

(51) International Patent Classification:
C12N 15/06 (2006.01)(21) International Application Number:
PCT/KR2008/006591(22) International Filing Date:
7 November 2008 (07.11.2008)

(25) Filing Language: Korean

(26) Publication Language: English

(30) Priority Data:
10-2008-0098659 8 October 2008 (08.10.2008) KR(71) Applicant (for all designated States except US): **INDUS-
TRY FOUNDATION OF CHONNAM NATIONAL
UNIVERSITY** [KR/KR]; 300, Yongbong-dong, Buk-gu,
Gwangju 500-757 (KR).

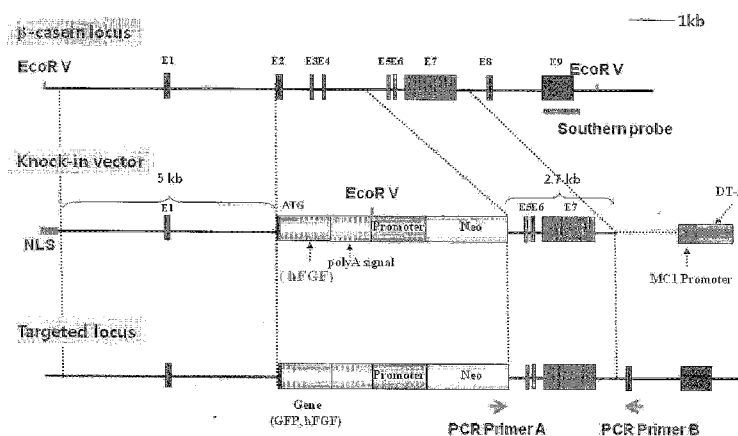
(72) Inventors; and

(75) Inventors/Applicants (for US only): **KANG, Man-Jong**
[KR/KR]; #101-1202, Samik Apt., 862-14, Ochi-dong,
Buk-gu, Gwangju 500-130 (KR). **MOON, Seung Ju**
[KR/KR]; #103-502, Hyundai I-Park Apt., Yongbong-
dong, Buk-gu, Gwangju 500-070 (KR). **KIM, Hye-Min**
[KR/KR]; #544-4, Geumho-dong, Seo-gu, Gwangju
502-803 (KR). **LEE, Sang Mi** [KR/KR]; #101, Yeollip,
761, Ilsim-ri, Hwasun-eup, Hwasun-gun, Jeollanam-do
519-809 (KR).(74) Agents: **KWON, Oh-Sig** et al.; 4F, Joeeunleaderstel, 921
Dunsan-dong, Seo-gu, Daejeon 302-120 (KR).(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ,
CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ,
EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KZ,
LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK,
SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW.(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).**Published:**

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: KNOCK-IN VECTORS FOR PRODUCING BIOACTIVE SUBSTANCES BY USING PORCINE BETA-CASEIN
GENOMIC DNA, AND PROCESSES FOR PRODUCING TRANSGENIC PORCINE SOMATIC CELLS USING THE SAME

[Figure 9]

(57) Abstract: The present invention relates to a knock-in vector for producing a bioactive substance using porcine β -casein genomic DNA and a method for preparing transgenic porcine somatic cells using the same. The transgenic pig expressing hFGF generated by the method of the invention expresses a bioactive substance mammary gland- specifically, more precisely it expresses a bioactive substance in milk at significantly higher level, compared with that according to the conventional method. In addition, the hFGF protein produced from the transgenic pig of the present invention demonstrates higher stability and more excellent bioactivity than the same kinds of proteins on the market. Therefore, the transgenic pig expressing hFGF of the present invention can be effectively used for the production of hFGF protein having superior physiological activity than the conventional hFGF.

【DESCRIPTION】**【Invention Title】**

KNOCK-IN VECTORS FOR PRODUCING BIOACTIVE SUBSTANCES BY
USING PORCINE BETA-CASEIN GENOMIC DNA, AND PROCESSES FOR
5 PRODUCING TRANSGENIC PORCINE SOMATIC CELLS USING THE SAME

【Technical Field】

The present invention relates to a knock-in vector using
porcine β -casein genomic DNA and a use of the same, more
10 precisely a knock-in vector for producing bioactive substances
by using 5'end fragment containing 2.65 kb promoter, exon 1
and intron 1 of porcine β -casein genomic DNA, nuclear
localization signal (nls) region, as an 5' arm and 3'-end
fragment containing exons 5, 6 and 7 of porcine β -casein
15 genomic DNA, as 3' arm, and processes for producing porcine
somatic cells and bioactive substances in transgenic animals
using the same.

【Background Art】

20 Biomedicines produced by genetic engineering techniques
are largely obtained from the culture of *E. coli* or animal
cells, so the production cost is very high. In particular, the
protein produced from the system using *E. coli* has a problem
of 'posttranslational modification'. When animal cells are

used, huge investment cost is required for the production but productivity is not that high.

To overcome the said problem fundamentally, scientists have tried to use transgenic animals as a bioreactor.

5 A method for producing a transgenic animal is exemplified by microinjection, retroviral vector method, and animal cloning using embryonic stem cells and somatic cells.

Microinjection is the method of injecting a foreign gene into the pronucleus of a fertilized egg by using micro-
10 manipulator to express the foreign gene, which is the most classical method widely used. However, according to this method, production efficiency is only 2-3% and the level of foreign gene expression is also very low because the foreign gene is integrated at random sites in any of the genome. In
15 addition, it is impossible with this method to insert a foreign gene into a specific target site or to eliminate expression of a specific endogenous gene by homologous recombination.

Retroviral vector method is the method for producing
20 animals by the processes of producing virus by linking a foreign gene to the virus vector; and infecting a fertilized egg with the virus to produce a transgenic animal. This method is superior to microinjection in the aspect of efficiency.

However, gene is inserted at random, and it is also

impossible to insert a foreign gene into a specific target area or to eliminate expression of a specific endogenous gene, like microinjection.

To overcome the above problems, gene targeting method
5 using homologous gene recombination technique is expected.

Gene targeting was first tried in a mouse, which is the technique to induce mutation of a specific gene in a specific target area. In the experiment with mice, the gene targeting vector was transfected in embryonic stem cells and then the
10 cells where homologous recombination occurred were selected, by which chimeric mouse was produced. Based on the above method, the gene-targeted mouse was produced and has been used for the studies on gene functions (Thomas, K.R. and Capecchi, M.R. Cell, 51, 3, 503-512, 1987).

15 Gene targeting is largely divided into knock-out and knock-in. Knock-out is the method that the specific endogenous gene expression is eliminated by inserting a mutant gene into the specific site of the endogenous gene. And knock-in is the method that a mutant gene is inserted into the specific site
20 of the endogenous gene, leading to knock-out of the endogenous gene and then a foreign gene is introduced into the specific site of the endogenous gene to express. So, knock-in is the system facilitating the insertion of a foreign gene into the specific site of the endogenous gene. A foreign gene inserted

by knock-in uses the whole parts of the gene regulatory sequence on the genome where it is inserted, suggesting that the foreign gene inserted is possibly expressed as high as the expression of original endogenous gene on the region.

5 Gene targeting was first reported as a technique to induce mutation of a specific gene region in a mouse (Thomas, K.R. and Capecchi, M.R. Cell, 51, 3, 503-512 (1987)). Such gene targeting technique was applied on knock-out of α -1,3-galactosyltransferase gene and Prion protein gene in sheep.

10 But, the efficiency was only 1.7% and 1% (Denning, C. et al., Nature Biotechnology, 19, 559-562 (2001)). α -1,3-galactosyltransferase knock-out pig was produced by gene targeting to produce a transgenic cloned pig for organ transplantation (Dai, Y. et al., Nature biotechnology, 20, 251-
15 155 (2002) and Lai, L. et al. Science, 295, 1089-1092 (2002)).

To produce a useful recombinant protein or a bioactive substance, milk and urine of a transgenic animal came into notice as a bioreactor.

20 Promoters of genes capable of producing milk and urine proteins have been used to produce transgenic animals. Casein which is one of major components of milk proteins includes calcium-sensitive caseins (α -S1-, β -, and α -S2-casein) and κ -casein. Particularly, β -casein is a major component of milk of a cow, a goat and a sheep and its promoter has been used as a

proper candidate for regulation of the expression of a transgenic gene in the mammary gland of a transgenic animal.

According to a report, when β -casein was knock-out by gene targeting in a mouse, the knock-out mouse survived normally and could be even pregnant and could nurse a new born baby (Kumar S et al., Proc. Natl. Acad. Sci. USA, 91, 6138-6142, 1994). The above result indicates that when β -casein gene is knock-out by gene targeting and a specific foreign gene is inserted into that place, the foreign gene can be expressed by β -casein gene regulatory sequence. By this method, a gene targeting vector has been developed using β -casein gene region of a goat (Yu, HQ et al., Sheng Wu Gong Cheng Xue Bao, 20, 21-24, 2004) and ht-PAm (human plasminogen activator mutant) (Shen, W. et al., Sheng Wu Gong Cheng Xue Bao, 20, 361-365, 2004), GFP (Shen, W. et al., Yi Chuan Xue Bao, 32, 366-371, 2005) or human lactoferrin (Li, L et al., Yi Chuan, 28, 1513-1519, 2005) knock-in cells have been developed by using the goat β -casein gene region.

Porcine β -casein cDNA sequence (1110 bp) has been registered at GenBank (X54974) and the sequence (5544 bp) containing porcine β -casein promoter and exon 1 region has also been registered (AY452035). 3.2 kb of porcine β -casein promoter sequence and a transgenic vector using the same were reported. However, the vector has to be improved because only

promoter region thereof has been used. Korean Patent No. 646222 (published on December 7, 2005) describes the sequencing and identification of porcine β -casein gene promoter and intron 1 along with the expression vector produced by linking 3,101 bp sized porcine β -casein gene promoter sequence and 2,398 bp sized intron 1 to luciferase reporter gene and a method of gene expression in the mammary gland of a transgenic animal using the same. However, the method also uses only a partial sequence of porcine β -casein gene, which still has to be improved. To construct a knock-in vector, approximately 5 kb of homologous sequence from the start codon (ATG) to the 5' upstream and 2-5 kb of homologous sequence to the 3' downstream are required. ATG of porcine β -casein gene is located at exon 2. The nucleotide sequences of the conventional promoter, intron 1, and exon 1 can be used for the construction of homologous region to the 5' upstream but the nucleotide sequence from exon 2 containing ATG to the 3' downstream has not been reported, yet. Therefore, it is very important to obtain a sequence from ATG to the 3' downstream of porcine β -casein gene.

A targeting vector for porcine β -casein gene locus is constructed by knock-in method, a gene targeting method. The vector is precisely targeted in porcine β -casein gene of a porcine somatic cell with high gene targeting efficiency. So,

it is requested to develop a novel system for producing a transgenic pig useful for the mass-production of a target protein by using the porcine β -casein gene specific targeting vector system.

5 So, it is an object of the present invention to construct a knock-in vector using porcine β -casein gene. More precisely, the present inventors identified the whole genomic DNA of porcine β -casein gene composed of a promoter, 9 exons and 8 introns, based on which the inventors constructed a knock-in
10 vector for the production of a bioactive substance using porcine β -casein genomic DNA. And the present inventors further confirmed that a bioactive substance could be efficiently and stably expressed using the vector. At last, the present inventors completed this invention by selecting
15 porcine β -casein gene targeted somatic cells after inducing homologous gene recombination by transfection knock-in vector using porcine β -casein into porcine somatic cells.

【Disclosure】

20 **【Technical Problem】**

It is an object of the present invention to provide a method for generating a transgenic pig comprising the following steps:

- (1) identifying the whole porcine β -casein genomic DNA

and analyzing the structure thereof;

(2) constructing a knock-in vector for producing a bioactive substance using the porcine β -casein genomic DNA;

(3) transforming porcine somatic cells with the knock-in
5 vector using porcine β -casein genome;

(4) inducing homologous recombination by culturing the porcine somatic cells;

(5) selecting the porcine somatic cells in which porcine β -casein gene is successfully targeted by homologous
10 recombination of knock-in vector;

(6) eliminating the nucleus of porcine egg cells and introducing gene targeted cells therein to prepare somatic cells nuclear transfer embryo; and

(7) implanting the embryo.

15 It is another object of the present invention to provide a method for producing a bioactive substance from the transgenic pig prepared by the above method.

【Technical Solution】

20 To achieve the above objects, the present invention provides a method for preparing transformed somatic cells applicable in somatic cell cloning by using a knock-in vector to produce a bioactive substance using porcine β -casein genomic DNA.

Hereinafter, the present invention is described in detail.

The present invention also provides a knock-in vector for producing a bioactive substance using porcine β -casein gene
5 which is composed of (1) porcine β -casein genomic DNA sequence; (2) nuclear localization signal (nls) region; (3) 5' homologous region of 5 kb containing porcine β -casein genomic DNA promoter, exon 1 and intron 1; (4) bioactive gene region (ex: human basic FGF or GFP); (5) polyA signal sequence region
10 for the expression of the bioactive gene; (6) positive selection marker region; (7) 3' homologous region of 2.7 kb containing exons 5, 6 and 7 and a part of intron 4 and introns 5 and 6 and a part of intron 7 of porcine β -casein genomic DNA; and (8) negative selection marker region and is
15 characterized by that those regions of (2) - (8) are arranged in order from 5' based on (1).

In this invention, 'knock-in vector' indicates the vector useful for elimination of a specific gene expression and instead insertion of a target gene for expression in that
20 place, which contains homologous sequence of the specific gene, the target of knock-in, to induce homologous recombination.

The knock-in vector of the present invention contains a gene inserted in the porcine β -casein genome in the direction of 5'→ 3' by homologous recombination, which can be circular

or linear.

The porcine β -casein genomic DNA region of the present invention contains at least 5 kb of upstream of ATG and at least 5 kb of down stream. The region is identified by genome library screening or PCR. The region contains 9 exons and 8 introns in total and the length from exon 1 to exon 9 is 8,410 bp and the length of promoter therein is 3,107 bp, suggesting that the whole length of the region is 11,517 bp (see SEQ. ID. NO 1).

