400> 36
40 tggcttccaa ttctctctgt ttatcaccca cggtgcatcc attgatgtag aggcacttca
60

gtcgggctat cagcaatggt accacctgtg aggagcctc ttgagatctt ttaggacaat
120

45 ttagaggttt tccccactgt ttctaaaat tacagattc cctgtccctt cttagtata
180

tagaactcca tccccttcca ccatcaaacc tagcttaagc tctggtaaat gatccagcca
240

50 gtctgcttcc caaaactctt gctcagttgc attctatctg atgcctacat gcccaacacc
300

ttcaaggcct gtccaagatc atagaacaca aaggcttgat catgacaaca tccatgcagc
 360
 caatcattca catgacccat tcctctcctc ctttccaggt cccaacggcc aactgagagt
 5 420
 tgtgctctga gctcttcagt gtccttccga gggatgtaaa gtccctttta atgttctgca
 480
 ttttctttgt cacagccttc tgagacccaa cttgaatgtg gaggagtgga gagtaatcat
 10 540
 cccgctctga ttatcagata cattagcctc ttcttggcac ctctggcaga cagaagactt
 600
 15 ccttggagag attacctgga tgataactgg gggcctcagt gcctcacagc agcaagtccc
 660
 cagttactaa gactctacac ctctttgggtg gcgctagttc tgatgtagtt cttcagttgc
 20 720
 ttaacatgct tgttgtttcc caagatttga ctgctatctg caccctccct ttcttccagc
 780
 25 cccagagcct tatatcatga gctgagaaca ttctgtgttc aaaatacaag acacatttat
 840
 gcaatttctg attctctacc catatctttg tgatagaaat tagggctact tatttcccat
 900
 30 agtgatttcc atgctgccta ccaggttcta agtcagggtca ccttgccagc ttcttccaaa
 960
 ccatgctcag gagtgcatca aatcacacgc agaccaagg aaaaaatgta ccaacaata
 35 1020
 ttaccacgca tgcaaactcc
 1040
 40
 <210> 37
 <211> 1024
 <212> DNA
 <213> Artificial Sequence
 45
 <220>
 <223> Left arm for gRNA10 integration to Chromosome Z
 TRIM36_PGGT1B_ENSGALG00000008188_ENSGALG00000008197_chrZ_78779927
 _1Chromosome Z
 50
 <400> 37
 ctaccttgcc ccaactccagc agctccacga ggtactcgaa gtgcccatgt ccggccatcc
 60

ccagcatggc acagctgtca cctgggagca gcgggggtcc ggcaggcata gccggcaccc
 120

5 acagcctggg ctgggctgag gatgggcttg gtgggcagct gggtgctgcc gtggcagctg
 180

agtgccgtcg tgcaggtggg ctcggtgctc tctgcagccc acggctcggg acccgtggtg
 240

10 ggcatactgc tgggcaaggg gacggctcgg atgtgcatcc cagagcaggg tgcatgatgt
 300

gtgctgcagcg gcacacagca gcgcacttgt ccccgaggca ccaccgggtc tgatgaaatt
 360

15 ccaggttttg ttttaagatga aatgcatgt gggcatacat ttgcacatga gaacacctgc
 420

20 taatgaaaca cacagagttg tatgtgcagt ctccctggct ggtggcatgt ctgtgggggg
 480

agttaggcag gaagaaaagc tccaccatct gtttctgtgt cccgggcttg gtgctgccac
 540

25 tagagatggg atgcacacag catgcctgcg tgctgcccag gcagctcagt acctgctgct
 600

gggggagagc agccctcagg gagctgctgc aaagaccccg tgggatggct ggtgaacagt
 660

30 gctggagaca tggcagatgt cgcacccctg ctccacacca ccctccgctg cagctgcctg
 720

35 cacgatgcca tttcagcgtg ccacggcttg tttcttctc ttgcagtctc aggttgccgc
 780

agcagcagct gcctgaggtc tgggtggcaaa gggcagagca cccagcccac tgctctccat
 840

40 gagtgagcgg gtagggggca cgccgtgctt cgcttctggt atcgggtggc ctgtggagca
 900

catcaccctt ggacagtga aacatgtcaa aggggtgtttt ataacagtat agcatatccc
 960

45 tttgaggctg aatttcctga gcttatggat atggtagtaa tgctatacca ggctctctct
 1020

50 gcag
 1024

<210> 38
 <211> 1022

<212> DNA

<213> Artificial Sequence

<220>

5 <223> Right arm for gRNA10 integration to Chromosome Z
TRIM36_PGGT1B_ENSGALG00000008188_ENSGALG00000008197_chrZ_78779927
_1Chromosome Z

