

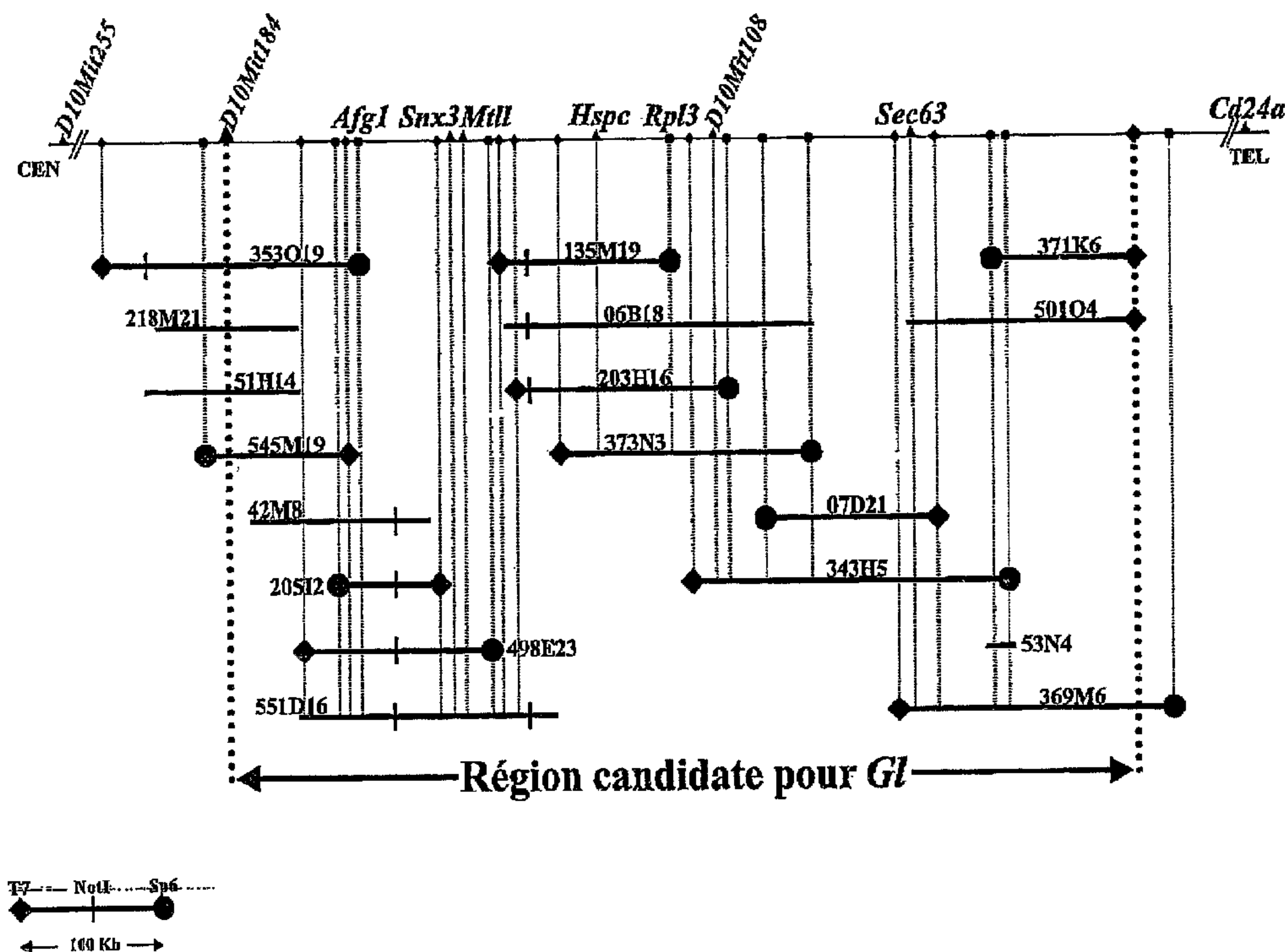


(86) Date de dépôt PCT/PCT Filing Date: 2002/09/25  
(87) Date publication PCT/PCT Publication Date: 2003/04/10  
(85) Entrée phase nationale/National Entry: 2004/03/29  
(86) N° demande PCT/PCT Application No.: EP 2002/010721  
(87) N° publication PCT/PCT Publication No.: 2003/029283  
(30) Priorité/Priority: 2001/09/27 (2,357,987) CA

(51) Cl.Int.<sup>7</sup>/Int.Cl.<sup>7</sup> C07K 14/435, A61K 38/17, A61P 19/10, C07K 14/47  
(71) Demandeur/Applicant: PROSKELIA PHARMACEUTICALS, FR  
(72) Inventeurs/Inventors: VACHER, JEAN-NOEL, CA; CHALHOUB, NADER, CA; BENACHENOU, NADIA, CA  
(74) Agent: OGILVY RENAULT

(54) Titre : SEQUENCE GENETIQUE LIEE A DES MALADIES OSSEUSES

(54) Title: GENETIC SEQUENCES RELATED TO BONE DISEASES IN THE OSTEOPETROTIC GREY-LETHAL (GL) MOUSE



(57) Abrégé/Abstract:

The identification, isolation and cloning of gl gene which when mutated is associated with bone related diseases as well as its transcript in gene products. A method of diagnostic and detection of potential carriers of this mutated gene, bone related diseases, diagnosis, gene therapy recombinant technology and therapy using the information derived from the DNA, protein and the function of the protein is also provided.

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

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
10 April 2003 (10.04.2003)

PCT

(10) International Publication Number  
**WO 03/029283 A3**(51) International Patent Classification<sup>7</sup>: **C07K 14/435**,  
14/47, A61P 19/10, A61K 38/17[CA/CA]; 10353 rue Paul Comtois, Montréal, Québec  
H4N 2Y3 (CA). **BENACHENOU, Nadia** [CA/CA]; 551  
Menneval, Longueuil, Québec J4L 4W6 (CA).

(21) International Application Number: PCT/EP02/10721

(74) Agent: **ROUSSEAU, Pierrick**; Aventis Pharma S.A., 102,  
route de Noisy, F-93235 Romainville Cedex (FR).(22) International Filing Date:  
25 September 2002 (25.09.2002)

(25) Filing Language: English

(81) Designated States (*national*): AE, AG, AL, AU, BA, BB,  
BR, BZ, CA, CN, CO, CR, CU, DM, DZ, EC, GD, GE, HR,  
HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MA,  
MG, MK, MN, MX, NO, NZ, OM, PH, PL, RO, SG, SI,  
TN, TT, UA, US, UZ, VC, VN, YU, ZA.

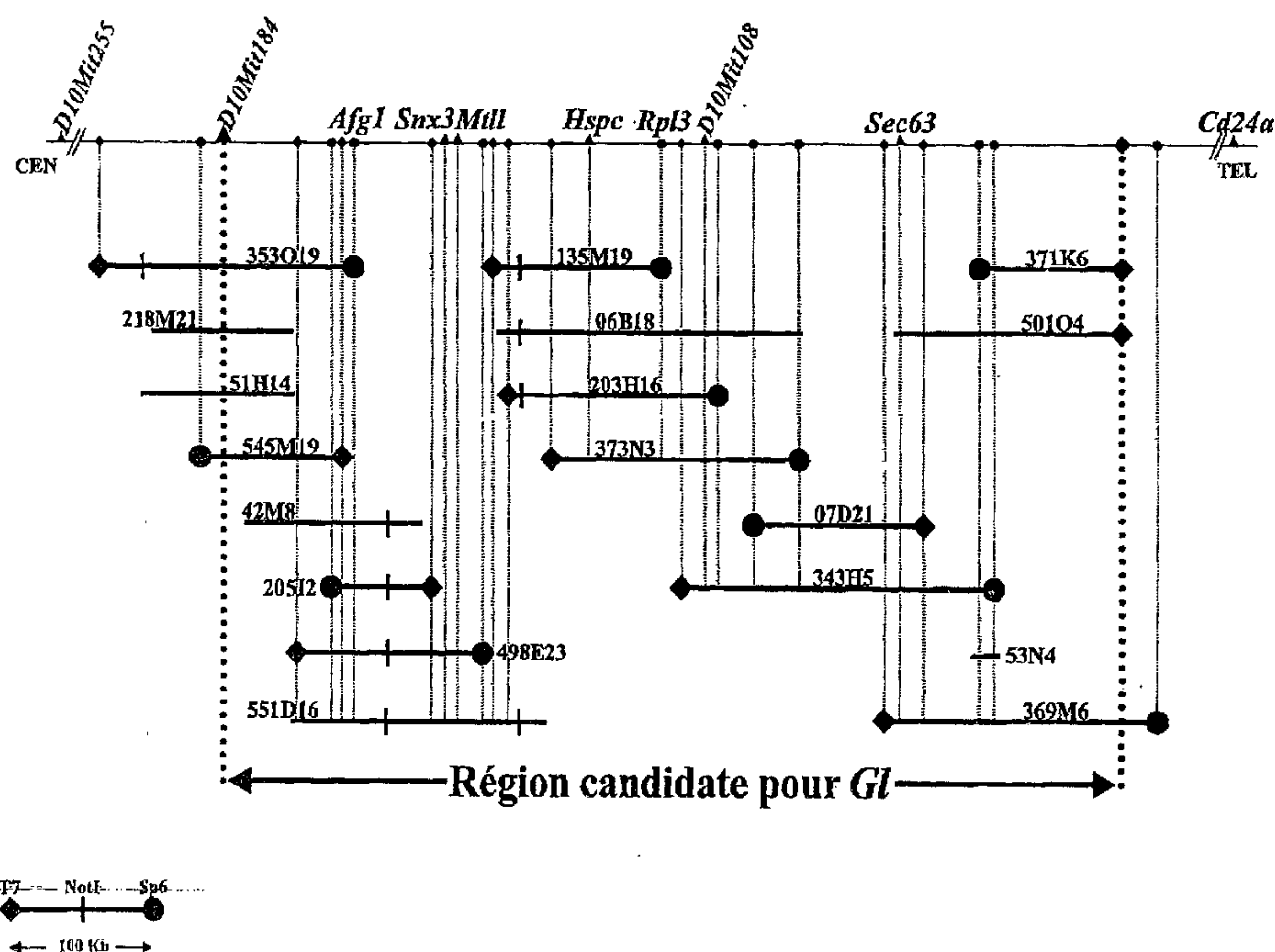
(26) Publication Language: English

(30) Priority Data:  
2,357,987 27 September 2001 (27.09.2001) CA(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,  
TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, ML, MR, NE, SN, TD, TG).(71) Applicant (*for all designated States except US*):  
**PROSKELIA PHARMACEUTICALS** [FR/FR]; 102,  
route de Noisy, F-93230 Romainville (FR).(72) Inventors; and  
(75) Inventors/Applicants (*for US only*): **VACHER,**  
**Jean-Noël** [CA/CA]; 403 Boulevard St-Joseph Ouest,  
Montréal, Quebec H2V 2P3 (CA). **CHALHOUB, Nader**

