



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
24.10.2001 Bulletin 2001/43

(51) Int Cl.7: **C12N 15/12, C07K 14/47,
C12Q 1/68, A61K 48/00,
A61P 27/00, A01K 67/027**

(21) Application number: **01201037.7**

(22) Date of filing: **20.03.2001**

(84) Designated Contracting States:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE TR**
Designated Extension States:
AL LT LV MK RO SI

(74) Representative: **Prins, Adrianus Willem et al
Vereenigde,
Nieuwe Parklaan 97
2587 BN Den Haag (NL)**

(30) Priority: **20.03.2000 US 528928**

Remarks:

The sequence listing, which is published as annex to the application documents, was filed after the date of filing. The applicant has declared that it does not include matter which goes beyond the content of the application as filed.

(71) Applicant: **UNIVERSITY OF HONG KONG
Hong Kong (CN)**

(72) Inventor: **Cheah, Kathryn S.E.
Hong Kong (CN)**

(54) **Gene locus involved in regulating hair pigmentation, vestibular function and fertility**

(57) This invention provides an isolated nucleic acid which defines a yellow submarine locus. This invention provides an isolated nucleic acid which defines a mutant yellow submarine locus, wherein the mutant yellow submarine locus is identical to a wildtype yellow submarine locus except for an integration of a pAA2 transgene into

at least one region on a chromosome. This invention provides an isolated nucleic acid which defines a human locus which corresponds to a yellow submarine locus. This invention provides an isolated nucleic acid which defines a human locus which corresponds to a mutant yellow submarine locus containing a pAA2 transgene integrated into at least one region the genome.

Description

[0001] This application is a continuation-in-part application of U.S. Serial No. 09/274,634, filed March 23, 1999, claiming the benefit of U.S. Provisional Application No. 60/079,020, filed March 23, 1998, the contents of which are hereby incorporated by reference into this application.

Background of the Invention

[0002] A mouse line bearing a recessive mutation, caused by insertion of a transgene, was created. These mice are characterized by a changed pigmentation of hair to give yellow coat color, circling behaviour and an inability to swim. It has been shown that the transgene has inserted into two sites on mouse chromosome 3. A search of the mouse genome database ascertained that the mutation is novel. The mutant locus has been named *yellow submarine* (*ysb*). DNA flanking the transgene insertion has been isolated and used to screen for the corresponding normal (unmutated) sequences. Genomic clones spanning 20kb of the unmutated *ysb* locus (+*ysb*) have been isolated and mapping experiments indicate that transgene integration has caused a deletion and chromosomal inversion resulting in two transgene integration sites. Circular behaviour can reflect abnormality in the inner ear or be due to an abnormality in hindbrain development. Studies show abnormal structure of the inner ears of *ysb* mice and a stunted acoustic nerve. *Ysb* mice are therefore a novel mutant showing both abnormal regulation of pigmentation and inner ear dysfunction. Molecular genetics, bioinformatics, developmental biology, transgenic and physiological approaches are used to 1) identify and characterize the gene(s) at the wild-type *ysb* (+*ysb*) locus; 2) determine the nature of the *ysb* mutation; 3) study the molecular and developmental bases underlying the defect(s) in *ysb* mice; and 4) characterize the neurophysiological changes in balance and hearing in *ysb* mice. Approximately 1/1000 infants are affected by hearing defects at birth, two thirds of which have a genetic basis. *Ysb* mice are a valuable model to identify molecules involved in controlling balance. These studies provide fundamental information on the development of the inner ear and the mechanisms by which hearing and balance defects may arise. Identification and characterization of the *ysb* gene(s) and the molecular defect in *ysb* mice also yield new insight into the complex regulatory pathways controlling agouti coat color in mice.

Summary of the Invention

[0003] This invention provides an isolated nucleic acid which defines a yellow submarine locus.

[0004] This invention provides an isolated nucleic acid which defines a mutant yellow submarine locus, wherein the mutant yellow submarine locus is identical to a wildtype yellow submarine locus except for an integration of a pAA2 transgene into at least one region on a chromosome.

[0005] This invention provides the above isolated nucleic acid further comprising a rearrangement of chromosomal sequences of a region of the chromosome designated A3 region.

[0006] This invention provides a replicable vector comprising the above nucleic acid. This invention provides a host cell comprising the vector.

[0007] This invention provides a nucleic acid of at least 14 nucleotides capable of specifically hybridizing with the nucleic acids of the subject invention.

[0008] This invention provides an isolated nucleic acid which defines a human locus which corresponds to a yellow submarine locus. This invention provides an isolated nucleic acid which defines a human locus which corresponds to a mutant yellow submarine locus containing a pAA2 transgene integrated into at least one region the genome.

[0009] This invention provides a method of diagnosing inner ear dysfunction in a subject comprising: a) obtaining a suitable sample from a subject; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to the mutated portion of the *ysb* locus; and d) detecting the labeled nucleic acid, thereby detecting mutant *ysb* and thereby diagnosing inner ear dysfunction in the subject.

[0010] This invention provides a method of diagnosing pigmentation dysfunction in a subject comprising: a) obtaining a suitable sample from a subject; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to the mutated portion of the *ysb* locus; and d) detecting the labeled nucleic acid, thereby detecting mutant *ysb* and thereby diagnosing pigmentation dysfunction in the subject.

[0011] This invention provides a method of diagnosing cell growth dysfunction, cell proliferation dysfunction, or cell death dysfunction in a subject comprising: a) obtaining a suitable sample from a subject; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to the mutated portion of the *ysb* locus; and d) detecting the labeled nucleic acid, thereby detecting mutant *ysb* and thereby diagnosing cell growth dysfunction or proliferation dysfunction in the subject.

[0012] The invention provides a polypeptide encoded by the nucleic acid of the subject invention.

[0013] The invention provides a method of repairing and regenerating nerve tissue in a subject comprising administering an effective amount of the above protein to the subject, wherein the protein is a wildtype protein, so as to

thereby repair and regenerate nerve tissue in the subject.

[0014] The invention provides a method of regulating cell migration in a subject comprising administering an effective amount of the above protein to the subject, wherein the protein is a wildtype protein, so as to thereby regulate cell migration in the subject.

[0015] The invention provides a method of regulating cell growth in a subject comprising administering an effective amount of the above protein to the subject, wherein the protein is a wildtype protein, so as to thereby regulate cell growth in the subject.

[0016] The invention provides an antibody which binds to the polypeptide of the subject invention. The invention provides a composition comprising the antibody of the subject invention.

[0017] The invention provides a method of producing a protein encoded by a nucleic acid in a wildtype *ysb* locus which comprises growing a host vector system under suitable conditions permitting production of the polypeptide and recovering the polypeptide so produced.

[0018] The invention provides a method determining vestibular dysfunction in a subject comprising: a) obtaining a suitable sample from a subject; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to a mutated portion of the *ysb* locus; d) detecting the labeled nucleic acid, thereby detecting gene responsible for yellow coat color, thereby indicating the presence of vestibular dysfunction in the subject.

[0019] The invention provides a method determining vestibular dysfunction using embryonal stem cells: comprising a) obtaining a suitable sample from the embryonal stem cells; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to a mutated portion of the *ysb* locus; d) detecting the labeled nucleic acid, thereby detecting gene responsible for yellow coat color, thereby indicating the presence of vestibular dysfunction in the embryonal stem cells.

[0020] The invention provides a method of determining successful deletion or inactivation of a gene which comprises examining coat color or pigmentation of a subject, wherein yellow coat color or pigmentation indicates that a normal gene has been inactivated, thereby resulting in yellow coat color or pigmentation, thereby indicating the successful deletion or inactivation of the gene.

[0021] The invention provides a method of determining successful deletion or inactivation of a gene comprising determining whether repairing the abnormal gene results in the disappearance of the yellow coat color or pigmentation.

[0022] The invention provides a method of treating vestibular dysfunction in a subject comprising introducing a nucleic acid comprising a gene or genes from the wildtype locus encoding a *ysb* into a suitable cell under conditions such that the nucleic acid expresses *ysb* so as to thereby treat vestibular dysfunction.

[0023] The invention provides a method of treating hearing impairment in a subject which comprises introducing a nucleic acid comprising a gene or genes from the wild-type locus into a suitable cell under conditions such that the nucleic acid expresses *ysb* so as to thereby treat hearing loss.

[0024] The invention provides a transgenic mouse line designated KM12.

[0025] The invention provides a method of determining whether a compound is mutagenic comprising: (a) examining the coat color or pigmentation of a subject; (b) administering the compound to the subject; (c) examining the coat color or pigmentation of a subject; (d) comparing the result obtained in step (c) with the result obtained in step (a); and (e) determining if the coat color or pigmentation of the subject is yellow so as to thereby determine whether the compound is mutagenic.

Brief Description of the Figures

[0026] Figure 1

(A) A homozygous *ysb* mouse showing yellow coat colour compared with (B) a heterozygous littermate, agouti and (C) a black C57BL.

[0027] Figure 2

Whole-mount X-gal staining (blue) of different pAA2 mouse embryos showing the typical *LacZ* expression pattern at different embryonic stages.

(A) 9.0 dpc embryo; (B) 12.5 dpc embryo; (C) 13.5 dpc, (D), 13.5 dpc embryo. Expression sites include branchial arch (ba), notochord (no), prevertebrae (pv), snout (sn) and digits (dg).

[0028] Figure 3

LacZ expression as seen by X-gal staining (blue) of KM12 transgenic embryos at different embryonic stages. *LacZ* is expressed in sites typical for pAA2 (Fig 2) and in extra sites including rhombomeres (r) 2,3 & 5 in hindbrain (hb), neural tube (nt), spinal cord (sc) and hair follicles (hf) which are not the normal expression sites of the *Col2a1* transgene. (A), whole-mount X-gal stained 9.5 dpc KM12 embryo shows *LacZ* expression in branchial arch (ba); heart (he); notochord (no); rhombomeres (r) 2.3 and 5; otic vesicle (ov); neural tube (nt). (B) shows sagittal section of 10.5 dpc KM12 transgenic embryo. (C), At 10.5 dpc., KM12 transgenic embryo has additional expression sites at prevertebrae (pv), forebrain (fb) and dorsal root ganglia (drg) (not shown in this view) but with the disappearance of r5 expression. (D) expression

pattern of 12.5 dpc fetus, at this stage staining is also in midbrain (mb), hindbrain (hb), spinal cord (sc). In (E), *LacZ* expression in 13.5 dpc embryos is also seen in snout (sn), digits (dg) and ribs (rb). (F) At 16.5 dpc, expression in the brain and vertebrate continue to be seen in sagittally-halved fetus. Expression is seen in hair follicles in whole-mount (G), and sectioned (H) 16.5 embryos, this continues to be present in 3.5 days postnatal skin which is whole-mount x-gal stained and cleared (I).

5 **[0029]** Figure 4

Swimming test: A), a wild-type mouse - floats and swims in water; B),C),D), and E) a homozygous yellow submarine (ysb) mouse - circles and submerges in water

[0030] Figure 5

10 Reaching response test. Mouse on left, a ysb homozygote with vestibular dysfunction, showing a tendency to curl up towards the belly when held by tail. Mouse on right, a normal reaching response, with outstretched body.

[0031] Figure 6

Southern blot hybridization of genomic DNA from mice heterozygous (he) and homozygous (ho) for the pAA2 transgene in the KM12 line and non-transgenic (nt) control littermates using (A) kreisler, (B) agouti, and (C) α -MSHR cDNA probes.

15 The structural genes of kreisler, agouti and α -MSHR have not been disrupted by COL2A1-lacZ transgene in *ysb* mice.

[0032] Figure 7

The *ysb* locus and chromosomal localisation of the pAA2 transgene. (A) Localisation of pAA2 (cheah et al, 1995) in KM12/*ysb* transgenic mice to chromosome 3 by fluorescence in situ hybridization (FISH) using pAA2 as probe. The transgene maps to A2 and B-C region on metaphase *ysb* chromosomes. (B) Chromosomal map of transgenic integration. (C) Genomic clones isolated by PCR probes for integration sites 1 and 2. (D) Chromosomal localisation of genomic clones for integration sites 1 and 2 by FISH to wild-type metaphase chromosomes. Both map to A3 on chromosomes 3. Reference: K.S.E. Cheah, A Levy, P.A. Trainor, A.W.K. Wai, T. Kuffner, C.L. So, K.K.H. Leung, R.H. Lovell-Badge and P.P.L. Tam. (1995) Human COL2A1-directed SV40 T-antigen expression in transgenic and chimeric mice results in abnormal skeletal development. *J.Cell Biol.* 128 223-237.

25 **[0033]** Figure 8

Fluorescence *in situ* hybridisation (FISH) of *ysb* chromosomes using transgene (pAA2) as probe. (A) The transgene has integrated into two insertion sites in the chromosome. Arrow shows position of fluorescence signals. (B) DAPI banding pattern shows that pAA2 insertion sites correspond to chromosome 3. (C) Corresponding bands that show fluorescence signal is shown by dots.

30 **[0034]** Figure 9

Whole-mount in situ hybridization of 9.5 dpc embryos using Krox-20 riboprobe, double-stained for β -galactosidase activity (reddish). Homozygous *ysb* embryos (E,F) and heterozygous *ysb* embryos (C,D) have normal Krox-20 expression pattern in rhombomeres (r) 3 and 5 at 9.5 dpc. Sagittal view (A), and dorsal view (B) of non-transgenic littermate showing normal pattern.

35 **[0035]** Figure 10

Whole-mount neurofilament immunostaining of 10.5 dpc non-transgenic control (A,B), heterozygous *ysb* (C,D), and homozygous *ysb* (E,F) embryos, using monoclonal antibody (2h3), double-stained for β -galactosidase activating. A reduction of the eighth nerve (VIII_n) is observed in homozygous *ysb* at 10.5 dpc.

[0036] Figure 11

40 3-D reconstructed images of inner ears of a) non-transgenic, b) heterozygous *ysb*, and c) homozygous *ysb* 16.5 dpc fetuses showing structural malformations in developing homozygous *ysb* inner ears. Note superior semicircular canal is obliterated (*) in homozygous *ysb*.

[0037] Figure 12

45 3-D reconstructed images of inner ears of 16.5 dpc. fetuses a) wild-type non-transgenic (+/+), b) heterozygous *ysb* (+/-) and c) homozygous *ysb* (-/-). Note that neuroepithelial structures are missing in developing inner ear of homozygous *ysb* fetuses. In 16.5 dpc homozygous *ysb* (-/-), superior semicircular canal is partially obliterated* and ends in a blind sac (c); the utricular macula, superior and lateral cristae are missing.

[0038] Figure 13

50 3-D reconstructed images of inner ears at 13.75 dpc showing structural malformations in developing homozygous *ysb* inner ears. Superior semicircular canal is obliterated* in homozygous *ysb*.

[0039] Figure 14

A) Sequence of PCR product at integration site 1

B) Sequence of PCR product at integration site 2

55 **[0040]** Figure 15

Southern analyses of *ysb* genomic DNA using flanking probes at integration sites 1 and 2. Different size bands are observed in the mutant alleles when probed with flanking probes from both integration sites (shown by corresponding

EP 1 148 128 A1

arrows). No segregation of the two integration sites are observed in the *ysb* mice studied (6th generation). Wt, wild-type; he, heterozygous *ysb*; ho, homozygous *ysb*

[0041] Figure 16

Deletion in *ysb* at integration site 2.

5 (A) 20 kb deletion in *ysb*. B-E; Autoradiograms of Southern blots of genomic DNA digested with EcoRI (B,D), SacI (C) and XbaI (E) from *ysb* heterozygote, homozygote and non-transgenic mice. Probes were: B), pSD14 (1.4kb EcoRI); C), pSD11 (1.9kb SacI); D), pSD9 (3.2kb BamHI) ; E) pSD8 (2.0kb Sac I)

(B) Note rearrangement band is observed when pSD14 (1.4kb EcoRI) is used to probe homozygous *ysb* genomic DNA. *signifies homozygous *ysb*

10 Note deleted sequences revealed by probes pSD11(C); pSD9 (D); and pSD8 (E)

[0042] Figure 17

Sequences information for the 8.1 kb within *ysb* locus deletion.

Locus: chromosome 3

Definition: pPL5

15 Keywords: genomic DNA

Source: house mouse

Organism: Mus Musculus

Vector: pBluescript

Inset: 8114 bp

20 Insertion sites: Hind III

Information:

A. Sequence of pPL5 containing 8114 bp DNA fragment (Hind III cut) of 20 kb deleted sequences in integration site 2

B. Sequence alignment showing pPL5 (1743 bp-1906 bp) contains sequences 90% identity to IMAGE 1196866

25 C. Sequence alignment showing pPL5 (106-395 bp) contains sequences with 100% identity to IMAGE 636095

[0043] Figure 18

Summary map of pPL5 showing location of EST-IMAGE sequences

30 Detailed Description of the Invention

[0044] This invention provides an isolated nucleic acid which defines a yellow submarine locus.

35 [0045] This invention provides an isolated nucleic acid which defines a mutant yellow submarine locus, wherein the mutant yellow submarine locus is identical to a wildtype yellow submarine locus except for an integration of a pAA2 transgene into at least one region on a chromosome.

[0046] In one embodiment, the chromosome is mouse chromosome 3. In one embodiment, the pAA2 transgene is inserted into a region of mouse chromosome 3 designated A2. In one embodiment, the pAA2 transgene is inserted into a region of mouse chromosome 3 designated B-C.

40 [0047] In one embodiment, the yellow submarine locus comprises a deletion of a nucleic acid segment having a sequence set forth in SEQ ID NO:1. See Figure 17.

[0048] In one embodiment, the yellow submarine locus comprises a deletion of a nucleic acid segment having a sequence set forth in SEQ ID NO:2. See Figure 17.

[0049] In one embodiment, the yellow submarine locus comprises a deletion of a nucleic acid segment having a sequence set forth in SEQ ID NO:3. See Figure 17.

45 [0050] This invention provides the above isolated nucleic acid further comprising a rearrangement of chromosomal sequences of a region of the chromosome designated A3 region. In one embodiment, the rearrangement comprises an inversion of a nucleotide segment.

[0051] The isolated nucleic acid of the subject invention includes genomic DNA, RNA, cDNA. The nucleic acid may be labeled with a detectable marker. The detectable marker includes but is not limited to a radioactive, a colorimetric, 50 a luminescent, or a fluorescent label.

[0052] This invention provides a replicable vector comprising the above nucleic acid. This invention provides a host cell comprising the vector. In one embodiment, the cell is a eukaryotic cell. In one embodiment, the cell is a bacterial cell. The vectors of the subject invention include but are not limited to a plasmid, cosmid, λ phage, YAC, BAC, or PAC.

55 [0053] This invention provides a nucleic acid of at least 14 nucleotides capable of specifically hybridizing with the nucleic acids of the subject invention.

[0054] In one embodiment of the invention, the nucleic acid encodes a growth factor. In another embodiment, the nucleic acid encodes a growth factor receptor. In another embodiment, the nucleic acid encodes an orphan receptor. In another embodiment, the nucleic acid encodes a signaling molecule. In another embodiment, the nucleic acid en-

codes a transcriptional regulator. In another embodiment, the nucleic acid encodes an intracellular transport protein. In another embodiment, the nucleic acid encodes a neural precursor cell which is an expressed and developmentally down-regulated 4 (NEDD4) family molecule. In another embodiment, the nucleic acid encodes a SOX protein. In another embodiment, the nucleic acid encodes a regulator of apoptosis. In another embodiment, the nucleic acid encodes a regulator of protein turnover. In another embodiment, the nucleic acid encodes a cell cycle regulator. In another embodiment, the nucleic acid encodes a calcium binding protein. In another embodiment, the nucleic acid encodes a potentiator of hormone dependent activation of transcription by progesterone or glucocorticoid receptors. In another embodiment, the nucleic acid encodes a membrane transport protein. In another embodiment, the nucleic acid encodes a co-activator of transcription.

[0055] This invention provides a nucleic acid, wherein the nucleic acid is isolated from a mouse. This invention provides an isolated nucleic acid which defines a human locus which corresponds to a yellow submarine locus. This invention provides an isolated nucleic acid which defines a human locus which corresponds to a mutant yellow submarine locus containing a pAA2 transgene integrated into at least one region the genome.

[0056] This invention provides a method of diagnosing inner ear dysfunction in a subject comprising: a) obtaining a suitable sample from a subject; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to the mutated portion of the ysb locus; and d) detecting the labeled nucleic acid, thereby detecting mutant ysb and thereby diagnosing inner ear dysfunction in the subject.

[0057] This invention provides a method of diagnosing pigmentation dysfunction in a subject comprising: a) obtaining a suitable sample from a subject; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to the mutated portion of the ysb locus; and d) detecting the labeled nucleic acid, thereby detecting mutant ysb and thereby diagnosing pigmentation dysfunction in the subject.

[0058] This invention provides a method of diagnosing cell growth dysfunction, cell proliferation dysfunction, or cell death dysfunction in a subject comprising: a) obtaining a suitable sample from a subject; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to the mutated portion of the ysb locus; and d) detecting the labeled nucleic acid, thereby detecting mutant ysb and thereby diagnosing cell growth dysfunction or proliferation dysfunction in the subject.

[0059] In one embodiment of the above method, the subject is mammal. In another embodiment, the subject is a non-mammal. The subject may be a human, a primate, an equine, an opine, an avian, a bovine, a porcine, a canine, a feline or a murine. The subject may be a vertebrate.