<400> 38

10 cagctgctgt acaatggctg tatggcagta tcgatgtgaa cttcctgata ataaacagaa
60

ctcagcagca aataaaacca gattgttctg atcagtaacg agaggctgtg caaaaagctg
120

15 agaaatgtac agcccttctt gtcaccact acagcccttc atgcagcggc ctgctgcact
180

20 agggcctgct tggccagagg cagggctgca gcttgagata cgcagcacc agtaccacag
240

caggctgcca tgtgcttctg tcaggccatg aattgcacaa gttggtatag ttttatacag
300

25 tgtcatgcag aagcacaact cttaccttgg cacactgata tggaaagcaa agtgtaaagg
360

aactgtatgt ttcttggctt gtgagttcca ttggttctcc caggacagca gtgacagctc
420

30 tagatthttct ggttgatctc ttccatctaa gthtttactt ctgaatacat ctgtgcagca
480

35 aagcctgatg cthttthttt tccccctccc ctgggatcgt gaaatattat tttcccatct
540

ccataataca gatgatccac tcaaacatgc atcctcagtc gcagtcctca gaagggcctt
600

40 tthccaccat acacctacac gctacaccag actthtattgg tcaccatgcc ttgagggatt
660

ctgcctaaca caggcacaca ggaagggctg ataccagtaa ttcattgctg aggagagaag
720

45 gggcacttct tttatagtag agccacagaa tcaccaggat tggaaaacac ctccaagatc
780

50 atccaatcca accaccacc taccaccagt atthccact aaaccatgcc ctttgggtacc
840

acatctaaaa gcatttctg caccttcaca gcttcattgc cthttctcagg acgtgctcca
900

gggccttggt gtctttcttg tattgagtgg cccaaaattg aacacaatac tcgaggtgca
 960

5 gcttcaccaa agctgagtgc agaggggtga tcacctctac taggccttgt tttaccctt
 1020

cc
 1022

10 <210> 39
 <211> 1003
 <212> DNA
 <213> Artificial Sequence

15 <220>
 <223> Left arm for gRNA11 integration to Chromosome Z
 RPS23_ATG10_ENSGALG00000015617_ENSGALG00000027714_chrZ_63364946_-
 1

20 <400> 39
 gggttgaaac tgcgtgatca ttgtagtctt tttcaaccca ggccattcta tgattccatg
 60

25 aagactctgt aaggacccca atgccatcag ctgtccccac agcttcagta cccactgcca
 120

ccacatTTTT actgagatag gtagtttggg agccttcctc ccatcttggg taggggaggg
 180

30 ttaactgctg gcatttttgt tctttctgc aatatgtggt ttctgcatgc attctctttt
 240

35 caccctaaaat tttaatgtgg acggactttg aggatattct gcatctgccc aatatttctt
 300

acagcctcac atctactgat ttatgcacag tatcttatat ataatgtata ttatattatg
 360

40 tatattatgc actgaatcac atctactgtg tatgcacagt aacttctcct atgggtactgt
 420

aagcccagaa atcccagata tgcactaca acgtgtctgc tatgtattgc ttctgtgaga
 480

45 cacagatgtg ataatcagag gctgtacaca gtagaaatgc atacatatcc tactgtgaaa
 540

50 tctctgaaat cttagccttaa tcttgaaca gaaatgaata tggtgacatc ttgatctgat
 600

agaattgggt gccagtagca cagctgtaat catccatgat atgaatcaaa ccaagcacag
 660

gtaaacaggt gagagaaatc atgaacaatt acatgcaaac ggaccagcta aaatgtgttt
 720

5 gtttgtgtgtt tttttttttt caggtgattc ttgattacag taagatcaga agctgctaca
 780

ttagcagacc agccactgca ctcaaggctg tgattcacag cttgcagacc tgacagcagt
 840

10 tctgtggaag aggcaggtcc ctgtcaacca gttaatacaa taaatcagtc tctgttacac
 900

aaataatgtt atcctgcacc actgctggcg ttacactatt tcaccaagt ttatcaccag
 960

15 caaactgagt cttatcgttc ctactgtgct ttgcttttct tgc
 1003

20 <210> 40
 <211> 1003
 <212> DNA
 <213> Artificial Sequence

25 <220>
 <223> Right arm for gRNA11 integration to Chromosome Z
 RPS23_ATG10_ENSGALG00000015617_ENSGALG00000027714_chrZ_63364946_-
 1

30 <400> 40
 cagaaacacc agagcttacg gtgtgaattc attgcacatc ttattagatc aacaaaatgg
 60