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv)) for US only

[Continued on next page]

(54) Title: GENETIC SEQUENCES RELATED TO BONE DISEASES IN THE OSTEOPETROTIC GREY-LETHAL (GL)  
MOUSE(57) Abstract: The identification, isolation and cloning of *gl* gene which when mutated is associated with bone related diseases as well as its transcript in gene products. A method of diagnostic and detection of potential carriers of this mutated gene, bone related diseases, diagnosis, gene therapy recombinant technology and therapy using the information derived from the DNA, protein and the function of the protein is also provided.

WO 03/029283 A3

**WO 03/029283 A3**



**Published:**

— *with international search report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**(88) Date of publication of the international search report:**

4 September 2003

## GENETIC SEQUENCE RELATED TO BONE DISEASES

### FIELD OF THE INVENTION

The present invention generally relates to the field of bone related diseases associated with osteoclast cells dysfunction. More particularly, the invention is concerned with the identification, isolation and cloning of a gene, which when mutated is associated with bone related diseases as well as its transcript in gene products. The present invention also relates to a method of diagnostic and detection of potential carriers of this mutated gene, bone related diseases, diagnosis, gene therapy recombinant technology and therapy using the information derived from the DNA, protein and the function of the protein.

### BACKGROUND OF THE INVENTION

#### A) BRIEF DESCRIPTION OF THE PRIOR ART

Bone homeostasis is dependent on two opposite and dynamic processes of bone formation and resorption in vertebrates and is regulated throughout adult life. Defective bone resorption (osteopetrosis or osteoporosis) results from a defect in bone resorption. More particularly, osteopetrosis results in accumulation of mineralised bone and cartilage due to a lack of bone remodelling activity. This activity is normally provided by osteoclast. Such fully differentiated cells are multinucleated and are formed by the fusion of myeloid cells from the monocyte-macrophage lineage.

Osteopetrosis results from a defect in the differentiation or the activation of the osteoclast, a specialized cell, which derives from the granulocyte-macrophage hematopoietic lineage. The role of the osteoclast is bone tissue resorption, a process that is counterbalanced by the osteoblast activity that results in bone tissue formation. When such balance is disrupted, major bone diseases as

osteoporosis and osteopetrosis can occur. Lazner, F. et al., Hum Mol Genet., 8:1839-1846 (1999).

The event of homologous recombination in association with gene targeting in the mouse, tremendously improved our understanding of osteoclastogenesis. The specific loss of osteoclast gene function resulted in osteopetrosis that is characterized by a general increase in bone mass. For example, PU-1, c-fos, NFk-B and RANKL gene activities are required for the differentiation/proliferation of osteoclast precursors, while the loss of c-src, TRAF6, V-ATPase and ClC-7 have been associated with defects in polarization/resorption of the osteoclast. Karsenty, G., Genes and Dev., 13:3037-3051 (1999); Teitelbaum, S.L., Science, 289:1504-1508 (2000).

In addition to these engineered mutations, four spontaneous mutations have been described in the mouse. The *op* gene encodes the hematopoietic colony stimulating factor 1 (CSF-1) Yoshida, H. et al., Nature, 345:442-445, (1990), *mi* encodes a transcriptional factor from the basic-loop-helix zipper (bHLH-zip) family, Hodgkinson, C.A., Cell, 74:395-404, (1993) and the *oc* mutation affects the 116KD subunit of the V-ATPase (Scimeca, J-C et al., Bone, 26:207-213 (2000). The fourth mutation, grey-lethal (*gl*), described for the first time by Grunberg, Gruneberg, H. J. Hered., 27:107-109 (1936), displays an osteopetrotic phenotype closely related to the most severe autosomal recessive form of the human disease. As in humans, early death occurs around three weeks of age in homozygous *gl/gl* mice, and functional rescue can be obtained following bone marrow transplantation demonstrating a cell-autonomous defect. Walker, D.G., Science, 190:784-785 (1975).

Therefore, there is a need to determine the nucleic acid sequence encoding for an osteoclast-related polypeptide having the biological activity of modulating the bone resorption.

The inventors have determined that the *gl* gene is required for osteoclast maturation/function. Rajapurohitam, V. et al., Bone, 28:513-523 (2001).

#### SUMMARY OF THE INVENTION

5 The present invention originates from the discovery of a *gl* gene encoding a polypeptide involved in the regulation of bone resorption in a mammal.

Accordingly, the present invention relates to an isolated or purified nucleic acid molecule encoding a mammalian osteoclast-related polypeptide (referred to hereinafter the *Gl* polypeptide) having the biological activity of modulating bone resorption in osteoclast cells.

10 The present invention also provides the following:

- an expression or cloning vector having the nucleic acid sequence of the *Gl* polypeptide mentioned above;
- a host cell having the above mentioned expression or cloning vector;
- 15 - a non-human mammal comprising a genetically modified nucleic acid molecule of the *Gl* polypeptide of the present invention;
- an isolated antibody that binds specifically to the *Gl* polypeptide and fragments thereof;
- 20 - a process for producing the *Gl* polypeptide of the present invention;

- a method for preventing or treating a bone resorption-related disease in a mammal subject by administering the polypeptide of the present invention to the subject; and
- 5       - a pharmaceutical composition containing the GI polypeptide of the present invention, for preventing or treating an osteoclast-related disease such as osteoporosis and osteopetrosis.

10       In summary, the work conducted in the context of the present invention has allowed the inventors to identify a novel gene with a specific function that is absolutely required for proper osteoclast maturation and bone tissue resorption.

This in turn, has allowed the inventors of the present application to provide methods, pharmaceutical compositions and diagnostic tools to treat and/or prevent bone related diseases such as osteopetrosis and osteoporosis.

## 15       BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the physical entrance cryptional map of BAC.

Figure 2 is a table regrouping information on the characterisation of BAC clones.

Figure 3 shows the expression of the GI gene in different tissue type.

20       Figure 4 shows the expression of the GI gene in transgenic mice in different tissue type.

Figure 5 A shows the result of the Western blot ANALYSIS OF GI polypeptide in Wild-type and *gl* osteoclasts.

Figure 5 B shows the specific cytoplasmic localisation of the Gl polypeptide.

Figure 6 shows the Kyte-Doolittle hydropathy plot for mouse GL polypeptide.

Figure 7 shows the TMpred-prediction of transmembrane Regions and  
5 Orientation for Mouse Gl polypeptide.

Figure 8 shows the nucleic acid sequence of the mouse Gl gene.

Figure 9 shows the nucleic acid sequence of the human GL gene homologue.

Figure 10 shows the amino acid sequence of the mouse Gl polypeptide.

10 Figure 11 shows the amino acid sequence of the human GL polypeptide.

## DETAILED DESCRIPTION OF THE INVENTION

### A) Definitions

In order to provide an even clearer and consistent understanding of the specification and the claims, including the scope given herein to such  
15 terms, the following definitions are provided:

Osteoclast broadly relates to a large multinucleated cell found in growing bone that resorbs bone tissue, as the renewal of bone matrix.

*gl* gene also relates to a gene, which encodes for an osteoclast-related polypeptide and which when mutated is associated with bone  
20 related diseases. This definition is understood to include the various sequence polymorphisms that exist wherein the codon substitutions or deletion in the gene sequence do not affect the essential function of the

gene product as well as functionally equivalence of the nucleotide sequences of SEQ. ID No. 1 and SEQ. ID No. 2. This term also relates to an isolated coding sequence, but can also include some or all of the flanking regulatory elements and/or introns. The term *g/* gene includes the  
5 gene and other species homologous to the human gene, which when mutated is associated with bone related diseases.

GI polypeptide refers to the polypeptide encoded by the *g/* gene. This polypeptide may be a natural or synthetic compound containing two or more amino acids having a specificity to osteoclast cells susceptible to  
10 modulate the activity of osteoclast cells. The preferred source of polypeptide is the mammalian polypeptide as isolated from humans or animals. The polypeptide may be produced by recombinant organisms or chemically or enzymatically synthesised. This definition is understood to include functional variance, such as the various polymorphic forms of the  
15 protein where the amino acids substitution or deletion within the amino acid sequence do not effect essential functioning of the protein or its structure. It also enclosed functional fragments of osteoclast related polypeptide.

Modulation refers to activation or inhibition of osteoclast cell activity in bone resorption.

20 Ruffled border broadly refers to the folded configuration of the osteoclast cell membrane through which the osteoclast may resorb the bone matrix.