[0060] The invention provides a polypeptide encoded by the nucleic acid of the subject invention. This invention provides a fusion protein comprising the above polypeptide. In one embodiment, the polypeptide is labeled with a detectable marker.

[0061] The invention provides a method of repairing and regenerating nerve tissue in a subject comprising administering an effective amount of the above protein to the subject, wherein the protein is a wildtype protein, so as to thereby repair and regenerate nerve tissue in the subject.

[0062] The invention provides a method of regulating cell migration in a subject comprising administering an effective amount of the above protein to the subject, wherein the protein is a wildtype protein, so as to thereby regulate cell migration in the subject.

[0063] The invention provides a method of regulating cell growth in a subject comprising administering an effective amount of the above protein to the subject, wherein the protein is a wildtype protein, so as to thereby regulate cell growth in the subject.

[0064] The invention provides an antibody which binds to the polypeptide of the subject invention. In one embodiment, the antibody is a monoclonal antibody. In one embodiment, the antibody is a polyclonal antibody. In one embodiment, the antibody is conjugated to a cytotoxic agent. In one embodiment, the antibody is labeled with a detectable marker.

[0065] The invention provides a composition comprising the antibody of the subject invention.

[0066] The invention provides a method of producing a protein encoded by a nucleic acid in a wildtype ysb locus which comprises growing a host vector system under suitable conditions permitting production of the polypeptide and recovering the polypeptide so produced.

[0067] The invention provides a method determining vestibular dysfunction in a subject comprising: a) obtaining a suitable sample from a subject; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to a mutated portion of the ysb locus; d) detecting the labeled nucleic acid, thereby detecting gene responsible for yellow coat color, thereby indicating the presence of vestibular dysfunction in the subject.

[0068] The invention provides a method determining vestibular dysfunction using embryonal stem cells: comprising a) obtaining a suitable sample from the embryonal stem cells; b) extracting nucleic acid from the sample; c) contacting the nucleic acid with the above nucleic acid which binds specifically to a mutated portion of the ysb locus; d) detecting the labeled nucleic acid, thereby detecting gene responsible for yellow coat color, thereby indicating the presence of vestibular dysfunction in the embryonal stem cells.

[0069] The invention provides a method of determining successful deletion or inactivation of a gene which comprises

examining coat color or pigmentation of a subject, wherein yellow coat color or pigmentation indicates that a normal gene has been inactivated, thereby resulting in yellow coat color or pigmentation, thereby indicating the successful deletion or inactivation of the gene.

[0070] The invention provides a method of determining successful deletion or inactivation of a gene comprising determining whether repairing the abnormal gene results in the disappearance of the yellow coat color or pigmentation.

[0071] The invention provides a method of treating vestibular dysfunction in a subject comprising introducing a nucleic acid comprising a gene or genes from the wildtype locus encoding a *ysb* into a suitable cell under conditions such that the nucleic acid expresses *ysb* so as to thereby treat vestibular dysfunction.

[0072] The invention provides a method of treating hearing impairment in a subject which comprises introducing a nucleic acid comprising a gene or genes from the wild-type locus into a suitable cell under conditions such that the nucleic acid expresses *ysb* so as to thereby treat hearing loss.

[0073] The invention provides a transgenic mouse line designated KM12.

[0074] The invention provides a method of determining whether a compound is mutagenic comprising: (a) examining the coat color or pigmentation of a subject; (b) administering the compound to the subject; (c) examining the coat color or pigmentation of a subject; (d) comparing the result obtained in step (c) with the result obtained in step (a); and (e) determining if the coat color or pigmentation of the subject is yellow so as to thereby determine whether the compound is mutagenic.

Experimental Details

[0075] A mouse line (KM12) bearing a recessive mutation, caused by insertion of a transgene was created. These mice are characterized by a change in coat colour to yellow and circular behaviour. (Figure 1)

[0076] The yellow coat of KM12 mice suggests abnormal regulation of pigment synthesis. Hair pigmentation in mice is mediated by many developmental and signalling processes involving enzymes, transcription factors, a growth factor and its receptor, a membrane transport protein, a hormone receptor and an antagonist of hormone binding. In mice, wild-type colour hair is agouti and contains two pigments in which the bases and tips contain the black pigment eumelanin and an intermediate band containing phaeomelanin. Two genes in mice, *agouti* (*A*) on chromosome 2 and *extension* (*E*) on chromosome 8 are involved in the regulation of the relative amounts of eumelanin and phaeomelanin in the mouse. Mutations in these two loci result in a yellow coat colour in mice. Dominant mutations in the *A* gene, such as lethal yellow and viable yellow, result in an obese mouse which is completely or almost all yellow because only phaeomelanin is made. In contrast mutations in *E* which result in yellow coat colour (*e*) are recessive. Both the *agouti* and *extension* loci have been cloned. The *A* gene encodes a 131 amino acid polypeptide with a structure consistent with its proposed paracrine function. *E* encodes the (-MSH receptor (MSH-R). It has been established that neither *A* nor *E* have been mutated in KM12 mice. In addition, the transgene has been mapped to two integration sites on mouse chromosome 3 (Figure 7), consistent with a mutation outside the *A* and *E* loci.

[0077] The circular behaviour of KM12 mice suggests improper inner ear function which affects balance. The mature inner ear consists of the auditory apparatus, which is responsible for the perception of sound (organ of Corti) and the vestibular apparatus which is responsible for the sense of balance. Many genes have been found to regulate the development and functions of the inner ear and mutations which cause hearing and balance defects have been found in both human and mouse. These genes encode diverse classes of proteins such as transcription factors, secreted growth factors, signalling molecules, receptors, cytoskeletal components, intracellular transporters, and ion channel proteins. Although pigmentation defects associated with deafness have been found (such as *microphthalmia*, *dilute*) to date no mutations causing both yellow coat and inner ear defects have been identified in mice.

[0078] The databases for candidate genes and known mutant loci on chromosome 3 which may account for the yellow coat and vestibular dysfunction were searched and none were found. The mutation is therefore a novel one. Based on the phenotypic features of KM12 homozygous mice, the mutant locus has been named *yellow submarine* (*ysb*).

[0079] To define the molecular defect(s) in *ysb* mice, a necessary prerequisite is to identify the mutated gene(s) and characterise the phenotypic abnormalities. In the current RGC project Analysis of the vestibular abnormalities in *ysb* mice has been started. Abnormalities in the acoustic nerve and inner ear have been found in *ysb* embryos. Inverse PCR was used to isolate flanking DNA from both the transgene insertion sites. These flanking DNA have been used to isolate genomic clones spanning a total of 49kb DNA. Southern analyses using these clones show that approximately 20kb DNA has been deleted at integration site 2. Clones for both integration sites have been mapped to the same site on chromosome 3. The data suggest that the transgene integration has caused a deletion and chromosomal inversion. It has also been determined that a chromosome 3 recessive mutant (named *lcc*: *light coat and circling*) with features similar to *ysb*, which arose as a result of X-ray irradiation and which may be allelic to *ysb* (Dr. C. Tease (MRC Mammalian Genetics Unit, Harwell)) and a comparison of the two mutants is being performed. Intercrosses between *ysb* and *lcc* mice show that the *ysb* mutants cannot complement the *lcc* mutation strongly suggesting the mutations are allelic.

[0080] Molecular genetics, bioinformatics, developmental biology, transgenic and physiological approaches are used to identify the gene(s) at the $+^{ysb}$ locus and study the molecular and developmental bases underlying the defect(s) in *ysb* mice. The following methods are used : 1) isolation and characterization of the $+^{ysb}$ gene(s); 2) determination of the nature of the *ysb* mutation; 3) performing genetic complementation tests between *ysb* and *lcc* mice, 4) performing transgenic rescue experiments, 5) characterization of the developmental defects in the inner ears of *ysb* mice using 3-D reconstruction analyses, molecular markers and chimera studies; 6) characterization of the neurophysiological changes in balance and hearing in *ysb* mice. Although some emphasis was placed on the inner ear defects of *ysb*, by cloning and characterizing the *ysb* gene(s) insight is gained into the biochemical and possible intracellular signal transduction defect(s) underlying the coat colour change. Approximately 1/1000 infants are affected by hearing defects at birth, two thirds of which have a genetic basis. About half of children with hearing impairment also have vestibular dysfunction. Mice are good models for studying auditory defects because of the similarity in structure and development between mouse and human inner ears. These studies provide fundamental information on the mechanisms of inner ear development and the molecular basis by which inner ear defects may arise in balance disorders of mice and human.

The KM12 mouse line

[0081] In making transgenic mice sometimes integration of the exogenous DNA disrupts the function of one or more genes. While studying the regulation of the *COL2A1* gene, using a recombinant plasmid (pAA2) which contained regulatory DNA sequences from *COL2A1* linked to the *lacZ* reporter gene (Cheah et al. 1995), a recessive mutation, caused by insertion of the transgene, was created. These mice are characterized by a change in coat colour to yellow and circular behaviour, only seen in offspring homozygous for the transgene (Fig. 1, Appendix). In developing KM12 embryos the transgene is expressed not only in the sites expected for *COL2A1* but also in additional expression domains such as specific rhombomeres of the hindbrain (r2,3,5), the spinal cord, dorsal root ganglia, and in the hair follicles of the skin. These neural and skin sites of expression are not typical of the endogenous *Co12a-1* gene but are consistent with the coat colour and behavioural phenotype of KM12 mice (Figures 2,3).

Agouti and yellow coat colour

[0082] The yellow coat of KM12 mice suggests abnormal regulation of pigment synthesis. In mice, wild-type color hair is agouti and contains two pigments in which the bases and tips contain the black pigment eumelanin and an intermediate band containing phaeomelanin. The regulation of hair pigmentation in mice is mediated by many developmental and signalling processes involving several enzymes, transcription factors, a growth factor and its receptor, a membrane transport protein, a G protein-coupled hormone receptor and an antagonist of hormone binding. These pigments are synthesized in melanocytes by tyrosinase (reviewed in). Two genes in mice, *agouti* (*A*) on chromosome 2 and extension (*E*) on chromosome 8 are involved in the regulation of the relative amounts of eumelanin and phaeomelanin in the mouse. Mutations in these two loci result in a yellow coat colour in mice. Dominant mutations in the *A* gene, such as lethal yellow and viable yellow, result in an obese mouse which is completely or almost all yellow because only phaeomelanin is made. In contrast mutations in *E* which result in yellow coat colour (*e*) are recessive. Both the *A* and *E* loci have been cloned. The *A* gene encodes a 131 amino acid polypeptide with a structure consistent with its proposed paracrine function. *E* encodes the (-MSH receptor (MSH-R), a 35kD polypeptide with seven transmembrane domains and is expressed in melanocytes. Activation of MSH-R promotes eumelanin synthesis while agouti protein enhances phaeomelanin synthesis .

Genes and inner ear defects

[0083] Many deaf mouse mutants are characterized by the classic shaker/waltzer behaviour of circling, head tossing and hyperactivity. The circular behavior of KM12 mice suggests improper inner ear function which affects balance and hearing. KM12 mice also show other characteristics of inner ear defects such as head-tossing, and an inability to swim (Figure 4). In addition KM12 mice show an abnormal reaching response. When picked up by the tail, KM12 mice do not stretch out their limbs as normal mice do, but rather curl up towards their bellies (Fig.5, Appendix). Although KM12 mice show some response to sharp sounds (Preyer reflex), this does not appear as strong as for wild-type mice. Therefore it is possible that KM12 mice are partially hearing impaired. Physiological tests which measure endocochlear potential (EP) and compound action potential (CAP) show that *ysb* mice are deaf and that *ysb/lcc* compound heterozygotes are profoundly deaf. These results suggest that the stria vascularis is not functioning properly and cochlear nerve activity in response to a sound is also abnormal.

[0084] The inner ear is a complex sensory organ which develops from a single-cell-layered epithelium (otic placode) which invaginates by a series of cell and tissue movements and closes to form the otic vesicle. Further morphogenetic movements and multi-step inductive events, which regulate differentiation and proliferation, lead to the formation of

the organ for hearing and balance. The mature inner ear consists of the auditory apparatus, which is responsible for the perception of sound (organ of Corti) and the vestibular apparatus which is responsible for the sense of balance. Many genes have been shown to be expressed in the developing inner ear. These include genes encoding transcription factors (e.g. *Nkx5.1/hmx3*, *Nkx5.2/hmx2*, *otx-1*, *msx1*, *pax2*, *kreisler*, etc); secreted factors (e.g. *Bmp4*, *fgf3*, *fgf2*, *bdnf*), receptors (e.g. *EphA4*, *trkA*, *trkB*, *trkC*, *Ednrb*, *PTHrpR*), cytoskeletal proteins such as unconventional myosins (e.g. *Myo7a*, *myo15*).

[0085] Inner ear defects have been traced to one or more of three types of abnormalities: morphogenetic, cochleo-saccular, and neuroepithelial. Morphogenetic defects are caused by developmental abnormalities in structure of the inner ear (labyrinth). Cochleo-saccular abnormalities result from defects in the secretory epithelium of the cochlear duct. Neuroepithelial defects arise from failure of the sensory epithelia to complete normal maturation. In addition abnormal hindbrain development may result in vestibular dysfunction. For example, *kreisler* mice are characterized by deafness and circular behaviour in adults, and in the embryo, by abnormalities in the positioning of the otic vesicles and segmentation of the hindbrain.

[0086] Mice are good models for the studying auditory defects because of the similarity in structure and development between mouse and human inner ears. Mutations in several genes expressed in the inner ear have been shown to cause hearing and/or balance disorders in human and mouse. These include *Myo7a*, *Myo15*, *kreisler*, *hmx3 (nkx5.1)*, *fgfr3* and others.

[0087] The following methods are performed: a) characterize abnormalities in hindbrain and ear development in KM12 mice and b) isolate the DNA sequences at the site(s) of integration of the transgene in KM12 mice.

Yellow submarine: a novel mutant locus

[0088] The recessive nature of the mutant phenotype suggests that the transgene has not integrated into the *A* gene.

[0089] Insertion of the transgene into the coding sequences of MSH-R, ACTH-R, agouti and Kreisler genes has been tested because they have been shown to be important for agouti coat colour or inner ear development. Southern blot analyses of DNA of KM12 mice, using probes for the *Kreisler*, *agouti*, *E* and ACTHR (*Mc2r*) genes, [gifts of Dr. G. Barsh (UCSF), Dr. R. Woychik (Oak Ridge National Laboratory, Oak Ridge) and Dr. R. Cone (Oregon Health Sciences University, Portland)], show that the coding sequences of these genes have not been interrupted by the transgene (Figure 6). These results therefore exclude chromosomes 2 (*agouti*, *kreisler*), 8 (*E*), and 18 (*Mc2r*).

[0090] Using the whole transgene and its end fragments as probes in Southern blot analyses, we have determined that 3 copies of the transgene have integrated into 2 sites in the genome. There are two copies of the transgene, head to tail, at one integration site (Int.1) and one copy at the other (Int2) (Summarized in Fig 7, Appendix). However, since the coat colour and behavioral phenotypes have not segregated over approximately 176 meioses (6 generations), the transgene may have caused a rearrangement and/or deletion of part of the chromosome. Therefore the coat colour and behavioral phenotypes of KM12 mice could be caused by mutation of one or more genes.

[0091] The chromosomal assignment of the KM12 transgene is an essential and informative first step in defining the nature of the mutation. The transgene has been mapped by FISH to two integration sites on mouse chromosome 3, consistent with the Southern analyses (Figure 7,8). The data are also consistent with the results excluding the *A*, *E*, *Kreisler* loci.

[0092] To date no mutations causing both yellow coat and inner ear defects have been identified in mice. The databases have been searched for candidate genes and known mutant loci on chromosome 3 which may account for the yellow coat and vestibular dysfunction and found none. The mutation is therefore a novel one. Based on the phenotypic features of KM12 homozygous mice, the mutant locus has been named *yellow submarine (ysb)* (and homozygous mutants hereafter referred to as *ysb*, wild-type as $+^{ysb}$; Leung, K.K, S. Dong, A. Tang, H. Heng, L.C. Tsui, P.P.L. Tam & K.S.E. Cheah *Yellow submarine (ysb)* a newly discovered locus regulating hair colour and inner ear function. Manuscript in preparation)

Characterisation of *ysb* phenotype

[0093] *Ysb* mice are a valuable model to identify molecules involved in controlling pigmentation and balance. The structural abnormalities in *ysb* mice have been analyzed. The phenotype and pattern of expression of the transgene in *ysb* mice indicate that integration of the transgene has caused a recessive mutation which affects the regulation of pigmentation and also causes abnormal inner ear development. The initial studies have focused on the inner ear abnormalities in *ysb* mice.

Hindbrain and cranial nerve structure

[0094] In vertebrate embryogenesis, the process of segmentation in which reiterating blocks of tissue form along the

anterior-posterior body axis, are fundamental to pattern formation and differentiation. In the development of the hind-brain, segmentation occurs with rhombomere formation in the neural epithelium. A detailed study of *kreisler* embryos using markers for hindbrain segmentation such as *Krox-20*, and *Hox* genes have shown that the *kreisler* mutation is probably due to the consequence of abnormal segmentation of the hindbrain resulting in the loss of rhombomeres (r) 5 and 6. *Krox-20* is a useful marker for hindbrain segmentation being normally expressed in r3 and 5 at 9.0-9.5 days, and *lacZ* expression was found in r3 of *ysb* embryos. In order to assess if the phenotype of *ysb* mice is associated with a problem in hindbrain development, initially the pattern of expression of *Krox-20* was studied in heterozygote and homozygote 9.0-9.5 day embryos by *in situ* hybridization in whole mount and on sections and compared to that for wild-type. Whole-mount *in situ* hybridization showed no alterations in the rhombomere expression of *Krox-20* mRNA in homozygotes and heterozygotes at 9.5 dpc (Figure 9). This is unlike the *Kreisler* mutant suggesting that the mutation has not caused gross alterations in rhombomeres 3 and 5.

[0095] Whole-mount immunostaining studies on 10.5 day *ysb* embryos using an antibody (2H3) against neurofilament, showed a reduction of the VIIIth nerve (which gives rise to the vestibular nerve) in homozygote mutants, while the other cranial nerves appeared normal (Fig.10, Appendix). Heterozygotes appeared normal, although the VIIIth nerve seemed to be thinner compared to wild-type. This result is consistent with the vestibular dysfunction in the *ysb* mice. The mutation could therefore have a) affected the ability of the VIIIth nerve to grow towards the otic capsule; b) caused a failure of the full number of progenitors to migrate from the otocyst initially; or c) caused abnormal cell death (apoptosis).

Inner ear morphology

[0096] To determine whether there are any inner ear malformations in homozygous *ysb* mice, 3-D reconstructions of the inner ear of the mice were generated using histological serial sections. 3D reconstructions of images of sections of the ears of wild-type, heterozygous and homozygous *ysb* fetuses at 16.5 dpc (days post coitum) were carried out. This was followed by painting in the lumen to give an impression of the whole structure of the labyrinth. The data show superior and lateral canal and ampulla defects in two *ysb* fetuses, with one more severely affected than the other (Appendix, Fig.11).

[0097] The reconstructions of the lumen in two 16.5 dpc homozygote *ysb* fetuses show the superior and lateral canal were truncated and their ampullae were absent. At 16.5dpc the sensory regions are distinguishable from non-sensory regions because the cells have differentiated into a pseudostratified epithelium. This epithelium appears thicker when compared to the surrounding epithelium because of the two or more layers of nuclei. Because this difference could be observed easily with the light microscope, these thickened areas were traced and painted onto the lumen of ears that had previously been reconstructed (Fig.12, Appendix). The different types of sensory areas including the organ of Corti (detector of sound), the maculae (detectors of linear motion), and the cristae (detectors of rotational motion), are represented in different colours to demonstrate more easily which sensory areas were affected.

[0098] Addition of the sensory information to these reconstructions showed that several different types of sensory areas were abnormal in the homozygote. The superior and lateral cristae which reside in the ampullae were absent (Fig 4, Appendix). This observation is consistent with the previous observation that the ampullae were absent. The detailed analysis of sensory epithelia also showed that there was no ectopic formation of these cristae. Both maculae also appeared to be affected by the mutation: the utricular macula showed almost no thickening in the expected region and while the saccular macula did demonstrate some thickening, it appeared abnormal. The only vestibular patch that appeared normal was the posterior crista. The hearing organ, the organ of Corti, also appeared normal at this time of development. No abnormalities were observed in the heterozygote using this method of analysis, and the wild-type (+/+) littermate was used as a control. Reconstruction of sections from 13.75 dpc fetuses showed a similar defect in the semicircular canals (Figure 13).

Isolation & characterisation of transgene integration sites

[0099] Fig. 7 summarizes the progress in isolating the wild-type locus at the site of transgene integration. By priming within the transgene, inverse-PCR was used to isolate DNA sequences flanking the integration sites (Fig 7). Two PCR products, 350bp (SpeI-1) and 600bp (SpeI-2), were obtained for integration sites 1 and 2 respectively. Sequence analysis (Figure 14) showed the presence of a possible polyA attachment signal in SpeI-1 but a BLAST search of the genomic and EST databases did not reveal significant matches. Longer 5' flanking sequence have been obtained for integration site 1 (1.8kb) and 3' flanking sequence for site 2 (500bp).