The knock-in vector herein to produce a bioactive substance using porcine β -casein gene region contains nuclear localization signal (nls) region at 5'-end. A foreign recombinant DNA containing the nuclear localization signal region plays a role in increasing introduction efficiency of a foreign recombinant DNA into the nucleus by combining with the nuclear localization signal during the migration of the transcription factor from cytoplasm to nucleus. Besides, by locating the nuclear localization signal region at 5'-end of the vector, the efficiency of the vector increases because it allows negative selection marker region to be located at 3'-end in the construction of the vector.

The nuclear localization signal region existing in SV40 promoter of the present invention can be cloned by PCR using Clontech pEGFP-N3 vector as a template.

The knock-in vector for producing a bioactive substance using porcine β -casein gene region contains 5' homologous and 3' homologous region of the porcine β -casein gene sequence.

The 5' homologous region contains 4-6 kb containing 2.65
5 kb of a promoter. The length of the 5' homologous region is an important factor determining gene targeting efficiency by homologous gene recombination, which is preferably 5 kb long.

And the 3' homologous area is located behind exon 2 ATG, which is preferably 1 - 5 kb long. Preferably, it includes
10 exons 5, 6 and 7 and a part of intron 4 and introns 5 and 6 and a part of intron 7, which would be then 2.7 kb long. This region preferably uses *Pst*I site in intron 4 and *Xba*I site in intron 7. The length of 2.7 kb of the 3' homologous region is proper for the efficient selection of cells with homologous
15 recombination by PCR. The 5' homologous region and the 3' homologous region have homology with porcine β -casein genomic DNA sequence, precisely at least 95% homology and preferably 100% homology.

The positive selection marker herein is the marker
20 facilitating positive selection by making those cells to express the selection marker alone in the presence of a selecting agent alive, which is exemplified by neomycin (neo) expressed by mouse phosphoglycerate kinase-I promoter, but not always limited thereto.

The bioactive substance in this invention is exemplified by human basic fibroblast growth factor (human bFGF) and green fluorescent protein (GFP), but not always limited thereto. Human basic fibroblast growth factor (human bFGF) sequence is shown in Sequence List 2.

PolyA in 3' region is involved in the expressions of the human basic fibroblast growth factor (human bFGF) and the green fluorescent protein (GFP) genes, precisely polyA of human FGF and GFP is generated during RNA transcription and then involved in the stabilization of RNA or migration from nucleus to cytoplasm. So, the knock-in vector of the present invention can contain polyA signal, for example SV40 polyA signal, for the stable expression of a bioactive substance.

The negative selection marker in this invention contains DT-A (Diphtheria Toxin-A) gene expressed by MC1 promoter, but not always limited thereto. In this invention, the positive selection marker and the negative selection marker are both used together, so that simultaneous positive-negative selection is possible. The negative selection marker of the present invention is located behind 3' homologous region, so that it can be easily eliminated during homologous recombination. When homologous recombination does not occur and thus expression of DT-A (Diphtheria Toxin-A) gene is still included in cells, the cells are going to be dead even without

the treatment of a specific selecting agent, making the selection of cells without homologous recombination easy.

Figure 9 is a schematic diagram illustrating the knock-in vector constructed by using porcine β -casein genomic DNA. In a preferred embodiment of the present invention, the location of ATG of β -casein gene coincides with the location of ATG of hFGF and GFP, targets for over-expression. So, the knock-in vector of the present invention is designed to use the whole gene regulatory sequence of porcine β -casein gene for the expression of a target gene. SV40 polyA signal is used herein for PolyA signal of the target gene. The positive selection marker and the negative selection marker are not contained polyA signal. When the knock-in vector is normally inserted into β -casein genome by homologous recombination, the positive selection marker will be expressed by using the original polyA of β -casein gene. The negative selection marker is designed to be selected by random insertion without homologous recombination.

According to the method described hereinbefore, 5'-hFGF(GFP)-polyA-Pneo-3' vector containing the positive selection marker alone, nls-5'-hFGF(GFP)-polyA-Pneo-3' vector composed of nls sequence known to be involved in the insertion of a vector into the nucleus of a cell and the positive selection marker, 5'-hFGF(GFP)-polyA-Pneo-3'-DT vector

containing both of the positive selection marker and the negative selection marker but not containing nls sequence, and nls-5'-hFGF(GFP)-polyA-Pneo-3'-DT vector containing all of nls sequence, the positive selection marker and the negative
5 selection marker were constructed for the expressions of hFGF and GFP (see SEQ. ID. NO 3 and 4). The composition of the knock-in vector constructed in a preferred embodiment of the present invention is shown in Table 2.

The knock-in vector constructed in this invention
10 contains an approximately 2.65 kb long promoter in the 5' upstream of exon 1 of β -casein gene. The location of ATG of GFP and hFGF, which would be expressed, coincided with the location of ATG of exon 2 of β -casein gene, suggesting that the β -casein gene regulatory sequence was used for the
15 expression.

The knock-in vector constructed to produce a bioactive substance using porcine β -casein gene region contains a 2.65 kb long promoter of porcine β -casein gene, so that the expression of a bioactive substance can be confirmed in
20 advance. When this vector is knock-in in the region of β -casein gene by homologous gene recombination, not only the 2.65 kb long promoter but also the gene regulatory sequence located in upstream of the promoter can be utilized for the gene expression. To examine whether or not the vector

constructed in this invention could be expressed in the mouse mammary gland cell line HC11 by the basic transcription factors, hFGF +DT or GFP+DT vector was introduced into the mouse mammary gland cell line HC11, followed by investigation
5 of the expression of hFGF +DT or GFP+DT after 48 hours. As a result, the expression of hFGF +DT or GFP+DT was confirmed in the HC11 cell line (see Figures 12 and 13).

In another preferred embodiment of the present invention, the present invention provides a method for preparing porcine
10 somatic cells targeted with the knock-in vector for produce a bioactive substance by using porcine β -casein genomic DNA.

The cells appropriate for the targeting by the knock-in vector of the present invention are preferably pig somatic cells. The cells can be separated from the tissues of liver,
15 kidney, spleen, muscle, lung, brain, bone marrow, ear and skin, but not always limited thereto. Particularly, ear fibroblasts are more preferred. To prepare the porcine somatic cells, ear tissues were taken from a pig at 10 days, which were cultured by tissue culture. At this time, it was more
20 preferred that the cells were fibroblasts.

The porcine ear cells were preferably cultured in DMEM (Dulbecco's Modified Eagle medium) supplemented with 20% FBS, 1mM sodium pyruvate, 1 $\times$ non-essential amino acid, 1 $\times$ β -mercaptoethanol, 100 Unit/ml penicillin and 100 μ g/ml

streptomycin.

The introduction of the vector constructed in this invention into the cells could be performed by the conventional method used for the introduction of nucleic acid
5 into a cell, and in this invention, electroporation was preferably used, but not always limited thereto.

When the vector were introduced into somatic cell by electroporation, the preferable density of the somatic cells was 4×10^6 cell/0.4ml - 6×10^6 cell/0.4ml, and 5×10^6
10 cell/0.4ml were more preferred. Electroporation was performed at 420V - 480V, and more preferably performed at the condition of 450V, 4 pulses and 1 ms.

To select the targeted cells, 100µg/ml - 600µg/ml of G418 was used and 300µg/ml was more preferred. The selection was
15 preferably performed in 12 - 18 days, and more preferably in 14 days.

The targeted cells introduced with the vector of the present invention were preferably identified by PCR and Southern blotting.

20 For the identification by PCR, G418-resistant colonies were selected, followed by sub-culture. During the sub-culture, cells were collected, followed by PCR. PCR was performed with a sense primer (neo)(GCCTGCTTGCCGAATATCATGGTGGAAAAT) and an antisense primer (SEQ. ID. NO: 21 :

GGCATGTGGGGAAATAATTGCACATAAGGA) as follows: denaturation at 94°C for 2 minutes, at 94°C for 30 seconds, at 66°C for 30 seconds, at 72°C for 5 minutes (35 cycles) and then at 72°C for 10 minutes.

5 For the identification by Southern blotting, the cells confirmed by the PCR above were sub-cultured and then genomic DNA was extracted. The genomic DNA was digested with *EcoRV* and then purified, followed by electrophoresis on 0.8% agarose gel.

The DNA was transferred on nitrocellulose membrane. The
10 introduction of the knock-in vector to the target cell and the insertion thereof into the precise target genome site by homologous recombination was confirmed using 600 bp/*SpeI-EcoRI* DNA fragment containing porcine β -casein exon 9 as probe.

The knock-in porcine somatic cell line prepared by the
15 method for preparing porcine somatic cells targeted with the knock-in vector of the present invention was named 'porcine β -casein human bFGF knock-in #939' which was deposited at KCTC (Korean Collection for Type Cultures, Korean Research Institute of Bioscience and Biotechnology, 111 Gwahangno,
20 Yuseong-gu, Daejeon, Korea) on September 11, 2008 (Accession No: KCTC 11389BP).

In another preferred embodiment of the present invention, the present invention provides a method for preparing a transgenic pig comprising the following steps: (1) eliminating

the nucleus of a porcine egg and introducing a gene targeted cell therein to prepare a clone embryo by somatic cell nuclear transfer; and (2) implanting the embryo, and a method for separation and purification of a bioactive substance from (3)
5 the transgenic pig prepared above.

The porcine somatic cell targeted with the knock-in vector of the present invention for producing a bioactive substance using porcine β -casein gene region was transplanted into in the nucleus-eliminated porcine egg. And the prepared
10 embryo is implanted to produce a knock-in pig producing a bioactive substance using porcine β -casein gene region.

The knock-in pig can produce a useful protein and the bioactive substance can be separated by the separation and purification method, particularly by filtration or
15 chromatography.

【Advantageous Effects】

The porcine β -casein genomic DNA of the present invention contains 3,101 bp promoter region, 9 exons and 8
20 introns, so that it can be effectively used for the construction of the knock-in vector for the porcine mammary gland specific expression.

The knock-in vector constructed by using the porcine β -casein genomic DNA of the present invention can induce gene

expression by using the whole gene regulatory sequence containing a promoter, so that it gives higher expression efficiency, compared with the method for inducing gene expression in the conventional transgenic animal.

5 Bioactive substances for example human basic fibroblast growth factor (human bFGF) and green fluorescent protein (GFP) can be efficiently and stably expressed and produced from porcine milk by using the knock-in vector using porcine β -casein genomic DNA of the present invention.

10 The transgenic pig expressing hFGF of the present invention expresses the human bioactive substance in mammary gland specifically and the expression level is very high, compared with that according to the conventional method.

 The hFGF produced from the transgenic pig of the present
15 invention demonstrates higher stability and more excellent bioactivity than the same kinds of proteins on the market.

 The hFGF expression vector and the transgenic pig of the present invention can be effectively used for the production of hFGF having higher physiological activity than the
20 conventional hFGF.

【Description of Drawings】

 The above and other objects, features and advantages of the present invention will become apparent from the following

description of preferred embodiments given in conjunction with the accompanying drawings, in which:

Figure 1 is a diagram illustrating the result of porcine
5 β -casein cDNA identification;

Figure 2 is a diagram illustrating the result of porcine β -casein genome library screening;

Figure 3 is a diagram illustrating the result of porcine β -casein genomic DNA cloning by LA-PCR;

10 Figure 4 is a diagram illustrating the structure and restriction enzyme map of porcine β -casein genomic DNA;

Figure 5 is a diagram illustrating the identification of human FGF cDNA and the nucleotide sequence thereof;

Figure 6 is a diagram illustrating the identification of
15 GFP gene and the nucleotide sequence thereof;

Figure 7 is a diagram illustrating the identification of NLS sequence and the nucleotide sequence thereof;

Figure 8 is a diagram illustrating the identification of SV40 polyA signal sequence and the nucleotide sequence
20 thereof;

Figure 9 is a schematic diagram illustrating the knock-in vector constructed by using porcine β -casein region;

Figure 10 is a schematic diagram illustrating the 4 kinds of knock-in vectors for inducing hFGF expression;

Figure 11 is a schematic diagram illustrating the 4 kinds of knock-in vectors for inducing GFP expression;

Figure 12 is a diagram illustrating the expression of 5'-GFP-SV40pA-neo- β 3'-DT knock-in vector in HC11 cells;

5 Figure 13 is a diagram illustrating the result of RT-PCR examining the expression of human basic FGF in HC11, the mouse mammary gland cell line, introduced with the knock-in vector. The number of colony expressing hFGF is marked as red;

10 Figure 14 is a photograph showing the somatic cells prepared from porcine ear cells.

Figure 15 is a diagram illustrating the introduction efficiency of the knock-in vector into porcine ear somatic cells, confirmed by using GFP gene;

15 Figure 16 is a schematic diagram illustrating the processes of introducing the knock-in vector into the somatic cells and sub-culturing the G418 resistance colonies;

20 Figure 17 is a diagram illustrating the result of PCR and Southern blotting of the somatic cells introduced with the knock-in vector. Name of each colony was presented on the upper part of the diagram as a number;

Figure 18 is a diagram illustrating the karyotypes of the knock-in somatic cells. Herein, the location of XY chromosome is marked with red box.

【Best Mode】

Practical and presently preferred embodiments of the present invention are illustrative as shown in the following examples.

5 However, it will be appreciated that those skilled in the art, on consideration of this disclosure, may make modifications and improvements within the spirit and scope of the present invention.

10 Example 1: Identification of porcine β -casein cDNA

Total RNA was extracted from the mammary gland of a female porcine and porcine β -casein cDNA was identified by RT-PCR as follows. First, 0.1 g of the mammary gland tissue was put in 1 ml of Trizol reagent (Gibco BRL), followed by
15 homogenization for one minute. Total RNA was recovered by ethanol precipitation, which was dissolved in RNase-free water and stored at -70°C . The first strand cDNA for RT-PCR was synthesized using 5 μg of the total RNA of the mammary gland with Superscript^{II} RNaseH-reverse transcriptase (Invitrogen)
20 and random primer (Takara). PCR was performed with 1 μl of the first strand cDNA using 20 pmol of sense(SEQ. ID. NO: 5: AAAGGACTTGATCGCCATGAAGCTCCTCAT) and antisense(SEQ. ID. NO: 6: AATCCTCTTAGACAAGGTTGTAACTGGGG) primers corresponding to porcine β -casein gene (Accession No. X54974), 1 $\times$ PCR buffer,

0.5 U Ex Taq polymerase (Takara) and 200 μ M of dNTP mixture as follows; denaturation at 94°C for 30 seconds, annealing at 68°C for 30 seconds, and extension at 72°C for 1 minute (33 cycles).