Functional homologues broadly refer to a protein/peptide or polypeptide sequence that possesses a functional biological activity that is  
25 substantially similar to the biological activity of the whole protein/peptide or polypeptide sequence. A functional derivative of a protein/peptide or polypeptide may or may not contain post-translational modifications such as covalently linked carbohydrate, if such modification is not necessary for the

performance of a specific function. The term "functional derivative" is intended to cover the "fragments", "segments", "variants", "analogs" or "chemical derivatives" of a protein/peptide or polypeptide.

5 Analog broadly refers to a peptide or polypeptide that is substantially similar in function to the polypeptide of the invention.

Derived broadly refers to a protein/peptide or polypeptide that is said to "derive" from a protein/peptide or from a fragment thereof when such protein/peptide comprises at least one portion, substantially similar in its sequence, to the native protein/peptide or to a fragment thereof.

10 Isolated or Purified refers to a state different from the natural state. More precisely, it is altered "by the hand of man" from its natural state, i.e., if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide naturally present in a living organism is not "isolated", the same polynucleotide separated from  
15 the coexisting materials of its natural state, obtained by cloning, amplification and/or chemical synthesis is "isolated" as the term is employed herein. Moreover, a polynucleotide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism.

20 The term peptide or polypeptide herein includes any natural or synthetic compounds containing two or more amino acids. Therefore, it comprises proteins, glycoproteins, and protein fragments derived from pathogenic organisms such as viruses, bacteria, parasites and the like, or proteins isolated from normal or pathogenic tissues, such as cancerous  
25 cells. It also includes proteins and fragments thereof produced through recombinant means that has been associated or not with other peptides coding for tumoral, viral, bacterial or fungic epitopes for forming a fusion protein.

30 Nucleic acid broadly refers to any DNA, RNA sequence or molecule having one nucleotide or more, including nucleotide sequences encoding a complete gene. The term is intended to encompass all nucleic acids

whether occurring naturally or non-naturally in a particular cell, tissue or organism. This includes DNA and fragments thereof, RNA and fragments thereof, cDNAs and fragments thereof, expressed sequence tags, artificial sequences including randomized artificial sequences.

- 5           Functional homologues broadly refer to any molecule, natural or synthetic, being able to carry out the same functions as the protein or polypeptide of interest.

10           The term "variant" as is generally understood and used herein, refers to a protein that is substantially similar in structure and biological activity to either the protein or fragment thereof. Thus two proteins are considered variants if they possess a common activity and may substitute each other, even if the amino acid sequence, the secondary, tertiary, or quaternary structure of one of the proteins is not identical to that found in the other.

15           Vector refers to a self-replicating RNA or DNA molecule, which can be used to transfer an RNA or DNA segment from one organism to another. Vectors are particularly useful for manipulating genetic constructs and different vectors may have properties particularly appropriate to express protein(s) in a recipient during cloning procedures and may comprise different selectable markers known by one skilled in the art. Bacterial  
20           plasmids are commonly used vectors.

          Probe or primer broadly refers to any DNA or RNA sequence that is marked with a fluorescent compound, a radioisotope or an enzyme and used for detecting homologues (complementary) sequences as by hybridization *in situ* or *in vitro*.

25           Osteoporosis relates to a disease in which the bones become extremely porous, are subject to fracture, and heal slowly, occurring especially in women following menopause and often leading to curvature of the spine from vertebral collapse.

30           Osteopetrosis relates to a disease in which the bones become extremely dense. There is absence of development of the bone marrow, of

teeth growth and of general growth. This disease also causes premature death of the subjects.

## B) OVERVIEW OF THE INVENTION

5 The present invention is concerned with the identification and sequencing of the mammalian *gl* gene in order to gain insight into the cause and etiology of bone related diseases. From this information, screening methods and therapies for the diagnosis and treatment of the diseases can be developed.

10 Although it is generally understood that bone related diseases are caused by osteoclast related polypeptide expressed most likely in the bones, expression of this polypeptide has been found in variety of mammalian tissue types such as the testis, the thymus, the heart, the kidney, the spleen, the brain and the liver.

15 The mutation identified in the context of the present invention has been related to bone diseases such as osteopetrosis and osteoporosis. With the identification of sequences of the gene and the gene products, probes and antibodies raised against the gene product can be used in a variety of hybridisation and immunological assays to screen for and detect the presence of either a normal or mutated gene or gene product.

20 Patient therapy through removal or blocking of the mutant gene product, as well as supplementation with the normal gene product by amplification, by genetic and recombinant techniques or by immunotherapy can now be achieved. Correction or modification of the defective gene product by protein treatment immunotherapy (for example using antibodies to the  
25 defective protein) or knock out of the mutated gene is now also possible.

The bone related disease aimed in the present invention could also be controlled by gene therapy in which the gene defect is corrected *in situ* or

by the use of recombinant or other vehicles to deliver a DNA sequence capable of expressing the normal gene product whose effect counter balances the deleterious consequences of the disease mutation to the affected cells of the patient.

- 5 Toward the isolation and characterization of the *gl* gene, the inventors of the present invention have used a positional cloning approach. A detailed physical map was established using yeast and bacterial artificial chromosome (YACs, BACs). Transgenic mice were then generated with different BAC clones to localise the *gl* gene based on functional rescue of  
10 the *gl* osteopetrotic defect. The candidate *gl* gene or region was isolated and sequenced. Finally, a large deletion in this candidate gene or region in *gl* mice that results in complete loss of gene expression was molecularly characterized.

#### Physical mapping of the *gl* gene

- 15 As an initial step in the positional cloning approach used by the inventors of the present application, the *gl* locus was localized genetically to the proximal portion of mouse chromosome 10 in a ~1cM interval. Vacher, J. and Bernard, H., Mammalian Genome, 10, 239-243, 1999.

20 Interestingly, this study allowed the inventors to define cosegregation of the *gl* locus transmission with a congenic polymorphic region, potentially of 129Sv origin, maintained by brother-sister matings for more than one hundred generations. These polymorphic markers were used to screen five YAC libraries and allowed the applicant to establish a YAC contig covering ~8.5Mb.

25 To obtain genomic clones that would most probably be non-chimeric, a BAC contig was isolated and established. The BAC contig was composed of eighteen overlapping clones covering the *gl* candidate region. The markers *D10 Mit184* and *Cd24a* were used as entry points and after several rounds of

chromosome walking, a minimal candidate genomic interval of ~500kb was covered by the contig. Complete characterization of these clones and end insert probes from the BACs 545 M19 and 343 H5 delineated the non-recombinant interval, showing that the *gl* locus must lie between these two markers.

#### Functional rescue in BAC transgenic mice

The strategy adopted by the inventor's of the present application was based on an *in vivo* biological activity test through functional rescue of the osteopetrotic *gl/gl* phenotype, using BAC transgenesis.

- 10 Three overlapping BACs (498 E23, 373 N3, 343 H5) covering ~75% of the candidate region were injected. In contrast to non-transgenic grey homozygous *gl/gl* osteopetrotic littermates, all transgenic *gl/gl* animals carrying the BAC 373 N3 displayed normal growth, an agouti coat color and appropriate bone marrow development as demonstrated by histological analysis.
- 15 Transgene transmission followed Mendelian distribution, and complete rescue was observed in all *gl/gl* transgenic mice. No detrimental phenotype was noticed with age in transgenic animals.

These results suggested that the *gl* mutation was linked to a decreased activity of a gene included in the BAC 373 N3.

#### 20 Identification and characterisation of the *gl* gene

To characterize the genes present on the BAC 373 N3, a shotgun M13 phage library was generated and sequenced. In parallel, BLAST searches against EST (Expressed Sequence Tags) databases and ORF (Open reading frame) prediction analyses were used to define transcription units and genes.

Northern blot and RT-PCR gene expression analyses showed loss of expression of a unique ~3 kb transcript in *gl/gl* animals.

Genomic structural characterization of the *gl* locus by Southern analysis defined six exons and five introns covering approximately 16780 base pairs for the wild-type *gl* locus. In contrast to the wild-type *gl* locus, genomic DNA from the *gl/gl* mice underwent a genomic rearrangement associated with a large ~8kb deletion, that included the gene promoter and a large part of the first exon.

This observation is consistent with the complete lack of detection of the *gl* messenger RNA.

#### Gl polypeptide structure and localization

The open reading frame corresponding to the *gl* mRNA encodes a 338 amino acid protein with no obvious similarity with known protein sequences represented in protein databases. Hydropathy and protein topology analysis suggested the presence of one putative transmembrane domain in a protein enriched in cysteine residues. Two specific *Gl* antibodies corresponding to two different epitopes were used to detect by Western blot a ~38KDa protein in wild-type osteoclast extracts. In contrast no protein was detected in *gl/gl* cell extracts. Immunofluorescence analysis on wild-type native osteoclasts (Fig. 5), demonstrated specific cytosolic localization for the *Gl* polypeptide in multinucleated osteoclast as confirmed following Hoechst staining.

The analysis of predicted protein topology suggested that the protein has a putative transmembrane domain. Thus, this polypeptide may act as a receptor which has a binding specificity to a ligand which in turn has the function of modulating the activity of the *Gl* polypeptide; a channel protein; or a structural membrane protein.

Expression pattern

Northern blot and RT-PCR analysis demonstrated a wide-spread expression pattern of a unique ~3Kb messenger RNA in several tissues including brain, spleen, liver, kidney, heart, thymus, testis and most importantly in osteoclast-like cells (OCLs) obtained in cocultures. Functional complementation was further correlated with detection of this specific transcript in rescued animals. Strong expression was detected in transgenic homozygous *gl/gl* tissues compared to the normal low level of expression in control non-transgenic wild-type littermate. This is in accordance with the high BAC transgene copy number (~6) in this transgenic line. Furthermore, bone *in situ* hybridization demonstrated *gl* specific expression in multinucleated wild-type osteoclasts with higher expression in transgenic osteoclasts.