[0100] Southern blot hybridization using the flanking sequences for both integration sites show that a deletion has occurred at integration site 2 but not at site 1 (Figure 15). These PCR fragments were then used to screen a normal 129 mouse genomic 1 phage library. Four overlapping genomic clones were obtained for integration site 1. Two genomic clones were obtained for integration site 2. Another two clones were isolated upon further screening. Altogether ge-

5 nomic sequences spanning 19kb and 30kb at integration sites 1 and 2 respectively, have been isolated. Southern analyses using these genomic clones as probes, have revealed that approximately 20kb DNA has been deleted at integration site 2 (Figure 16). One clone for integration site 2 hybridized to 15.5 dpc mouse fetus mRNA in Northern analyses. The genomic clones for integration sites 1 and 2 co-localise to the same region of chromosome 3, suggesting that a chromosomal inversion may also have occurred (Figure 7). These results would explain why the two transgene insertions have not segregated over so many meioses. The sequence of 8114bp (cloned into a plasmid pPL5) within the 20kb region deleted at integration site 2 in *ysb* has been determined (Figure 17). This region contains sequences with 90% and 100% homology respectively to IMAGE clones 1196866 and 636096 in the EST database (Figure 17). The positions of these IMAGE clone sequences within pPL5 is summarized in Figure 18.

10 **[0101]** The *ysb* gene(s) are being identified and characterized, and the molecular and developmental bases underlying the defect(s) in *ysb* mice are being studied. Towards these aims, the current results are being built on, and molecular genetics, bioinformatics, developmental biology, transgenic and physiological approaches are used to a) identify and characterise the $+^{ysb}$ gene(s) and the encoded transcripts; b) determine the nature of the *ysb* mutation; c) perform genetic complementation tests and transgenic rescue experiments, d) further characterise the developmental defects in the inner ears of *ysb* mice using 3-D reconstruction analyses, molecular markers and chimera studies; e) characterise neurophysiological changes in balance and hearing.

15 **[0102]** These studies provide fundamental information on the mechanisms by which inner ear defects may arise.

20 Methods

Molecular cloning and characterization of sequences of the wild-type and mutant *ysb* locus

1. Gene discovery and characterisation at the *ysb* locus

25 Several approaches are used to identify genes at the *ysb* locus.

30 **[0103]** a. Bioinformatics: The isolated genomic clones are sequenced and the data analysed for the presence of potential exons using bioinformatics tools. These predictions are very important for identifying regions to follow up.

[0104] b. Expressed gene sequences are identified using a combination of approaches. First DNA fragments containing exon sequences are screened for by Northern analyses. Once such fragments are identified, exon trapping approaches are used to identify regions with transcribed sequences. Potential exon containing fragments are used in situ hybridization studies and the pattern of expression compared with that of the transgene in *ysb* mice.

35 **[0105]** In addition, to gain insight into the function of the gene(s), we use bioinformatic tools to screen for homologous sequences in other model genome databases e.g. yeast, fly, worm, fish and human.

2. Cloning the mutant locus: genomic library construction:

40 **[0106]** In order to isolate the site of transgene integrations a cosmid genomic library will be constructed from DNA isolated from homozygous *ysb* mice. The cosmid, pCos8 (gift of Dr. A-M Frischauf, Imperial Cancer Research Fund, London), is used as a vector in constructing the library. This vector does not contain *lacZ* sequences and its use helps minimize isolating false positives during library screening. Primary embryo fibroblast cultures from KM12 homozygote mice are established and used as a source of high molecular weight genomic DNA. The cosmid library is constructed by standard methods. To ensure high efficiency of cosmid packaging, tested packaging mixes (e.g. Stratagene's Gigapack Gold) are obtained from commercial sources. DNA probes covering the 5 and 3' ends of the transgene and *lacZ* sequences are used to screen the library by colony hybridization. Cosmid libraries and isolated genes have been successfully constructed in the past.

50 3. Characterization of genomic clones and the mutant locus:

[0107] The Cosmids are further characterized for the presence of the transgene by restriction enzyme mapping and Southern blot analyses using the appropriate probes. Once cosmid(s) containing the transgene are identified, sub-clones are made and the nature of the sequences flanking both 5' and 3' ends of the insertion site are determined by DNA sequencing. Sequences obtained are then analyzed using bioinformatic tools to find potential open reading frames (ORFs) and also are used to scan the nucleotide and amino acid sequence databases for homologous sequences. The *ysb* clones are used to analyze genomic DNA from the Harwell mutants. To determine whether the *ysb* and Harwell mutants are allelic, a complementation test is carried out by crossing the two mutants. If the two mutations do not

complement, this information is useful for the identification of the responsible gene, because two different alleles would be available. Furthermore, both mutants involve two possible mutation sites and non-complementation will immediately cut down the number of sites to be investigated at a molecular level, as there is only one chromosomal region of overlap between the two mutations.

5

4. Genetic rescue experiments

[0108] Expression of BACs in transgenic mice have been successfully used to correct deafness in *shaker-2* mice. To establish if the sequences characterized are those which have been mutated in *ysb*, bacterial artificial chromosome clones (BAC) will be screened for, using the genomic clones for integration site 2 as probe. We will test the ability of BAC clones spanning the *ysb* locus to rescue the phenotype of *ysb* mice by transgenesis. It would also be important to compare the *ysb* locus with that of the Harwell mutant.

10

5. Isolation of human YSB

15

[0109] The aim is to be able to test if any human vestibular disorder is caused by mutations in the human homologue of *ysb*. Therefore it is important to isolate human BAC clones which cover the equivalent *ysb* locus. The mouse genomic clones are used to screen a human BAC library. Once these are isolated they are characterized as for the mouse clones in terms of sequence etc. Comparison is made between the human and mouse clones. Once candidate human clones are isolated, their chromosomal locations are mapped and the mutant databases (e.g. OMIM- Online Mendelian Inheritance in Man) scanned for possible associated human disorders. The availability of human clones is a resource for future studies on human patients with hearing and balance problems.

20

6. Morphological and developmental analyses

25

[0110] From the reconstructions of the lumen of 13.75 and 16.5 day fetuses, it was clear that in the homozygote mutant the superior and lateral canal were truncated and their ampullae were absent. To trace the timing of the developmental abnormality, the heads of 11, 12.5 day fetuses and newborns from non-transgenic control and *ysb* mice are fixed and processed for 3-D reconstruction. The structure of the inner ears are reconstructed in 3-D using software.

30

[0111] To determine when the abnormal canal morphogenesis first becomes apparent, the morphology of the *ysb* embryonic ears at 11, 13 (a critical time point in canal formation) and 16.5 dpc are carefully dissected out, cleared and the lumen filled with white paint and then examined for abnormality. This technique provides a quick and accurate visualization of the 3-dimensional structure of the ear to assess variability in the homozygote phenotype and detect subtle defects in heterozygotes. The otoconia in the extracellular matrix, which lies atop the hair cells of the maculae, are easily observed after clearing, before paint injection. Whether this matrix forms normally in the mutant is studied, since the maculae appear abnormal at 16.5dpc. The inner ears at 18dpc are also examined by scanning electron microscopy to determine if there are any abnormalities in the distribution and structure of the sensory cells.

35

[0112] In addition, a careful in situ hybridization analysis is made to compare, in *ysb* and wild-type embryos, the expression of genes known to have a role in inner ear development, using molecular markers such as Hox genes (*Hoxa3*, *Hoxb1* and *Hoxb2*), *Fgf3*, *nkx 5.1*, *otx1*, *Bmp4*, *Trkb*, *Trkc*, *NT-3*, *BDNF*, *neurogenin1* for the following reasons. In *kreisler* mice the defect was shown to be the consequence of abnormal segmentation of the hindbrain. The *hox* gene and *fgf3* probes are used to examine hindbrain development in *ysb* mice. In vitro fate mapping has shown the lateral half of the otocyst gives rise to the canals and cristae. *Nkx5.1* is expressed in the dorso-lateral portion of the otocyst (and later in the semi-circular canals) and inactivation of the gene affects the semicircular canals. *Otx-1* is expressed postero-laterally in the otocyst and the lateral semicircular canal is missing when the gene is "knocked out" in mice. Since Xgal staining was found in the ventro-lateral part of the otic epithelium in 9.5day *ysb* embryos, expression of these genes is studied to determine if the *ysb* gene(s) are acting upstream or downstream of these two genes. BMP4 will be used as marker to study the cristae, sensory areas of the ear.

45

[0113] Several neurotrophic factors (e.g. BDNF and NT-3) and their receptors (*Trkb*, *Trkc*) are important for the early development of the inner ear. They regulate survival of vestibular and cochlear neurons and the neurons which innervate the inner ear. Since the VIIIth nerve in *ysb* mice is stunted, the expression of these genes will also be studied. The *neurogenin1* gene has recently been shown to be essential for the determination of neuronal precursors for proximal cranial sensory ganglia and will be a good marker for studying the prospective ganglion cells in *ysb* mice. The same markers are used to compare the Harwell and *ysb* mutants should the complementation tests prove them to be allelic (see above).

55

7. Origin of developmental abnormalities

[0114] Developmental analysis of marked ES cells in mouse chimeras is a powerful approach to studying cell-fate and lineage-specific gene function. The fact that the *ysb* mutant locus is marked by *lacZ* is exploited to dissect lineage-specific gene function as well as to test whether the gene defect is cell-autonomous or non-autonomous.

[0115] Embryonal stem cells are derived from *ysb* blastocysts and used to generate chimeras by blastocyst injection. Chimeric embryos analyzed (5-10 per stage) are collected at stages of development between 9.5 days and birth and the relative contribution of ES cells in the developing inner ear as judged by *lacZ* expression in the chimeras is studied. The preferential loss or under-representation of *ysb* ES cell contribution in particular sites in the developing inner ear is indicative of specific changes in lineage potency. The lineages are characterised by in situ hybridizations using the appropriate molecular markers described above.

8. Neurophysiology of balance and hearing in *ysb* mice

[0116] The neurons that innervate the inner ear are derived from the otic placode. These neurons arise from the placode early in development of the inner ear to form the acoustic and vestibular ganglia. The neurofilament staining experiments had shown a reduction of the eighth cranial nerve in 10.5 days *ysb* embryos, suggesting failure to innervate the inner ear could be a cause of the balance and/or hearing problem in *ysb* mice. It is important to assess whether or not the *ysb* mice can effectively convey (a) head movement signals and (b) auditory signals to the central nervous system. The expression of *c-fos* is used as an indicator of functionally activated neuronal activity in the brainstem. Fos expression as an indicator of postsynaptic stimulation is an established method for identifying functional connections between peripheral sensory receptors and central neurons. Fos immunostaining of brain sections is therefore used to map the central neurons involved in the functional neural pathway.

9. Vestibular Experiments

[0117] Natural vestibular stimulations are used, viz. sinusoidal rotations on the yaw or pitch plane and constant velocity off-vertical axis rotations. The former stimulates the respective pairs of semicircular canals while the latter selectively activates the utricular hair cells in a sequential manner. With these modes of stimulation, secondary neurons in the vestibular nucleus that have functional connection with hair cells on the respective canal pairs or the utricular maculae should be excited and should express Fos. Immunohistochemical techniques involving *c-fos* are used to map the pattern of postsynaptic vestibular stimulation within the vestibular nuclei and other parts of the brainstem. To determine the spinal projection pattern of Fos-expressing central vestibular neurons, brain sections are examined for neurons where Fos immunostaining and retrogradely transported rhodamine-labeled latex beads (previously injected in the spinal cord) are co-localized.

[0118] As anesthetics have been known to influence the levels of Fos expression, conscious animals will be used. Control experiments are performed to ensure that the observed results are specifically due to the activation of canal and otolith receptors: (a) intact mice mounted but without rotation, (b) acute labyrinthectomized mice mounted but without rotation, and (c) acute labyrinthectomized mice mounted and then subjected to rotation. As the expression of *c-fos* can be induced by sensory stimuli other than that intentionally delivered in this study, care is taken to minimize uncontrolled variables and other sensory inputs. Control mice with sham operation on the spinal cord are also prepared.

10. Auditory Experiments

[0119] These studies are carried out in conjunction with the vestibular experiments. The auditory pathway is activated by repetitive stimulation paradigms using tone bursts. The sound stimulation is conducted in a darkened soundproof chamber. Each freely moving mouse will be presented with pure tone bursts delivered from the ceiling of the chamber. Functional connection between central neurons and the peripheral auditory receptors is indicated by Fos expression. These experiments provide information on the activation pattern of neurons in the central auditory pathway.

11. Additional tests of vestibular/hearing function and behavioural studies

[0120] The following studies complement the neurophysiological studies and greatly enhance the scope of the investigation, facilitating the definition of the underlying defect(s) in *ysb* mice. Using the *ysb* mice, a breeding colony is established to provide mutants for study. The behavioural consequences of the vestibular defects are described using a standard battery of simple tests of balance, including, contact righting response, elevated platform test, and open field test, reaching response, Preyer's reflex, and quantify the extent and type of the behavioural abnormality. The function of the cochlea in young adult *ysb* and littermate controls is by measuring thresholds for detection of a compound

action potential from the cochlea in response to tonebursts at various frequencies and intensities. This approach gives an indication of any hearing impairment in the mutant.

[0121] The molecular cloning and characterization of the potential *ysb* clones is performed. 3-D reconstructions on *ysb*, heterozygous and wild-type 11 day fetuses and newborns is performed. In situ hybridization experiments are performed. The detailed restriction map of the genomic clones isolated for the *ysb* locus is made and then clones are tested for the presence of expressed sequences by Northern analyses. Bioinformatics is used to analyse the DNA sequences that are obtained.

References

- [0122]** Leung, K.K., Ng, L.J., Ho, K.K.Y., Tam, P.P.L. & Cheah, K.S.E. *J.Cell Biol.* 141, 1291-1300 (1998).
- [0123]** Jackson, I.J. *Curr.Biol.* 3, 518-521 (1993).
- [0124]** Conklin, B.R. & Bourne, H.R. *Nature* 364, 110-110 (1993).
- [0125]** Bultman, S.J., Michaud, E.J. & Woychik, R.P. *Cell* 71, 1195-1204 (1992).
- [0126]** Miller, M.W., Duhl, D.M.J., Vrieling, H., et al. *Genes Dev.* 7, 454-467 (1993).
- [0127]** Mountjoy, K.G., Robbins, L.S., Mortrud, M.T. & Cone, R.D. *Science* 257, 1248-1251 (1992).
- [0128]** Robbins, L.S., Nadeau, J.H., Johnson, K.R., et al. *Cell* 72, 827-834 (1993).
- [0129]** Hughes, D.C. *Audiol Neurootol* 2, 3-11 (1997).
- [0130]** Steel, K.P. & Brown, S.D.M. *Trends Genet.* 10, 428-435 (1994).
- [0131]** Steel, K.P. *Annu.Rev.Genet.* 29, 675-701 (1995). Steel, K.P. & Hardisty, R. *Neuroscience Short Course Syllabus* 1, 26-38 (1996).
- [0132]** Torres, M. & Giraldez, F. *Mech.Dev.* 71, 5-21 (1998).
- [0133]** Hertwig, P. *Z.KonstLehnre* 28, 327-354 (1994).
- [0134]** Deol, M.S. *JEEM* 12, 475-490 (1964).
- [0135]** Frohman, M.A., Martin, G.R., Cordes, S.P., Halamek, L.P. & Barsh, G.S. *Development* 117, 925-936 (1993).
- [0136]** Kalatzis, V. & Petit, C. *Hum.Mol.Genet.* 7, 1589-1597 (1998).
- [0137]** Petit, C. *Nat.Genet.* 14, 385-391 (1996).
- [0138]** Probst, F.J., Fridell, R.A., Raphael, Y., et al. *Science* 280, 1447 (1998).
- [0139]** Wang, A., Liang, Y., Fridell, R.A., et al. *Science* 280, 1447-1451 (1998).
- [0140]** Hadrys, T., Braun, T., Rinkwitz-Brandt, S., Arnold, H.-H. & Bober, E. *Development* 125, 33-39 (1998).
- [0141]** Wang, W., Van de Water, T.R. & Lufkin, T. *Development* 125, 621-634 (1998).
- [0142]** McKay, I.J., Muchamore, I., Krumlauf, R., Maden, M., Lumsden, A. & Lewis, J. *Development* 120, 2199-2211 (1994).
- [0143]** Birren, B. Green, E.D. Klapholz, S. Myers, R.M. & Roskams, J. *Genome Analysis: A laboratory manual. Volume 2: Detecting Genes* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY, 1998).
- [0144]** Krizman, D.B. (eds Birren, B., Green, E.D., Klapholz, S., Myers, R.M. & Roskams, J.) *Exon trapping. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. (1998). 191 p. (2): Genome analysis: a laboratory manual Volume 2 Detecting genes.*
- [0145]** Martin, P. & Swanson, G.J. *Dev.Biol.* 159, 549-558 (1993).
- [0146]** Kiernan, A.E., Nunes, F., Wu, D.K. & Fekete, D.M. *Dev.Biol.* 191, 215-229 (1997).
- [0147]** Rinkwitz-Brandt, S., Arnold, H.-H. & Bober, E. *Hear.Res.* 99, 129-138 (1996).
- [0148]** Acampora, D., Mazan, S., Avantaggiato, V., et al. *Nat.Genet.* 14, 218-222 (1996).
- [0149]** Simeone, A., Acampora, D., Mallamaci, A., et al. *EMBO J.* 12, 2735-2747 (1993).
- [0150]** Morsli, H., Choo, D., Ryan, A., Johnson, R. & Wu, D.K. *J.Neurosci.* 18, 3327-3335 (1998).
- [0151]** Fritsch, B., Farinas, I. & Reichardt, L.F. *J.Neurosci.* 17, 6213-6225 (1997).
- [0152]** Fritsch, B., Silos-Santiago, I., Bianchi, L.M. & Farinas, I. *Trends Neurosci.* 20, 159-164 (1997).
- [0153]** Fritsch, B., Silos-Santiago, I., Bianchi, L.M. & Farinas, I. *Seminars in Cell Dev.Biol* 8, 277-284 (1997).
- [0154]** Ma, Q., Chen, Z., del Barco Barrantes, I., De la Pompa, J.L. & Anderson, D.J. *Neuron* 20, 469-482 (1998).
- [0155]** Rossant, J. & Spencer, A. *Trends Genet.* 14, 358-363 (1998).
- [0156]** Robertson, E.J. in *Teratocarcinomas and embryonic stem cells* (ed Robertson, E.J.) 71-112 (Oxford IRL press, Oxford, 1987).
- [0157]** Morgan, J.I. & Curran, T. *Annu.Rev.Neurosci.* 14 , 421-451 (1991).
- [0158]** Kaufman, G.D., Anderson, J.H. & Beitz, A.J. *J.Neurosci.* 12, 4489-4500 (1992).
- [0159]** Steel, K.P. & Smith, R.J. *Nat.Genet.* 2, 75-79 (1992).