The PCR product was confirmed by electrophoresis on 0.8% agarose gel, followed by sub-cloning to pGEM T-easy vector (Promega) for sequencing. Sequencing was performed by using ABI PRISM 377 sequencer (Applied Biosystems). The sequenced nucleotide sequence was analyzed by Genetyx-win (version 4.0).

The results are shown in Figure 1. As shown in Figure 1, the identified porcine β -casein cDNA demonstrated 98% homology with the gene registered at NCBI, suggesting that β -casein cDNA was normally identified. The identified porcine β -casein cDNA was used as a probe for the identification of β -casein genomic DNA.

Example 2: Construction of genome library from porcine fetal fibroblasts

Genomic DNA was extracted from porcine fetal fibroblasts (female) and purified by the conventional method, followed by partial digestion with *Sau3A*I. 50 μ g of the genomic DNA was treated with 1 unit of the restriction enzyme at 37°C for 5, 10, 15, 20, 25, 30, 40, 50, 60, and 90 minutes. After recovering 5 μ g of the DNA, the restriction enzyme reaction was terminated.

A part of the DNA was electrophoresed, resulting in

proper sized fragments. The partially digested group (15 minutes) was used for the preparation of an insert. The insert was treated with CIAP (calf intestine-alkaline phosphatase).

The CIAP treated insert was ligated to γ DASHII vector (Stratagene) pre-digested with *Bam*HI. After packaging the ligated γ DNA using Gigapack III XL packaging extract, which infected XL1-blue MRA (P2) host, followed by measurement of titer. This library was amplified again and the titer of the amplified library was measured and used for screening.

Example 3: Identification and interpretation of β -casein genomic DNA

For the first screening, the above library was plated on NZCYM plate (50,000 pfu/plate), so total 500,000 libraries were used for screening. Plaque was lifted by using two nitrocellulose membranes per each plate. The membrane was treated in denaturation solution (1.5M NaCl, 0.5M NaOH) for 2 minutes and then treated in neutralization solution for 5 minutes (1.5M NaCl, 0.5M Tris-HCl (pH8.0)), followed by washing in 2x SSC buffer containing 0.2 M Tris-HCl (pH 7.5) for 30 seconds. DNA was cross-linked onto the membrane by using UV crosslinker. The membrane was hybridized in a solution containing 5 $\times$ SSPE, 5 $\times$ Denhardt's, 1% SDS (w/v) and 50% formamide (w/v) at 42 $^{\circ}$ C for 15 hours. A probe was prepared by

using the cloned β -casein cDNA/HindIII 160 bp DNA fragment, random labeling kit (Amersham) and [α -32P]dCTP (110 TBq/mmol, Amersham). The concentration of the probe for the hybridization was used one million cpm/ml hybridization
5 solution. After hybridization, the membrane was washed three times with 0.2% SSC and 0.1% SDS (w/v) at 68°C for 30 minutes.

The membrane was exposed on X-ray film at -80°C for 72 hours, and then developed.

For the second screening, the positive signal obtained
10 from the first screening was distributed on NZCYM plate (1,000 pfu/plate), and then plaque was lifted by using two nitrocellulose membranes per each plate. The membrane was treated by the same manner as described in the above for the first screening, followed by hybridization, washing and
15 autoradiography under the same conditions as described above.

For the third screening, the positive signal obtained from the second screening was distributed on NZCYM plate (1,000 pfu/plate), and then plaque was lifted by using two nitrocellulose membranes per each plate. The membrane was
20 treated by the same manner as described in the above for the first or second screening, followed by hybridization, washing and autoradiography under the same conditions as described above. As a result, two positive clones were obtained and the results are shown in Figure 2.

To purify phage DNA, XL1-blue MRA(P2) host was infected with the positive phage and the phage DNA was purified by using Lamda midi kit (Qiagen). The purified phage DNA was digested with the restriction enzymes *NotI* and *SalI*, followed
5 by electrophoresis on 0.8% agarose gel. The size of the insert was measured. The DNA was transferred onto zeta-probe membrane (Bio-Rad) for Southern blotting. Hybridization was performed by the same manner as described above using β -casein cDNA/*HindIII* 160 bp as probe.

10 Phage DNA was treated with *SalI* for sub-cloning the band demonstrating positive signal after hybridization into pBluescript SK, followed by electrophoresis on 0.5% agarose gel. The insert was prepared by using Gene clean kit II. The pBluescript SK vector was digested with the same restriction
15 enzyme *SalI* and then treated with alkaline phosphatase. The vector and the insert (molar ratio = 1:2) were ligated by using DNA ligation kit ver2 (Takara) at 16°C for overnight, followed by transformation using XL10-Gold Ultracompetent cells. After transformation, colonies were recovered and
20 plasmid was purified and digested with the restriction enzyme *SalI* to confirm the introduction of the insert. The sub-cloned DNA was treated with each restriction enzyme, leading to the construction of restriction enzyme map.

The identified genomic DNA was digested with the

restriction enzymes *SalI* and *NotI*, followed by electrophoresis on 0.8% agarose gel. The size of the insert was measured and confirmed to be approximately 16 kb. The 16 kb sized insert was sub-cloned into pBluescript-SK, followed by analysis using
5 restriction enzyme map and sequencing. As shown in Figure 4, the insert was identified as the clone containing exon 6 - exon 9. To construct a knock-in vector, upstream and downstream of exon 2 containing methionine site are necessary.

Thus, β -casein genomic DNA was identified by LA-PCR (Long
10 & Accurate PCR) using NCBI genome and cDNA information.

Example 4: Identification of porcine β -casein genomic DNA by LA-PCR

Porcine β -casein genomic DNA was extracted from the
15 porcine fetal fibroblasts and LA-PCR was conducted for identify of porcine β -casein genomic DNA as follows. Long PCR was performed using primers prepared based on porcine β -casein promoter region (Accession No. AY452035) reported to NCBI and the result of cDNA analysis thereof. Primers for PCR were as
20 follows; porcine β -casein promoter upstream sense primer (SEQ. ID. NO: 7: CCCACTATTTCTGATTCTTGATTAACCTT), antisense primer presumed to be β -casein exon 6 (SEQ. ID. NO: 8: TGTGTTCTCTCCCGCTTTAGCTTCTCAATT), sense primer containing exon 2 (SEQ. ID. NO: 9: GACTTGATCGCCATGAAGCTCCTCATCCTT) and

antisense primer containing exon 9 (SEQ. ID. NO: 10: GCCTAAGGATTAATTTATTGAAATGACTGG). PCR was performed using 100 ng of the genome DNA extracted from porcine fetal fibroblasts with 10 pmol of sense and antisense primers, 0.5U i-Max II DNA
5 polymerase (Intron), 1×PCR buffer and 200μM dNTP mixture as follows; denaturation at 94°C for 30 seconds, annealing at 63°C for 30 seconds, and extension at 72°C for 7 minutes and 30 seconds (30 cycles). The PCR product was confirmed by electrophoresis on 0.8% agarose gel. As a result, as shown in
10 Figure 3, bands in sizes of approximately 6.8 kb and 5.5 kb were detected.

Sub-cloning was performed using pGEM T-easy vector (Promega) for sequencing. Sequencing was performed by using ABI PRISM 377 sequencer (Applied Biosystems). The sequenced
15 nucleotide sequence was analyzed by Genetyx-win (version 4.0).

As a result, it was identified as β-casein DNA (see SEQ. ID. NO: 1). The identified porcine β-casein genome included 9 exons and 8 introns, which was consistent with the report on human and rodent β-casein genome. The entire nucleotide
20 sequence was analyzed and exon-intron joint region was investigated and the results are shown in Table 1.

【Table 1】

exon No	exon size (bp)	sequence of exon-intron linked part		length of intron
		5' splice doner	3'splice acceptor	
1	45	GGA GAA AAG gtaagaattt	ccatttcacag GAC TTG ATC	2398
2	63	GCA AGA GCG gtaagtacag A R A	ttctctatag AAG GAA GAA K E E	735
3	27	TCT GGT GAG gtaagatatt S G E	tccttttcag ACT GTG GAA T V E	118
4	27	AGC AGT GAG gtaagccaat S S E	ttttctaaag GAA TCT ATT E S I	1325
5	24	ATC AGC AAG gtaaagactt I S K	tgttttctag GAG AAA ATT E K I	94
6	45	CAA ACA GAG gtaatttggt Q T E	ttctttccag GAT GAA CGC D E R	1307
7	519	TAC AAC CCT gtaagtccaa Y N P	aatttttaag GTC TAA GAG V	606
8	42	TCA CTT TTG gtaagcttta	tattccgcag AAT TGA CTG	784
9	305			

As shown in Table 1, exon 1 was 45 bp, exon 2 was 63 bp, exon 3 was 27 bp, exon 4 was 27 bp, exon 5 was 24 bp, exon 6 was 45 bp, exon 7 was 519 bp, exon 8 was 42 bp, and exon 9 was 306 bp. Except exon 7 and exon 9, all the other exons were very short. In the meantime, intron 1 was 2398 bp, intron 2 was 735 bp, intron 3 was 118 bp, intron 4 was 1325 bp, intron 5 was 94 bp, intron 6 was 1307 bp, intron 7 was 606 bp, and intron 8 was 734 bp. It was confirmed from the above results that β -casein genome comprising exon 1 - exon 9 was 8,410 bp in length. The promoter region was 3,107 bp, so the identified β -casein region containing the promoter was 11,517 bp in full length. This result indicates that the β -casein gene contains 5 kb long upstream and downstream of exon 2 necessary for the

construction of the β -casein knock-in vector.

Example 5: Production of *E. coli* containing porcine β -casein genomic DNA

5 pGEM T-easy vector (Promega) containing porcine β -casein genomic DNA was mixed with *E. coli* DH5 α , which hold on ice for 30 minutes, followed by reaction at 42°C for 30 seconds to introduce the target gene into *E. coli*. *E. coli* transformed with the gene was cultured in LB liquid medium at 37°C for 1
10 hour and then further cultured on LB plate for overnight, followed by selection. The selected *E. coli* was cultured in liquid medium again for overnight and then preserved as frozen.

Example 6: Identification and analysis of human FGF gene

15 Total RNA was extracted from human liver cancer cell line HepG2 (ATCC HB-8065), followed by RT-PCR (Reverse Transcription-Polymerase Chain Reaction) to identify human FGF gene. To extract total RNA, the recovered cells were put in 1
20 ml of Trizol reagent (Gibco BRL), followed by homogenization for 1 minute. Ethanol precipitation was performed and the precipitate was dissolved in RNase-free water, which was stored at -70°C until use. The first strand cDNA for the cloning of human FGF gene was synthesized using 5 μ g of the total RNA with Superscript II RNaseH-reverse transcriptase

(Invitrogen) and random primer (Takara). PCR was performed with 1 μ l of the synthesized first strand cDNA using 20 pmol of sense (SEQ. ID. NO: 11: GCCCCGCAGGGACCATGGCA) and antisense (SEQ. ID. NO: 12: AAATCAGCTCTTAGCAGACA) primers corresponding to the reported human FGF (Accession No. NM_002006), 1 $\times$ PCR buffer, 0.5 U Taq polymerase (Promega) and 200 μ M of dNTP mixture as follows; denaturation at 94 $^{\circ}$ C for 30 seconds, annealing at 50 $^{\circ}$ C for 30 seconds, and extension at 72 $^{\circ}$ C for 1 minute (40 cycles). The PCR product was confirmed by electrophoresis on 2% agarose gel, followed by sequencing and analyzing thereof. The identified FGF2 was 484 bp. After determining the nucleotide sequence, homology test was performed. As a result, the sequence was 100% identical with the nucleotide sequence of FGF2 reported on NCBI and the results are shown in Figure 5.

Example 7: Identification and analysis of GFP gene

GFP gene was cloned used pEGFP-N3 vector (Clontech) as template by PCR. PCR was performed with 10pg of DNA using 20 pmol of sense (SEQ. ID. NO: 13: AAGCTTCGAATTCTGCAGTC) and antisense (SEQ. ID. NO: 14: CTCGAGTTACTTGTACAGCT) primers, 1 $\times$ PCR buffer, 0.5 U Taq polymerase (Promega) and 200 μ M of dNTP mixture as follows; denaturation at 94 $^{\circ}$ C for 30 seconds, annealing at 60 $^{\circ}$ C for 30 seconds, and extension at 72 $^{\circ}$ C for 1

minute (35 cycles). The PCR product was confirmed by electrophoresis on 1.5% agarose gel, followed by sequencing and analyzing thereof. As shown in Figure 3, 778 bp sized fragment was identified. After determining the nucleotide
5 sequence, homology test was performed. As a result, it was confirmed to be GFP gene and the results are shown in Figure 6B.

Example 8: Identification of NLS and sequencing thereof

10 NLS was cloned used pEGFP-N3 vector (Clontech) as template by PCR. PCR was performed with 10pg of DNA using 10 pmol of sense (SEQ. ID. NO: 15: CGGAGTTAGGGCGGGACTA) and antisense (SEQ. ID. NO: 16: CCAGCTGTGG AATGTGTGTC) primers, 1×PCR buffer, 0.5 U Taq polymerase (Promega) and 200μM of
15 dNTP mixture as follows; denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds, and extension at 72°C for 30 seconds (40 cycles). The PCR product was confirmed by electrophoresis on 2% agarose gel, followed by sequencing and analyzing thereof. The identified nls had 72bp tandem repeat
20 structure. The nucleotide sequence was consistent (100%) with that shown in Figure 7.