35 aataaaaagaa tacagaagat tactactcca ttgggcatgt ggacgtttta caggcctgga
 120

taaattagat cttaaaaaca aacaacaaa caaacaaaa aactccttgc aaagataatg
 180

40 ttatgtaata ttagttgcaa gaagtaagca aacacacaga actgggagca gaagcaaagc
 240

actaagttat taaagcaagt tgcacatttt gagttgcatt ttgccactgg tttataaac
 300

45 atgttttagca tgtctggcga gaatttgggc accaggatgc ttttaagatg tctgtctatg
 360

50 gaacctgtca gtgctcaaga ataacttctg ttatttggat gctgcaccaa agaattcaga
 420

ggaagacgag ccaagccaga cgttatcata gtcactagta aagtggttct aagcctaatt
 480

aagacatgtc agaactatgt gttgtgcaac caaatcctcc aaaagagaaa tcagaggtga
 540

5 acttgtgcaa taaatataga agacacgtaa atcctgaggc agttagctaa ccatatgaag
 600

ccaatcatac ctgactgctg gacgcaggag actgaacctt acagaccctg gagaatcact
 660

10 gtttggccta gttaggcctg aatggaatth accaagattc atgactttaa catggatcag
 720

gtgcaaagaa aagaaggctc agttagttcc tacaggctac cagatctttt tccacctctg
 780

15 ctcagcctgg agctgtgggc tctacctgcc cctgaagtga agtggcatca actgcaacac
 840

20 tttttgcaga ggcaaaaatc atcaagtcgt gcctctgcat ttttaagggtga tggattccaa
 900

gagatagatt ccaaacccaa cactgagatt cctctgggtga tgcagatgac tgcttactgg
 960

25 gatccttctc tctctatagc ctaatcccat gcagcaccaa aag
 1003

30 <210> 41
 <211> 200
 <212> DNA
 <213> Artificial Sequence

35 <220>
 <223> For cloning using the gRNA3, Z chromosome, "Left arm"

<400> 41
 cctgcaggac atgtggacac tctgccaggt actgggggag gatccccctcc actgctcct
 60

40 gggagatggc aaacatctag agaggacac tgcagggatg ctgcatgaac acctccagcc
 120

45 ttctgcaggc tgtgatgggg tcagggctga aggtcaacac aagcaatgac tgcttgctga
 180

caatgtgggc attgccgacc
 200

50 <210> 42
 <211> 258
 <212> DNA
 <213> Artificial Sequence

<220>

<223> For cloning using the gRNA3, Z chromosome, "Right arm"

5 <400> 42
 tggtagtgagg gccaggaga gcaggacagg agtgtggact cggaggcgcg ggcagaggag
 60

10 gtagctcaga gcatgagata tattcgccag gtgtcagtgg acttctggca gtgcatactg
 120

cagagagctt tggttgctgg ccatctggat gatgacaaaa tctcagccac agaccatgtg
 180

15 ggttggaaga gtcctcagga gatcaccttt ccacctcctg ctaaagcaca ttccttacag
 240

tatgtttcac atgacacc
 258

20

<210> 43

<211> 436

<212> DNA

25 <213> Artificial Sequence

<220>

<223> For cloning using the gRNA4, Z chromosome, "Left arm"

30 <400> 43
 tcaacagctg tacgaaaagt ccatgcttcc tcattataaa cggaggaaaa aaagttgttt
 60

35 acagctgtaa tgggattata aaaggcaaat ggggatagta cagtgggtgag aaccagatct
 120

atgaaggaaa agccctaaga aaaagagagt gcacagatac tcttaacat ctaataactg
 180

40 tttcctccat cctacagctc agagttaaga cttacagagg actctagtac ttagtaagat
 240

gaatacgagc taatagtggc aaaaataatc ccagtgcctc aacctgacc tgggaaaaag
 300

45 gggcatgtat agaccttctg atattgtgat gctgtgtttg tacacttatt atctacattt
 360

50 tcagaaatta ggttaaactt cagaaaactg aagatctcca gggcttgtag cagaccctga
 420

ccaccagact ggtccc
 436

<210> 44
<211> 521
<212> DNA
5 <213> Artificial Sequence

<220>
<223> For cloning using the gRNA4, Z chromosome, "Right arm"