GI polypeptide may be expressed using eukaryotic and prokaryotic expression systems. Eukaryotic expression system can be used for many studies of the *gl* gene and gene product including the determination of proper expression and post-translational modification for full biological activity, the identification of regulatory elements located in the 5' region of the production of large amounts of the normal and mutant protein for isolation and purification, to use cells expressing the GI polypeptide as a functional assay system for antibodies generated against the protein and to test effectiveness of pharmacological agents or as a component of a signal transcription system to study the function of the normal complete protein, specific portion of the protein or of spontaneously occurring and genetically engineered mutant proteins.

One example of the prokaryotic expression system that may be used in the context of the present invention is the PET vector (Novagen).

Cloning of a human homologues of the *gl* gene

Database searches with the full length murine Gl polypeptide sequence identified homologous sequences in *C. elegans* and *D. melanogaster*. In contrast no human homologues were directly detected. However, highly  
5 conserved human EST clones were found and using genomic sequence of a PAC (Sanger center) combined with the mouse gene intron/exon structure, a *gl* human cDNA was assembled. The human sequence displayed high degree of conservation and close protein sequence identity with a 334 amino acids protein instead of 338 for the mouse protein.

10 The Gl polypeptide of the present invention comprises an amino sequence at least 89% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4 and functional homologues thereof, exclusive of a NH2-terminal signal peptide (Target Program).

15 The Gl polypeptide of the present invention is also defined to comprise a nucleic acid sequence at least 90% identical to a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2 and functional homologues thereof.

Antibodies for detecting Gl polypeptide

20 The present invention further provides an antibody that has a binding specificity to the Gl polypeptide of the present invention and fragments thereof.

Gl polypeptide antibodies can provide information on characteristic of the protein. For instance, generation of antibodies will enable the visualisation of the protein in cells and tissues using Western blotting.

25 In this technique, proteins are run on polyacrylamide gel and then transferred onto nitrocellulose membranes. These membranes are then incubated in the presence of the antibody (primary), then following washing are incubated with a secondary antibody which is used for detection of the

protein-primary antibody complex. Following repeated washing, the entire complex is visualised using colorimetric or chemiluninescent assays.

GI polypeptide antibodies also allow for the use of immunocytochemistry in immunofluorescent techniques in which the proteins can be visualised  
5 directly in cells and tissues. This is most helpful in order to establish the subcellular location of the protein and the tissues specificity of the protein.

In order to prepare polyclonal antibodies, fusion proteins containing defined portions or all of the GI polypeptide can be synthesised in bacteria or in fungi by expression of corresponding DNA sequences in a suitable cloning  
10 vehicle. The protein can then be purified, coupled to a carrier protein and mixed with an adjuvant known by one skilled in the art suitable and injected into laboratory animals such as mice.

Alternatively, protein can be isolated from cultured cells expressing the protein. Following busters injections at bi-weekly intervals, the mice or other  
15 laboratory animals are then bled and the protein isolated. These sera can be used directly or purified prior to use, by various methods including affinity chromatography, protein A-sepharose, antigene sepharose, antimouse Ig-sepharose. The sera can then be used to probe protein extract run on a polyacrylimide gel to identify the GI polypeptide. Alternatively, synthetic  
20 peptide can be made to the antigenic portion of the protein in use to inoculate the animals.

To produce monoclonal GI polypeptide antibodies are prepared according to standard techniques known by one skilled in the art. For instance, cells actively expressing the protein are cultured or isolated from tissues and the  
25 cells membranes isolated. The membranes, extracts or recombinant protein extracts, containing the GI polypeptide, are injected with an adjuvant into mice. After been injected nine times over a three weeks period, the mice spleens are removed and resuspended in phosphate saline buffer PSB. The spleen cells serve as a source of lymphocytes some of which are  
30 producing antibody of the appropriate specificity. These are then fused with a permanently growing myeloma partner cell and the product of the fusion are plated under a number of tissue culture wells in the presence of

selective agents, such as HAT. The wells are then screened to identify those containing cells making useful antibody by ELISA. These are then freshly plated. After a period of growth, these wells are again screened to identify antibody-producing cells. Several cloning procedure are then  
5 carried out until over 90% of the wells containing single clones, which are positive for antibody production. From this procedure to stable the line of clones is established which produce the antibody. The monoclonal antibody can then be purified by affinity chromatography using protein A sepharose, ion-exchange chromatography, as well as variation and combinations of  
10 these techniques.

*In situ* hybridisation is another method used to detect the expression of GI polypeptide. *In situ* hybridisation relies upon the hybridisation of specifically labelled nucleic acid probe to the cellular RNA in individual cells or tissues. Therefore, it allows the identification of mRNA within intact tissues such as  
15 the brain. In this method, oligonucleotide corresponding to unique portions of the *gl* gene are used to detect specific mRNA species in the tissue of interest.

Antibodies may also be used coupled to compounds for diagnostic and/or therapeutic uses such as radionucleic for imaging and therapy and  
20 liposome for the delivering of compound to a specific tissue location.

#### Process for producing the GI polypeptide

According to a preferred embodiment of the present invention, the GI polypeptide is produced with a process comprising the step of culturing a host cell that is transformed or transfected with an expression vector  
25 comprising the nucleic acid or amino acid sequence of any one of SEQ ID NO:1 to 4, under condition suitable for the expression of the polypeptide.

In a preferred embodiment of the present invention, the host cell is a colony forming unit granulocyte macrophage selected from the group consisting of granulocyte macrophage lineage and monocyte.

In the alternative, the GI polypeptide can be expressed in other cells such as insect cells using baculoviral vectors, or in mammalian cells using vaccinia virus or a specialised eukaryotic expression vectors. For expression in mammalian cells, the cDNA sequence may be ligated to  
5 heterologous promoters such as the simian virus (SV 40) promoter in the pSV2 vector or other similar vectors and introduced into cultured eukaryotic cells such as COS cells to achieve transit or a long term expression. The stable integration of the chimeric gene construct may be maintained in mammalian cells by biochemical selections such as neomycin and  
10 mycophoenolic acid.

Vectors are introduced into recipient cells by various methods including calcium phosphate, strontium, electroporation, lipofection, DEAE dextran, microinjection, or by photoplast fusion. Alternatively, the cDNA can be introduced by infection using viral vectors.

15 Using the techniques mentioned, the expression vectors containing the *gl* gene or portion thereof can be introduced into a variety of mammalian cells from other species or into non mammalian cells.

The recombinant cloning vector, according to this invention, comprises selected DNA of the DNA sequences of this invention for expression in a  
20 suitable host. The DNA is operatively linked in the vector to a promoter sequence in recombinant vehicle so that normal and/or mutant GI polypeptide can be expressed. The expression controlled sequence will be selected from the group consisting of sequences that control the expression of genes of prokaryotic or eukaryotic cells and the viruses and combination  
25 therefore. The expression controlled sequence may be selected from the group consisting of the lac system, the trp system, the tac system, the trc system, major operator and promoter regions of phage lambda, the control region of the fd coat protein, promoter of SV 40, promoters derived from polyoma, adenovirus, baculovirus, 3-phosphoglycerate kinase promoter,  
30 yeast promoters, combinations thereof.

The host cell which may be transfected with the vector of the present invention may be selected from the group consisting of bacteria, the yeast, fungi, insects, mouse or other animals, plant hosts or human tissue cells.

5 This process may further have a recovering and/or purifying step, wherein the polypeptide is recovered and/or recovered from the host cell through standard and well known procedures.

10 The GI polypeptide may be isolated and purified by methods selected on the basis of properties revealed by its sequence. Since the protein processes properties of a membrane-spanning protein, a membrane fraction of cells in which the protein is highly expressed would be isolated and the proteins removed by extraction and the protein solubilised using a detergent.

15 Purification can be achieved using protein purification procedures, such as chromatography methods (gel, filtration, ion-exchange and immune affinity), by high performance liquid chromatography (RP-HPLC, ion exchange HPLC, size-exclusion HPLD and high performance chromatofocusing and hydrophobic interaction chromatography) or by precipitation (immuno precipitation). Polyacrylamide gel electrophoresis can also be used to isolate the GI polypeptide based on its molecular weight, charge properties  
20 and hydrophobicity.

Similar procedures to those just mentioned could be used to purify the protein from cells transfected with vectors containing the GI polypeptide (e.g. baculovirus systems, yeast expression systems, eukaryotic expression systems). Purified protein can be used in further biochemical analysis to  
25 establish secondary and tertiary structure, which may aid in the design of pharmaceuticals to interact with the protein or charge interaction with other proteins, lipid or saccharide moieties, alter its function in membranes as a

transportor channel or receptor and/or in cells as an enzyme or structural protein in treated disease.

The protein may be in the form of a fusion protein GI polypeptide-GST, which will facilitate its purification. For example, a fusion protein may be created by ligating the GI cDNA sequence to a vector, which contains  
5 sequence for another peptide (e.g. GST-glutathione succinyl transferase). The fusion protein is expressed and recovered from a prokaryotic (e.g. bacterial or baculovirus) or an eukaryotic cell. The fusion protein can then be purified by affinity chromatography based upon the fusion vector  
10 sequence. The GI polypeptide can then be further purified from the fusion protein by enzymatic cleavage of the fusion protein.

### THERAPIES

Methods for preventing or treating a bone-related disease in a mammal are provided.

15 An important aspect of the biochemical studies using the genetic information of this invention is the development of therapies to circumvent or overcome the *gl* gene defects and thus prevent, treat, control serious symptoms or cure the disease. In view of expression of the *gl* gene in a variety of tissues, one has to recognise that other defects than osteoporosis  
20 and/or osteopetrosis may be caused by mutation in the *gl* gene in other tissues. Hence, in considering various therapies, it is understood that such therapies may be targeted at tissue other than the bone marrow, such as the heart, the testis, the spleen and the kidneys, where GI polypeptide is also expressed.