Example 9: Identification of SV40 polyA signal sequence and sequencing thereof

SV40 polyA was cloned used pCMV-Tag1 vector (Stratagene) as template by PCR. PCR was performed with 10pg of DNA using 20 pmol of sense (SEQ. ID. NO: 17: AAGCTTATCGATACCGTCGA) and antisense (SEQ. ID. NO: 18: GGGCCCTTAAGATACATTGATGAG) primers, 5 1×PCR buffer, 0.5 U Taq polymerase (Promega) and 200μM of dNTP mixture as follows; denaturation at 94°C for 30 seconds, annealing at 50°C for 30 seconds, and extension at 72°C for 40 seconds (40 cycles). The PCR product was confirmed by electrophoresis on 2% agarose gel, followed by sequencing and 10 analyzing thereof. As a result, 504 bp sized fragment was identified, and the nucleotide sequence was confirmed that the sequence was SV40 polyA signal sequence as shown in Figure 8.

Example 10: Construction of knock-in vector by using
15 porcine β-casein genomic DNA

According to a preferred embodiment of the present invention, the composition of the knock-in vector of the invention is as show in Table 2.

【Table 2】

Name	Structure	Description
hFGF-DT	5'-hFGF-polyA-Pneo-3'	Containing positive selection marker alone
nls+hFGF-DT	nls-5'-hFGF-polyA-Pneo-3'	Containing nls sequence and positive selection marker

hFGF+DT	5'-hFGF-polyA-Pneo-3'-DT	Containing positive selection marker and negative selection marker but not containing nls sequence
nls+hFGF+DT	nls-5'-hFGF-polyA-Pneo-3'-DT	Containing nls sequence, positive selection marker and negative selection marker
GFP-DT	5'-GFP-polyA-Pneo-3'	Containing positive selection marker alone
nls+GFP-DT	nls-5'-GFP-polyA-Pneo-3'	Containing nls sequence and positive selection marker
GFP+DT	5'-GFP-polyA-Pneo-3'-DT	Containing positive selection marker and negative selection marker but not containing nls sequence
nls+GFP+DT	nls-5'-GFP-polyA-Pneo-3'-DT	Containing nls sequence, positive selection marker and negative selection marker

The knock-in vector was constructed by using 5kb fragment of upstream of porcine β -casein genome as the left arm and using approximately 2.7 kb of 3' region containing porcine β -casein exon 5 - exon 7 as the right arm. As a positive selection marker, neo gene expressed by mouse phosphoglycerate kinase-I promoter was used. As a negative selection marker, DT-A gene was used. As a result, the knock-in vector facilitating simultaneous positive-negative selection was constructed. The start codon of β -casein was linked to be consistent with the start codon of GFP and hFGF. As a result, nls sequence+5'+hFGF(GFP)+polyA+Pneo+3'+DT was constructed as follows.

5' region of human basic FGF, the target for over-expression, was inserted with *NcoI* restriction enzyme site and 3' region thereof was inserted with *SalI* restriction enzyme site to construct primers (sense: GCCCCGCAGGGACCATGGCA, anti-sense: GTCGACCCATTAAAATCAGC). PCR was performed with the primers using the human basic FGF cDNA as a template. The amplified PCR product was sub-cloned into pGEM-T-easy vector.

SV40 polyA was sub-cloned by PCR using pCMV-Taql DNA as a template with a sense primer (CTCGAGACTCGATCGCCCTT) inserted with *XhoI* restriction enzyme site and an antisense primer (GATATCAATTTACGCGTTAA) inserted with *EcoRV* restriction enzyme site. The nls sequence known to be involved in the introduction of a vector into the nucleus of a cell was amplified by PCR using pEGFP-N3 DNA as a template with nls sequence specific primers (sense primer: CGGAGTTAGGGGCGGGACTA, anti-sense: CCAGCTGTGGAATGTGTGTC), followed by sub-cloning into pGEM-T easy vector.

To construct 5' arm of the knock-in vector, *NotI* restriction enzyme site was inserted into the promoter region of the sequence ranging from porcine β -casein promoter to exon 2 and *NcoI* restriction enzyme site was inserted into the place of exon 2, resulting in the construction of primers (sense: GCGGCCGCGATATCTAGGGTCTCTTCTAGT, anti-sense: CCATGGCGATCAAGTCCTGTGAATGGGGAA) for PCR amplification.

The amplified 5' arm region of the knock-in vector was digested with *NotI*-*NcoI* to prepare a 5 kb insert. Human basic FGF gene, the target for over-expression, was digested with *NcoI*-*SallI* to prepare an insert. 3' fragment ligation was performed using pBSK+/*NotI*-*SallI* vector. Whether the fragment of the knock-in vector, 5'arm+hFGF, was normal was examined by sequencing. Once it was confirmed to be normal, nls sequence was inserted in front of the promoter. To insert nls sequence, nls ligated pGEM T-easy vector was digested with *SpeI*, followed by blunting of sticky end. *NotI* site on the opposite was digested with the restriction enzyme, leading to the preparation of the insert from the nls fragment (*NotI*-Blunting). The plasmid containing 5'arm+hFGF fragment of the knock-in vector was digested with *NotI*-*EcoRV*, resulting in the vector capable of ligation of nls fragment (*NotI*-Blunting).

The two fragments were ligated, resulting in the construction of nls sequence+5'arm+hFGF. To link SV40 poly A to behind hFGF gene, SV40 poly A plasmid sub-cloned by PCR was digested with *XhoI*-*EcoRV*, resulting in the preparation of an insert. The plasmid containing nls sequence+5'arm+hFGF fragment was digested with *NotI*-*SallI*, resulting in the preparation of another insert. 3' fragment ligation was performed to pBSK+/*NotI*-*EcoRV* vector. As a result, the plasmid containing nls sequence+5'arm+hFGF+SV40 poly A was constructed.

To use β -casein 3' region for the construction of the knock-in vector, *Pst*I site located in intron 4 of β -casein genomic DNA and *Xba*I site located in intron 7 thereof were used. When β -casein 3' region was digested with *Pst*I and *Xba*I, a fragment of approximately 2.7 kb was obtained. This fragment was sub-cloned into pBSK+/*Pst*I-*Xba*I vector. The sub-cloned DNA was digested with *Pst*I, followed by blunting. After ligating *Sal*I linker thereto, the DNA was digested with *Xba*I and the resultant fragment was ligated to pBSK+/*Sal*I-*Xba*I vector.

After confirming that the restriction enzyme *Pst*I site was substituted with *Sal*I, *Xba*I site was substituted with *Xho*I.

At last, the plasmid in which β -casein 3' region was ligated to pBSK+/*Sal*I-*Xho*I vector was prepared. PGK-neo gene was ligated to pBSK+/*Eco*RV-*Sal*I vector by enzyme site substitution. To link β -casein 3' region to PGK-neo, the plasmid containing β -casein 3' region was digested with *Eco*RV-*Sal*I. And PGK-neo plasmid was digested with *Eco*RV-*Sal*I, resulting in the preparation of an insert. At last, the plasmid in which PGKneo + β -casein 3' region was linked to pBSK+/*Eco*RV-*Xho*I was prepared by ligation of the two fragments.

To ligate pBSK+/*Not*I-*Eco*RV (nls sequence+5'arm+hFGF+SV40 poly A) to pBSK+/*Eco*RV-*Xho*I (PGK-neo+ β -casein 3' region), pBSK+/*Eco*RV-*Xho*I (PGK-neo+ β -casein 3' region) plasmid was digested with *Eco*RV-*Xho*I, resulting in the preparation of an

insert. pBSK+/NotI-EcoRV (nls sequence+5'arm+hFGF+SV40 poly A) plasmid was digested with EcoRV-XhoI, preparing a vector for ligation. The prepared two fragments were ligated to obtain pBSK+/NotI-XhoI(nls sequence+5'arm+hFGF+SV40poly A+PGKneo+ β -casein 3' region) plasmid. The plasmid was digested with NotI-XhoI, resulting in the preparation of an insert, which was introduced in NotI-XhoI site of pMCDT-A (A+T/pau) vector. As a result, the final knock-in vector nls sequence+5'+hFGF+polyA+Pneo+3'+DT was constructed. The result is shown in Figure 9 as a schematic diagram (see Figures 9, 10 and 11). Figure 9a - Figure 9e illustrate the full length nucleotide sequence (SEQ. ID. NO: 3) and location of each part of nls sequence+5'+hFGF+polyA+Pneo+3' vector, one of the knock-in vectors of the present invention. Figure 10a - Figure 10e illustrate the full length nucleotide sequence (SEQ. ID. NO: 4) and location of each part of nls sequence+5'+GFP+polyA+Pneo+3' vector, one of the knock-in vectors of the present invention.

Example 11: Expression of the constructed knock-in vector

To confirm whether the constructed knock-in vector could be expressed in mammary gland cells, the vector 5'+GFP+polyA+Pneo+3'+DT designed to express GFP was transfected in the mouse mammary gland cell line HC11 (derived

from COMMA-1D mouse mammary gland cell line (Danielson, K. et al. (1984), Proc. Nat. Acad. Sci. USA 81, 3756-3760). The HC11 cells were cultured in RPMI1640 containing 10% FBS and antibiotics. Transfection was performed as follows. 16 hours
5 before the transfection, the HC11 cells were inoculated on 3 cm dishes at the density of 2.4×10^5 cells/dish and culture medium was replaced with 1.8 ml of fresh medium on the next day. 4 μ g of GFP DNA (control) and the knock-in vector DNA (0.2 μ g/ μ l) were used for transfection. 8 μ l of jetPEI™ transfection
10 reagent (Polyplus-transfection) was added to each sample as a transfection reagent. 20 μ l (4 μ g) of DNA prepared for transfection and 80 μ l of 150 mM NaCl were mixed to make the total volume 100 μ l. The mixture was well mixed for 15 seconds, to which the mixture of 8 μ l of the transfection reagent
15 jetPEI™ and 92 μ l of 150 mM NaCl (100 μ l in total volume) was added. 100 μ l of the mixed DNA solution and 100 μ l of the transfection reagent solution were put in a tube, followed by vortexing for 15 seconds. The mixture holds at room temperature for 30 minutes. Transfection was performed by
20 inoculating the DNAjetPEI™ mixture on the culture dish (200 μ l/dish). After 24 hours, the culture medium was replaced with fresh one and 48 hours later, GFP expression was investigated under fluorescent microscope. Transfection efficiency and GFP expression as a positive control were examined by using pEGFP-

N3 DNA.

As shown in Figure 12, fluorescence was detected in HC11 cells transfected with the knock-in vector 5'-GFP-SV40pA-neo- β 3'-DT. Compared with the control vector pEGFP-N3, the number of cells expressing GFP was low, which seems to be because the size of the knock-in vector was approximately 14 kb which was three times longer than the control vector pEGFP-N3 (4.7 kb), suggesting that gene transfection efficiency decreased. And, GFP can be extracellular secreted, which might be another reason of the low GFP expression. However, the result herein definitely indicates that GFP can be stably expressed by using ATG in exon 2 of β -casein gene in the construction of the knock-in vector.

Example 12: Expression of human basic FGF knock-in vector in HC11 cells

To introduce the human basic FGF knock-in vector in HC11 cells, the cells cultured in RPMI1640 supplemented with 10% FBS and antibiotics and inoculated on 3.5 cm dishes at the density of 2×10^5 cells/dish. After 24 hours, DNA was transfected using jetPEI transfection reagent. To select the cells transfected with the DNA, the knock-in vector and SV2 neo gene were diluted at the molar ratio of 10:1, followed by transfection with 4 μ g of the DNA.

In order to selection cell, the cells were inoculated in a 96-well plate, followed by selection using 300 ug/ml of G418.

The selected colonies were sub-cultured in 24 well, 12 well, 6 well and 10 cm dish, which were frozen or used for experiments. The expression of human basic FGF was examined by RT-PCR as follows. Total RNA was extracted from the HC11 cells transfected with the knock-in vector by using Trizol Reagent (Gibco BRL). The first strand cDNA for RT-PCR was synthesized using 2 μ g of the total RNA with Superscript II RNaseH-reverse transcriptase (Invitrogen) and random primer (Takara). PCR was performed with 1 μ l of the first strand cDNA using human basic FGF sense primer (SEQ. ID. NO: 19: CCCCACGGCCGAGTTGAC) and human basic FGF antisense primer (SEQ. ID. NO: 20: AGATTCCAATCGTTCAAAA) as follows; predenaturation at 94°C for 2 minutes, denaturation at 94°C for 30 seconds, annealing at 55°C for 20 seconds, and extension at 72°C for 40 seconds, 33 cycles from denaturation to extension, and final extension at 72°C for 10 minutes. The amplified PCR product was confirmed by electrophoresis on 2% agarose gel. As a result, 200 bp sized PCR product expressed by the gene producing human basic FGF specifically by the knock-in vector was confirmed and the results are shown in Figure 13.

Example 13: Preparation of porcine ear somatic cells and

culture of the same

Porcine ear cells were prepared by using the ear of a pig at 10 days. Hair was eliminated from both sides of the porcine ear by using surgical scalpel and the both sides of the ear was sterilized by rubbing the ear with 2% cresol cotton. After sterilizing with cresol cotton, the ear was sterilized again with 70% alcohol cotton. After sterilization, ear tissues were cut into 2x2 cm by using sterilized scissors. The ear tissues were washed with PBS containing antibiotics 2 - 3 times, which were then put in a 50 ml falcon tube containing 15 ml medium supplemented with DMEM (WelGENE), 15% FBS, 100 unit/ml penicillin and 100 µg/ml streptomycin, which was transferred in a Lab with keeping the temperature of the tube as 4°C. The tissues were transferred on Petri dish and the remaining hair was eliminated. Then, the tissues were dipped in 70% alcohol for 1 -2 seconds, followed by washing with PBS containing antibiotics.

Next, the following processes were carried out under stereoscopic microscope. Epidermis and cartilage were eliminated by using scalpel and pincette to expose basal epithelium. The obtained tissues were cut into 2x2 mm sections.

The sections were put in 6 cm dish containing 2 ml of culture medium, followed by culture for 2 days. When cells were grown from the tissue sections and attached to the dish,

3 ml of culture medium was added thereto, followed by further culture until 80-90% confluent. The cells were sub-cultured once and then treated with trypsin. The cells were for cryopreservation. For the culture, DMEM (WelGENE) containing 15% FBS, 100 unit/ml penicillin and 100 µg/ml streptomycin was used. 10% DMSO was added for freezing the cells. The ear cells prepared by the tissue culture method exhibited typical morphology of fibroblasts. The result is shown in Figure 14.