10 <400> 44
tcaaaccgat tgtggctggt tctcagaggg aactctttgc gctctagccc ctctataaca
60

15 tggggcaact tcccctgccc caccttcctt tctgtatct tctgaaaagc ttgtagcct
120

caattgtcac gctccagcca tatgaatcat cccaccacgt cctgtgttc acaattacct
180

20 tctagtcttc taattgcacc atggcttcca attcctcctg tttatcacc acgttgcatc
240

cattgatgta gaggcacttc agtcgggcta tcagcaatgt taccacctgt gaggagcct
300

25 cttgagatct ttaggacaa ttagagggtt ttcccactg tttcctaaaa ttacagcatt
360

30 cctgtcctt tcttagtgat atagaactcc atccccttc accatcaaac ctagcttaag
420

ctctggtaaa tgatccagcc agtctgcttc caaaactct tgctcagttg cattctatct
480

35 gatgcctaca tgcccaacac cttcaaggcc tgtccaagat c
521

<210> 45
40 <211> 505
<212> DNA
<213> Artificial Sequence

<220>
45 <223> For cloning using the gRNA5, Z chromosome, "Left arm"

<400> 45
gcctgtccaa actgaccttg gggccccggg cacagctgct cccgagcaag gagaggcaga
60

50 gaattgtgga ggaaaccctc ttccacctag ctccaagggt gcacgctcaa tgcttgatg
120

aagattcaag ttctcaatga agaactgcat tcaaagacat ctttgaggt gaccagact
 180

5 gcctctggga tgaatccaga acaaacagaa tcttcccagc aacaccttg tttatcatac
 240

acgttcagaa atgcagtgcc tgcttggtat ttttttaaac tcatcacaga ggaatctctg
 300

10 agtggagcag aggagtgttt cctcttgctt ttttcttccc cctttctgta agaatgcat
 360

atgccagttt cccctaaat gttttcaaac tccaaattgt ttctgctgg tgacattctc
 420

15 ctctccagca ctgcctacac ccagctgcg tctgatagga aaagcacca gtaccatgct
 480

20 ggcatggcat ggtcagtgc accca
 505

<210> 46
 <211> 599
 25 <212> DNA
 <213> Artificial Sequence

<220>
 <223> For cloning using the gRNA5, Z chromosome, "Right arm"
 30

<400> 46
 ctgccagagc agagggctgc tgaaacgccg agtgccaact gacactgttc agctgacagc
 60

35 ctcaagagag tgctgagagg gttaagtgag gcaacaaaat acaagtactc aaaaatagaa
 120

tggaatggaa tggaataaca gagggaatgg ggaacaggct gcagtgagaa ggaaaccccc
 180

40 gtgcaccgac caaccgctc ccagtagcct ggtccctacc ttgccccact ccagcagctc
 240

caagaggtac tcgaagtgcc catgtccggc catccccagc atggcacagc tgtcacctgg
 300

45 gagcagcggg ggtccggcag gcatcgccgg caccacagc ctgggctggg ctgaggatgg
 360

50 gcttgggtggg cagctgggtg ctgccgtggc agctgagtgc cgtcgtgcag gtgggctcgg
 420

tgctctctgc agcccacggc tcgggaccgg tgggtgggcat actgctgggc aaggggacgg
 480

ctcggatgtg catcccagag caggggtgcat gatgtgtgcg cagcggcaca cagcagcgca
540

5 cttgtccccg aggcaccacc cggctctgatg aaattccagg ttttgtttaa gatgaaaat
599

<210> 47
10 <211> 453
<212> DNA
<213> Artificial Sequence

<220>
15 <223> For cloning using the gRNA6, Z chromosome, "Left arm"

<400> 47
ttgcagtctc aggttgccgc agcagcagct gcctgaggtc tggaggcaaa gggcagagca
60

20 cccagcccac tgctctccat gagtgagcgg gtagggggca cgccgtgctt cgcttctggt
120

25 atcgggtggc ctgtggagca catcaccctt ggacagtga aacatgtcaa aggggtgttt
180

ataacagtat agcatatccc tttgaggctg aatttctga gcttatggat atggtagtaa
240

30 tgctatacca ggctctctct gcaggttgct cactgaaaca atataccctt tctttctcaa
300

gaaatgggac tcatgaatag ctaccaggcc tttctgctac tagatactgt agaattacat
360

35 atatcagatg gctcattaat cacatagctt gttattggga cagaggcagg ttttcagcat
420

ttccacact gcttttctgc aaatggatta gga
40 453

<210> 48
45 <211> 439
<212> DNA
<213> Artificial Sequence

<220>
50 <223> For cloning using the gRNA6, Z chromosome, "Right arm"