Annex to the application documents - subsequently filed sequences listing

[0160]

5

10

15

20

25

30

35

40

45

50

55

EP 1 148 128 A1

SEQUENCE LISTING

<110> The University of Hong Kong

5 <120> New gene locus containing a gene involved in regulating hair pigmentation, vestibular function and fertility

<130> P55601EP00

<140> EP 01201037.7

10 <141> 2001-03-20

<150> US 09/528,928

<151> 2000-03-20

<160> 6

15 <170> PatentIn Ver. 2.1

<210> 1

<211> 8114

<212> DNA

20 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: sequence of pPL5 containing 8114 bp DNA fragment (Hind III cut) of 20 kb deleted sequences in integration site 2

25

<220>

<221> misc_feature

<222> Complement((1)..(8114))

30 <400> 1

aagcttaaaa tgagccaagc gcagagtga gctgaagttg agagacaccg gcgctgtgcc 60
cagaacacag tccgaacctt aacatttga tggagcaaca caaaagtggg aaccagatg 120
cccagacatc tagcgacaca gttggtaaag ccttcttaca taacaatgat aaaataactt 180
aacatttatt taacattttt acacaggagt gcatgagaag gttgacagat gttatctggg 240
35 gagattctca aagccgatgt gagaccagtc acctctttac aaattatctc tcttcacaaa 300
catctacatg ggctttgcag gctgagcagt ctgcccaaag gaaccattta aatagttttc 360
tcatgccaaa gtaaaagtta aaaagacaac ttttttttaa aaaaaaaaaa aaagtgaagt 420
aaatgactta aggtttagtt tacttctggt aggagagatt caaccggttg ggaaaggggt 480
gtgcgcttgt gtgtgagacc agtgagagcg aaatgctgga tacgattggt cctcagtaaa 540
agagaggaac tggaaatgac taatattcca tagaatgata agctccaac ttttagcatt 600
cttctctatt tgtctattat gtttctctac aagacaaaaa aataattttc aacttacaag 660
40 tcttattggt tacacttaat tttcaagact tgtacttagc tgatgaaaca gatgttcccc 720
aacaaccaga tgatctcagg tggacttact ggacagtgtg tggagatctg tgcccttct 780
aaggtcatcc gttggacca ttgccgtgga ggctaagaga tgacacatag acagacgcag 840
aatctcatga gtcagactgg tgcccttgcc tcctcaattt ttgtttttca tttttttgag 900
aagagatctc atactcccag gctggccttg aactcacttt gtagctgagg ctggctttat 960
actgatgatc ctctgcctc caccctcctca catgctggga ttccggcagt gtgtcacctg 1020
45 cccaggggat agttccctct tagaagaagc accagggagt tagtttctcc tttatcctat 1080
gtgagggttt gctgaggctg tcttctatga atcataaagt aggccagatg ctgaaactgc 1140
cactgcccct gtccacgttt ggaactagaa gaaactattt tctggggttt gtatgctacc 1200
cagcctgtgg tgtttgttg tagcaacttg gatggctaaa gacaatggat tgagctgcct 1260
tctgttctac aactgcactg accacatgtg accatttagc acttgacact tgactatagt 1320
ggactgaggg aatgattctt gcttttcttt tctttctttc ggtttttttt tttttttttt 1380
50 tttcgagaca gggtttctct gtgtagcctt ggctgtcctg gatctcactc tgtagagtat 1440
gctggccaat tctaattaga ttagtgtctt atgactagta actactacag tagcaaactt 1500
tggttgtcaa cttgactaca tctggaatca actaaaacct aagataatag gtgtcacaca 1560
cacagacaca cacacacaca cagatagaga catatgttta acacatagtt aaaaataaac 1620
cttttaaaag aacatgtatc acggggcaag ttaaatcata attctgaata ctgtacttct 1680
ctgtaaatta gagatggcac tgttcatcca ctgggagaag atattaagag ttgctagtgt 1740
55 gggctggaga gatggctcag cggttaagag caccgactgc tcttcttag gtctgagtt 1800
caaatcccag caaccacatg tagctacaaa ccatccgtaa tgagatccga caacttcttc 1860
tgggtgtgtct gaagtcagct acagtgtact tagatataat aataaataaa tctaaaaaat 1920

	aaaaacattt	gctaatagcca	aaacagacat	atcaagagta	caaaataaaa	acttccaac	1980
	taaaaagatg	ccaccagaaa	tagagtatct	gggttgcac	gtggcatagt	agcacacctg	2040
	ccaaacctgg	caggtatgag	gccctgcac	ccccacatga	gaactaaaat	tttccaatat	2100
	gagatgaaac	atattaaaac	aagatgaaac	tcaccaaaaac	atgaaggagt	ttctcctcaa	2160
5	ctttaaggat	ttggtaattg	atctcccaca	taaaaatatg	tttgtgggct	agcaagatgg	2220
	ctcagagcaa	aagcacttgt	catgcaggcc	cgacgacctg	aattcagtg	caagagcatg	2280
	caacggaagg	cgaaggctgc	aaatcattct	ctgacctcta	catatgcatc	atggcatggg	2340
	tacctacggt	cataaaaatg	tcataagctc	atctacaata	ataaattata	tatgcacatg	2400
	cacacacaca	tatacacttg	tcaaatctat	aatgttggaa	acctttgcgt	gtcatatagc	2460
	tgatctgtct	aggtagaaat	atatcaatta	acaatgctga	ttgaaaaacc	tcaaatatta	2520
10	tttttgaatg	cttattttga	gtcattttcc	aaggatgtg	agcccgccat	atctgtggtt	2580
	tgcccatacca	tgaatttggg	taatcacata	tgcatcaaaa	taagcccttc	aagagacatg	2640
	ttttcttcag	ggctgtagct	tctggtagt	catccacacc	cctatgcatg	ctcctgtaag	2700
	gaccccaatt	aaactcagta	ggccactgtc	ctagtttgct	ttcttctact	gtgataaaaac	2760
	actgagctaa	gcagcttgag	ggacaagggg	ttactatagc	ttactgggta	cagtccatca	2820
	ctgaatgagc	tcaggacaag	agctcaaggg	aggaatctga	agcagagacc	atggaggaac	2880
	actgctgact	ggctcactcc	cacatgaggc	attaatcatg	caaatgcttc	agagacatgt	2940
15	acacaggtca	gtcttattga	ggcagtttct	caactgaagc	accctctttc	ctgggtgactc	3000
	tagagtccag	ctgactaaaa	ctaactaagt	gctactgatc	tctcctcaag	ttcatgcaca	3060
	aacacatcac	taatggacca	tagtcatttt	tcttgtttat	accaaaggcc	tcacaatata	3120
	aaacatgggt	tagcctttca	atttcacag	tccttaaaaa	tttcaacata	tagaagacac	3180
	aggtcttctt	caaagtccaa	agccccctaa	ctgtgggctt	tttgcaaaat	aaaaacaagt	3240
20	aacattatct	cctgatctta	gaggaagaa	ccagggcaca	gttacaataa	gaacaaagca	3300
	aaaccaacat	ccagatgtgt	aaatcaaatc	ctatagctca	atgtccagca	tccggggctt	3360
	atggtctccc	agactccaaa	ggtcttgggc	agctccactc	tgctcttagc	tccagtacat	3420
	aggctcaggc	tggctccaat	ccacagcacc	ccactgtcct	ggctcctcta	acatgttgca	3480
	atatctgctg	caactgaagc	tgaccttcca	ccaatgaacg	acctctccag	acttctcaca	3540
	gtgacaatcc	tctacttctc	tatgtgactt	cctcaatcct	ggggtctcca	ctgcaactga	3600
25	ggctgtacct	tcatcaataa	cctgtcctta	ggctctcttc	agagactctg	tggccttgat	3660
	ttatggtacc	aagccttaac	tgttctcaat	ggtccctttc	aagaccagta	ccaccaatga	3720
	gactgtgctt	ggtaggagct	aattattctg	tggatggttt	cccaagccct	ttggaagagc	3780
	ccagaagacc	atgagtgact	tccagaagtt	agacactaag	attcactggt	gaactttggt	3840
	ttcactctga	tttaattgtg	actgtgccct	tttggagttc	ttccctcttg	gagtaagaaa	3900
	atatttaact	tattttgact	ttgcaagagc	ccacagctgt	gagagtttgg	agtttgggag	3960
30	atattttgga	gtgttagaga	gatcagggat	attttagaga	ggcatgaact	ttgaaagaga	4020
	ctttggatgt	tctaagggga	tttaactttt	aaagtgtctg	gatttgtaaa	gacttttaa	4080
	agttggaatg	ttttatattg	tgatatgaat	attgatattg	aatcttggca	gtgaacaaga	4140
	aaggggaaggt	tataatttaa	ctgtaatgtg	tttgtgtggt	acattaacaa	gcagtcatat	4200
	ttttatgggt	ttatggaaag	ttttcataag	ctcctcaagg	ccttggaaat	ccaaaatcca	4260
	ttcctttggt	aagagggaaac	ctcaaatgag	aaaatgccct	tactaagatt	ggcccaggag	4320
35	aaagcctgca	gtgttttgt	tttgtttgt	ttttaaatt	gaggcagatg	tgggaggtt	4380
	ctagcccact	gtaggcagtg	tgacacctgg	gctggtggtc	ctctgtgcca	acagaaagca	4440
	ggctactagc	taggctagcc	atgagaagca	agccagtaag	catcaggttc	ctcctgtctt	4500
	gctccagctt	ccctcattga	tggatttaat	cagctgtaag	gtgaaatgaa	ttggttactt	4560
	ttatcacagc	aaaagaaatt	cttttttaa	aattttatt	aggtattttc	ctcatttaca	4620
	tttccaatgc	tatcccaaaa	gtccccata	ccctcccccc	cccactcccc	tactcaccca	4680
40	ctcccacttc	ttggccctgg	cgttcccctg	tactgaggca	cataaagttt	gcaagaccaa	4740
	tgggcctctc	ttccactga	tggccgacta	ggccatcttt	tgatacatac	gcagctcata	4800
	tagtcaagag	ctccggggta	ttggttagtt	cataatgttg	ttccacctat	agggttgcag	4860
	atccccttag	caacttggat	actttctcta	gctcctccat	tgggggccct	gtgatccatc	4920
	caatagctga	caatgagcat	ccacttctgt	ttttgctagg	ccccggccta	gtctcacaag	4980
	agacagctac	atctgggtcc	tttcagcaaa	atcttgctag	tgatgcaac	ggtgtctgcg	5040
45	tttggcggct	gattatggca	tggatccccg	gatatggtag	tctctagatg	gtccatcctt	5100
	tcgtctcagc	tccaacittt	gtctctgtaa	ctccttccat	gggtgttttg	ttcccaattc	5160
	taagaagggg	caaagtgtcc	acactttggt	cttcgttctt	cttgagtttc	actaactaag	5220
	acagagactc	ttaaataatta	tcaaattcag	ctgccagcct	tatgtggtct	tgacaccctt	5280
	ggactacaac	ttctgtgtgg	gagctctgag	gagacacttc	gcagaagatt	ttgcctgtta	5340
	ggcaggtctc	gtctcaatca	cactgagttt	gcagctccag	ttgagcaccg	ttcaattggc	5400
	ctagtaaagc	acaggtttca	cttaggcgat	gctggagctt	tagcagaagc	cttacctga	5460
50	acttcacgag	ccaagcctga	gtgtcggca	ttgcttccaa	caccaccatc	caagccccct	5520
	acactagtat	atcttcaagt	tctaagccct	cagagctcaa	agctccaaat	gcttccgcaa	5580
	ccacacttcc	aacccccaaa	caacacacct	gtgtccatag	tagcaacaac	acactctctg	5640
	gtaccaatct	ctgtttcgat	agatgatgga	caggcagaca	gacaagaaag	tagaaagggg	5700
	accagagga	aagaagaaga	gagtccacaa	gagtgagggg	gacaagaaag	agtaattggg	5760
	gtgagtatga	tcaaaatgaa	catatatgaa	atatcattat	gaaatgtata	taattaatat	5820
55	atactaacaa	gaaacctttt	ttaaaatttt	caaaaatcat	gctgtcactg	aacatgggca	5880
	gcttttcttg	tcggtacccc	ctaagcagca	cattaacact	atttacataa	catttaataa	5940
	tgtaagcat	tttaactaac	tagagatggt	ttgagtgagt	atagggagat	gtgtgcacat	6000

EP 1 148 128 A1

catgagcaaa tgctgccgtt ttccactggg gcccgaggac ttgttttggc aacttccctg 6060
 agcactgggtg tggtcctctc atagatacct agcgatagct gatttccatt tttctccatg 6120
 cctatcttta aggtaagtag gtctttttat tgtctttaca actttatgaa ctacatacag 6180
 ctctctcgt gtagttctc cgcaggaaaa atgggtgatt cagaaaaaac tctgctgttc 6240
 5 ttgccaaatg cagccgccat ggagcagaga gctgcagaaa cagagatggg agctcatcc 6300
 tgcactgcga aggcttcgag gaaggagaat tccatcaggc tgaatgtgtc atgaacttta 6360
 agtaataata ataaaatata aataatata ataactcact tggcaatggg taaagagaca 6420
 ccttgttggg gacattgcct tatatcagtc agggacagag cagatgactg tcttttgc 6480
 ttctagctt gatggcaaa gctgtcttat gtctaggata catgaatatac aaatcaaatc 6540
 aaatctaggc agcacagcac aagaagagaa atacagagat ttgcatacca ccatgctacc 6600
 10 caggatctg ccctcccaca catggcacag acaccactt gccacctcat cgtggagcat 6660
 ggatgaaggg tgttaccacc agacgtgtgg tcgggcttca gccagagtca ccacaatgtg 6720
 caaatatag gtgggcaact gggagagaga gaaaaaaaag tgacaacttg agcattatga 6780
 agagagaggg aactggagac tgaagggtag aaaggagtgc cctgttctga gtgtccagcc 6840
 tatggttaagg ccccagcctg ggctgccact gagggtgta tctgagtcca tgagtcaca 6900
 gcagcagggc tcagtgtgga tatccgtgac acatgttatc actagagaac acggggaact 6960
 ccctggtcag gacagctgta gggaccaca tggacgtaca ggggctgtgc ataactagcc 7020
 15 ttgctctca caggatgagg cctttgggag agctgggtggc agcacttggg agagagggcc 7080
 ctacacctac ccagtcagca cagtggagct ggccctgggg gtgggctgg ggtgactcac 7140
 cccaagggca tgagtgtgg agagccgacc tcatcactca tctgctgctg ggtggcaca 7200
 gtgcagaggt gatgtcacc ccaccccc tgcaaccct tgtcacctct ggaagtggg 7260
 aaagctgccc atagggccat gaactccgga gaactagccc tgccctcacc attcagcact 7320
 20 aggaggatcc tgcacctctc ctggatagca cggcagagct gaccctggca gcaggggtac 7380
 ggggtaacca gagggagaga gagaacgaga gcatgggaga gctggtcca ccaccctca 7440
 gccatgagag gtgacctggg acagctgacc tggcagtcag gagagtaggt gagctggccc 7500
 tacctctcac tcaggagtgg ctgcagcact tgggagagtg ggtcctacat cttgcttgg 7560
 caacacagca cagtgtgct ggttctggta gcaaatgcat ggggtggcca gccccgaggg 7620
 tataggagca ggagagatgg cctagcccct cacaggatgc atcacttggg acattgggtc 7680
 ccacgccatg actgggcagc acagtggcgc tggctttgga ggcttgggtg tgggtgctg 7740
 25 ggcccaagg gcaaggaac aggagatcta aacctacctc ctgactatgg cggtagtatt 7800
 gactagccta gccaaaacag tgctggagtg tggagagccg tgccacaagc aatcgcgagc 7860
 cgtgaacaat cgccattata agatggcgt agcttccact gtgcctaact agtaacaag 7920
 ccttatgcac aagtgaaga gtaaattcac gccagtcac tgccatctt gggacgtagt 7980
 aatgggggtga tgggcgagca actaatcagg tgctgtcag ccacatcagg tgctgaaatg 8040
 30 tcacactgtg ggctatataa gcagcgcgat tttccgggtt cgggtcttc ctgagaagta 8100
 agcaataaaa gctt 8114

<210> 2
 <211> 97
 <212> PRT
 35 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: protein
 sequence of pPL5

40 <220>
 <221> UNSURE
 <222> (37)..(97)
 <223> /note="Xaa on positions 37, 40, 92, 97 is
 unsure/undefined"

45 <400> 2
 Val Gly Thr Gln Met Pro Arg His Leu Ala Thr Gln Leu Val Lys Pro
 1 5 10 15
 Ser Tyr Ile Thr Met Ile Lys Ile Leu Asn Ile Tyr Leu Thr Phe Leu
 20 25 30
 50 His Arg Ser Ala Xaa Glu Gly Xaa Gln Met Leu Ser Gly Glu Ile Leu
 35 40 45
 Lys Ala Asp Val Arg Pro Val Thr Ser Leu Gln Ile Ile Ser Leu His
 50 55 60
 55 Lys His Leu His Gly Leu Cys Arg Leu Ser Ser Leu Pro Lys Gly Thr
 65 70 75 80

EP 1 148 128 A1

Ile Xaa Ile Val Phe Ser Cys Gln Ser Lys Ser Xaa Lys Asp Asn Phe
 85 90 95

5 Xaa

<210> 3
 <211> 55
 <212> PRT
 <213> Artificial sequence

10

<220>
 <223> Description of Artificial sequence: protein sequence of pPL5

15

<220>
 <221> UNSURE
 <222> (16)..(55)
 <223> /note="Xaa on positions 16, 27, 55 is unsure/undefined"

20

<400> 3
 Ala Gly Glu Met Ala Gln Arg Leu Arg Ala Pro Thr Ala Leu Pro Xaa
 1 5 10 15

Val Leu Ser Ser Asn Pro Ser Asn His Met Xaa Leu Thr Thr Ile Arg
 20 25 30

25

Asn Glu Ile Arg Gln Leu Leu Leu Val Cys Leu Lys Ser Ala Thr Val
 35 40 45

Tyr Leu Asp Ile Ile Ile Xaa
 50 55

30

<210> 4
 <211> 335
 <212> DNA
 <213> Artificial sequence

35

<220>
 <223> Description of Artificial Sequence: 5' part of sequence of PCR product at integration site 1

40

<400> 4
 tgaaccacac aactaaattc ctgcttttca caaagtgggt cacagaacaa aatgtagaat 60
 aaaaaaaagc agcaacttct ctagaaaaat cttcgaggag aaaaatccag gtggccctgg 120
 gtaaagcaag gaaccgtgaa ggacagcaca gagaatgagg gaatgagggg agccaaaact 180
 gacaattcta ggagcgtaa agacttctgc tctgcctgac acagaagtaa gggataaaga 240
 aaaccacaga gtggagaggg tattgcaaga gacactcata aaggtggttg agacctggac 300
 aatgttcgtg agatgtataa tgtatgaact gaagg 335

45

<210> 5
 <211> 34
 <212> DNA
 <213> Artificial sequence

50

<220>
 <223> Description of Artificial Sequence: 3' part of sequence of PCR product at integration site 1

<400> 5
 ggcgaactca tgcagacatg gggataaact agtt 34

55

EP 1 148 128 A1

<210> 6
<211> 328
<212> DNA
<213> Artificial Sequence

5

<220>
<223> Description of Artificial Sequence: sequence of
PCR product at integration site 2

<400> 6

10 gttattgaga ggtcgtatga aaacctattc ttgtagaagc ttcttaaaat atatccacgt 60
atgaaagaaa tctaaataga gtcaccaaata tacaaggaag acaattttca ccaagtgaaa 120
cttccagtgc caggaatagg ttgcatctaa ctgagtcatt ggcaaaaggg ccccatggaa 180
agccccaat aatccaggct gttgccaagg ctactggttg ctctctacca acggaaggta 240
agtccctatt gctaaaacat ttatgtcctt attgaacatg gagagctcag ctggtgccta 300
actaaggcct ttgcccctat tgactagt 328

15

20

25

30

35

40

45

50

55

Claims

1. An isolated nucleic acid which defines a yellow submarine locus.
- 5 2. An isolated nucleic acid which defines a mutant yellow submarine locus, wherein the mutant yellow submarine locus is identical to a wildtype yellow submarine locus except for an integration of a pAA2 transgene into at least one region on a chromosome.
- 10 3. The isolated nucleic acid of claim 2, wherein the chromosome is mouse chromosome 3.
4. The isolated nucleic acid of claim 3, wherein the pAA2 transgene is inserted into a region of mouse chromosome 3 designated A2.
- 15 5. The isolated nucleic acid of claim 3, wherein the pAA2 transgene is inserted into a region of mouse chromosome 3 designated B-C.
6. The isolated nucleic acid of claim 2, further comprising a rearrangement of chromosomal sequences of a region of the chromosome designated A3 region.
- 20 7. The isolated nucleic acid of claim 6, wherein the rearrangement comprises an inversion of a nucleotide segment.
8. The nucleic acid of claim 1 or 2 wherein the nucleic acid is genomic DNA.
9. The nucleic acid of claim 1 or 2 wherein the nucleic acid is RNA.
- 25 10. The nucleic acid of claim 1 or 2 wherein the nucleic acid is cDNA.
11. The nucleic acid of claim 1 or 2, wherein the nucleic acid is labeled with a detectable marker.
- 30 12. The nucleic acid of claim 6, wherein the marker is a radioactive, a colorimetric, a luminescent, or a fluorescent label.
13. A replicable vector comprising the nucleic acid of claim 1 or 2.
14. A host cell comprising the vector of claim 8.
- 35 15. The host cell of claim 9 wherein the cell is a eukaryotic cell.
16. The host cell of claim 9 wherein the cell is a bacterial cell.
- 40 17. The vector of claim 8 wherein the vector is a plasmid.
18. The vector of claim 8 wherein the vector is a cosmid.
19. The vector of claim 8 wherein the vector is a λ phage.
- 45 20. The vector of claim 8 wherein the vector is a YAC.
21. The vector of claim 8, wherein the vector is a BAC.
- 50 22. The vector of claim 8, wherein the vector is a PAC.
23. A nucleic acid of at least 14 nucleotides capable of specifically hybridizing with the nucleic acid of claim 1.
24. A nucleic acid of at least 14 nucleotides capable of specifically hybridizing with the nucleic acid of claim 2.
- 55 25. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a growth factor.
26. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a growth factor receptor.

EP 1 148 128 A1

27. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes an orphan receptor.
28. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a signaling molecule.
- 5 29. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a transcriptional regulator.
30. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes an intracellular transport protein.
31. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a neural precursor cell which is an expressed
10 and developmentally down-regulated 4 (NEDD4) family molecule.
32. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a SOX protein.
33. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a regulator of apoptosis.
15
34. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a regulator of protein turnover.
35. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a cell cycle regulator.
- 20 36. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a calcium binding protein.
37. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a potentiator of hormone dependent activation
of transcription by progesterone or glucocorticoid receptors.
- 25 38. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a membrane transport protein.
39. The nucleic acid of claim 1 or 2, wherein the nucleic acid encodes a co-activator of transcription.
40. The nucleic acid of claim 1 or 2, wherein the nucleic acid is isolated from a mouse.
30
41. An isolated nucleic acid which defines a human locus which corresponds to a yellow submarine locus.
42. An isolated nucleic acid which defines a human locus which corresponds to a mutant yellow submarine locus
containing a pAA2 transgene integrated into at least one region the genome.
35
43. A method of diagnosing inner ear dysfunction in a subject comprising:
- a) obtaining a suitable sample from a subject;
 - b) extracting nucleic acid from the sample;
 - 40 c) contacting the nucleic acid with the nucleic acid of claim 24 which binds specifically to the mutated portion
of the ysb locus; and
 - d) detecting the labeled nucleic acid, thereby detecting mutant ysb and thereby diagnosing inner ear dysfunction
in the subject.
- 45 44. A method of diagnosing pigmentation dysfunction in a subject comprising:
- a) obtaining a suitable sample from a subject;
 - b) extracting nucleic acid from the sample;
 - 50 c) contacting the nucleic acid with the nucleic acid of claim 24 which binds specifically to the mutated portion
of the ysb locus;
 - d) detecting the labeled nucleic acid, thereby detecting mutant ysb and thereby diagnosing pigmentation dysfunction
in the subject.
- 55 45. A method of diagnosing cell growth dysfunction, cell proliferation dysfunction, or cell death dysfunction in a subject
comprising:
- a) obtaining a suitable sample from a subject;
 - b) extracting nucleic acid from the sample;

EP 1 148 128 A1

- c) contacting the nucleic acid with the nucleic acid of claim 24 which binds specifically to the mutated portion of the ysb locus; and
- d) detecting the labeled nucleic acid, thereby detecting mutant ysb and thereby diagnosing cell growth dysfunction or proliferation dysfunction in the subject.