Figure 14A is a photograph illustrating the cells during preparation and Figure 14B is a photograph illustrating the somatic cells under the general cell culture after preparation.

Example 14: Transfection of the constructed human basic FGF knock-in vector into somatic cells, selection and sub-culture thereof

Electroporation was used for the transfection of the gene targeting vector into porcine somatic cells. For the electroporation, the cultured cells were recovered by treating with trypsin and the cells were resuspended in F10 medium at the density of 5×10^6 cell/0.4 ml. 10 µg of the linear gene targeting vector dissolved in 0.1 ml of F10 medium was added thereto. The cell-vector mixture was put in a 4 mm gap cuvette.

The cuvette was loaded in BTX Electro-cell manipulator (ECM 2001), followed by electroporation at 450 V, 4 pulses and

1 ms. Then, the cuvette was placed on ice for 10 minutes. The cells in the cuvette were resuspended in 10 ml of medium, which was distributed in a 48-well plate, followed by further culture. After 48 hours, selection was performed for 12 days
5 using 300 $\mu\text{g/ml}$ of G418. During this selection, DMEM supplemented with 20% FBS, 1 mM sodium pyruvate, 1 $\times$ non-essential amino acid, 1 $\times$ β -mercaptoethanol, 100 Unit/ml penicillin and 100 $\mu\text{g/ml}$ streptomycin was used as a medium.

The selected colony was treated with trypsin, followed by
10 sub-culture with a 24-well plate. 3-4 days later, when the cells were confluent, sub-culture was performed again with a 12-well plate. Sub-culture was performed phase by phase with a 6-well, a 60 mm dish and a 100 mm dish, by the same manner as described above, and the cells were frozen for preservation.

15 To confirm the transfection, the cells introduced with GFP gene were analyzed by FACS (BECKMAN COULTER-cytomics FC500) at 48 hours after the transfection. When the gene was introduced into porcine somatic cells by the above method, the transfection efficiency was 44.2%. And it was confirmed that
20 the GFP gene was expressed 24 hours after the transfection.

The results are shown in Figure 15.

Figure 15 is a graph illustrating the introduction efficiency analyzed by FACS and the cells introduced with the gene show GFP fluorescence.

Example 15: Analysis of cells transfected with human basic FGF knock-in vector

Identification of knocked-in cells was performed by the following two steps. First step is the identification by PCR.

5 Precisely, the colony showing resistance against G418 was selected. When the cells were sub-cultured from 24-well plate to 12-well plate, 1/5 of the cells proceeded to PCR. Cell suspension of 1/5 of the cells was centrifuged at 10,000 rpm at 4°C for 10 minutes and then the supernatant was discarded.

10 40 μ l of deionized water was added thereto, followed by vortexing and heating at 100°C for 10 minutes. 1 μ l of proteinase K (10 mg/ml) was added thereto, followed by incubation at 55°C for 130 minutes and then at 100°C for 10 minutes. 20 μ l of the total cell extract was used for PCR. PCR
15 was performed with sense (neo) (GCCTGCTTGCCGAATATCATGGTGGAAAAT) and antisense (SEQ. ID. NO: 21: GGCATGTGGGGAAATAATTGCACATAAGGA) primers as follows; predenaturation at 94°C for 2 minutes, denaturation at 94°C for 30 seconds, annealing at 66°C for 30 seconds, and extension at
20 72°C for 5 minutes, 35 cycles from denaturation to extension, and final extension at 72°C for 10 minutes. The PCR product was confirmed by electrophoresis on 0.8% agarose gel. As a result, approximately 3.2 kb sized band was confirmed. And the results are shown in Figure 17A.

Second step is the identification by Southern blotting.

The cells confirmed by PCR above were sub-cultured and then genomic DNA was extracted, followed by Southern blotting.

10 μ g of the genomic DNA was digested with *EcoRV*,
5 followed by purification. The DNA was electrophoresed on 0.8% agarose gel, and transferred on nitrocellulose membrane. The membrane was washed with 2×SSC buffer for 30 seconds. The DNA was cross-linked onto the membrane by using a UV crosslinker, followed by hybridization in the solution containing 5×SSPE,
10 5×Denhardt's, 1% SDS (w/v) and 50% formamide (w/v) at 42°C for 15 hours. A probe was prepared by using 600 bp/ *SpeI-EcoRI* DNA fragment containing porcine β -casein exon 9, random labeling kit (Amersham) and [α -32P]dCTP (110TBq/mmol, Amersham). The concentration of the probe for hybridization was one million
15 cpm/ml hybridization solution.

After hybridization, the membrane was washed three times with 0.2% SSC and 0.1% SDS (w/v) at 68°C for 30 minutes, exposed on X-ray film at -80°C for 72 hours and then developed.

As shown in Figure 17B, the wild type allele was
20 confirmed by 11.6 kb band and the knock-in allele was confirmed by 6.1 kb band.

Table 3 illustrates that introduction of the knock-in vector into porcine ear somatic cells and the confirmation of the knocked-in somatic cells.

【Table 3】

Exp.	No. of cells transfected	No. of G418 ^R colonies	No. of G418 ^R colonies analyzed by PCR	No. of PCR – Positive colonies	No. of Southern blot-positive colonies
1	5 x 10 ⁶ cells	436	387	5	2

【Sequence List Text】

SEQ. ID. NO: 1 is the nucleotide sequence of porcine β -
5 casein genomic DNA.

SEQ. ID. NO: 2 is the nucleotide sequence of human FGF
cDNA.

SEQ. ID. NO: 3 is the nucleotide sequence of hFGF knock-
in vector.

10 SEQ. ID. NO: 4 is the nucleotide sequence of GFP knock-in
vector.

SEQ. ID. NO: 5 is the artificial sequence of the sense
primer for the identification of porcine β -casein cDNA.

15 SEQ. ID. NO: 6 is the artificial sequence of the
antisense primer for the identification of porcine β -casein
cDNA.

SEQ. ID. NO: 7 is the artificial sequence of the sense
primer of the porcine β -casein promoter upstream.

20 SEQ. ID. NO: 8 is the artificial sequence of the
antisense primer containing porcine β -casein exon 6.

SEQ. ID. NO: 9 is the artificial sequence of the sense

primer containing porcine β -casein exon 2.

SEQ. ID. NO: 10 is the artificial sequence of the antisense primer containing porcine β -casein exon 9.

5 SEQ. ID. NO: 11 is the artificial sequence of the sense primer for the identification of human FGF gene.

SEQ. ID. NO: 12 is the artificial sequence of the antisense primer for the identification of human FGF gene.

SEQ. ID. NO: 13 is the artificial sequence of the sense primer for the cloning of GFP gene.

10 SEQ. ID. NO: 14 is the artificial sequence of the antisense primer for the cloning of GFP gene.

SEQ. ID. NO: 15 is the artificial sequence of the sense primer for the cloning of NLS gene.

15 SEQ. ID. NO: 16 is the artificial sequence of the antisense primer for the cloning of NLS gene.

SEQ. ID. NO: 17 is the artificial sequence of the sense primer for the cloning of SV40 polyA gene.

SEQ. ID. NO: 18 is the artificial sequence of the antisense primer for the cloning of SV40 polyA gene.

20 SEQ. ID. NO: 19 is the artificial sequence of the sense primer for the examination of the expression of hFGF knock-in vector.

SEQ. ID. NO: 20 is the artificial sequence of the antisense primer for the examination of the expression of hFGF

knock-in vector.

SEQ. ID. NO: 21 is the artificial sequence of the antisense primer for the confirmation of the introduction of hFGF knock-in vector in porcine somatic cells.

5

Those skilled in the art will appreciate that the conceptions and specific embodiments disclosed in the foregoing description may be readily utilized as a basis for modifying or designing other embodiments for carrying out the same purposes of the present invention. Those skilled in the art will also appreciate that such equivalent embodiments do not depart from the spirit and scope of the invention as set forth in the appended claims.

10

【CLAIMS】**【Claim 1】**

A method for preparing porcine somatic cells targeted with knock-in vector using porcine β -casein gene comprising
5 the following steps:

(1) transforming porcine somatic cells with knock-in vector using porcine β -casein gene;

(2) inducing homologous recombination by culturing the porcine somatic cells; and

10 (3) selecting the porcine somatic cells in which knock-in vector using porcine β -casein gene is successfully targeted by homologous recombination.

【Claim 2】

15 The method for preparing porcine somatic cells targeted with knock-in vector using porcine β -casein gene according to claim 1, wherein the porcine β -casein gene targeting knock-in vector is to produce a bioactive substance which is composed of nls nucleotide sequence, 5'end fragment containing 2.65 kb
20 promoter, exon 1 and intron 1 of porcine β -casein genomic DNA as 5' arm; a gene encoding a bioactive substance; and 3'end fragment containing exons 5, 6 and 7 of porcine β -casein genomic DNA as 3' arm.

【Claim 3】

The method for preparing porcine somatic cells targeted with knock-in vector using porcine β -casein gene according to claim 2, wherein the bioactive substance is human basic
5 fibroblast growth factor (bFGF) or green fluorescent protein (GFP).

【Claim 4】

The method for preparing porcine somatic cells targeted
10 with knock-in vector using porcine β -casein gene according to claim 1, wherein the porcine somatic cells are originated from porcine ear cells.

【Claim 5】

15 The method for preparing porcine somatic cells targeted with knock-in vector using porcine β -casein gene according to claim 4, wherein the ear cells are fibroblasts.

【Claim 6】

20 The method for preparing porcine somatic cells targeted with knock-in vector using porcine β -casein gene according to claim 1, wherein the number of somatic cells of step (1) is 4×10^6 cell/0.4ml - 6×10^6 cell/0.4ml.

【Claim 7】

The method for preparing porcine somatic cells targeted with knock-in vector using porcine β -casein gene according to claim 1, wherein the transformation of porcine somatic cells
5 are induced by electroporation.

【Claim 8】

The method for preparing porcine somatic cells targeted with knock-in vector using porcine β -casein gene according to
10 claim 1, wherein the transformation of step (1) is induced by using the linear vector or the linear vector in which plasmid backbone is eliminated.

【Claim 9】

15 A method for generating a transgenic pig comprising the following steps:

(1) transforming porcine somatic cells with the knock-in vector using porcine β -casein gene;

(2) inducing homologous recombination by culturing the
20 porcine somatic cells;

(3) selecting the porcine somatic cells in which knock-in vector using porcine β -casein gene is successfully targeted by homologous recombination;

(4) eliminating the nucleus of a porcine egg and

introducing a gene targeted cells therein to prepare somatic cells nuclear transfer embryo; and

(5) implanting the embryo.

5 **【Claim 10】**

A method for obtaining a target protein from the transgenic pig prepared by the method of claim 9.

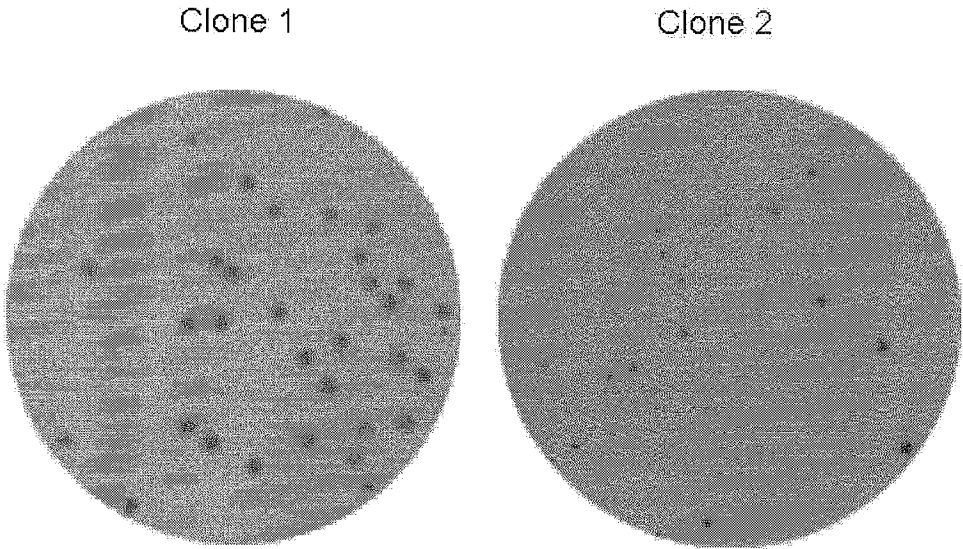
【Claim 11】

10 The method for obtaining a target protein according to claim 10, wherein the target protein is a bioactive substance.

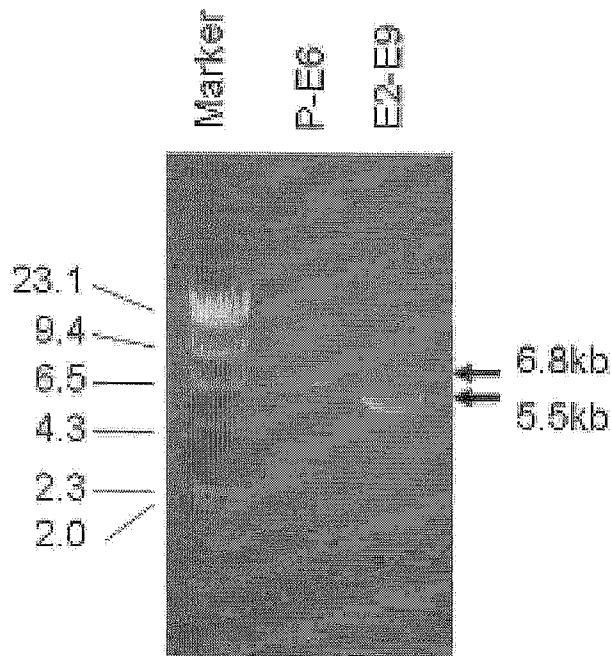
[DRAWINGS]
[Figure 1]