<400> 48
gggaccatca ctcaacttct aggccatgct tcgtagttta actccatgca aggtattttc
60

ttgctctgat ctaactctag atcactaaat ggcacttgtg caggacactt ttccactgtc
 120

5 attttggggg gtagaagtgt ggtcagagcc agggaggctt gcagggcca caccagctgt
 180

ggccaagca gctgctgtac aatggctgta tggcagtatc gatgtgaact tcctgataat
 240

10 aaacagaact cagcagcaaa taaaccaga ttgttctgat cagtaacgag aggctgtgca
 300

aaaagctgag aaatgtacag cccttctgc tcaccactac agcccttcat gcagcggcct
 360

15 gctgcactag ggctgcttg gccagaggca gggctgcagc ttgagatagc cagcaccag
 420

20 taccagca ggctgcat
 439

25 <210> 49
 <211> 22
 <212> DNA
 <213> Artificial Sequence

30 <220>
 <223> LEFT PRIMER

<400> 49
 ttggtgtgtg ctaataggca gt
 22

35 <210> 50
 <211> 20
 <212> DNA
 <213> Artificial Sequence

40 <220>
 <223> RIGHT PRIMER

45 <400> 50
 tagtcctcgt tgtgggaggt
 20

50 <210> 51
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> LEFT PRIMER
 <400> 51
 gcattgatct gtccagttgc
 5 20

<210> 52
 <211> 20
 10 <212> DNA
 <213> Artificial Sequence

<220>
 <223> RIGHT PRIMER
 15 <400> 52
 tactgcaaaa accgcatcac
 20

<210> 53
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 25

<220>
 <223> ChZgRNA_FORWARD PRIMER
 30 <400> 53
 tgtatgagac cacgtaatac agagctaaac cag
 33

<210> 54
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 35

<220>
 <223> ChZgRNA_Reverse PRIMER
 40 <400> 54
 aaacctgggt tagctctgta ttacgtggtc tca
 45 33

<210> 55
 <211> 51
 <212> DNA
 50 <213> Artificial Sequence

<220>
 <223> Primer Forward for cloning of the 'Left arm' of gRNA 7

<400> 55
acggggtcat tagttcatag cccatatatg gtgaacggat aggcagcaag c
51

5
<210> 56
<211> 52
<212> DNA
<213> Artificial Sequence

10
<220>
<223> Primer Reverse for cloning of the 'Left arm' of gRNA 7

<400> 56
15 ggcgggccat ttaccgtaag ttatgtaacg tgaggaagg tctgttactg ga
52

20
<210> 57
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
25 <223> Primer Forward for cloning of the 'Right arm' of gRNA 7

<400> 57
ttctatcgcc ttcttgacga gttcttctga gtcattgctgg tggaggagaa a
51

30
<210> 58
<211> 53
<212> DNA
35 <213> Artificial Sequence

<220>
<223> Primer Reverse for cloning of the 'Right arm' of gRNA 7

40 <400> 58
gtgatggcag gttgggcgct gcttggtcgg gcatgggatg ttaaagagaa gct
53

45 <210> 59
<211> 16
<212> DNA
<213> Artificial Sequence

50 <220>
<223> Repetition of 16 nucleotides

<400> 59

cggtaaattgg cccgcc
 16

5 <210> 60
 <211> 261
 <212> DNA
 <213> Artificial Sequence

10 <220>
 <223> Flanking left arm of Z chromosome

<400> 60
 tggagatccc ttattggtga agctttcccc tgcaaaccta gaatcttctg atgcaagtga
 15 60

gcaatcttct tcccttcctt agcagctctc tgttgcacca gtccttccat ccttgcacct
 120

20 ttgttgtaat gcttttccca tggcattgat ctgtccagtt gctacagtgt agaaacaaga
 180

aagtgttttc tcaaaacatt atacatgctt actatcagat cttacttaac attgataaga
 240

25 agtacacctt tcattacaac a
 261

30 <210> 61
 <211> 245
 <212> DNA
 <213> Artificial Sequence

35 <220>
 <223> Flanking right arm of Z chromosome

<400> 61
 caaactttta atcactttca atcaatgttg gaaatctgca gtacataaat tttattttga
 40 60

ttcattatlc agagaatcag agaaataagg acttttttag aggtttatgt ttattttaga
 120

45 aaatataaat aaatcttgta tacatagatt tccaactagt aaaattttta gctatcgctg
 180

tagcatttct atgaggagtt ttacttttta ctgtctctta tgaaacttag tggaacaaaa
 240

50 gctag
 245

<210> 62
 <211> 1014
 <212> DNA
 <213> Artificial Sequence
 5
 <220>
 <223> Left arm of Z chromosome