5

46. The method of any one of claims 43-45, wherein the subject is mammal or non-mammal.

47. The method of any one of claims 43-45, wherein the subject is a human, a primate, an equine, an opine, an avian, a bovine, a porcine, a canine, a feline or a murine.

10

48. The method of any one of claims 43-45, wherein the subject is a vertebrate.

49. A polypeptide encoded by the nucleic acid of claim 1, 2, 41 or 42.

15

50. A fusion protein comprising the polypeptide of claim 49.

51. The polypeptide of claim 49, wherein the polypeptide is labeled with a detectable marker.

20

52. A method of repairing and regenerating nerve tissue in a subject comprising administering an effective amount of the protein of claim 49 to the subject, wherein the protein is a wildtype protein, so as to thereby repair and regenerate nerve tissue in the subject.

25

53. A method of regulating cell migration in a subject comprising administering an effective amount of the protein of claim 49 to the subject, wherein the protein is a wildtype protein, so as to thereby regulate cell migration in the subject.

54. A method of regulating cell growth in a subject comprising administering an effective amount of the protein of claim 49 to the subject, wherein the protein is a wildtype protein, so as to thereby regulate cell growth in the subject.

30

55. An antibody which binds to the polypeptide of claim 49.

56. The antibody of claim 55, wherein the antibody is a monoclonal antibody.

57. The antibody of claim 55, wherein the antibody is conjugated to a cytotoxic agent.

35

58. The antibody of claim 55, wherein the antibody is labeled with a detectable marker.

59. A composition comprising the antibody of claim 55.

40

60. A method of producing a protein encoded by a nucleic acid in a wildtype ysb locus which comprises growing a host vector system under suitable conditions permitting production of the polypeptide and recovering the polypeptide so produced.

45

61. A method determining vestibular dysfunction in a subject comprising:

- a) obtaining a suitable sample from a subject;
- b) extracting nucleic acid from the sample;
- c) contacting the nucleic acid with the nucleic acid of claim 24 which binds specifically to a mutated portion of the ysb locus;
- d) detecting the labeled nucleic acid, thereby detecting gene responsible for yellow coat color, thereby indicating the presence of vestibular dysfunction in the subject.

50

62. A method determining vestibular dysfunction using embryonal stem cells:

55

- a) obtaining a suitable sample from the embryonal stem cells;
- b) extracting nucleic acid from the sample;
- c) contacting the nucleic acid with the nucleic acid of claim 24 which binds specifically to a mutated portion of the ysb locus;

EP 1 148 128 A1

d) detecting the labeled nucleic acid, thereby detecting gene responsible for yellow coat color, thereby indicating the presence of vestibular dysfunction in the embryonal stem cells.

5 **63.** A method of determining successful deletion or inactivation of a gene which comprises examining coat color or pigmentation of a subject, wherein yellow coat color or pigmentation indicates that a normal gene has been inactivated, thereby resulting in yellow coat color or pigmentation, thereby indicating the successful deletion or inactivation of the gene.

10 **64.** A method of determining successful deletion or inactivation of a gene comprising determining whether repairing the abnormal gene results in the disappearance of the yellow coat color or pigmentation.

65. A method of treating vestibular dysfunction in a subject comprising introducing a nucleic acid comprising a gene or genes from the wildtype locus encoding a ysb into a suitable cell under conditions such that the nucleic acid expresses ysb so as to thereby treat vestibular dysfunction.

15 **66.** A method of treating hearing impairment in a subject which comprises introducing a nucleic acid comprising a gene or genes from the wild-type locus into a suitable cell under conditions such that the nucleic acid expresses ysb so as to thereby treat hearing loss.

20 **67.** A transgenic mouse line designated KM12.

68. A method of determining whether a compound is mutagenic comprising:

- 25 (a) examining the coat color or pigmentation of a subject;
(b) administering the compound to the subject;
(c) examining the coat color or pigmentation of a subject;
(d) comparing the result obtained in step (c) with the result obtained in step (a); and
(e) determining if the coat color or pigmentation of the subject is yellow so as to thereby determine whether the compound is mutagenic.

30

35

40

45

50

55

Figure 1



Figure 3
KM12 Line (*ysb*)