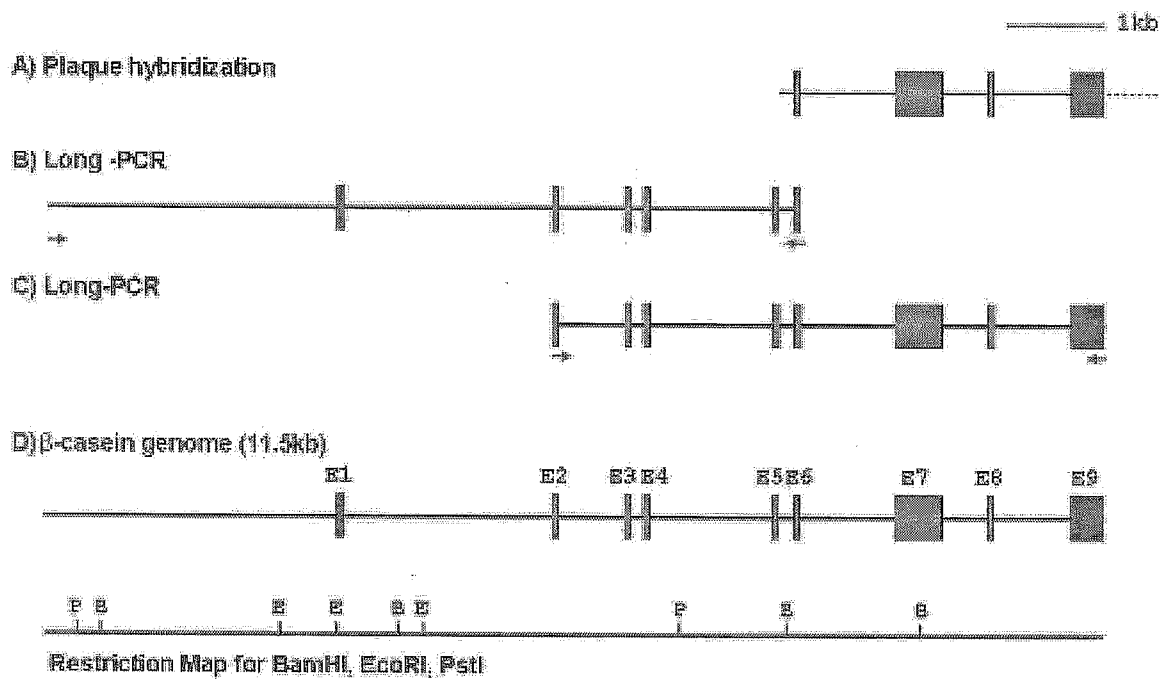
[Figure 2]