 <400> 62
 10 gtgaacggat aggcagcaag cattcagaac aatgaaaaag atctgcgttc tattgtaaag
 60

 caaaciaaagg tgatgacgga acgtttatth gagatacagg caccatagta caagatatca
 120
 15 aagaggaatt aattcaaccc tccaacaac agttgggaaa aaaaaaggat aagctgagga
 180

 atgcagcctt ctccgatttt tactctggtg ccggaaggag gaggaggagc gggggaaaaac
 20 240

 ccagcggact cagactgcat acgcactctc ggaagtacct ctgaggcggg gtgagaggaa
 300
 25 ggggtggatct aggcctcccc cgctccacac actcacgtgc tcgtgtactg gttcaggctg
 360

 tgacttgga actcctctac atctatccat ctagttagac tcaacaagac catcgtgatt
 420
 30 caggagtctg aagatgctta actctctcag aacgtctagt tctcctggcc tggtaaattg
 480

 tatgtttctc ataactgcctc tctgaagaat gcttcgaatg cttctagtga tgctctaaag
 35 540

 ttctaaacag aaaaatctgg agacagttct ggtctttaga tagaaaaaat gccaacatgc
 600
 40 caaaggatgg ttacatcctt caagcaacct tgttgcatgc tgtacaatag actcatgtaa
 660

 taacttagcc gtagtcatcg tatctcttat ttacctgttc gttattacat tttcctggta
 720
 45 ctgctttata tttagtcagt tgcctttta agacaaattt tttgggtgtg gctaattaggc
 780

 agttacaaa tgttctagag ggaggaata taatcagtga gtatgtagat gtatatagat
 50 840

 gtataaccag taagtataca gagaagttag ggtgggtcttt tgagagcata tagctggaaa
 900

gattatcaca tgggaaaagg taatcagaaa taaatggaaa aatgctcacc agtgtcatag
 960

5 gctcacaaga aaactcccca aggagcaatc tccagtaaca gacccttctt cagt
 1014

<210> 63
 <211> 629
 10 <212> DNA
 <213> Artificial Sequence

<220>
 <223> CMV Enhancer

15 <400> 63
 tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc
 60

20 cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc
 120

attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt
 180

25 atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt
 240

atgccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca
 30 300

tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg
 360

35 actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc
 420

aaaatcaacg ggactttcca aaatgtcgt acaactccgc cccattgacg caaatgggocg
 480

40 gtaggcgtgt acggtgggag gtctatataa gcagagctgg tttagtgaac cgtcagatcc
 540

gctagcgcta ccggactcag atctcgagct caagcttcga attctgcagt cgacgggtacc
 45 600

gcgggcccgg gatccaccgg tcgccacca
 629

50 <210> 64
 <211> 714
 <212> DNA
 <213> Artificial Sequence

<220>

<223> dsRED2 reporter gene

5 <400> 64
 tggcctcctc cgagaacgtc atcaccgagt tcatgcgctt caaggtgcgc atggagggca
 60

10 ccgtgaacgg ccacgagttc gagatcgagg gcgagggcga gggccgcccc tacgagggcc
 120

acaacaccgt gaagctgaag gtgaccaagg gcggccccct gcccttcgcc tgggacatcc
 180

15 tgtcccccca gttccagtac ggctccaagg tgtacgtgaa gcaccccgcc gacatccccg
 240

actacaagaa gctgtccttc cccgagggct tcaagtggga gcgctgatg aacttcgagg
 300

20 acggcggcgt ggcgaccgtg acccaggact cctccctgca ggacggctgc ttcatttaca
 360

25 aggtgaagtt catcggcgtg aacttcccct ccgacggccc cgtgatgcag aagaagacca
 420

tgggctggga ggcctccacc gagcgcctgt acccccgcga cggcgtgctg aagggcgaga
 480

30 cccacaaggc cctgaagctg aaggacggcg gccactacct ggtggagttc aagtccatct
 540

acatggccaa gaagcccgtg cagctgcccg gctactacta cgtggacgcc aagctggaca
 600

35 tcacctcca caacgaggac tacaccatcg tggagcagta cgagcgcacc gagggccgcc
 660

40 accacctgtt cctgtagcgg ccgcgactct agatcataat cagccatacc acat
 714

45 <210> 65
 <211> 1026
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Right arm of Z chromosome