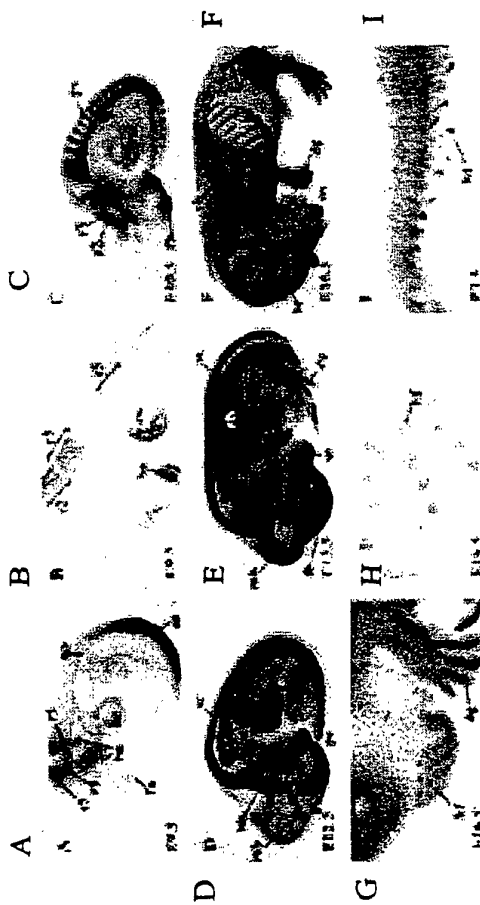


Figure 2
Typical pAA2 line



Figure 4

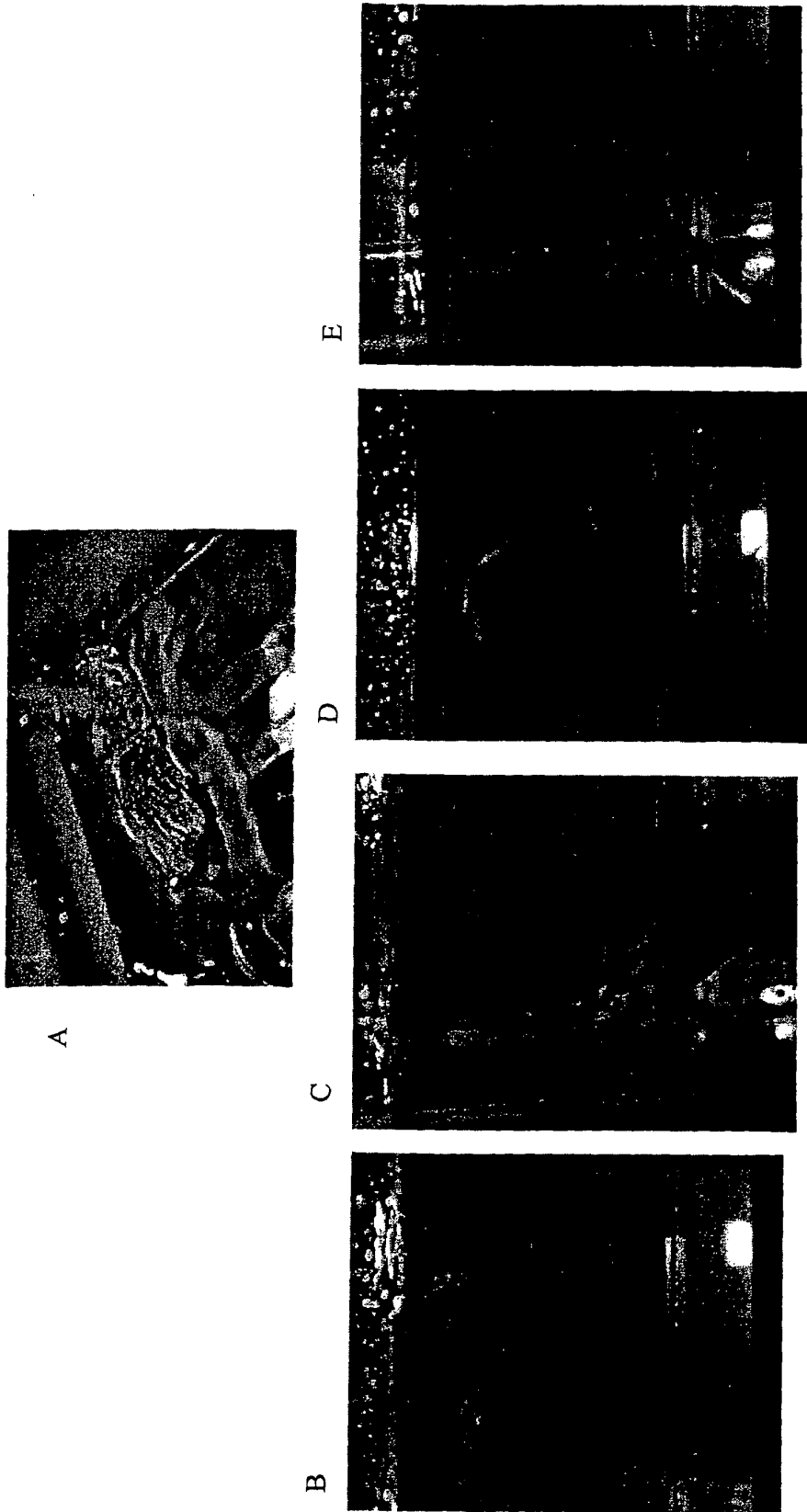


Figure 5



Figure 7

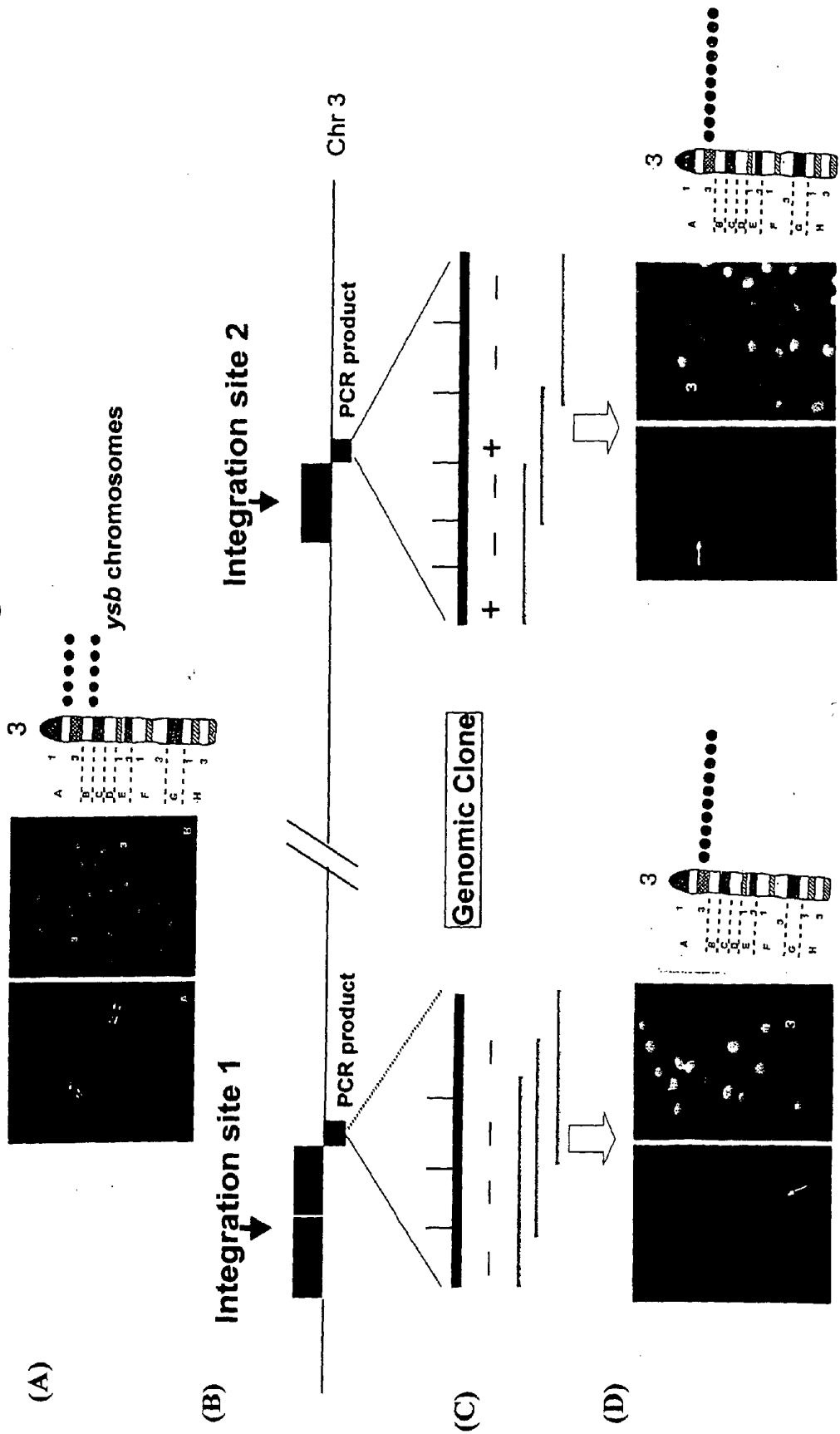


Figure 8

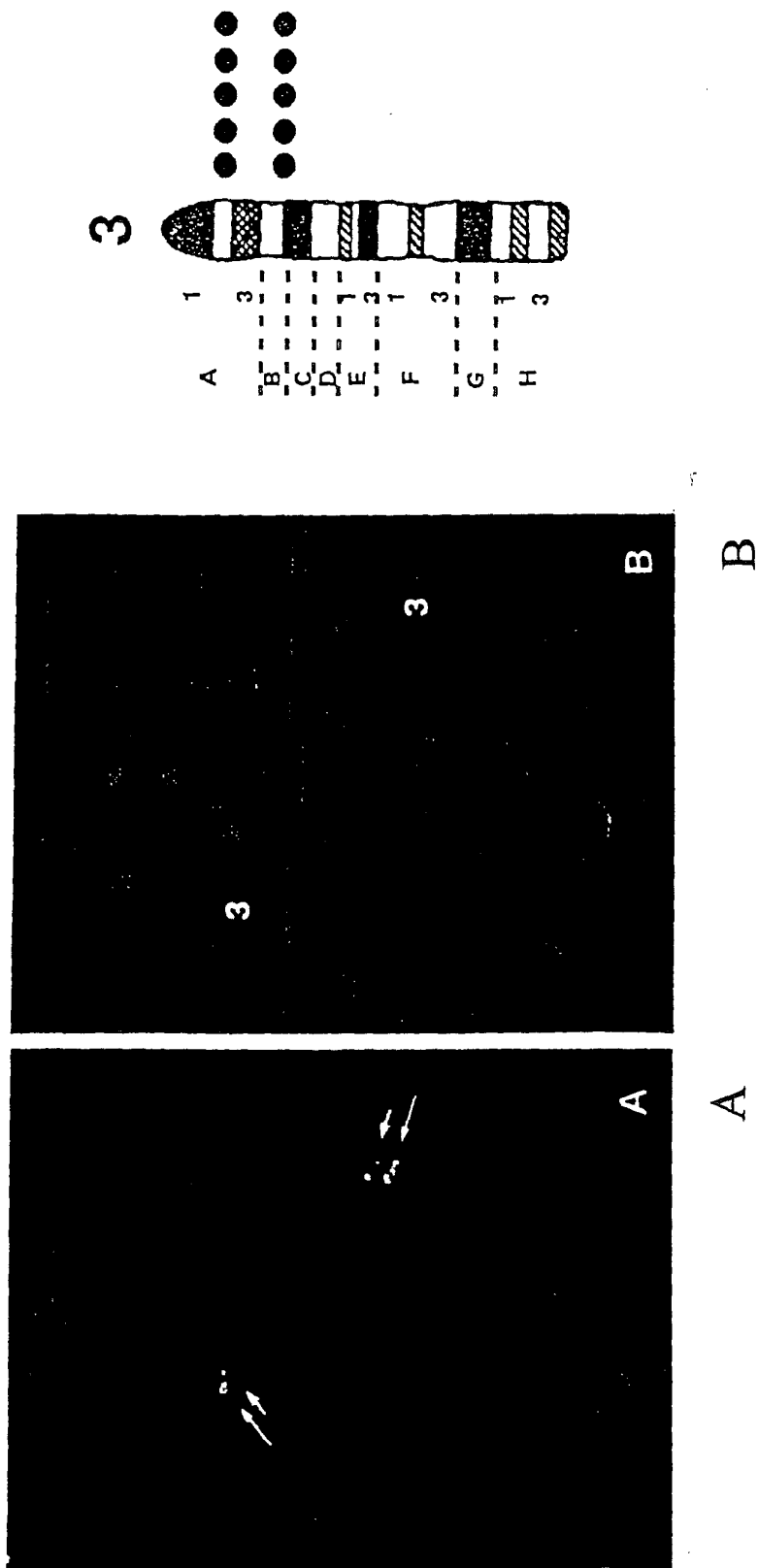


Figure 9

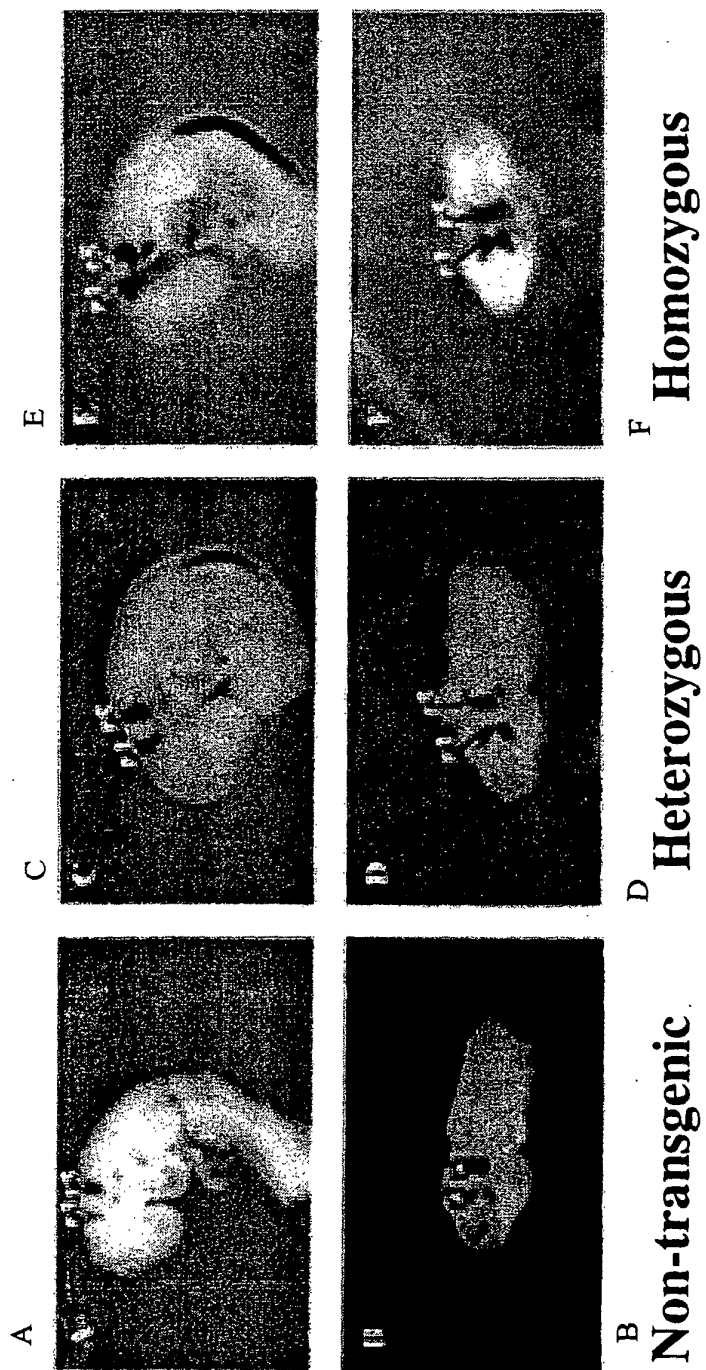


Figure 10

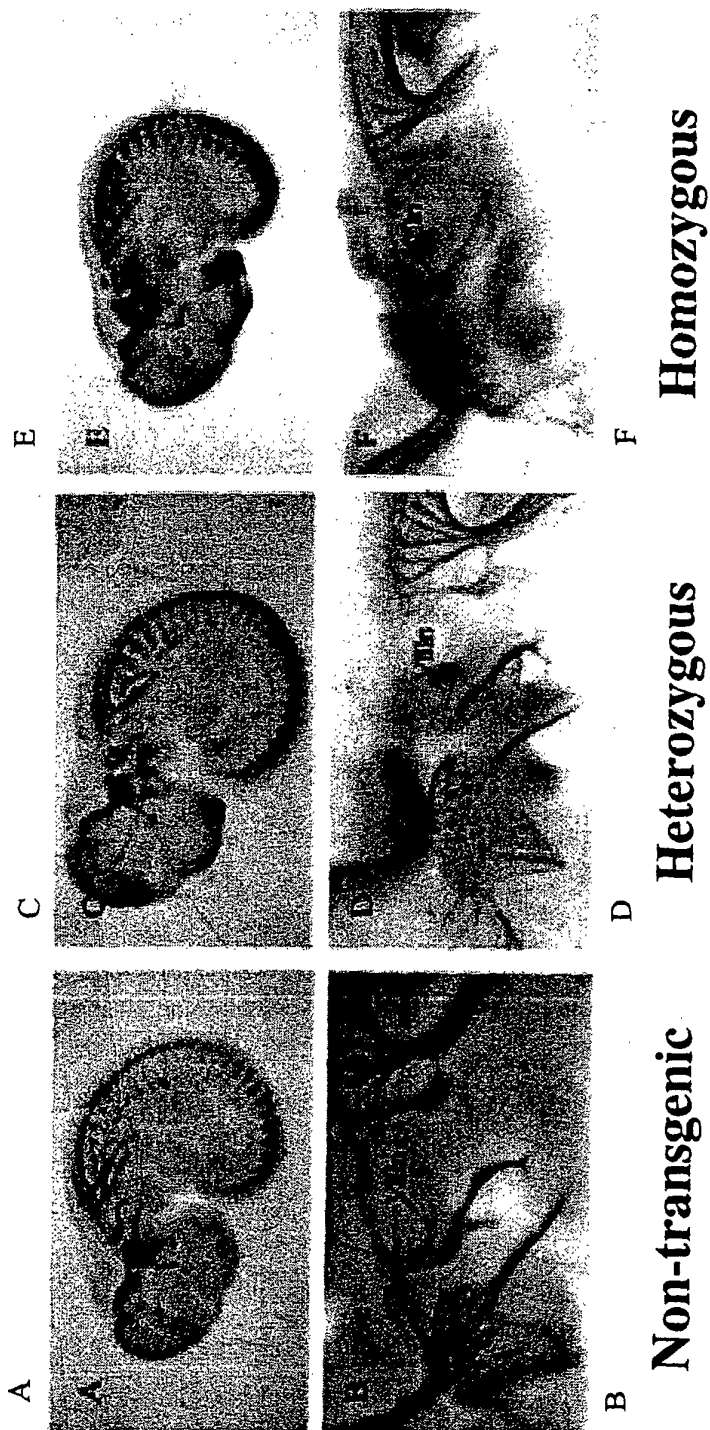


Figure 11

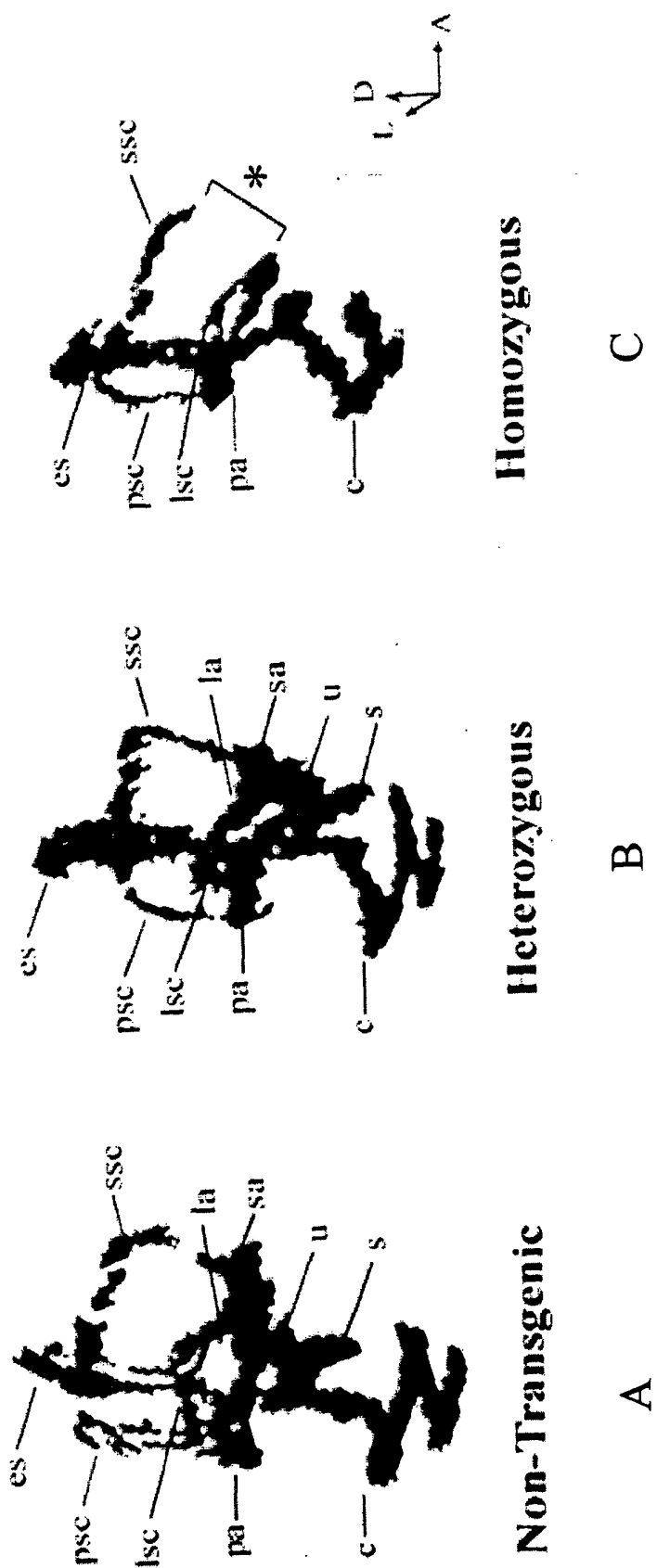


Figure 12

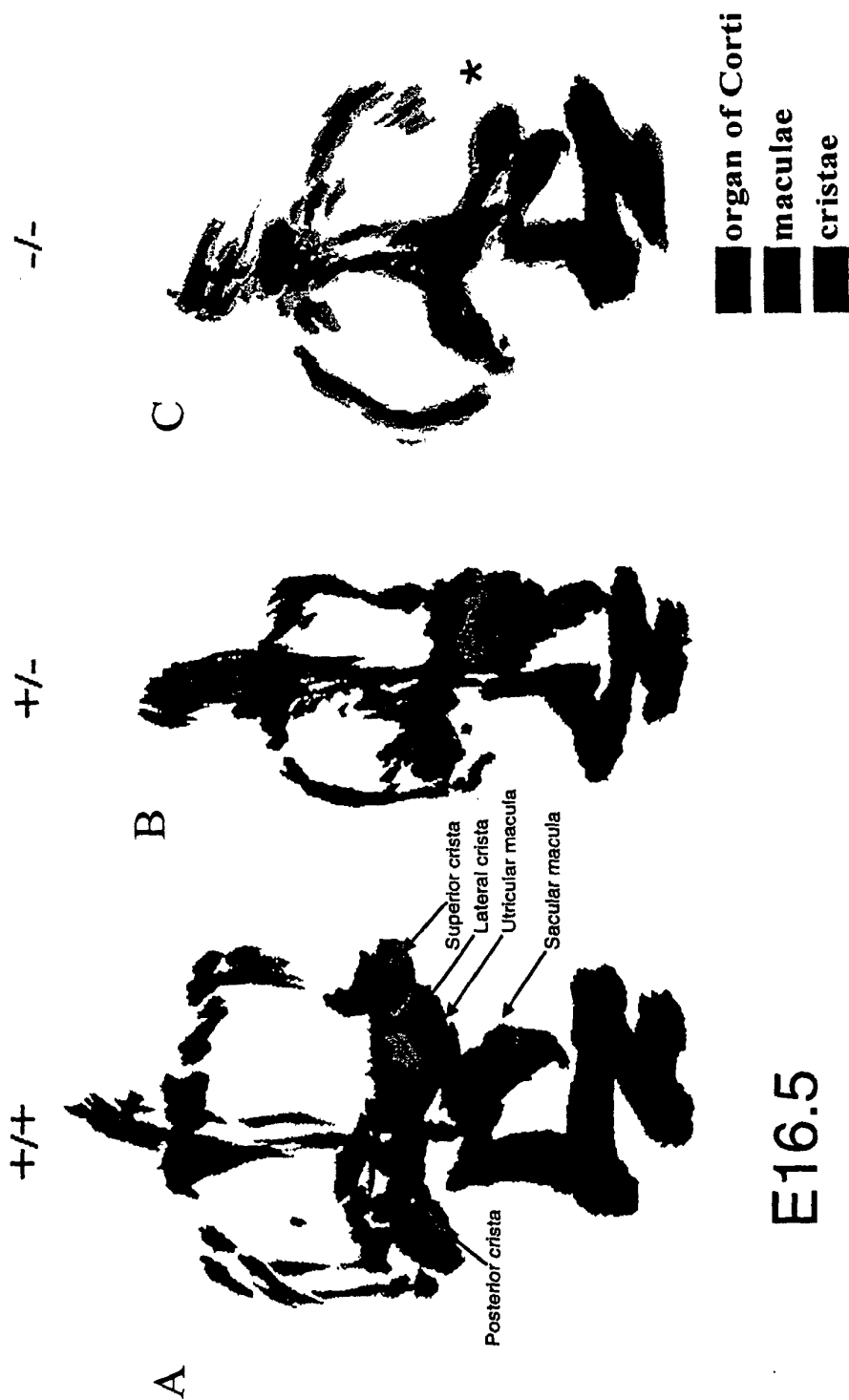


Figure 13

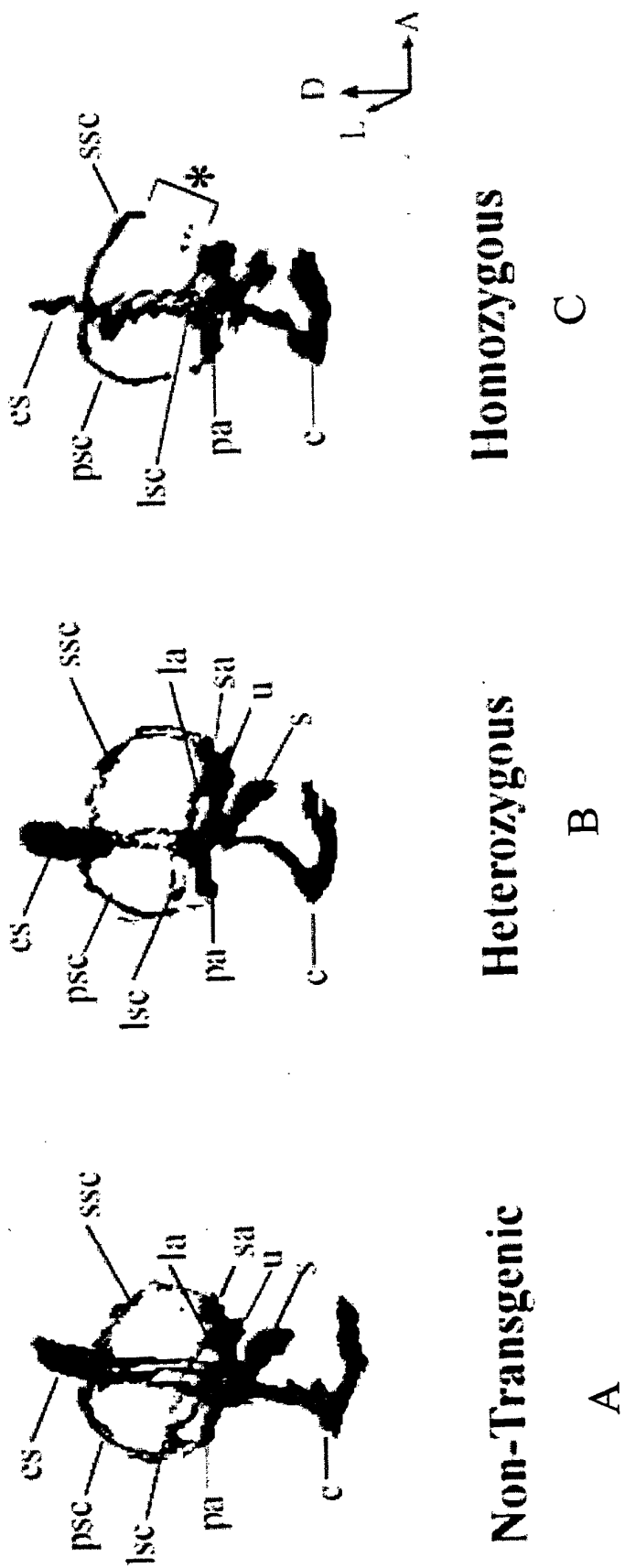


Figure 14

Sequence of PCR product at integration site 1

TGAACCAACAATAATTCC TGCTTTTCACAAAAGTGGGTACAGAACAAAATGTAGAAATAAAAAAAGCAGCAACTTCTC
 TAGAAAATCTTCGAGGAGAAAAATCCAGGTGGCCCTGGGTAAAGCAAGAACCCGTGAAGGACAGCACAGAGAATGAGGGA
 ATGAGGGAA GCCAAA ACTGACAA TTTCTAGGAGCGTTAAAGACTTCTGCTCTGACACAGAAAGTAAGGGATAAAGAAAA
 CCACAGAGTGGAGAGGGTATTGCAAGAGACACTCA TAAAGGTGGTTGAGACCTGGACAATGTTCGTGAGATGTATAATGTA
 TGAAC TGAAGG.....GGCGAACTCATGCAGACATGGGGATAAACTAGTT

A

Sequence of PCR product at integration site 2

GTATTGAGAGGTCGTATGAAAAACCTATTCTTGTAGAAGCTTCTTAAATAATATCCACGTTATGAAAAGAAATCTAAAATAGAGT
 CACCAAATTACAAGGAAGACAAATTTTCA CCAAGTGAAA CTTCAGTGGCCAGGAAATAGGTTGCATCTAACTGAGTCA TTGGCA
 AAAGGGCCCATGGAAAGCCCAAAATAATCCAGGCTGTTGCCAAGGCTACTGGTTGCTCTCTACCAACGGAAGGTAAGTCCC
 TATTGCTAAAACA TTTATGTCTTATTGAACATGGAGAGCTCAGCTGGTGCCTAAC TAAAGGCC TTTGGCCCTATTGACTAGT

B

Figure 15

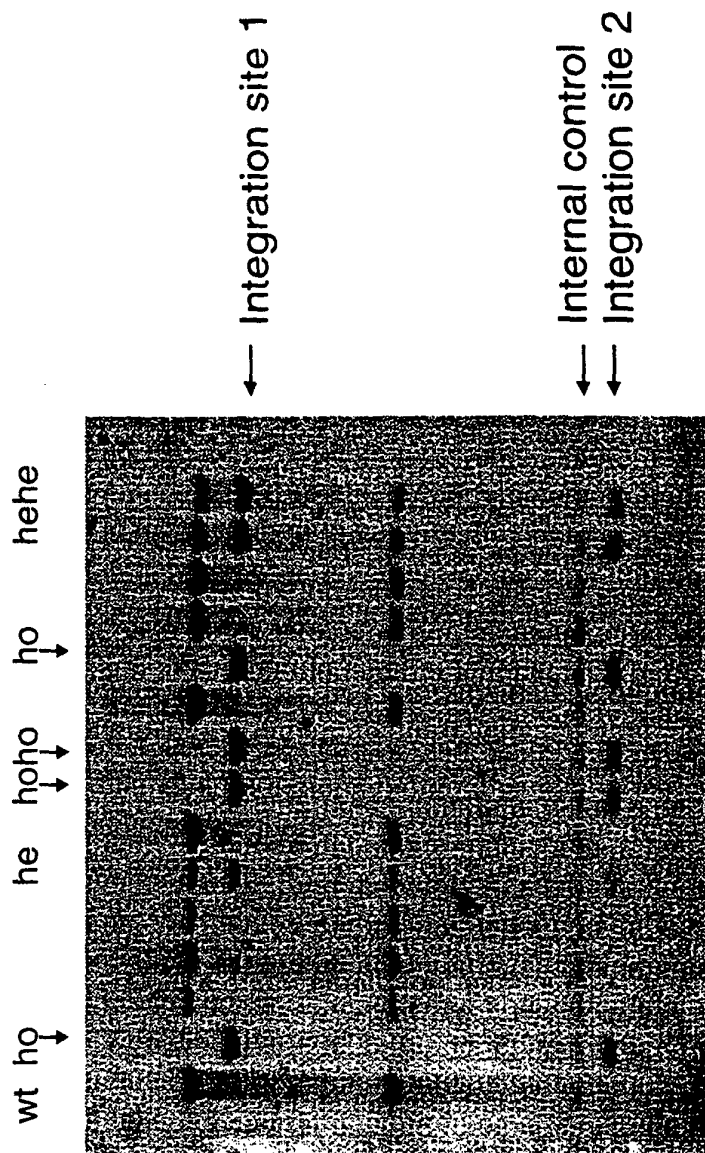


Figure 16

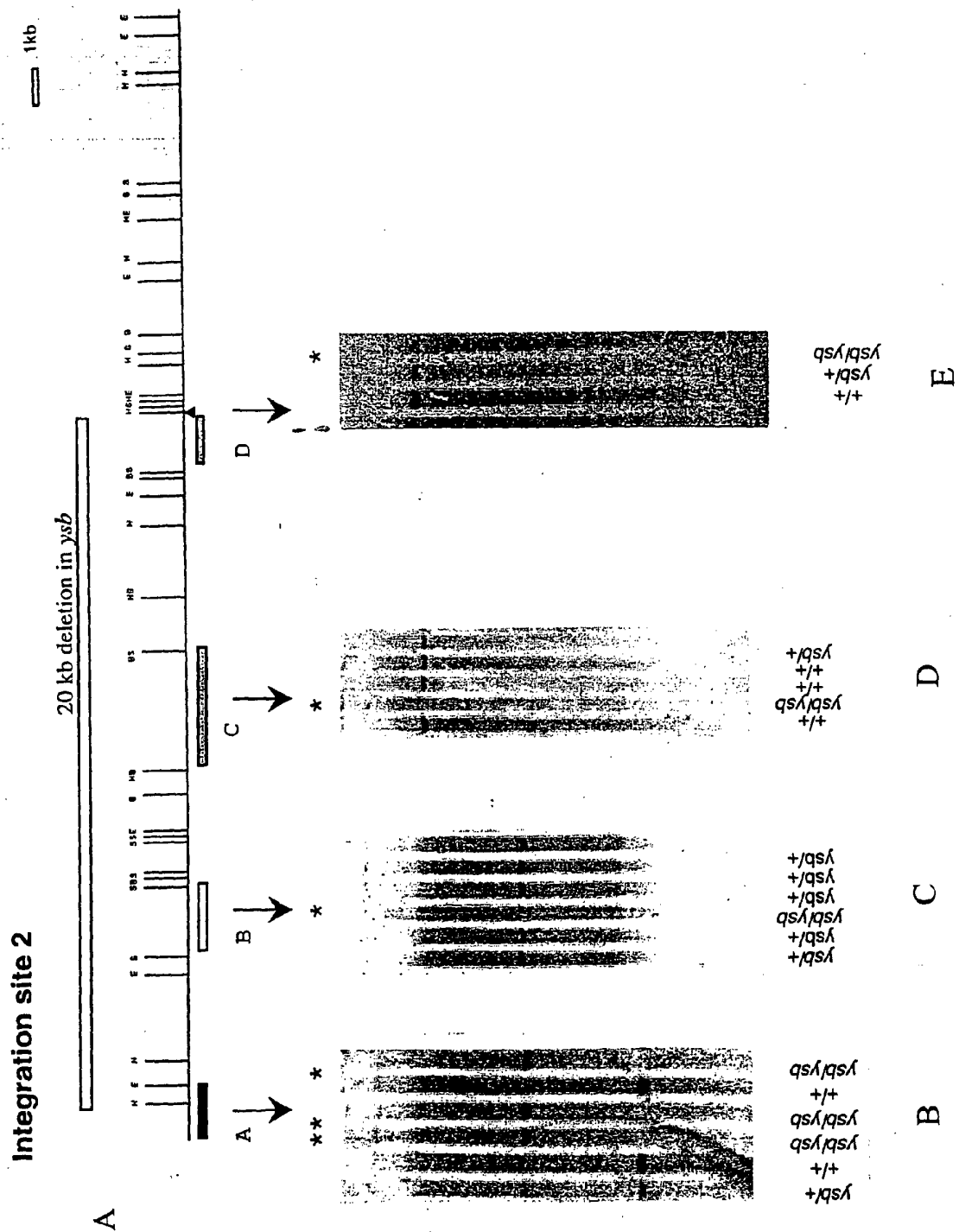


Figure 17-1

Sequences of pPL5(8.1kb HindIII fragment of 20kb deleted sequences in integration site 2)