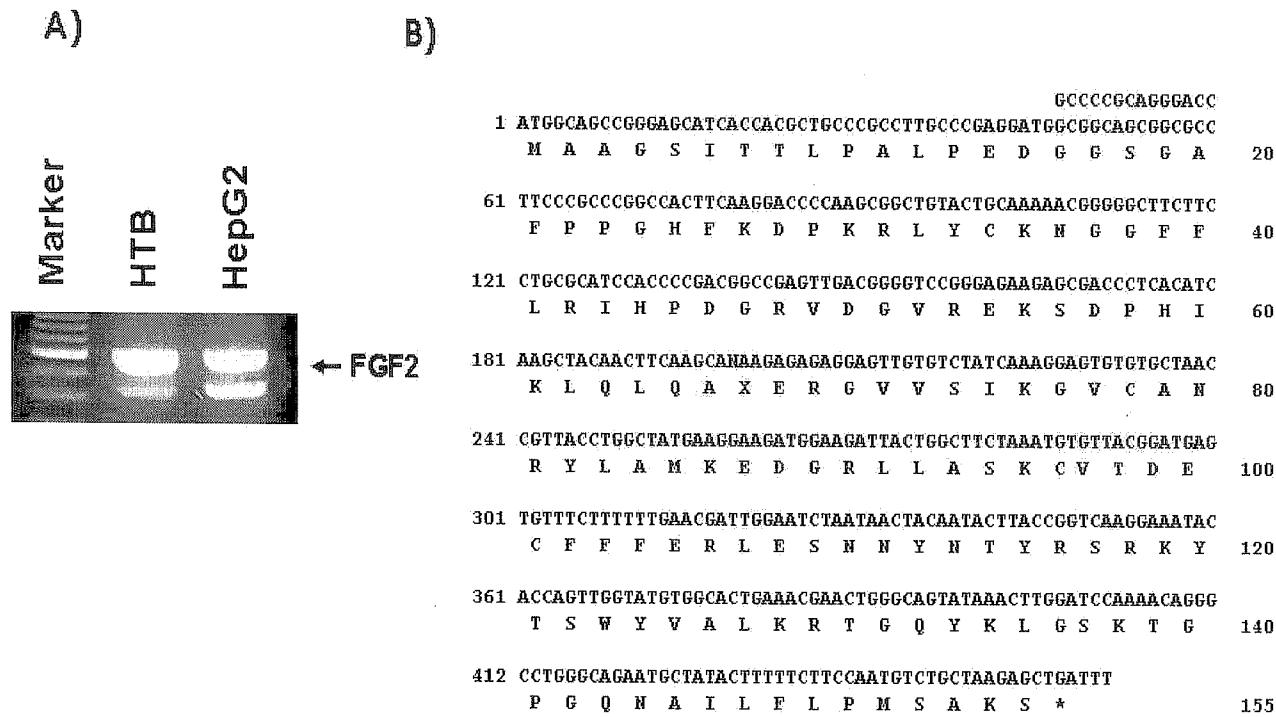
【Figure 3】



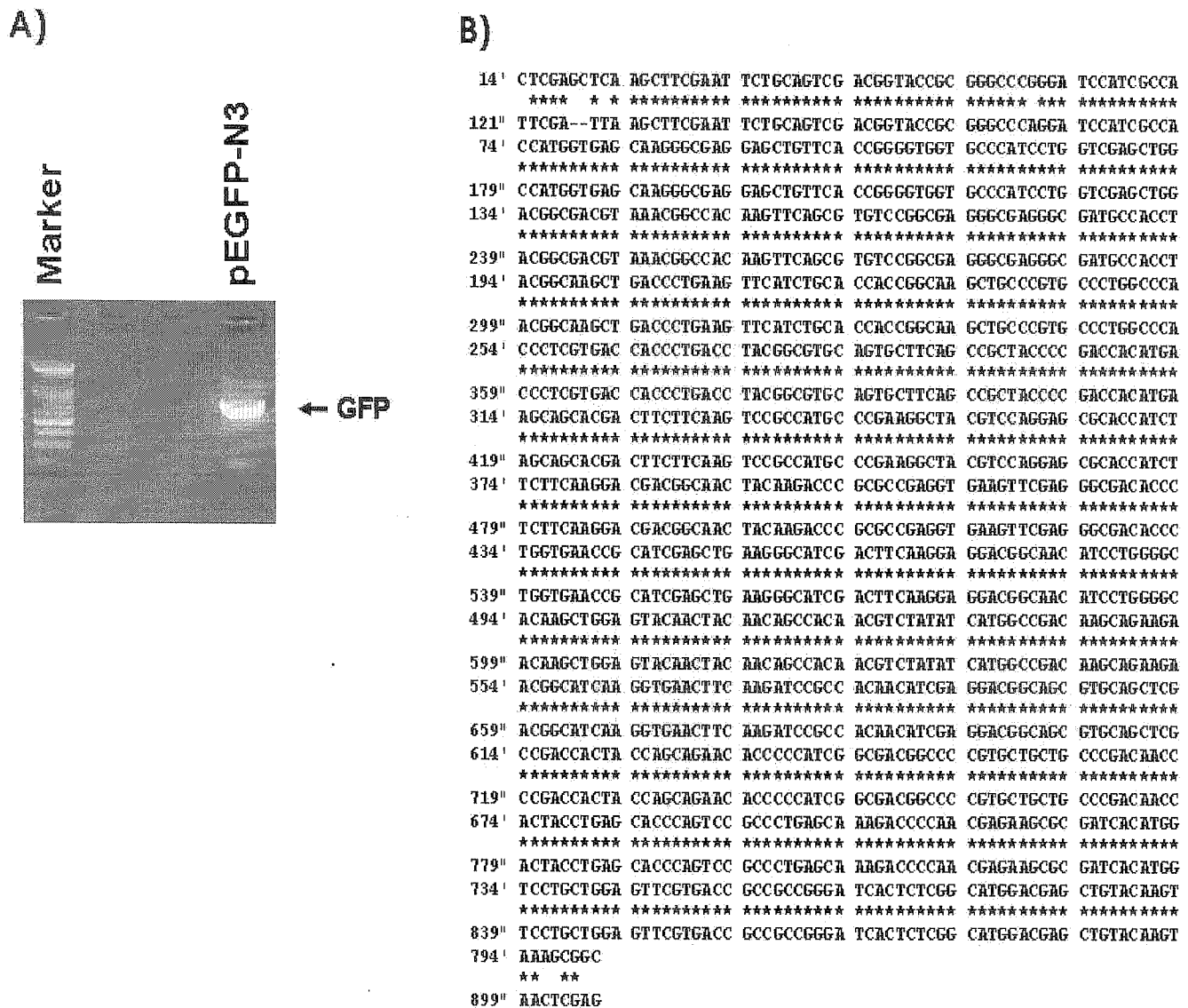
【Figure 4】



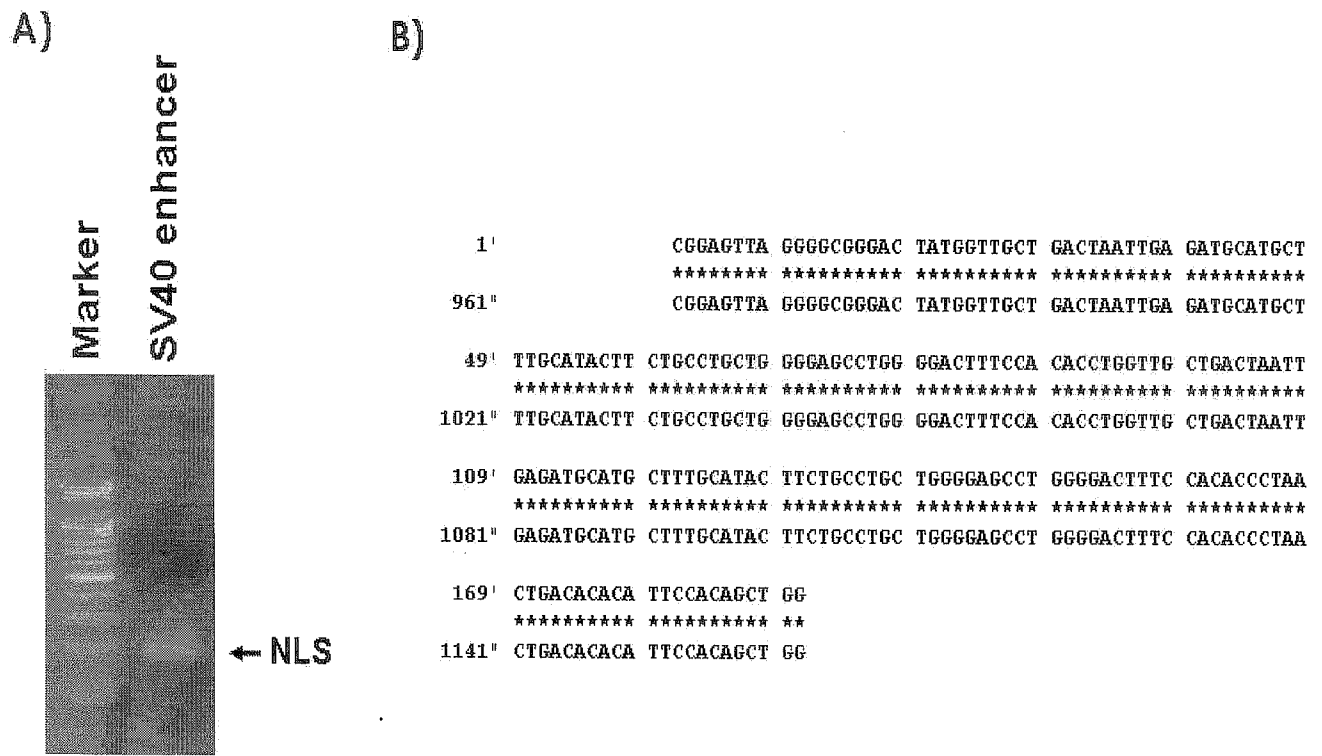
【Figure 5】



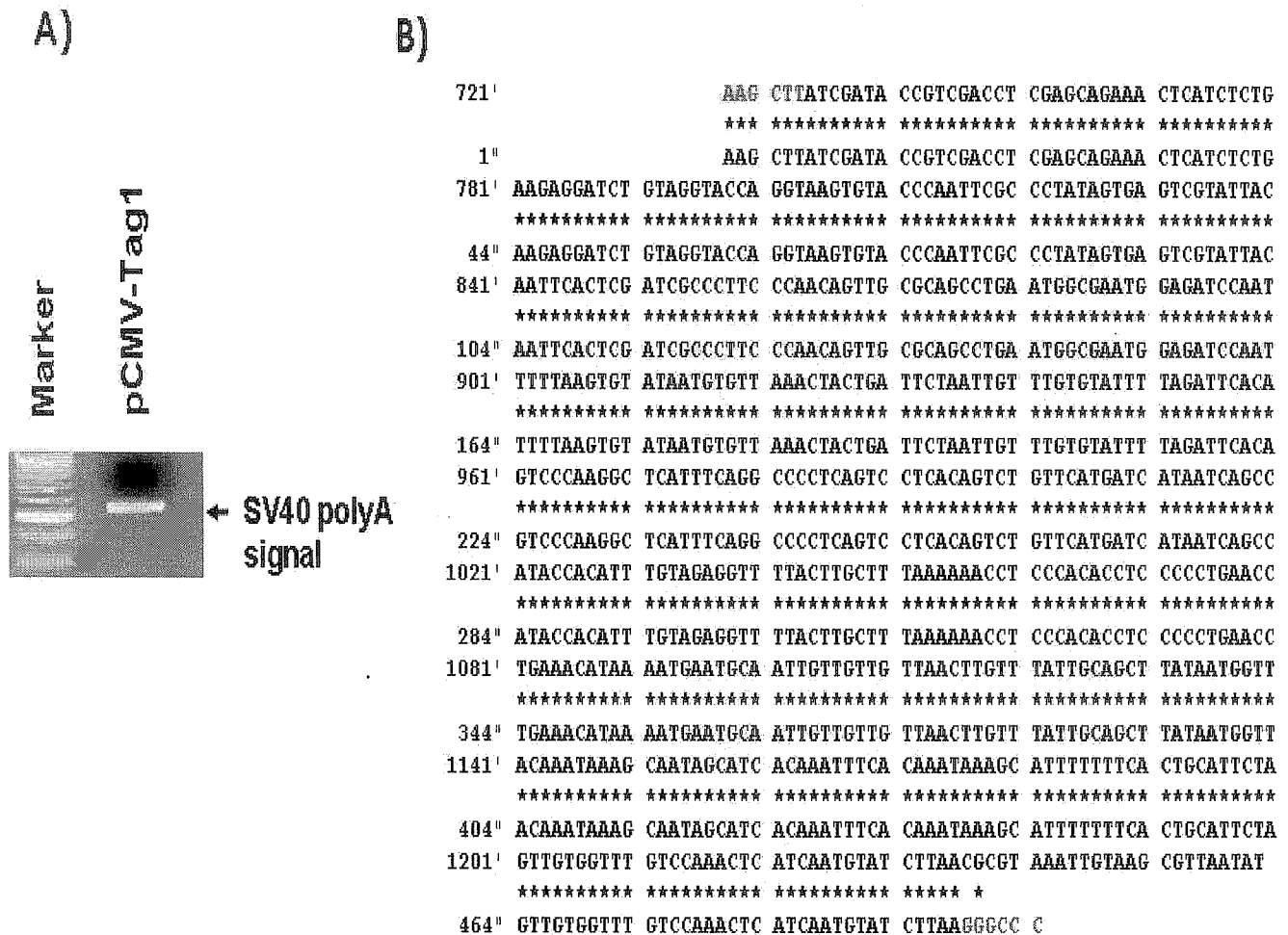
【Figure 6】



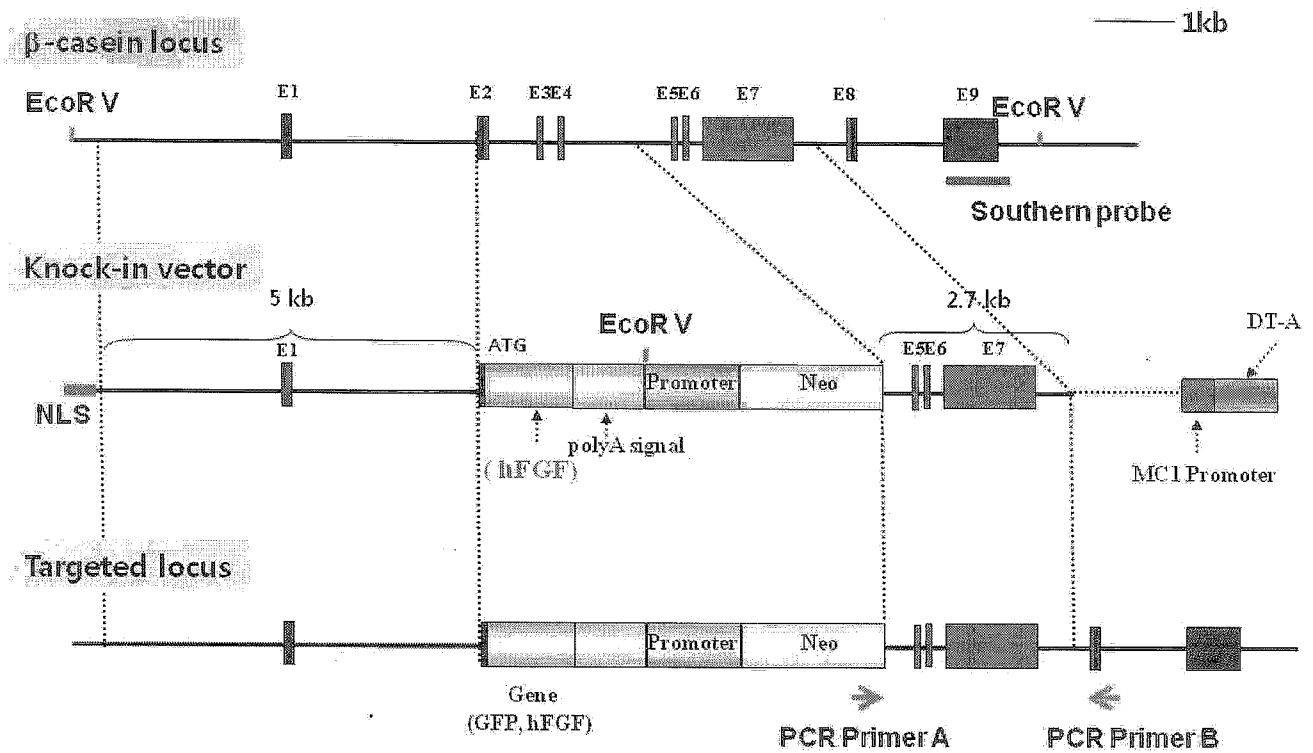
【Figure 7】



【Figure 8】

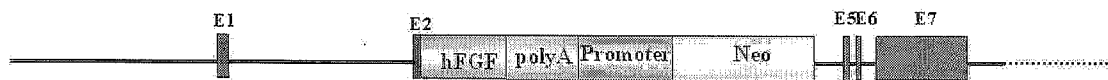


【Figure 9】

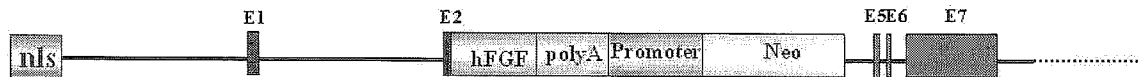


【Figure 10】

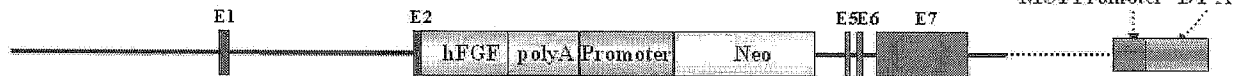
1) 5'-hFGF-polyA-Pneo-3' (hFGF-DT vector)



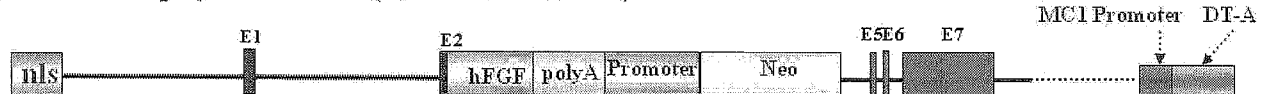
2) nls-5'-hFGF-polyA-Pneo-3' (nls+hFGF-DT vector)



3) 5'-hFGF-polyA-Pneo-3'-DT (hFGF+DT vector)



4) nls-5'-hFGF-polyA-Pneo-3'-DT (nls+hFGF+DT vector)

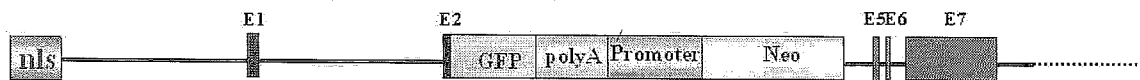


【Figure 11】

1) 5'-GFP-polyA-Pneo-3' (GFP-DT vector)



2) nls-5'-GFP-polyA-Pneo-3' (nls+GFP-DT vector)



3) 5'-GFP-polyA-Pneo-3'-DT (GFP+DT vector)

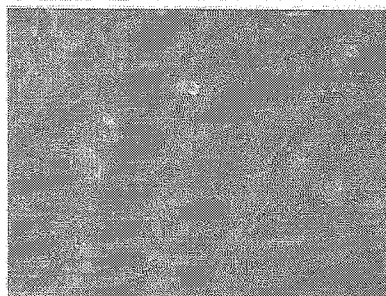


4) nls-5'-GFP-polyA-Pneo-3'-DT (nls+GFP+DT vector)

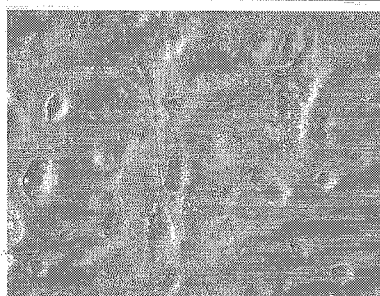
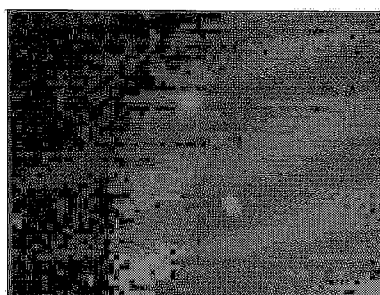


【Figure 12】

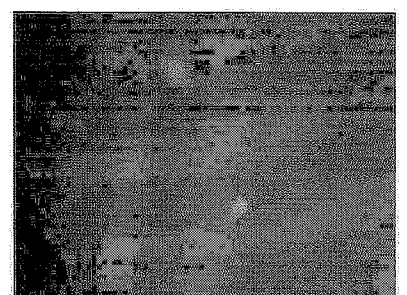
pEGFP-N3 vector



Knock-in GFP vector (Ex 1)



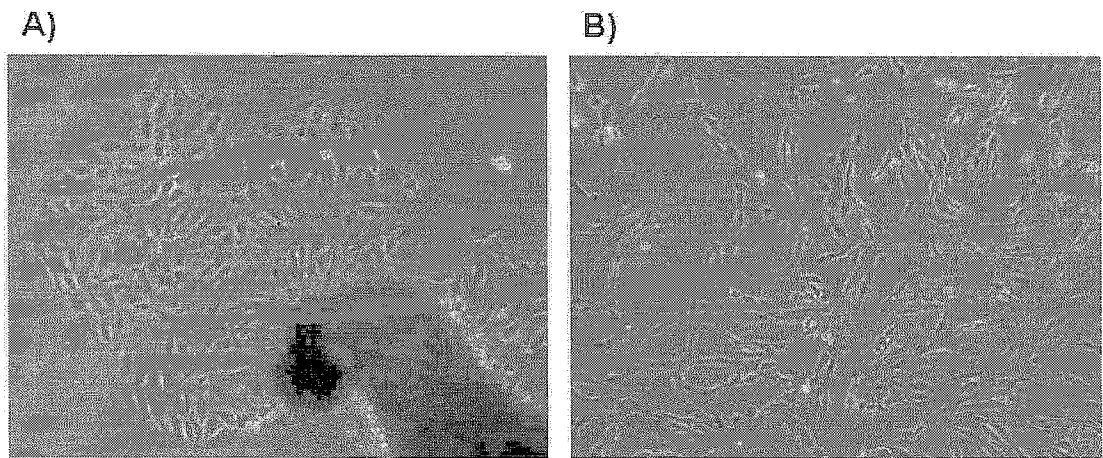
Knock-in GFP vector (Ex 2)