25 In a particular embodiment, the method comprises modulating the expression of the nucleic acid and/or the concentration of the GI polypeptide of the present invention. The expression and/or concentration

of the osteoclast-related polypeptide may be increased thereby preventing or treating a lack of bone resorption, such as osteopetrosis, in the mammal subject.

In this embodiment, the expression of the nucleic acid or the concentration  
5 of the polypeptide is increased by administering to the mammal subject at least one of the following: a functional nucleic acid molecule of the present invention; an expression or cloning vector having the nucleic acid molecule of the present invention; a host cell comprising the latter; a molecule for activating in said mammal the expression of the above mentioned nucleic  
10 acid molecule; an GI polypeptide of the present invention; a molecule for activating the production or increasing the concentration of the GI polypeptide of the invention; and a viral vector having the nucleic acid sequence of the invention.

In another embodiment, the expression and/or concentration of the GI  
15 polypeptide of the present invention is reduced, thereby treating or preventing an excess of bone resorption, such as osteoporosis in a mammal subject.

In this embodiment, the expression and/or concentration of the GI  
20 polypeptide is reduced by administering to the mammal subject at least one of the following molecule: a molecule having the function of inhibiting the expression of the nucleic acid sequence encoding for the GI polypeptide of the invention; a molecule for inhibiting the production of the GI polypeptide of the invention; and/or a molecule having the function of reducing the concentration of the GI polypeptide of the invention.

25 The molecule which has the function of inhibiting the expression of the nucleic acid mentioned herein above may be one that, for instance, binds to the nucleic acid thereby blocking the transcription and/or the translation steps of the polypeptide thus inhibiting its production.

In the case where the Gl polypeptide of the present invention is a receptor or an ion channel protein, a test for osteoporosis or osteopetrosis can be produced to detect an abnormal receptor or an abnormal function related to abnormalities that are inquired or inherited in the *gl* gene and its product, or  
5 in one of the homologues genes and their products. This test can be accomplished either *in vivo* or *in vitro* by measurements of ion channel fluxes and/or transmembrering voltage or current fluxes using patch, clamp, voltage clamp and fluorescent dies since it is to intracellular calcium or transmembrane voltage. Defective ion channel or receptor function can also  
10 be assayed by measurements of activation of second messengers such as cyclic AMP, cGMP kinases, phosphates, increases in intracellular  $Ca^{2+}$  levels, etc. Recombinantly made protein may also be reconstructed in artificial membrane systems to study ion channel conductance.

Therapies which affect bone related diseases can be tested by analysis of  
15 their ability to modify an abnormal ion channel or receptor function mutation in the *gl* gene in one of its homologues. Therapies could also be tested by their ability to modify the normal function of an ion channel or receptor capacity of the *gl* gene products and its homologues. Such assays can be performed on cultured cells expressing endogenous normal or mutant *gl*  
20 genes/gene products (or its homologues). Such studies can be performed in additional cells transfected with vectors capable of expressing Gl polypeptide, parts of the *gl* gene and gene product, mutant Gl polypeptide or of its homologues (abnormal or mutant form).

Therapies for bone related diseases could be divided to modify an  
25 abnormal ion channel or receptor function of the *gl* gene or its homologues. Such therapies can be conventional drugs, peptides, sugars or lipids as well as antibodies or other agents, which affect the properties of the *gl* gene product. Such therapies can also be performed by direct replacement of the *gl* gene and/or its homologues by gene therapy. In the case of an ion  
30 channel, the gene therapy could be performed using either many genes

(cDNA + a promoter) or a genomic construct bearing genomic DNA sequences for parts or all of the *gl* gene. Mutant Gl polypeptide or homologous gene sequences might also be used to counter the effect of the inherited or acquired abnormalities of the *gl* gene. The therapy may also  
5 be directed at augmenting the receptor Gl channel function of the homologous genes in order that it may potentially take over the functions of the *gl* gene rather defective by acquired or inherited defects. Therapies using antisense oligonucleotides to block the expression of the mutant *gl* gene co-ordinated with gene replacement with normal Gl polypeptide or a  
10 homologue gene can also be applied using standard techniques or either gene therapy or protein replacement therapy.

#### Pharmaceutical preparation

A pharmaceutical composition for preventing or treating osteopetrosis is provided. The composition comprises in a pharmaceutically effective  
15 amount a molecule that has the function of increasing the expression and/or concentration of the Gl polypeptide of the present invention. This molecule may be selected from the group of molecule used to prevent or treat osteopetrosis mentioned in the previous section.

A pharmaceutical composition for preventing or treating osteoporosis is  
20 also provided. The composition comprises in a pharmaceutically effective amount of a molecule having the function of reducing the expression and/or concentration of the Gl polypeptide of the present invention. This molecule may be selected from the group of molecule used for treating or preventing osteoporosis mentioned in the previous section.

25 The term "pharmaceutically effective amount" means an amount, which provides a therapeutic effect for a specified condition and route of administration.

According to various embodiments of the present invention, the pharmaceutical composition may further comprise pharmaceutically acceptable diluant, carrier, solubiliser, emulsifier, preservative and/or adjuvant.

- 5 The composition may be in a liquid or lyophilised form and comprises a diluant (Tris, acetate or phosphate buffers) having various pH values and/or ion exchange; solubiliser such as Tween or polysorbate; carriers such as human serum, albumin or gelatine; preservatives such as thimerosal or benzyl alcohol and antioxidants such as ascorbic acid or sodium  
10 metabisulfite.

The composition of the invention may be in solid or liquid form or any suitable form for a therapeutic use. They may be formulated for a rapid or slow release of its components. The composition of the invention may be prepared according to conventional methods known in the art.

15 Kit for screening the Gl polypeptide molecule of the present invention

Screening for a human related disease such as osteopetrosis and/or osteoporosis as link to chromosomes 6 may now be really carried out because of the knowledge of the location of the gene.

- 20 People with high risk for osteopetrosis or osteoporosis (person in family pedigree) or individuals not previously known to be at high risk, or people in general may be screened routinely using probes to detect the presence of a mutant *gl* gene by a variety of techniques.

- 25 Genomic DNA used for the diagnosis may be obtained from body cells, such as those present in the blood, tissue biopsy, and surgical specimens or autopsy material. The DNA may be isolated and used directly for

detection of its specific sequence or may be amplified part to analysis. RNA or cDNA may also be used.

To detect a specific DNA sequence, hybridisation using a specific oligonucleotide, direct DNA sequencing, restriction enzyme digest, RNase  
5 protection, chemical cleavage and ligase-mediated detection are all methods, which can be utilised.

Oligonucleotides specific to mutant sequences can be chemically synthesised and labelled radioactively with isotopes or non-radioactive using biotin tags and hybridised to individual DNA samples immobilised on  
10 membranes or other solid supports by dot-blot or transfer from gels after electrophoresis. The presence or absence of these mutant sequences is then visualised using methods such as autoradiography, fluorometry or colormetric reaction.

Direct DNA sequencing reveals sequence differences between normal and  
15 mutant OR polypeptide DNA. Cloned DNA segments may be used as probes to detect specific DNA segments. PCR can be used to enhance the sensitivity of this method by exponentially increasing of the target DNA.

Other nucleotide sequence simplification techniques may be used such as ligation-mediated PCR, anchored PCR and enzymatic amplification as  
20 would be understood by those skilled in the art.

Sequence alteration may also generate fortuitous restriction enzyme recognition sites, which are revealed by the use of appropriate enzyme digestion followed by gel electrophoresis and blot hybridisation. DNA fragments carrying the site (normal or mutant) are detected by their  
25 increased reduction size or by the increase of corresponding restriction family numbers. Genomic DNA samples may also be amplified by PCR prior to treatment with appropriate restriction enzyme and the fragments of

different sizes are visualised under UV light in the presence of ethidium bromide after gel electrophoresis.

- Genetic test is based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels.
- 5 Small sequence deletion and insertion can be visualised by high resolution gel electrophoresis. Small deletions may also be detected as changes in the migration pattern of DNA heteroduplexes in non denaturing gel electrophoresis. Alternatively a single base substitution or deletion mutation may be detected based on differential PCR product length in PCR. The
- 10 PCR products for the normal and mutant gene could be differentially detected in acrylamide gels.

- A kit for screening a nucleic acid sequence encoding for the GI polypeptide of the present invention is provided. The kit comprises a nucleic acid probe or primer complementary to the nucleic acid sequence of the present
- 15 invention; reagents for hybridization of the probe or primer to a complementary nucleic acid sequence; and means for detecting hybridization.

The present invention also provides a kit for detecting the presence of the GI polypeptide of the present invention.

- 20 In this embodiment, the kit comprises a probe or primer having a binding specificity to the GI polypeptide of the present invention; reagents for hybridization of the probe or primer to the GI polypeptide; and means for detecting hybridization.

EXAMPLE 1Mice

The mouse strain GL/Le *dl/+gl* was purchased from the Jackson Laboratory (Bar Harbor, ME). Homozygous *gl/gl* mice were generated by breeding  
5 heterozygous *gl/+* animals, and displayed a typical grey coat color instead of agouti, a major growth retardation and a lack of tooth eruption. All animals produced from these matings were genotyped at the *gl* locus by using cosegregating polymorphic markers that we have previously described. Vacher, J and Bernard, H., Mamm. Genome, 10, 239-243, 1999.

10

EXAMPLE 2BAC library screening and contig establishment

The 129/Sv CITB mouse BAC library (Research Genetics, Huntsville, AL) was screened by PCR using the markers D10Mit184. Amplification reaction was  
15 performed in a 20 µl of 10mM Tris-hydroxychloride, pH 8.3, 50mM KCl and 1.5mM MgCl<sub>2</sub>. Each reaction contained 10ng of DNA, 0.5 µM of each primer, 0,2 mM dNTP and 1U Taq polymerase (GIBCO-BRL). Thermal cycler conditions were 94°C, 5 min and 30 cycles (94°C, 1 min; 55°C, 1 min; 72°C, 2 min). PCR reactions were analyzed by gel electrophoresis on 10% acrylamide  
20 slab gel and specific products detected by ethidium staining.