50 <400> 65
 gagtcatgtc ggtggaggag aaagtggctg tagaaaaggt aacggacaag aaatccctca
 60

ctggcttagc tgtaaaggtc agaggaaaat catcttctgt tctccataca catctcctgt
 120

5 aattaggcac ctgtgccctt taaaaaagta aagtaatcaa agagtcattt ggtgaaaata
 180

gaagccaaac agttacaaaa tttagttaat attcaaagac aaaacaactc tggttccaaa
 240

10 tataaaacc tttgtgtaaa cactgcggga aggtgccaaa gagccaccta cacaaagaga
 300

tgccacagag taatttgctg accccaacaa taggatcata tatgggcaaa ccaaagacc
 360

15 aacaagacc ccttgcaata atcccttgat tggcagaagc aagtgcagcg ttagatttag
 420

20 tatgtgtgac aattctgaga ataggaagga atttactga aactgtgtga agatttgtat
 480

gttaaataca tataatgagg gttatttagc tctgggggtg gcatgctatg tggagagatc
 540

25 cccatgtgcc cagcactgca atagactaat gtcagctttt taaaccatca tttggtttga
 600

agagtttatt atgattttca gaaacagaat tatgtattca gtgtctgtac attcttacia
 660

30 gcctgctggt ctgtacacat gaaaaagcta cttatgggtg aatcagtat agagaagcag
 720

35 ttttgacata cagtggatc gactgggtga tctgctgtca gaaggtaaga cactgggata
 780

cagaattgca aacgtgagat gaagacctca agtgcaactc tgtcatcata ccacttctct
 840

40 tgctattatc tattcagtgt ttgttctaaa tattcaggca gaaggcttgg tttcacatca
 900

ggtgttacat ttaagtattc tttgccatt ttttgotggt tgtatgtgta acttcagatt
 960

45 ctcacattga ctgtgtagtg taaatttga aatagatgta acagcttctc ttaacatcc
 1020

50 catgca
 1026

<210> 66
 <211> 2618

<212> DNA

<213> Artificial Sequence

<220>

5 <223> Right arm of Z chromosome of Figure 15

<400> 66

tggagatccc ttattggtga agctttcccc tgcaaaccta gaatcttctg atgcaagtga
60

10

gcaatcttct tcccttcctc agcagctctc tgttgcaacca gtccttccat ccttgcacct
120

15

ttgttgtaat gcttttccca tggcattgat ctgtccagtt gctacagtgt agaaacaaga
180

aagtgttttc tcaaaacatt atacatgctt actatcagat cttacttaac attgataaga
240

20

agtacacctt tcattacaac agtgaacgga taggcagcaa gcattcagaa caatgaaaaa
300

gatctgcggt ctattgtaaa gcaaacaaag gtgatgacgg aacgtttatt tgagatacag
360

25

gcaccatagt acaagatata aaagaggaat taattcaacc ctccaacaa cagttgggaa
420

30

aaaaaaagga taagctgagg aatgcagcct tctccgattt ttactctggt gccggaagga
480

ggaggaggag cgggggaaaa cccagcggac tcagactgca tacgactct cggaagtacc
540

35

tctgaggcgg ggtgagagga aggggtggatc taggctcccc ccgctccaca cactcacgtg
600

ctcgtgtact ggttcaggct gtgacttggc aactcctcta catctatcca tctagttaga
660

40

ctcaacaaga ccatcgtgat tcaggagtct gaagatgctt aactctctca gaacgtctag
720

45

ttctcctggc ctggtaaata ctatgtttct catactgcct ctctgaagaa tgcttogaat
780

gcttctagtg atgctctaaa gttctaaaca gaaaaatctg gagacagttc tggctcttag
840

50

atagaaaaaa tgccaacatg ccaaggatg gttacatcct tcaagcaacc ttggtgcatg
900

ctgtacaata gactcatgta ataacttagc cgtagtcatc gtatctctta tttacctggt
960

cgttattaca ttttcctggt actgctttat atttagtcag ttgtcctttt aagacaaatt
1020

5 ttttgggtg tgctaatagg cagttaccaa atgttctaga gggaggggaat ataatcagtg
1080

agtatgtaga tgtatataga tgtataacca gtaagtatac agagaagtta ggggtggtcct
1140

10 ttgagagcat atagctggaa agattatcac atgggaaaag gtaatcagaa ataaatggaa
1200

aaatgctcac cagtgtcata ggctcacaag aaaactcccc aaggagcaat ctccagtaac
1260

15 agacccttcc tcagttggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct
1320