Sequence Range: 1 to 8114

```

>HindIII
|
|      10      20      30      40      50      60
1 AAGCTTAAATGAGCCAAGCGCAGAGTGAAGCTGAAGTTGAGAGACACCGGCGCTGTGCC
2 TTCGAATTTTACTCGGTTGCGCTCACTTCGACTTCAACTCTCTGTGGCCGCGACACGG

      70      80      90      100     110     120
CAGAACACAGTCCGAACCTTAACATTTGAATGGAGCAACACAAAAGTGGGAACCCAGATG
GTCTTGTGTGAGCTTGGAAATTGTAACCTACCTCGTTGTGTTTTCCACCTTGGGTCTAC
      V G T Q M>
      IMAGE 63  >

      130     140     150     160     170     180
CCGAGACATCTAGCGACACAGTTGGTAAAGCCTTCTTACATAACAATGATAAAAAATACTT
GGGTCTGTAGATCGCTGTGTCAACCATTTGGAAGAATGTATTGTTACTATTTTTATGAA
P R H L A T Q L V K P S Y I T M I K I L>
      IMAGE 636095 >

      190     200     210     220     230     240
AACATTTATTTAACATTTTACACAGGAGTGCATGAGAAGGTTGACAGATGTTATCTGGT
TTGTAATAAATTTGTAATAAATGTGTCCTCAGTACTCTTCCAAGTGTCTACAATAGACCA
N I Y L T F L H R S A * E G * Q M L S G>
      IMAGE 636095 >

      250     260     270     280     290     300
GAGATTCTCAAAGCCGATGTGAGACCAGTCACTCTTTACAAATTATCTCTTCAAAA
CTCTAAGAGTTTCGGCTACACTCTGGTCAAGTGGAGAAATGTTAATAGAGAGAAGTGT
E I L K A D V R P V T S L Q I I S L H K>
      IMAGE 636095 >

      310     320     330     340     350     360
CATCTACATGGGCTTTGCGAGGCTGAGCAGTCTGCCCAAAGGAACCATTTAAATAGTTTT
GTAGATGTACCCGAAACGTCGACTCGTCAGACGGGTTTCCCTGGTAAATTTATCAAAG
H L H G L C R L S S L P K G T I * I V F>
      IMAGE 636095 >

      370     380     390     400     410     420
TCATGCCAAAGTAAAGTAAAAAGACAACCTTTTTTTTAAAAAAAAAAAAAAAAAGTGAAGT
AGTACGGTTTCATTTTCAATTTTCTGTTGAAAAAAATTTTTTTTTTTTTTTTCACTTCA
S C Q S K S * K D N F X>
      IMAGE 636095 >

      430     440     450     460     470     480
AAATGACTTAAGGTTTAGTTTACTTCTGTTAGGAGAGATTCAACCGGTTGGGAAAGGGGT
TTTACTGAATTCAAATCAAATGAAGACAATCCTCTCTAAGTTGGCCAACCCCTTTCCCCA

      490     500     510     520     530     540
GTGCGCTTGTGTGTGAGACCAGTGAAGCGAAATGCTGGATACGATTGTCTCCTCAGTAAA
CACGCGAACACACACTCTGGTCACTCTCGCTTTACGACCTATGCTAACAAGGAGTCATT

      550     560     570     580     590     600
AGAGAGGAACTGGAAATGACTAATATTCCATAGAATGATAAGCTCCCAACTTTTAGCATT
TCTCTCCTTGACCTTTACTGATTATAAGGTATCTTACTATTTCGAGGGTTGAAATCGTAA

      610     620     630     640     650     660
CTTCTCTATTTGTCTATTATGTTTCTCTACAAGACAAAAAATAATTTTCAACTTACAAG

```


17-2

GAAGAGATAAACAGATAATACAAAGAGATGTTCTGTTTTTTTATTTAAAAGTTGAATGTTT
 670 680 690 700 710 720
 TCTTATTGTTTACACTTAATTTTCAAGACTTGTACCTAGCTGATGAAACAGATGTTCCCC
 AGAATAACAAATGTGAATTTAAAAGTTCTGAACATGGATCGACTACTTTGTCTACAAGGGG
 730 740 750 760 770 780
 AACACCAGATGATCTCAGGTGGACTTACTGGACAGTGTGGAGATCTGTGTCCCTTCT
 TTGTTGGTCTACTAGAGTCCACCTGAATGACCTGTCACACACCTCTAGACACAGGGAAGA
 790 800 810 820 830 840
 AAGGTCATCCGTTGGCACCATTGCCGTGGAGGCTAAGAGATGACACATAGACAGACCGCAG
 TTCCAGTAGGCAACCGTGGTAACGGCACCTCCGATTCTACTGTGTATCTGTCTGCGTC
 850 860 870 880 890 900
 AATCTCATGAGTCAGACTGGTGCCTTGCCTCCTCAATTTTTGTTTTTCATTTTTTTGAG
 TTAGAGTACTCAGTCTGACCACGGGAACGGAGGAGTTAAAAACAAAAGTAAAAAACTC
 910 920 930 940 950 960
 AAGAGATCTCATACTCCCAGGCTGGCCTTGAACCTCACTTTGTAGCTGAGGCTGGCTTTAT
 TTCTCTAGAGTATGAGGGTCCGACCGGAACCTGAGTGAACATCGACTCCGACCGAAATA
 970 980 990 1000 1010 1020
 ACTGATGATCCTCCTGCCTCCACCCCTCACATGCTGGGATTCCGGCAGTGTGTCACCTG
 TGACTACTAGGAGACGGAGGTGGGGAGTGTACGACCTAAGGCCGTACACAGTGGAC
 1030 1040 1050 1060 1070 1080
 CCCAGGGGATAGTTCCCTCTTAGAAGAAGCACCAGGGAGTTAGTTTCTCCTTTATCCTAT
 GGGTCCCCTATCAAGGGAGAATCTTCTTCGTGGTCCCTCAATCAAAGAGGAAATAGGATA
 1090 1100 1110 1120 1130 1140
 GTGAGGGTTTGTGAGGCTGTCTTCTATGAATCATAAAGTAGGCCAGATGCTGAAACTGC
 CACTCCCAAACGACTCCGACAGAAGATACTTAGTATTTTCATCCGGTCTACGACTTTGACG
 1150 1160 1170 1180 1190 1200
 CACTGCCCTGTCCACGTTTGGAACTAGAAGAACTATTTTCTGGGGTTTGTATGCTACC
 GTGACGGGGACAGGTGCAAACCTTGATCTTCTTTGATAAAAGACCCCAAACATACGATGG
 1210 1220 1230 1240 1250 1260
 CAGCCTGTGGTGTGTTGTGTAGCAACTTGGATGGCTAAAGACAATGGATTGAGCTGCCCT
 GTCGGACACCACAAAACAACATCGTTGAACCTACCGATTCTGTGTACTAACTCGACGGA
 1270 1280 1290 1300 1310 1320
 TCTGTTCTACAACCTGCACTGACCACATGTGACCATTTAGCACTTGACACTTGACTATAGT
 AGACAAGATGTTGACGTGACTGGTGTACACTGGTAAATCGTGAACCTGTGAACCTGATATCA
 1330 1340 1350 1360 1370 1380
 GGACTGAGGGAATGATTCTTGCTTTTCTTTTCTTTCTTTTCGGTTTTTTTTTTTTTTTTT
 CCTGACTCCCTTACTAAGAACGAAAAGAAAAGAAAAGAACCCAAAAAATAAAAAAAAAA
 1390 1400 1410 1420 1430 1440
 TTTGAGACAGGGTTTCTGTGTAGCCCTGGCTGTCCCTGGATCTCACTCTGTAGAGTAT
 AAAGCTCTGTCCCAAAGACACATCGGGACCGACAGGACCTAGAGTGAGACATCTCATA
 1450 1460 1470 1480 1490 1500
 GCTGGCCAATTCTAATTAGATTAGTGTCTTATGACTAGTAACTACTACAGTAGCAAATCT
 CGACCGGTTAAGATTAATCTAATCACAGAATACTGATCATTGATGATGTCATCGTTTTAGA
 1510 1520 1530 1540 1550 1560
 TGGTTGTCAACTTGACTACATCTGGAATCAACTAAAACCCAAGATAATAGGTGTCACACA
 ACCAACAGTTGAACTGATGTAGACCTTAGTTGATTTTTGGGTTCTATTATCCACAGTGTGT

17-3

1570 1580 1590 1600 1610 1620
CACAGACACACACACACACAGATAGAGACATATGTTTAAACATAGTTAAAAATAAAC
GTGTCTGTGTGTGTGTGTCTATCTCTGTATACAAATTGTGTATCAATTTTTTATTTG

1630 1640 1650 1660 1670 1680
TCTTTAAAAGAACATGTATCACGGGGCAAGTTAAATCATAATTCTGAATACTGTACTTCT
AGAAATTTTCTTGTACATAGTGCCCCGTTCAATTTAGTATTAAGACTTATGACATGAAGA

1690 1700 1710 1720 1730 1740
CTGTAAATTAGAGATGGCACTGTTCCACTGGGAGAAGATATTAAGAGTTTGCTAGTG
GACATTTAATCTCTACCGTGACAAGTAGGTGACCCTCTTCTATAATTCTCAAACGATCAC

1750 1760 1770 1780 1790 1800
GGGCTGGAGAGATGGCTCAGCGGTTAAGAGCACCGACTGCTCTTCCTTAGGTCCTGAGTT
CCCGACCTCTCTACCGAGTCGCCAATTCCTGGCTGACGAGAAGGAATCCAGGACTCAA
A G E M A Q R L R A P T A L P * V L S>

IMAGE 1196866 >

1810 1820 1830 1840 1850 1860
CAAATCCAGCAACCACATGTAGCTCACAAACATCCGTAATGAGATCCGACAACCTCTTC
GTTTAGGGTCGTTGGTGTACATCGAGTGTGGTAGGCATTACTCTAGGCTGTTGAAGAAG
S N P S N H M * L T T I R N E I R Q L L>

IMAGE 1196866 >

1870 1880 1890 1900 1910 1920
TGGTGTGTCTGAAGTCAGCTACAGTGTACTTAGATATAATAATAAAATAAACTAAAAAT
ACCACACAGACTTCAGTCGATGTCACATGAATCTATATTATTATTATTAGATTTTTTA
L V C L K S A T V Y L D I I I X>

IMAGE 1196866 >

1930 1940 1950 1960 1970 1980
AAAAACATTTGCTAATGCCAAAACAGACATATCAAGAGTACAAAATAAAAACCTCCAAAC
TTTTGTAAACGATTACGGTTTTGTCTGTATAGTTCTCATGTTTTATTTTTGAAGGTTTG

1990 2000 2010 2020 2030 2040
TAAAAAGATGCCACCAGAAATAGAGTATCTGGGTGTGCATGTGGCATAGTAGCACACCTG
ATTTTTCTACGGTGGTCTTTATCTCATAGACCCAACAGTACACCGTATCATCGTGTGGAC

2050 2060 2070 2080 2090 2100
CCAAACCTGGCAGGTATGAGGCCCTGCATCCCCACATGAGAACTAAAATTTTCCAATAT
GGTTTGACCGTCCATACTCCGGGACGTAGGGGGTACTCTTGATTTTAAAAGGTTATA

2110 2120 2130 2140 2150 2160
GAGATGAAACATATTTAAACAAGATGAAACTCACAAAACATGAAGGAGTTTCTCCTCAA
CTCTACTTTGTATAATTTTGTCTACTTTGAGTGGTTTTGACTTCTCAAAGAGGAGTT

2170 2180 2190 2200 2210 2220
CTTTAAGGATTTGGTAATTGATCTCCACATAAAAATATGTTTGTGGGCTAGCAAGATGG
GAAATTCCTAAACCATTAAGTAGAGGGTGTATTTTATACAAAACCCCGATCGTTCTACC

>EcoRI

2230 2240 2250 2260 2270 2280
CTCAGAGCAAAGCACTTGTCATGCAGGCCCGACGACCTGAATTCAGTGCCAAGAGCATG
GAGTCTCGTTTTCGTGAACAGTACGTCCGGGCTGCTGGACTTAAGTCACGGTTCTCGTAC

2290 2300 2310 2320 2330 2340
CAACGGAAGGCGAAGGCTGCAAATCATTCTCTGACCTCTACATATGCATCATGGCATGGG
GTTGCCTCCGCTCCGACGTTTAGTAAGAGACTGGAGATGTATACGTAGTACCGTACCC

17-4

2350 2360 2370 2380 2390 2400
 TACCTACGGTCATAAAAATGTCATAAGCTCATCTACAATAATAAATTATATATGCACATG
 ATGGATGCCAGTATTTTTACAGTATTCGAGTAGATGTTATTATTTAATATATACGTGTAC

2410 2420 2430 2440 2450 2460
 CACACACACATATACACTTGTCAAATCTATAATGTTGGAAACCTTTGCGGTGCATATAGC
 GTGTGTGTGTATATGTGAACAGTTTAGATATTACAACCTTTGGAAACGCACAGTATATCG

2470 2480 2490 2500 2510 2520
 TGATCTGTCTAGGTAGAAATATATCAATTAACAATGCTGATTGAAAACCTCAAATATTA
 ACTAGACAGATCCATCTTTATATAGTTAATTGTTACGACTAACTTTTTGGAGTTATAAT

2530 2540 2550 2560 2570 2580
 TTTTGAATGCTTATTTGAGTCATTTTCCAAGGGATGTGAGCCCGCCATATCTGTGGTT
 AAAAAGTACGAATAAACTCAGTAAAAGGTTCCCTACACTCGGGCGGTATAGACACCAA

2590 2600 2610 2620 2630 2640
 TGCCCATCCATGAATTTGGTTAATCACATATGCATCAAAATAAGCCCTTCAAGAGAGCTA
 ACGGGTAGGTAATAACCAATTAGTGTATACGTAGTTTATTTCGGGAAGTTCTCTCGAT

2650 2660 2670 2680 2690 2700
 TTTTCTTCAGGGCTGTAGCTTCTGGTAAGTCATCCACACCCCTATGCATGCTCCTGTAAG
 AAAAGAAGTCCCGACATCGAAGACCATTAGTAGGTGTGGGGATACGTACGAGGACATTC

2710 2720 2730 2740 2750 2760
 GACCCCAATTAAACTCAGTAGGCCACTGTCCTAGTTTGCTTCTTCTACTGTGATAAAAC
 CTGGGGTTAATTTGAGTCATCCGGTGACAGGATCAAACGAAAGAAGATGACACTATTTTG

2770 2780 2790 2800 2810 2820
 ACTGAGCTAAGCAGCTTGAGGGACAAGGGGTTACTATAGCTTACTGGTTACAGTCCATCA
 TGACTCGATTTCGTCGAACCTCCTGTCCCAATGATATCGAATGACCAATGTCAGGTAGT

>SacI >SacI
 | |
 2830 | 2840 | 2850 2860 2870 2880
 CTGAATGAGCTCAGGACAAGAGCTCAAGGCAGGAATCTGAAGCAGAGACCATGGAGGAAC
 GACTTACTCGAGTCTGTTCTCGAGTTCGGTCCCTAGACTTCGTCTCTGGTACCTCCTTG

2890 2900 2910 2920 2930 2940
 ACTGCTGACTGGCTCACTCCACATGAGGCATTAATCATGCAAATGCTTACAGAGACATGT
 TGACGACTGACCGAGTGGGGTGTACTCCGTAATTAGTACGTTTACGAAGTCTCTGTACA

2950 2960 2970 2980 2990 3000
 ACACAGGTCAGTCTTATTGAGGCAGTTTCTCAACTGAAGCACCTCTTCTCCTGGTGACTC
 TGTGTCCAGTCAGAATAACTCCGTCAAAGAGTTGACTTCGTGGGAGAAAAGGACCACTGAG

3010 3020 3030 3040 3050 3060
 TAGAGTCACGCTGACTAAACTAACTAAGTGCTACTGATCTCTCCTCAAGTTCATGCACA
 ATCTCAGTGGGACTGATTTTGATTGATTACAGTACTAGAGAGGAGTTCAAGTACGTGT

3070 3080 3090 3100 3110 3120
 AACACATCACTAATGGACCATAGTCATTTTTCTTGTATTATACCAAAGGCCTCACAATATA
 TTGTGTAGTGATTACCTGGTATCAGTAAAAGAACAATAATGGTTTCCGGAGTGTATAT

3130 3140 3150 3160 3170 3180
 AAACATGGTGTAGCCTTTCAATTTCCACAGTCCTTAAAAATTTCAACATATAGAAGACAC
 TTTGTACCACATCGGAAAGTTAAGGTGTCAGGAATTTTAAAGTTGTATATCTTCTGTG

3190 3200 3210 3220 3230 3240
 AAGTCTTCTCAAAGTCCAAAGCCCTTAACTGTGGGCTTTTGGCAAAATAAAAACAAGT
 TCCAGAAGAAGTTTCAGGTTTCGGGGAATTGACACCCGAAAACGTTTTATTTTGTTCAT

17-5

3250 3260 3270 3280 3290 3300
 AACATTATTTCTGATCTTAGAGGGAAGAACCAGGGCACAGTTACAATAAGAACAAGCA
 TTGTAATAAAGGACTAGAATCTCCCTTCTTGGTCCCGTGTCAATGTTATTCTTGTTCGT

3310 3320 3330 3340 3350 3360
 AAACCAACATCCAGATGTGTAATCAAATCCTATAGCTCAATGTCCAGCATCCGGGGCTT
 TTTGGTTGTAGGTCTACACATTTAGTTTAGGATATCGAGTTACAGGTCGTAGGCCCCGAA

3370 3380 3390 3400 3410 3420
 ATGGTCTCCCAGACTCCAAGGTCTTGGGCAGCTCCACTCTGCCTCTAGCTCCAGTACAT
 TACCAGAGGGTCTGAGGTTTCCAGAACCCGTCGAGGTGAGACGGAGATCGAGGTCAATGTA

3430 3440 3450 3460 3470 3480
 AGGCTCAGGCTGGCTCCAATCCACAGCATCCCACTGTCTGGCTCCTAACATGTTGCA
 TCCGAGTCCGACCGAGGTTAGGTGTCTGAGGGTGACAGGACCGAGGAGATTGTACAACGT

3490 3500 3510 3520 3530 3540
 ATATCTGCTGCAACTGAAGCTGCACCTTACCAATGAACGACCTCTCCAGACTTCTCACA
 TATAGACGACGTTGACTTCGACGTGGAAGTGTTACTTGCTGGAGAGGTCTGAAGAGTGT

3550 3560 3570 3580 3590 3600
 GTGACAATCCTCTACTTCTATGTGACTTCTCAATCCTGGGGTCTCCACTGCAACTGA
 CACTGTTAGGAGATGAAGAGATACACTGAAGGAGTTAGGACCCAGAGGTGACGTTGACT

3610 3620 3630 3640 3650 3660
 GGCTGTACCTTCAATAACCTGTCTCTAGGCTCTCTTCAAGACTCTGTGGCCTTGAT
 CCGACATGGAAGTAGTTATTGGACAGAGATCCGAGAGAAGTCTCTGAGACACCCGAACTA

3670 3680 3690 3700 3710 3720
 TTATGGTACCAAGCCTTAACTGTCTCAATGGTCCCTTCAAGACCAGTACCACCAATGA
 AATACCATGGTTCCGAATTGACAAGAGTTACCAGGAAAGTTCTGGTCATGGTGGTACT

3730 3740 3750 3760 3770 3780
 GACTGTGCTTGGTAGGAGCTAATTATTCTGTGGATGGTTTCCCAAGCCCTTGGGAAGAGC
 CTGACACGAACCATCCTCGATTAATAAGACACCTACCAAGGGTTCGGGAAACCTTCTCG

3790 3800 3810 3820 3830 3840
 CCAGAAGACCATGAGTACTTCCAGAAGTTAGACACTAAGATTCACTGTTGAACTTTGGT
 GTCTTCTGGTACTCACTGAAGGTCTTCAATCTGTGATTCTAAGTGACAACCTGAAACCA

3850 3860 3870 3880 3890 3900
 TTCACTCTGATTTAATTGTGACTGTGCCCTTTGGAGTTCTCCCTCTGGAGTAAGAAA
 AAGTGAGACTAAATTAACACTGACACGGGAAACCTCAAGAAGGGAGAACCTCATTCTTT

3910 3920 3930 3940 3950 3960
 ATATTTAACTTATTTTGACTTTGCAAGAGCCACAGCTGTGAGAGTTTGGAGTTTGGGAG
 TATAAATTGAATAAACTGAAACGTTCTCGGGTGTGACACTCTCAAACCTCAAACCTC

3970 3980 3990 4000 4010 4020
 ATATTTTGAGTGTGAGAGATCAGGGATATTTAGAGAGGCATGAACTTTGAAAGAGA
 TATAAAACCTCACAACTCTCTAGTCCCTATAAAATCTCTCCGTACTTGAACCTTCTCT

4030 4040 4050 4060 4070 4080
 CTTTGGATGTTCTAAAGGGATTTAACTTTTAAAGTGCTGGGATTTGTAAAGACTTTTTAA
 GAAACCTACAAGATTTCCCTAAATTGAAAATTCACGACCCTAAACATTTCTGAAAAATT

4090 4100 4110 4120 4130 4140
 AGTTGGAATGTTTTATATTGTGATATGAATATTGATATGTAATCTTGGCAGTGAACAAGA
 TCAACCTTACAAAATATAACTATACTTATAACTATACATTAGAACCCTCACTTGTCT

17-6

4150 4160 4170 4180 4190 4200
 AAGGGAAGTTATATTTAACTGTAATGTGTTTGTGTGTACATTAACAAGCAGTCATAT
 TTCCTTCCAATATAAAATTGACATTACACAAACACACAATGTAATTGTTTCGTCAGTATA