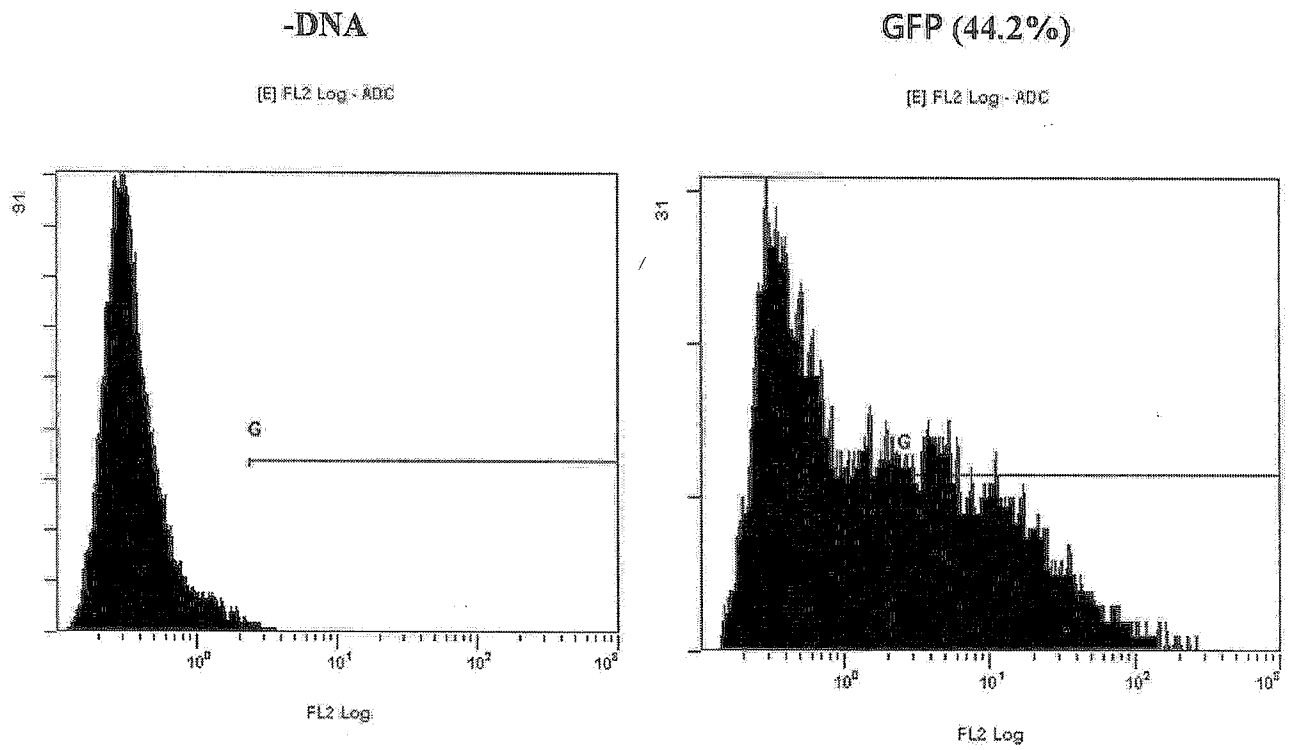
【Figure 13】



【Figure 14】

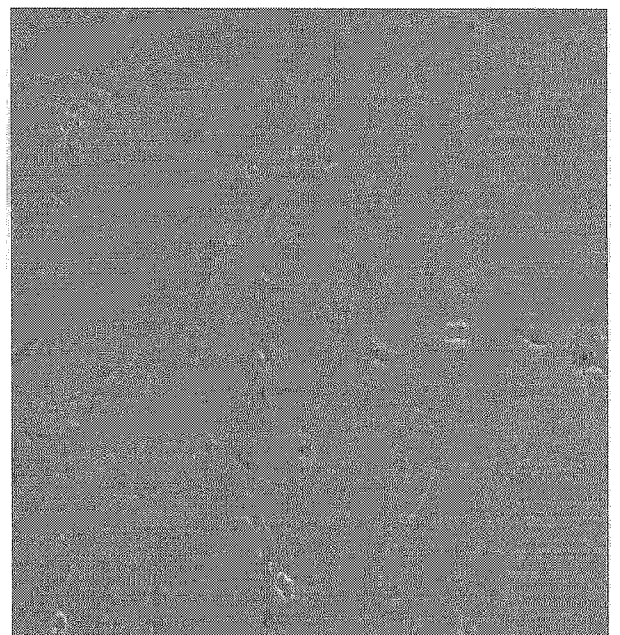


【Figure 15】

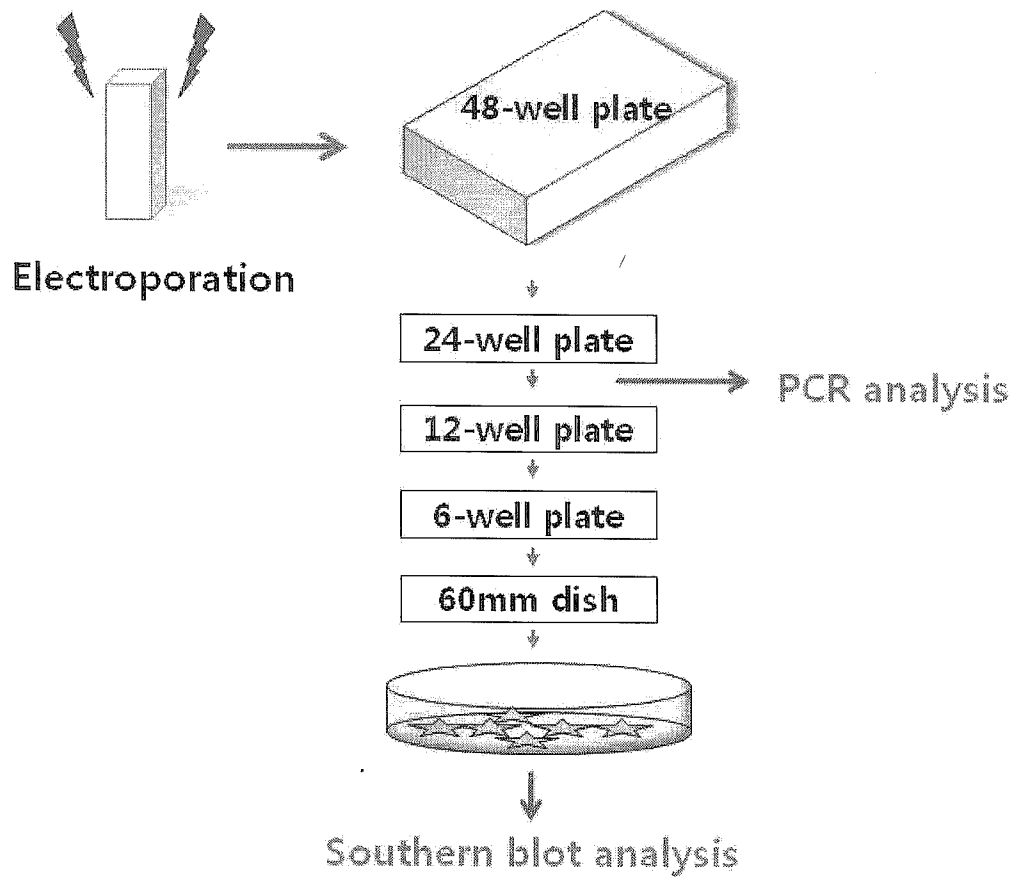


A)

B)

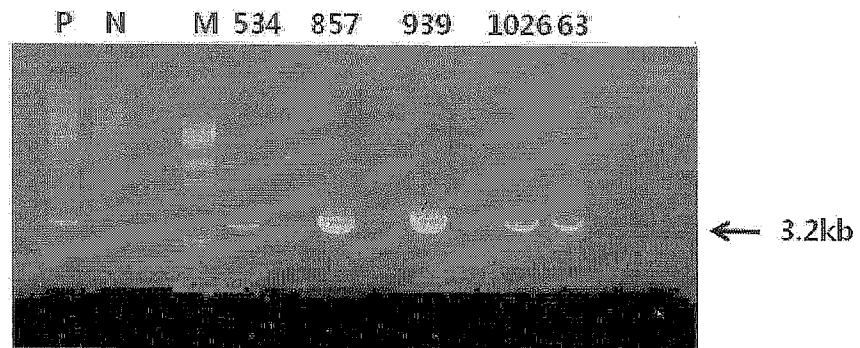


【Figure 16】

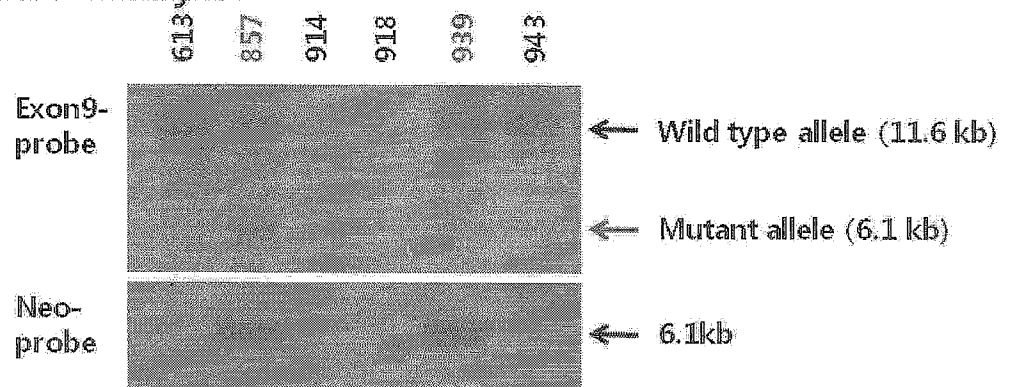


【Figure 17】

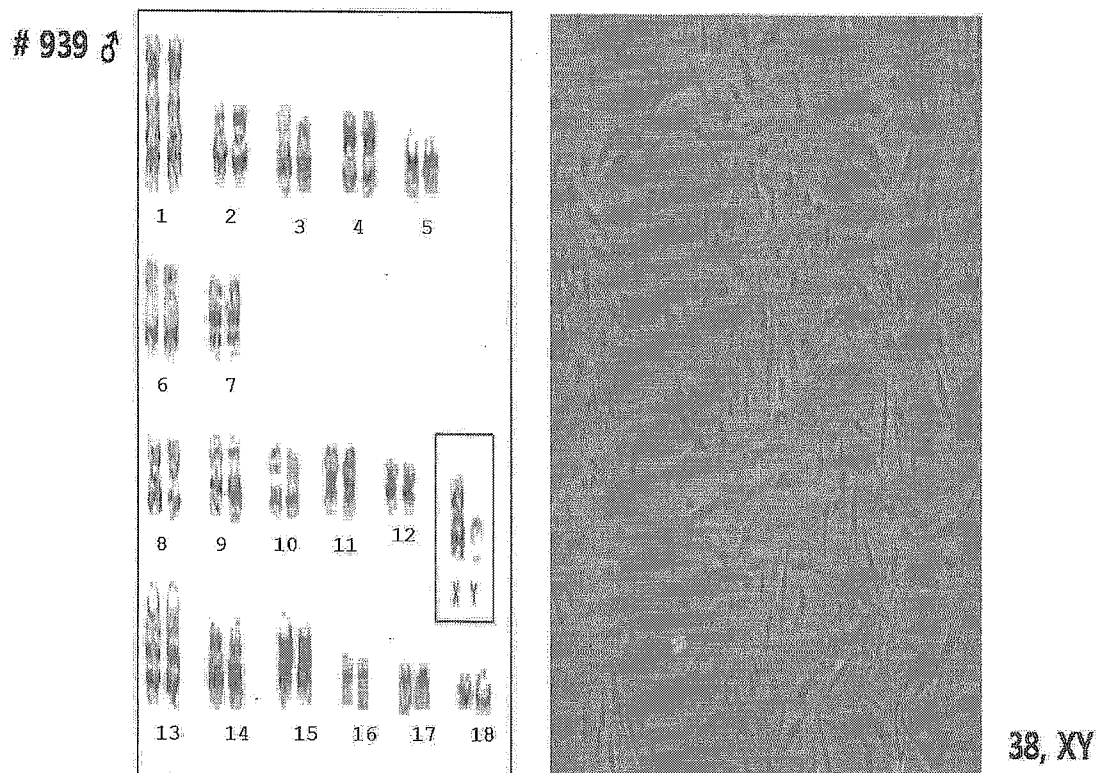
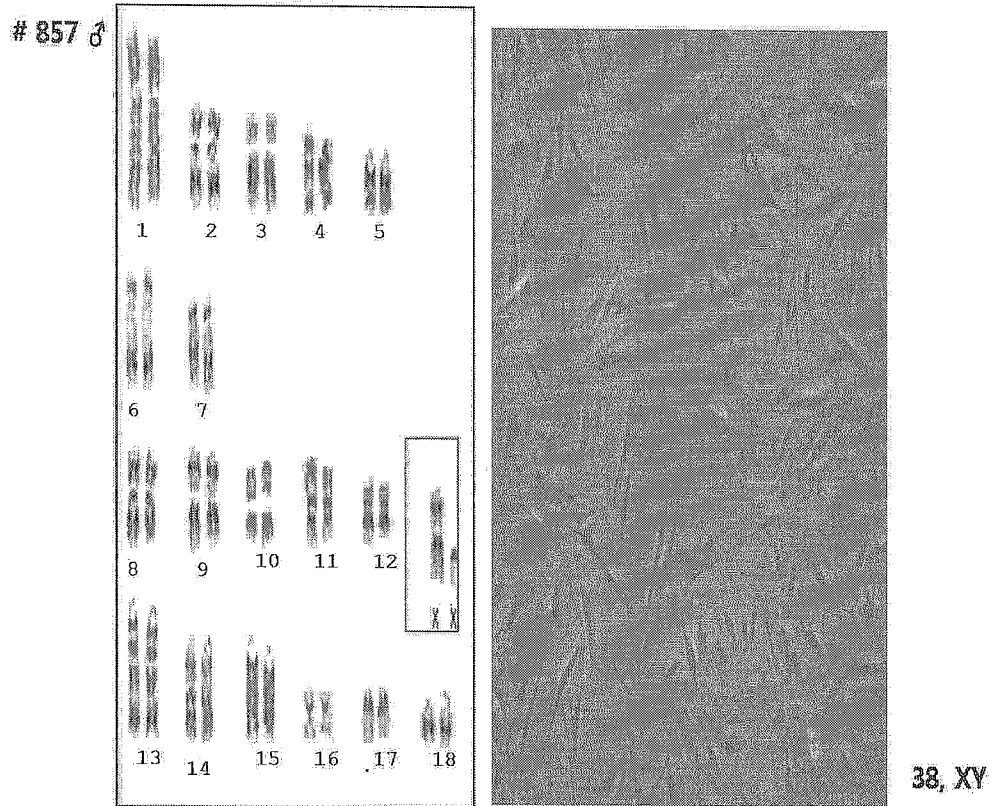
A) PCR analysis



B) Southern blot analysis



【Figure 18】



A. CLASSIFICATION OF SUBJECT MATTER*C12N 15/06(2006.01)i*

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC C12N 15/06, 15/85, 15/63

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

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

eKIPASS(KIPO internal), PubMed(NCBI), JPO, USPTO, Google

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Sang Mi Lee et al. 'Construction and Expression Analysis of Knock-in Vector for EGFP Expression in the Porcine beta-casein Gene Locus.' Repro Dev Biol. Vol. 32, No. 3, pp. 205-209 (September 30, 2008). See abstract, Abstract, Pages 206-207(Material and Methods), 208-209(Discussion), Fig 2 and 3.	1-11
A	Poongyeon Lee et al. 'Cloning and characterization of 5'-untranslated region of porcine beta casein gene (CSN2).' Domest Anim Endocrinol. Vol. 35, No. 3, pp. 245-253 (online publication, May 21, 2008). See Abstract, Page 247	1-11
A	Van Cott KE et al. 'Recombinant human protein C expression in the milk of transgenic pigs and the effect on endogenous milk immunoglobulin and transferrin levels.' Transgenic Res. Vol. 10, pp. 43-51., (2001) See Abstract	1-11
A	KR 10-0646222 B1 (REPUBLIC OF KOREA(MANAGEMENT : RURAL DEVELOPMENT ADMINISTRATION)) November 11, 2006. See Abstract, Fig. 1	1-11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

08 JULY 2009 (08.07.2009)

Date of mailing of the international search report

08 JULY 2009 (08.07.2009)

Name and mailing address of the ISA/KR

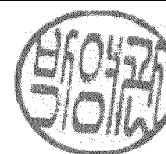
Korean Intellectual Property Office
Government Complex-Daejeon, 139 Seonsa-ro, Seo-
gu, Daejeon 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

PARK, Yeong-Gwan

Telephone No. 82-42-481-8407



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2008/006591

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

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

a. type of material



a sequence listing



table(s) related to the sequence listing

b. format of material



on paper



in electronic form

c. time of filing/furnishing



contained in the international application as filed



filed together with the international application in electronic form



furnished subsequently to this Authority for the purposes of search

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2008/006591Patent document
cited in search reportPublication
datePatent family
member(s)Publication
date

KR 10-0646222 B1

23.11.2006

(None)