20 gaccgcccaa cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc
1380

caatagggac tttccattga cgtcaatggg tggagtattt acggtaaact gcccaacttg
1440

25 cagtacatca agtgtatcat atgccaaagta cccccctat tgacgtcaat gacggtaaat
1500

ggccccgctg gcattatgcc cagtacatga ccttatggga ctttctact tggcagtaca
1560

30 tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc
1620

35 gtggatagcg gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga
1680

gtttgttttg gcacaaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat
1740

40 tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctggtttag
1800

tgaaccgtca gatccgctag cgctaccgga ctcagatctc gagctcaagc ttcgaattct
1860

45 gcagtcgacg gtaccgcggg cccgggatcc accggtcgcc accatggcct cctccgagaa
1920

50 cgtcatcacc gagttcatgc gcttcaaggt gcgcatggag ggcaccgtga acggccacga
1980

gttcgagatc gagggcgagg gcgagggccg ccctacgag ggccacaaca ccgtgaagct
2040

gaaggtgacc aagggcggcc ccctgccctt cgcttgggac atcctgtccc cccagttcca
 2100

5 gtacggctcc aaggtgtacg tgaagcaccc cgccgacatc cccgactaca agaagctgtc
 2160

cttccccgag ggcttcaagt gggagcgcgt gatgaacttc gaggacggcg gcggtggcgac
 2220

10 cgtgaccag gactcctccc tgcaggacgg ctgcttcac tacaaggtga agttcatcgg
 2280

15 cgtgaacttc ccctccgacg gccccgtgat gcagaagaag accatgggct gggaggcctc
 2340

caccgagcgc ctgtaccccc ggcagcggcgt gctgaagggc gagaccaca aggccctgaa
 2400

20 gctgaaggac ggcggccact acctggtgga gttcaagtcc atctacatgg ccaagaagcc
 2460

cgtgcagctg cccggctact actacgtgga cgccaagctg gacatcacct cccacaacga
 2520

25 ggactacacc atcgtggagc agtacgagcg caccgagggc cgccaccacc tgttcctgta
 2580

30 gcggccgcga ctctagatca taatcagcca taccacat
 2618

<210> 67
 <211> 1271
 35 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Left arm of Z chromosome of Figure 15

40 <400> 67
 gagtcatgtc ggtggaggag aaagtggctg tagaaaaggt aacggacaag aaatccctca
 60

45 ctggcttagc tgtaaaggtc agaggaaaat catcttctgt tctccataca catctcctgt
 120

aattaggcac ctgtgccctt taaaaaagta aagtaatcaa agagtcattt ggtgaaaata
 180

50 gaagccaaac agttacaaaa tttagttaat attcaaagac aaaacaactc tggttccaaa
 240

tataaaaccc tttgtgtaaa cactgcgggga aggtgccaaa gagccaccta cacaaagaga
300

5 tgccacagag taatttgcgt accccaacaa taggatcata tatgggcaaa ccaaagacc
360

aacaaagacc ccttgcaata atcccttgat tggcagaagc aagtgcagcg ttagatttag
420

10 tatgtgtgac aattctgaga ataggaagga atttcactga aactgtgtga agatttgtat
480

gttaaataca tataatgagg gttattttagc tctgggggtgt gcatgctatg tggagagatc
540

15 cccatgtgcc cagcactgca atagactaat gtcagctttt taaaccatca tttggtttga
600

20 agagtttatt atgattttca gaaacagaat tatgtattca gtgtctgtac attcttacia
660

gcctgctggt ctgtacacat gaaaaagcta cttatgggtgc aaatcagtat agagaagcag
720

25 ttttgacata cagtgggtcat gactgggtga tctgctgtca gaaggtaaga cactggtata
780

cagaattgca aacgtgagat gaagacctca agtgcaactc tgatcatcata ccacttctct
840

30 tgctattatc tattcagtgt ttgttctaaa tattcaggca gaaggcttgg tttcacatca
900

35 ggtgttacat ttaagtattc tttgcccatt ttttgcgtgt tgtatgtgta acttcagatt
960

ctcacattga ctgtgtagtg taaatttgca aatagatgta acagcttctc tttaacatcc
1020

40 catgcacaaa cttttaatca ctttcaatca atgttggaat tctgcagtac ataaatttta
1080

ttttgattca ttattcagag aatcagagaa ataaggactt ttttagaggt ttatgtttat
1140

45 tttagaaaat ataaataaat cttgtatata tagatttcca actagtaaaa ttttagctta
1200

tcgctgtagc atttctatga ggagttttac tttttactgt ctcttatgaa acttagtgga
50 1260

aacaaagcta g
1271

ABSTRACT

The present invention relates to methods of fertilization and gender determination and identification in avian subjects. More specifically, the invention provides non-invasive methods using transgenic avian animals that comprise at least one reporter gene, specifically, RFP, integrated into at least one gender chromosome Z or W. The
5 transgenic avian animals of the invention are used for gender determination and selection of embryos in unhatched avian eggs.