A last round of screening on filter was carried out using BAC insert ends as probes. In brief, the membrane was prehybridized 2h at 65°C in 5X SSC 5X Denhardt's solution, 0.5% SDS and 10 mg/ml of sonicated denatured salmon  
25 sperm DNA, followed by hybridization with the BAC end probes overnight at 65°C in the same solution. The membrane was washed twice at 65°C in 1XSSC/0.1%SDS for 20 min and exposed to X-ray film. The size of each clone was determined by pulsed-field gel electrophoresis.

By this approach, we have established a contig of 18 adjacent clones using overlapping PCR assays derived from BAC end sequences and polymorphic markers.

### EXAMPLE 3

#### 5 Library screening and cDNA isolation

To isolate the full-length *gl* cDNA, we screened a C57BL/6 spleen cDNA mouse library (Stratagene) by PCR. Takumi, T. and Lodish, H.F. BioTechniques, 17:443-444 (1994). This library was divided in 16 pools, each of which contained approximately 100,000 clones. PCR assays were conducted with *gl* forward 5'-GGCGAGCTATCTGTTACAGTCC-3' and *gl* reverse 5'-TTACTGGCACAACGTGAGGTC-3' primers. PCR amplification conditions were 94°C, 5 min and 30 cycles (94°C, 1 min; 63°C, 1 min; 72°C, 2 min) in 20mM Tris-HCl, pH 8.4, 50mM KCl, 2mM MgCl<sub>2</sub>, 5% DMSO with 0.5mM dNTPs, 0.5µM primers and 1U Taq polymerase in 20µl volume. The last step of screening consisted of filter hybridization in the same conditions as described above. The cDNA was then sequenced (Thermosequenase, Amersham) and the protein open reading frame deduced.

### EXAMPLE 4

#### 20 Expression analysis

Expression analysis of the *gl* gene was carried out by both Northern blots and RT-PCR analysis.

25 First we have isolated total RNAs from adult mouse whole brain, liver, spleen, kidney, heart, thymus and testis tissues by a standard LiCl/Urea method as previously described. Vacher, J. and Tilghman, S.M. Science, 250:1732-1735 (1990). Total RNAs from osteoclast-like cells (OCLs) were isolated by TRIzol

(Gibco BRL) as previously described. Rajapurohitam, V. et al. Bone, 28:513-523 (2001).

For Northern analysis 15µg of total RNA or 2µg of polyA<sup>+</sup> RNA were  
5 fractionated by 1.5% agarose/2.2M formaldehyde gel electrophoresis and  
transferred onto membrane. The membrane was prehybridized 2h at 65°C in  
5X SSC, 5X Denhardt's solution, 0.5% SDS and 10 mg/ml sonicated denatured  
salmon sperm DNA, and hybridized overnight at 65°C in the same solution with  
a radiolabelled 1.9kb gl cDNA probe. The membrane was washed twice at  
10 65°C in 2X SSC/0.1% SDS for 20 min. and exposed to X-ray film.

For RT-PCR analysis, reverse transcription (Superscript II, Gibco BRL) of 1µg  
of RNA with oligo dT primer was conducted in 20mM Tris-HCl, pH 8.4, 50mM  
KCl, 1.5mM MgCl<sub>2</sub> with 0.5mM dNTPs, 0.5mM primers and 1U Taq  
15 polymerase in a 20µl volume. PCR amplification conditions of 1µl of cDNA  
were 94°C, 5 min and 30 cycles (94°C, 1min; 60°C, 1min; 72°C, 2min). The gl  
primers were: Forward 5'-CCTGCTTTGAGCATAACCTGC-3' and Reverse 5'-  
TTACTGGCACAACGTGAGGTC-3' and for beta-actin control were Forward 5'-  
TGACGATATCGCTGCGCTG-3' and Reverse 5'-  
20 ACATGGCTGGGGTGTGAAG-3'. PCR products were analyzed on 1%  
agarose gels and detected by ethidium bromide staining. Generation of BAC  
transgenic mice and histologic analysis

### EXAMPLE 5

#### 25 Generation of BAC transgenic mice and histologic analysis

Circular BAC DNA (1ng/µl) was injected into fertilized mouse oocytes isolated  
from F1 (C3H x C57BL/6) x C57BL/6 crosses. Transgenic founders were  
identified by PCR using specific BAC end sequence assay and internal  
polymorphic markers. Each founder was first crossed with heterozygote gl/+  
30 mice, and gl/+ transgenic progeny were intercrossed. The gl/gl transgenic mice

were then identified by homozygosity at the polymorphic *D10Mit184*, *D10Mit108* and *D10Mit255* loci. Vacher, J and Bernard, H., Mammalian. Genome, 10, 239-243, 1999.

5 Histology was done on bone samples fixed in 10% phosphate-buffered formalin, decalcified in 14% EDTA, and embedded in paraffin. Adjacent sections (6µm) were stained with hematoxylin and eosin.

### EXAMPLE 6

#### Gl gene structure

10 Intron-exon boundaries were characterized following alignment of the complete mouse cDNA sequence against mouse genomic sequences obtained by BLAST searches from NCBI Genomic Survey Sequence (GSS) and NCBI mouse Trace archive. Each intron-exon junction corresponds to the loss of alignment between cDNA and genomic sequences (usually at splicing consensus sites GT/AG).

15

Introns size was estimated by restriction mapping of genomic DNA, followed by membrane transfer and Southern blot hybridization (conditions described above) with various parts of the *gl* cDNA as probes.

20 Genomic DNA was prepared from tail biopsies as previously described. Laird, P W et al. Nucl Ac Res 19:4293 (1994). After restriction digests (BamHI, BglII, EcoRI), Southern blots were hybridized with *gl* cDNA probes in the same conditions as described above. The membrane was washed twice at 65°C in 1X SSC/0.1% SDS for 20 min. and exposed to X-ray film.

25

EXAMPLE 7Protein extracts, GL Antibodies and Western blotting

- OCLs were obtained by co-culturing one-day-old FVB/NJ calvarial osteoblasts and spleen cells of either +/+ or *gl/gl* mouse as previously described. Rajapurohitam, V. et al., Bone, 28:513-523 (2001). Cultured OCLs were washed twice with phosphate buffered saline and lysates were prepared in ice cold cell lysis buffer (50 mM sodium pyrophosphate, 50 mM sodium fluoride, 50 mM NaCl, 5 mM EDTA, 5 mM EGTA, 2 mM sodium ortho vanadate, 10 mM HEPES, 0.1% Triton X-100, 0.05% NP-40) in the presence of protease inhibitors. Lysates were sonicated for 30 sec, incubated on ice for 30 min and centrifuged at 12,000 x g for 20 min at 4°C. Supernatants were collected and protein concentrations determined by Bradford assay using BSA as the standard.
- Rabbit polyclonal antibodies Ab1 and Ab2 were raised against multiple-antigen peptides MAP1: LHSEQKKRKLILPKR-MAP and MAP2 LNGLENKAEPETHLC-MAP respectively. Protein extracts (25µg) were resolved on 12% SDS-PAGE gels and transferred onto nitrocellulose membranes. Following transfer, membranes were stained with Ponceau red to confirm uniform transfer and proteins integrity. Membranes were incubated in 5% milk for 1 h and then washed twice 10 min in Tris buffered saline Tween (TBS-T). Membranes were probed with polyclonal antisera (Ab1, 1:100 dilution; Ab2, 1:100 dilution or 31kDa V-ATPase subunit, 1:500 dilution in TBS-T, 3% BSA) for 1 h at room temperature. Membranes were washed twice 10 min in TBS-T, incubated with horseradish peroxidase-protein-A (HRP-A) secondary antibodies for 1 h at room temperature. After TBS-T washing, the signal was revealed by the ECL western blotting detection reagent (Amersham) and exposed on film.

EXAMPLE 8In situ hybridization and immunofluorescence

5 In situ hybridization was done as described previously. Emerson et al. Dev. Dynamics 195: 55-66, 1992. Bone samples were fixed in 10% phosphate-buffered formalin, decalcified in 14% EDTA, and embedded in paraffin. Paraffin was removed in xylenes and sections were fixed in 4% paraformaldehyde and hybridized to  $\alpha$ -S<sup>35</sup>UTP-labeled riboprobes overnight at 55°C. The *gl*  
10 antisense riboprobe was generated by T7 polymerase transcription of the 0.5kb 3'UTR fragment cloned into Bluescript and linearized by SpeI. The sense riboprobe was generated by T3 transcription of the same template linearized by KpnI. Hybridized sections were dipped in K2 photographic emulsion, exposed 2-3 weeks at 4°C, developed using D-19 developer and general fixer  
15 (Kodak) and stained with hematoxylin and eosin.

Immunofluorescence was conducted on isolated wild-type osteoclasts isolated from three-days old pups and cultured overnight on slides in  $\alpha$ -MEM with 10% fetal calf serum, in 5% CO<sub>2</sub>. Slides were washed in  
20 phosphate-buffered saline (PBS) and the cells were fixed in 4% paraformaldehyde in PBS for 10 min. Samples were then incubated at room temperature for 1hr in PBS containing 0.1%BSA, 0.05% saponin and 5% normal goat serum to block non-specific binding, and subsequently for 1hr with Gl primary antibody (1:50). Slides were then washed in PBS and  
25 incubated with secondary AlexaFluor 488-conjugated goat anti-rabbit IgG antibodies (1:100; Molecular Probes) for 1hr in the dark. For Hoechst staining, slides were incubated in 1:1500 dilution in water of a 0.5mg/ml Hoechst 33258 at room temperature for 10 min. After washing with PBS, samples were mounted in FluorSave (Calbiochem) and cells were  
30 visualized by confocal laser scanning microscopy (Axiophot, Zeiss).