4210 4220 4230 4240 4250 4260
 TTTTATGGTTTTATGGAAAGTTTTTCATAAGCTCCTCAAGGCCTTGGAAATCCAAAATCCA
 AAAATACAAAATACCTTTCAAAGTATTCGAGGAGTTCGGGAACCTTTAGGTTTTAGGT

4270 4280 4290 4300 4310 4320
 TTCCTTGTGGAAGAGGGAACCTCAAATGAGAAAATGCCCTACTAAGATTGGCCCAGGAG
 AAGGAACACCTTCTCCCTTGAGTTTACTCTTTTACGGGAATGATTCTAACCGGGTCCCTC

4330 4340 4350 4360 4370 4380
 AAAGCCTGCAGTTGTTTTGTTTTGTTTTGTTTTTAAATTGAGGCAGATGTGGGAGGTTT
 TTTCCGACGTCAACAAAACAAAACAAAACAAAATTTAACTCCGTCACACCCTCCAAA

4390 4400 4410 4420 4430 4440
 CTAGCCACTGTAGGCAGTGTGACACCTGGGCTGGTGGTCTCTGTGCCAACAGAAAGCA
 GATCGGTGACATCCGTCACTGTGGACCCGACCACCAGGAGACACGGTTGTCTTTCGT

4450 4460 4470 4480 4490 4500
 GGCTACTAGCTAGGCTAGCCATGAGAAGCAAGCCAGTAAGCATCAGGTTCCCTCCTGTCTT
 CCGATGATCGATCCGATCCGTACTCTTCGTTCCGTCATTTCGTAGTCCAAGGAGGACAGAA

4510 4520 4530 4540 4550 4560
 GCTCCAGCTTCCCTCATTGATGGATTTAATCAGCTGTAAGGTGAAATGAATTGGTTACTT
 CGAGGTGGAAGGAGTAACCTAAATTAGTCGACATTCCACTTTACTTAACCAATGAA

4570 4580 4590 4600 4610 4620
 TTATCACAGCAAAGAAATCTTTTTTAAAATTTTATTAGGTATTTTCCTCATTTACA
 AATAGTGTGTTTTCTTTAAGAAAAAATTTTAAAAATAATCCATAAAAGGAGTAAATGT

4630 4640 4650 4660 4670 4680
 TTTCCAATGCTATCCCAAAGTCCCCCATACCCTCCCCCCCCACTCCCCTACTCACCCA
 AAAGTTACGATAGGGTTTTTCAGGGGGTATGGGAGGGGGGGGTGAGGGGATGAGTGGGT

4690 4700 4710 4720 4730 4740
 CTCCCCTTCTTGGCCCTGGCGTTCCTGTACTGAGGCACATAAAGTTTGCAAGACCAA
 GAGGGTGAAGAACCGGGACCGCAAGGGGACATGACTCCGTGTATTTCAAACGTTCTGGTT

4750 4760 4770 4780 4790 4800
 TGGGCCTCTCTTCCACTGATGGCCGACTAGGCCATCTTTTGATACATACGCAGCTCATA
 ACCCGAGAGAAAGGTGACTACCGGCTGATCCGGTAGAAAACCTATGTATGCGTCGAGTAT

>SacI
 |
 4810 | 4820 4830 4840 4850 4860
 TAGTCAAGAGCTCCGGGGTATTGGTTAGTTTCATAATGTTGTTCCACCTATAGGGTTGCAG
 ATCAGTTCTCGAGGCCCCATAACCAATCAAGTATTACAACAAGGTGGATATCCCAACGTC

4870 4880 4890 4900 4910 4920
 ATCCCTTTAGCACCTTGGATACTTTCTCTAGCTCCTCCATTGGGGGCCCTGTGATCCATC
 TAGGGAAATCGTGAACCTATGAAAGAGATCGAGGAGGTAACCCCGGGACACTAGGTAG

4930 4940 4950 4960 4970 4980
 CAATAGCTGACAATGAGCATCCACTTCTGTTTTTGTAGGCCCGGCCTAGTCTCACAAG
 GTTATCGACTGTTACTCGTAGGTGAAGACAAAACGATCCGGGGCCGGATCAGAGTGTTTC

4990 5000 5010 5020 5030 5040
 AGACAGCTACATCTGGGTCCTTTCAGCAAAATCTTGCTAGTGTATGCAACGGTGTCTGCG
 TCTGTGATGTAGACCAGGAAAGTCGTTTTAGAACGATCACATACGTTGCCACAGACGC

17-7

>BamHI

```

      5050      5060 |      5070      5080      5090      5100
TTTGGCGGCTGATTATGGCATGGATCCCCGGATATGGTAGTCTCTAGATGGTCCATCCCT
AAACCGCGACTAATACCGTACCTAGGGGCCTATACCATCAGAGATCTACCAGGTAGGAA

      5110      5120      5130      5140      5150      5160
TCGTCTCAGCTCCAACTTTGTCTCTGTAACCTCCATGGGTGTTTTGTTCCCAATTC
AGCAGAGTCGAGGTTTGAACAGAGACATTGAGGAAGGTACCCACAAAACAAGGGTTAAG

      5170      5180      5190      5200      5210      5220
TAAGAAGGGCAAAGTGTCCACACTTTGGTCTTCGTCTTCTTGAGTTTCACTAACTAAG
ATTCCTCCCCGTTTCACAGGTGTGAAACCAGAAGCAAGAAGAACTCAAAGTGATTGATTC

      5230      5240      5250      5260      5270      5280
ACAGAGACTCTTAAATATTATCAAATTCAGCTGCCAGCCTTATGTGGTCTTGACACCCCT
TGCTCTGAGAATTTATAATAGTTTAAAGTCGACGGTCGGAATACACCAGAACTGTGGGGA

```

>SacI

```

      5290      5300 |      5310      5320      5330      5340
GGACTACAACCTTCTGTGTGGGAGCTCTGAGGAGACACTTCGCAGAAGATTTTGCCTGTTA
CCTGATGTTGAAGACACACCCTCGAGACTCCTCTGTGAAGCGTCTCTAAACCGGACAA

      5350      5360      5370      5380      5390      5400
GGCAGGTCTCGTCTCAATCACACTGAGTTTGCAGCTCCAGTTGAGCACCGTTCAATTGGC
CCGTCCAGAGCAGAGTTAGTGTGACTCAAACGTCGAGGTCAACTCGTGGCAAGTTAACCG

      5410      5420      5430      5440      5450      5460
CTAGTAAAGCACAGGTTTCACTTAGGCGATGCTGGAGTCTTAGCAGAAGCCTTACACTGA
GATCATTTCGTGTCCAAAGTGAATCCGCTACGACCTCAGAATCGTCTTCGGAATGTGACT

      5470      5480      5490      5500      5510      5520
ACTTCACGAGCCAAGCCTGAGTTGTGGCATTGCTTCCAACACCACCATCCAAGCCCCCT
TGAAGTGCTCGGTTCCGACTCAACAGCCGTAAACGAAGTTGTGGTGGTAGGTTCCGGGGGA

```

>SacI

```

      5530      5540      5550      5560      5570      5580
ACACTAGTATATCTTCAAGTTCTAAGCCCTCAGAGCTCAAAGCTCCAAATGCTTCCGCAA
TGTGATCATATAGAAGTTCAAGATTCGGGAGTCTCGAGTTTCGAGGTTTACGAAGGCGTT

      5590      5600      5610      5620      5630      5640
CCACACTTCCAACCCCCAAACAACACACCTGTGTCCATAGTAGCAACAACACACTCTCTG
GGTGTGAAGGTTGGGGGTTTGTGTGTGGACACAGGTATCATCGTTGTTGTGTGAGAGAC

      5650      5660      5670      5680      5690      5700
GTACCAATTTCTGTTTCGATAGATGATGGACAGGCAGACAGACAAGAAAGTAGAAAGGGA
CATGGTTAAAGACAAAGCTATCTACTACCTGTCCGTCGTCTGTTCTTTTCATCTTCCCT

      5710      5720      5730      5740      5750      5760
ACCAGCAGGAAAGAAGAAGAGAGTCCACAAGAGTGGAGGGGACAAGAAAGAGTAATGGGG
TGGTCGTCCTTTCTTCTCTCAGGTGTTCTCACCTCCCCTGTTCTTTCTCATTACCCC

      5770      5780      5790      5800      5810      5820
GTGAGTATGATCAAAATGAACATATATGAAATATCATTATGAAATGTATATAATTAATAT
CACTCATACTAGTTTACTTGTATATACTTTATAGTAATACTTTACATATATTAATTATA

      5830      5840      5850      5860      5870      5880
ATACTAACAAGAAACCTTTTTTAAATTTTCAAAAATCATGCGTGCACCTGAACATGGGCA

```


17-9

6730 6740 6750 6760 6770 6780
 CAAATATATGGTGGGCAACTGGGAGAGAGAGAAAAAAGTGACAACTTGAGCATTATGA
 GTTTATATACCACCCGTTGACCCTCTCTCTTTTTTTTCACTGTTGAACTCGTAATACT

 6790 6800 6810 6820 6830 6840
 AGAGAGAGGGAAGTGGAGACTGAAGGGTAGAAAGGAGTCGCCTGTTCTGAGTGTCCAGCC
 TCTCTCTCCCTTGACCTCTGACTTCCCATCTTCCCTCAGCGGACAAGACTCACAGGTCGG

 6850 6860 6870 6880 6890 6900
 TATGGTAAGGCCCCAGCCTGGGCTGCCACTGAGGGCTGTATCTGAGTCCATGAGTCCACA
 ATACCATTCCGGGGTCGGACCCGACGGTACTCCCGACATAGACTCAGGTACTCAGGTGT

 6910 6920 6930 6940 6950 6960
 GCAGCAGGGCTCAGTGTGGATATCCGTGACACATGTTATCACTAGAGAACACGGGGAAGT
 CGTCGTCCCGAGTCACACCTATAGGCACTGTGTACAATAGTGATCTCTTGTGCCCTTGA

 6970 6980 6990 7000 7010 7020
 CCCTGGTCAGGACAGCTGTAGGGACCCACATGGACGTACAGGGGCTGTGCATAACTAGCC
 GGGACCACTCTGTCGACATCCCTGGGTGTACCTGCATGTCCCGACACGTATTGATCGG

 7030 7040 7050 7060 7070 7080
 TTGCCCTCTCACAGGATGAGGCCTTTGGGAGAGCTGGTGGCAGCACTTGGGAGAGAGGGCC
 AACGGAGAGTGTCTACTCCGAAACCCTCTCGACCACCGTCGTGAACCCTCTCTCCCGG

 7090 7100 7110 7120 7130 7140
 CTACACCTACCCAGTCAGCACAGTGGAGCTGGCCCTGGGGGTGGGGCTGGGGTACTCAC
 GATGTGGATGGGTCACTGTGTACCTCGACCGGGACCCCAACCCGACCCCACTGAGTG

 7150 7160 7170 7180 7190 7200
 CCCAAGGGCATGAGTGTGGGAGAGCCGACCTCATCACTCATCTGCTGCGGGGTGGCACCA
 GGGTCCCCTACTCACACCCTCTCGGCTGGAGTAGTGAGTAGACGACGCCCCACCGTGGT

 7210 7220 7230 7240 7250 7260
 GTGCAGAGGTGATGCTCACCCCATCCCCCTGCAACCCCTTGTACCTCTGGAAGTTGGG
 CACGCTCCACTACGAGTGGGGTAGGGGGGACGTTGGGGAACAGTGGAGACCTTCAACCC

 7270 7280 7290 7300 7310 7320
 AAAGCTGCCCATAGGGCCATGAACTCCGGAGAAGTAGCCCTGCCCTCACCATTGAGCACT
 TTTGACGGGTATCCCGTACTTGAAGCCTCTTGTATCGGGACGGGAGTGGTAAGTCGTGA

 >BamHI
 |
 | 7330 7340 7350 7360 7370 7380
 AGGAGGATCCTGCACCTCTCCTGGATAGCACGGCAGAGCTGACCCTGGCAGCAGGGGTAC
 TCCTCCTAGGACGTGGAGAGGACCTATCGTGCCGCTCTGACTGGGACCGTCGTCCCATG

 7390 7400 7410 7420 7430 7440
 GGGTGAACCAGAGGGAGAGAGAACGAGAGCATGGGAGAGCTGGTCCCACCACCCCTCA
 CCACCTTGGTCTCCCTCTCTCTTGGCTCTCGTACCCTCTCGACCAGGGTGGTGGGGAGT

 7450 7460 7470 7480 7490 7500
 GCCATGAGAGGTGACCTGGGACAGCTGACCTGGCAGTCATGAGAGTAGGTGAGCTGGCCC
 CGGTACTCTCCACTGGACCCCTGTCGACTGGACCGTCAGTACTCTCATCCACTCGACCGGG

 7510 7520 7530 7540 7550 7560
 TACCTCTCACTCAGGAGTGGCTGCAGCACTTGGGAGAGTGGGTCCATCATCTTGCCTTGG
 ATGGAGAGTGAAGTCTCACCGACGTCGTGAACCCTCTCACCCAGGATGTAGAACCGAACC

 7570 7580 7590 7600 7610 7620
 CAACACAGCACAGTAGTGCTGGTCTGGTAGCAAATGCATGGGTGGGCCAGCCCCGAGGG

17-10

GTTGTGTCGTGCATCAGACCAAGACCATCGTTTACGTACCCACCCGGTCGGGGCTCCC

7630 7640 7650 7660 7670 7680
 TATAGGAGCAGGAGATGGCCTAGCCCCTCACAGGATGCATCACTTGGGACATTGGGTC
 ATATCCTCGTCTCTTACCGGATCGGGGAGTGTCTACGTAGTGAACCCTGTAACCCAG

7690 7700 7710 7720 7730 7740
 CCACGCCATGACTGGGCAGCACAGTGGCGCTGGCTTGGAGGCTTGGGTGTGGGTGCGCT
 GGTGCGGTACTGACCCGTCGTGCACCGCGACCGAAACCTCCGAACCCACACCCACGCGA

7750 7760 7770 7780 7790 7800
 GGCCCAAGGGCAAGGGAACAGGAGATCTAAACCTACCTCTGACTATGGCGGTAGTATT
 CCGGGGTCCCGTCCCTTGTCTCTAGATTTGGATGGAGGACTGATACCGCCATCATAA

7810 7820 7830 7840 7850 7860
 GACTAGCCTAGCCAAAACAGTGTGGAGTGTGGAGAGCCGTGCCACAAGCAATCGCGAGC
 CTGATCGGATCGGTTTTGTACGACCTCACACCTCTCGGCACGGTGTTCGTTAGCGCTCG

7870 7880 7890 7900 7910 7920
 CGTGAACAATCGCCATTATAAGATGGCGCTAGCTTCCACTGTGCCTAACTAGTAAACAAG
 GCATTGTTAGCGGTAATATTCTACCGCGATCGAAGGTGACACGGATTGATCATTGTTC

7930 7940 7950 7960 7970 7980
 CCTTATGCACAAGTGCAAGAGTAAATTCACGCCCAGTCACTGCCATCTTGGGACGTAGT
 GGAATACGTGTTACGTTCTCATTTAAGTGCGGGTCACTGACGGGTAGAACCCTGCATCA

7990 8000 8010 8020 8030 8040
 AATGGGGTGATGGGCGAGCAACTAATCAGGTGCTGTACGCCACATCAGGTGCTGAAATG
 TTACCCCACTACCCGCTCGTTGATTAGTCCACGACAGTGCGGTGTAGTCCACGACTTTAC

8050 8060 8070 8080 8090 8100
 TCACACTGTGGGCTATATAAGCAGCGCGATTTCCGGGTTCGGGTCTTCTGAGAAGTA
 AGTGTGACACCCGATATATTCGTGCGCTAAAAGGCCAAGCCCCAGAAGGACTCTTCAT

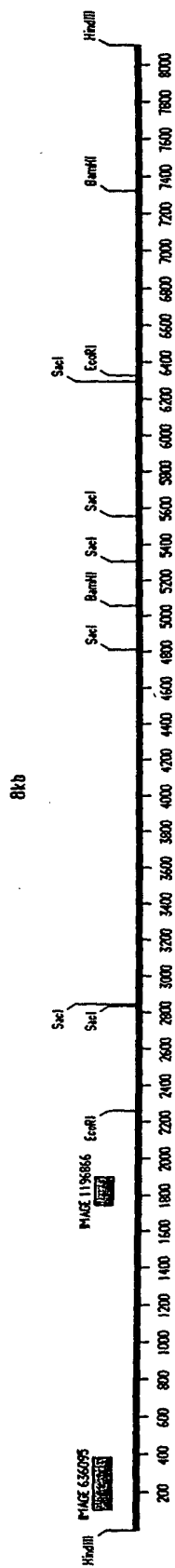
>HindIII

|

8110

AGCAATAAAAGCTT
 TCGTTATTTTCGAA

Figure 18





European Patent
Office

PARTIAL EUROPEAN SEARCH REPORT

Application Number

which under Rule 45 of the European Patent Convention EP 01 20 1037 shall be considered, for the purposes of subsequent proceedings, as the European search report

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	MICHAUD EDWARD J ET AL: "The embryonic lethality of homozygous lethal yellow mice (A-y/A-y) is associated with the disruption of a novel RNA-binding protein." GENES & DEVELOPMENT, vol. 7, no. 7A, 1993, pages 1203-1213, XP002076781 ISSN: 0890-9369 * abstract *	63,64	C12N15/12 C07K14/47 C12Q1/68 A61K48/00 A61P27/00 A01K67/027
X	--- LANG R: "THE MAMMALIAN SPOT TEST AND ITS USE FOR TESTING OF MUTAGENIC AND CARCINOGENIC POTENTIAL EXPERIENCE WITH THE PESTICIDE CHLORDIMEFORM ITS PRINCIPAL METABOLITES AND THE DRUG LISURIDE HYDROGEN MALEATE" MUTATION RESEARCH, vol. 135, no. 3, 1984, pages 219-224, XP000926604 AMSTERDAM NL ISSN: 0027-5107 * abstract * * page 221 *	68	---
	---	---	TECHNICAL FIELDS SEARCHED (Int.Cl.7) C07K A01K
-/--			
INCOMPLETE SEARCH			
<p>The Search Division considers that the present application, or one or more of its claims, does/do not comply with the EPC to such an extent that a meaningful search into the state of the art cannot be carried out, or can only be carried out partially, for these claims.</p> <p>Claims searched completely :</p> <p>Claims searched incompletely :</p> <p>Claims not searched :</p> <p>Reason for the limitation of the search: see sheet C</p>			
Place of search BERLIN		Date of completion of the search 10 August 2001	Examiner De Kok, A
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ----- & : member of the same patent family, corresponding document	

EPO FORM 1503 03 82 (P04/C07)

European Patent
OfficeINCOMPLETE SEARCH
SHEET CApplication Number
EP 01 20 1037

Claim(s) searched completely:
63, 64, 67 and 68

Claim(s) searched incompletely:
1-51,55-62

Claim(s) not searched:
52-54, 65 and 66

Reason for the limitation of the search:

Present claims 1-48, 61, 62 relate to an isolated nucleic acid defined by reference to a desirable characteristic or property, namely that it defines a 'yellow submarine' locus, in stead of by reference to a specified nucleotide sequence.

The claims cover all nucleic acids having this characteristic or property, whereas the application provides support within the meaning of Article 84 EPC and/or disclosure within the meaning of Article 83 EPC for only ONE of such a nuclei acid. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the nucleic acid having the nucleotide sequence identified in SEQ.ID.No. 1 (see description page 14 and figure 17).

Present claims 49-51 and 55-59 relate to a protein encoded by the above-mentioned nucleic acid. Since in the application two putative protein fragments (as defined in SEQ.ID.No. 2 and 3) have been identified encoded by SEQ.ID.No.1, the search for those claims has been restricted to SEQ.ID. No. 2 and 3.

Present claims 52-54 and 65, 66 relate to methods of treatment using the protein respectively the nucleic acid identified above. Since the application does not provide any evidence that the nucleic acid nor the putative proteins encoded thereby is involved in regeneration of nerve tissue, cell migration, cell growth, vestibular dysfunction or hearing, the claims lack disclosure within the meaning of article 83 EPC. Consequently, no search has been carried out for those claims



European Patent
Office

PARTIAL EUROPEAN SEARCH REPORT

Application Number
EP 01 20 1037

DOCUMENTS CONSIDERED TO BE RELEVANT		CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim
A	<p>DATABASE SWALL [Online] 1 November 1999 (1999-11-01) TAKADA, S. ET AL: "Mouse cDNA similar to 5' region of human EXML1 cDNA" retrieved from EMBL-EBI, accession no. Q9WTN9 Database accession no. Q9WTN9 XP002174568 * abstract *</p>	1,49
A	<p>--- US 5 723 719 A (MULLINS JOHN J ET AL) 3 March 1998 (1998-03-03) * abstract *</p>	67
A,D	<p>--- CHEAH KATHRYN S E ET AL: "Human COL2A1-directed SV40 T Antigen Expression in Transgenic and Chimeric Mice Results in Abnormal Skeletal Development." JOURNAL OF CELL BIOLOGY, vol. 128, no. 1-2, 1995, pages 223-237, XP000926594 NEW YORK US ISSN: 0021-9525 * page 224, column 1, last paragraph * * page 236, column 1; table 1 * -----</p>	67
		TECHNICAL FIELDS SEARCHED (Int.Cl.7)

EPO FORM 1503 03.82 (P/AC10)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 01 20 1037

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

10-08-2001

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5723719 A	03-03-1998	NONE	

EPO FORM P0489

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82