Genomic structure of the mouse gl gene

| Exon<br>No. | Exon<br>Length (bp) | CDNA <sup>a</sup><br>Position | <u>Exon/Intron junction sequences<sup>c</sup></u> |                 | Intron<br>No. | Intron <sup>b</sup><br>Length (Kbp) |
|-------------|---------------------|-------------------------------|---------------------------------------------------|-----------------|---------------|-------------------------------------|
|             |                     |                               | Splice donor                                      | splice acceptor |               |                                     |
| 1           | 458                 | 1-458                         | ATCGGGgtgggt                                      | TtgcagAATACC    | 1             | ~3.9                                |
| 2           | 115                 | 459-573                       | GCGCAAgtagt                                       | TtacagATTGCC    | 2             | ~3.6                                |
| 3           | 98                  | 574-671                       | CTGCAGgtcagt                                      | TtttagGGGCAC    | 3             | ~2.8                                |
| 4           | 168                 | 672-839                       | GATGCAGtagt                                       | TtctagATGAAC    | 4             | ~1.7                                |
| 5           | 166                 | 840-                          | TTCTACgtaagt                                      | CcccagCCAAA     | 5             | ~1.8                                |
| 6           | 1973                | 1006-<br>2978                 |                                                   |                 |               |                                     |

a : cDNA sequence was obtained from clones isolated by screening the STRATAGENE and CLONETECH spleen libraries and comparing their sequences to the corresponding ESTs from GenBank and Riken database.

b : intron size was estimated from restriction mapping analysis.

c : Exon sequences are in uppercase letters, intron sequences are in lowercase.

Genomic structure of the human homologue of the mouse gl gene

| Exon<br>No. | Exon<br>Length (bp) | cDNA <sup>a</sup><br>Position | <u>Exon/Intron junction sequences<sup>c</sup></u> |                 | Intron<br>No. | Intron <sup>b</sup><br>Length (Kbp) |
|-------------|---------------------|-------------------------------|---------------------------------------------------|-----------------|---------------|-------------------------------------|
|             |                     |                               | Splice donor                                      | splice acceptor |               |                                     |
| 1           | 493                 | 1-493                         | GCGGGGgtggg                                       | TtacagAATACT    | 1             | 9.95                                |
| 2           | 115                 | 494-608                       | GTGCAAgtaagt                                      | TgacagATTGTT    | 2             | 9.597                               |
| 3           | 98                  | 609-706                       | CTTCAGgtattt                                      | TtttagGGGAAT    | 3             | 3.291                               |
| 4           | 168                 | 707-874                       | GATGCAgtaagt                                      | TtccagATGAAC    | 4             | 1.612                               |
| 5           | 166                 | 875-1040                      | TTCTGCgtaagt                                      | AtctagCCAAAC    | 5             | 4.412                               |
| 6           | 2023                | 1041-3069                     |                                                   |                 |               |                                     |

a : cDNA sequence was obtained by aligning the mouse cDNA sequence against the Genbank ESTs database.

b : Genomic sequence was obtained from the sequenced human PAC (RP1-111B22) at Sanger center (Acc No. : Z98200).

c : Exon sequences are in uppercase letters, intron sequences are in lowercase.

## SEQUENCE LISTING

<110> Aventis Pharma S.A

<120> genetic sequence related to bone diseases

<130> 11425-0408

<160> 4

<170> PatentIn version 3.0

<210> 1

<211> 2997

<212> DNA

<213> Mus musculus

<220>

<221> CDS

<222> (45)..(1061)

<400> 1

gtcgggaagca ccgggcgagc tatctgttac agtccggccc gggg atg gct cgg gac 56  
Met Ala Arg Asp  
1

gcg gag ctg gcg cgc agt agc ggg tgg ccg tgg cgg tgg ctg ccg gcg 104  
Ala Glu Leu Ala Arg Ser Ser Gly Trp Pro Trp Arg Trp Leu Pro Ala

|                                                                 |     |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|-----|
| 5                                                               | 10  | 15  | 20  |     |
| ctg ctg ctg ctg cag ctg ctg cgg tgg agg tgc gcc ctg tgc gcg ctc |     |     |     | 152 |
| Leu Leu Leu Leu Gln Leu Leu Arg Trp Arg Cys Ala Leu Cys Ala Leu |     |     |     |     |
|                                                                 | 25  | 30  | 35  |     |
| ccc ttc acc agc agt cgg cac cca ggc ttt gcg gac ctg ctg tcg gag |     |     |     | 200 |
| Pro Phe Thr Ser Ser Arg His Pro Gly Phe Ala Asp Leu Leu Ser Glu |     |     |     |     |
|                                                                 | 40  | 45  | 50  |     |
| cag cag ctg ttg gag gtg cag gac ttg acc ctg tct ttg ctg cag ggc |     |     |     | 248 |
| Gln Gln Leu Leu Glu Val Gln Asp Leu Thr Leu Ser Leu Leu Gln Gly |     |     |     |     |
|                                                                 | 55  | 60  | 65  |     |
| gga ggt cta ggg ccg ctg tca ctg cta cct ccg gac ctg ccg gat ctg |     |     |     | 296 |
| Gly Gly Leu Gly Pro Leu Ser Leu Leu Pro Pro Asp Leu Pro Asp Leu |     |     |     |     |
| 70                                                              | 75  | 80  |     |     |
| gag cct gag tgc cgg gag ctg ctg atg gac ttc gcc aat agc agc gcc |     |     |     | 344 |
| Glu Pro Glu Cys Arg Glu Leu Leu Met Asp Phe Ala Asn Ser Ser Ala |     |     |     |     |
| 85                                                              | 90  | 95  | 100 |     |
| gag ctg acc gcc tgt atg gtg cgc agc gct cgg ccc gtg cgc ctc tgc |     |     |     | 392 |
| Glu Leu Thr Ala Cys Met Val Arg Ser Ala Arg Pro Val Arg Leu Cys |     |     |     |     |
|                                                                 | 105 | 110 | 115 |     |
| cag acc tgc tac ccg ctc ttc caa cag gtc gca atc aag atg gac aac |     |     |     | 440 |
| Gln Thr Cys Tyr Pro Leu Phe Gln Gln Val Ala Ile Lys Met Asp Asn |     |     |     |     |

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 120                                                             | 125 | 130 |     |
| atc agc cga aac atc ggg aat acc tcc gag ggc ccg cgc tgt ggc gga |     |     | 488 |
| Ile Ser Arg Asn Ile Gly Asn Thr Ser Glu Gly Pro Arg Cys Gly Gly |     |     |     |
| 135                                                             | 140 | 145 |     |
| agt ctc ctg acg gca gac aga atg cag ata gtt ctc atg gtc tct gag |     |     | 536 |
| Ser Leu Leu Thr Ala Asp Arg Met Gln Ile Val Leu Met Val Ser Glu |     |     |     |
| 150                                                             | 155 | 160 |     |
| ttt ttc aac agc acg tgg cag gag gcg aac tgc gca aat tgc cta aca |     |     | 584 |
| Phe Phe Asn Ser Thr Trp Gln Glu Ala Asn Cys Ala Asn Cys Leu Thr |     |     |     |
| 165                                                             | 170 | 175 | 180 |
| aac aat ggt gag gat ttg tca aac aac aca gag gac ttc ctc agt ctg |     |     | 632 |
| Asn Asn Gly Glu Asp Leu Ser Asn Asn Thr Glu Asp Phe Leu Ser Leu |     |     |     |
| 185                                                             | 190 | 195 |     |
| ttt aac aag act ttg gcc tgc ttt gag cat aac ctg cag ggg cac aca |     |     | 680 |
| Phe Asn Lys Thr Leu Ala Cys Phe Glu His Asn Leu Gln Gly His Thr |     |     |     |
| 200                                                             | 205 | 210 |     |
| tac agt ctc ctc cca cca aaa aat tac tcc gaa gtg tgc aga aac tgt |     |     | 728 |
| Tyr Ser Leu Leu Pro Pro Lys Asn Tyr Ser Glu Val Cys Arg Asn Cys |     |     |     |
| 215                                                             | 220 | 225 |     |
| aaa gag gca tat aaa aac ctg agc ctc ctg tac agt caa atg cag aaa |     |     | 776 |
| Lys Glu Ala Tyr Lys Asn Leu Ser Leu Leu Tyr Ser Gln Met Gln Lys |     |     |     |

|                                                                    |     |     |      |
|--------------------------------------------------------------------|-----|-----|------|
| 230                                                                | 235 | 240 |      |
| ctg aac ggg ctt gag aac aag gct gag cct gag acg cac ttg tgc atc    |     |     | 824  |
| Leu Asn Gly Leu Glu Asn Lys Ala Glu Pro Glu Thr His Leu Cys Ile    |     |     |      |
| 245                                                                | 250 | 255 | 260  |
| gat gtg gag gat gca atg aac att act cgg aag ctt tgg agt cga acc    |     |     | 872  |
| Asp Val Glu Asp Ala Met Asn Ile Thr Arg Lys Leu Trp Ser Arg Thr    |     |     |      |
|                                                                    | 265 | 270 | 275  |
| ttc aac tgt tcg gtc acc tgc agc gac acg gtg tcc gtg gtt gct gtg    |     |     | 920  |
| Phe Asn Cys Ser Val Thr Cys Ser Asp Thr Val Ser Val Val Ala Val    |     |     |      |
|                                                                    | 280 | 285 | 290  |
| tct gtg ttc att ctc ttc ctg cct gtc gtc ttc tac ctc agt agc ttc    |     |     | 968  |
| Ser Val Phe Ile Leu Phe Leu Pro Val Val Phe Tyr Leu Ser Ser Phe    |     |     |      |
|                                                                    | 295 | 300 | 305  |
| ctt cac tca gag caa aag aaa cgc aaa ctc att cta ccc aaa cgt ctc    |     |     | 1016 |
| Leu His Ser Glu Gln Lys Lys Arg Lys Leu Ile Leu Pro Lys Arg Leu    |     |     |      |
| 310                                                                | 315 | 320 |      |
| aag tcg agc acc agt ttt gcc aac att caa gaa aat gcc acc tga        |     |     | 1061 |
| Lys Ser Ser Thr Ser Phe Ala Asn Ile Gln Glu Asn Ala Thr            |     |     |      |
| 325                                                                | 330 | 335 |      |
| agcctgcaaa acggggactc gacctcacgt tgtgccagta agtggttagac cacagcacag |     |     | 1121 |

tcgagaagaa gatgagccaa ggtcggacaa gttgcattct cacgaaatgt tgggactgca 1181

gacctataat ttattctgaa taagggttct caaattccct tttcctgagc accccttttt 1241

ttttttttga agatttctgt attttttagtt ttcaaacata gcaatgttac atattttaag 1301

gtatatctgt tacaataaca agtgagggtt tttttctcag gcatatgaat gactactgga 1361

cacttctgat ttatcctcgt tagcagaagt acacaaagca gaaaaggctg aggtctgcta 1421

tttacacatt agtcactggg agcccactct gaaaaagaaa catacttgcc aaatggtagc 1481

aggctcagtg attaacttaa gtgaattccc attgtagtat tgttgtatgt atatacatc 1541

atgcacacac acacacatat atatatacac gtatacatag atgtatatat gtaatgtata 1601

cttaatatat catacattaa aataatgttc tctagttccc tgaagtcctt tttgaaacca 1661

ctagttgatt ataaacctcc ttaacagttt tcagagagtg attccacatt atgcatttat 1721

ccttgttaaa ggtttacagt aactgagggt ctaatatgac tttataaat actattttac 1781

atcttatttt tgtctttatt tagtaagtaa tttataatca ctggactgct taattacctt 1841

tgaggacaag atggattcat cttatgccag ggatttgcac catgaatttc attaagttat 1901

ttggcaacct gtaacttggt agtagttcaa gtcgaatgtc acccaagtgt gtcatactgt 1961

gttttaaattt gtgattttttt ttaatgaaaa ttttatcttg gaatatattgg agatttgggg 2021

agaaacaagg acaaacacaa gagcttaaata ttcagaaaaat agacagggac ctgagggatg 2081

ctcacggtga gacagctgcg tggtttacac tggagatgac tcggttgaca ggctcgcaca 2141

ggaagcctcc cagttacggg aaagatgaaa gtcacatgac tgaaacgaaa ttacccatct 2201

cactgtcagg aaactagttc ttctttggca tatttctagc aacctttaaa accatgcttg 2261

tttcagtgtc actcagttgt atttctcaag atgtagaagt tgatggtttt gttggttaat 2321

ccggtggaaa cgggctttgt tgtaaaggta atgaatagga aactcctcag attcaatggg 2381

taagaaaatg tgactccctt cacaacctgt aattgcccta caggaaggca ggagtgtttg 2441

ggtatttttt gtatgtttcc cacatatgca gagtgtgaga gcaggctagt cttagtccca 2501

gagtgtgtca caccgggtat gtgacaatca gacgacgctg tgatccacta gatgtgccgg 2561

ggttcattgt gctgtcattg ttctgtctt gatttgaagc acatgggtga gggtcattgg 2621

aagccatctt catcagtaca tgtaaagctt atttacatgt gcaaagtgag tgaagtgaca 2681

tatttaaact gtgagtagcg actcctcggg tacctttcag tactgtgtgt acaaaccact 2741

gcttttggct aagaagctgg agagcacttt aacaagccag ccatctctgt tcctgatcag 2801

WO 03/029283

PCT/EP02/10721

ggctctggctc tctagagggt gcattagaaa tatatttgaa aatgtgccaa agaatttcat 2861

cttgtggtca tattaaaaaa atgtacatag ttctgaatcc tgaggcacat agggttatgt 2921

gtgtgcacaa gaaaacctgt tttttcctta tgctttacaa taaaggaaat aacaaggaaa 2981

aaaaaaaaaa aaaaaa 2997

<210> 2

<211> 338

<212> PRT

<213> Mus musculus

<400> 2

Met Ala Arg Asp Ala Glu Leu Ala Arg Ser Ser Gly Trp Pro Trp Arg  
1 5 10 15

Trp Leu Pro Ala Leu Leu Leu Leu Gln Leu Leu Arg Trp Arg Cys Ala  
20 25 30

Leu Cys Ala Leu Pro Phe Thr Ser Ser Arg His Pro Gly Phe Ala Asp  
35 40 45

Leu Leu Ser Glu Gln Gln Leu Leu Glu Val Gln Asp Leu Thr Leu Ser  
50 55 60

Leu Leu Gln Gly Gly Gly Leu Gly Pro Leu Ser Leu Leu Pro Pro Asp  
65 70 75 80

Leu Pro Asp Leu Glu Pro Glu Cys Arg Glu Leu Leu Met Asp Phe Ala  
85 90 95

Asn Ser Ser Ala Glu Leu Thr Ala Cys Met Val Arg Ser Ala Arg Pro  
100 105 110

Val Arg Leu Cys Gln Thr Cys Tyr Pro Leu Phe Gln Gln Val Ala Ile  
115 120 125

Lys Met Asp Asn Ile Ser Arg Asn Ile Gly Asn Thr Ser Glu Gly Pro  
130 135 140

Arg Cys Gly Gly Ser Leu Leu Thr Ala Asp Arg Met Gln Ile Val Leu  
145 150 155 160

Met Val Ser Glu Phe Phe Asn Ser Thr Trp Gln Glu Ala Asn Cys Ala  
165 170 175

Asn Cys Leu Thr Asn Asn Gly Glu Asp Leu Ser Asn Asn Thr Glu Asp  
180 185 190

Phe Leu Ser Leu Phe Asn Lys Thr Leu Ala Cys Phe Glu His Asn Leu  
195 200 205

Gln Gly His Thr Tyr Ser Leu Leu Pro Pro Lys Asn Tyr Ser Glu Val  
210 215 220

Cys Arg Asn Cys Lys Glu Ala Tyr Lys Asn Leu Ser Leu Leu Tyr Ser  
225 230 235 240

Gln Met Gln Lys Leu Asn Gly Leu Glu Asn Lys Ala Glu Pro Glu Thr  
245 250 255

His Leu Cys Ile Asp Val Glu Asp Ala Met Asn Ile Thr Arg Lys Leu  
260 265 270

WO 03/029283

PCT/EP02/10721

Trp Ser Arg Thr Phe Asn Cys Ser Val Thr Cys Ser Asp Thr Val Ser  
275 280 285

Val Val Ala Val Ser Val Phe Ile Leu Phe Leu Pro Val Val Phe Tyr  
290 295 300

Leu Ser Ser Phe Leu His Ser Glu Gln Lys Lys Arg Lys Leu Ile Leu  
305 310 315 320

Pro Lys Arg Leu Lys Ser Ser Thr Ser Phe Ala Asn Ile Gln Glu Asn  
325 330 335

Ala Thr

<210> 3

<211> 3082

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (92)..(1096)

&lt;400&gt; 3

cgctcgcgga aaccggaagc ggcggctgtc cgcggtgccg gctggggggcg gagaggcggc 60

ggtgggctcc ctgggggtgtg tgagcccggg g atg gag ccg ggc ccg aca gcc 112

Met Glu Pro Gly Pro Thr Ala

1

5

gcg cag cgg agg tgt tcg ttg ccg ccg tgg ctg ccg ctg ggg ctg ctg 160

Ala Gln Arg Arg Cys Ser Leu Pro Pro Trp Leu Pro Leu Gly Leu Leu

10

15

20

ctg tgg tcg ggg ctg gcc ctg ggc gcg ctc ccc ttc ggc agc agt ccg 208

Leu Trp Ser Gly Leu Ala Leu Gly Ala Leu Pro Phe Gly Ser Ser Pro

25

30

35

cac agg gtc ttc cac gac ctc ctg tcg gag cag cag ttg ctg gag gtg 256

His Arg Val Phe His Asp Leu Leu Ser Glu Gln Gln Leu Leu Glu Val

40

45

50

55

gag gac ttg tcc ctg tcc ctc ctg cag ggt gga ggg ctg ggg cct ctg 304

Glu Asp Leu Ser Leu Ser Leu Leu Gln Gly Gly Gly Leu Gly Pro Leu

60

65

70

tcg ctg ccc ccg gac ctg ccg gat ctg gat cct gag tgc cgg gag ctc 352

Ser Leu Pro Pro Asp Leu Pro Asp Leu Asp Pro Glu Cys Arg Glu Leu

75

80

85

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| ctg ctg gac ttc gcc aac agc agc gca gag ctg aca ggg tgt ctg gtg | 400 |
| Leu Leu Asp Phe Ala Asn Ser Ser Ala Glu Leu Thr Gly Cys Leu Val |     |
| 90 95 100                                                       |     |
|                                                                 |     |
| cgc agc gcc cgg ccc gtg cgc ctc tgt cag acc tgc tac ccc ctc ttc | 448 |
| Arg Ser Ala Arg Pro Val Arg Leu Cys Gln Thr Cys Tyr Pro Leu Phe |     |
| 105 110 115                                                     |     |
|                                                                 |     |
| caa cag gtc gtc agc aag atg gac aac atc agc cga gcc gcg ggg aat | 496 |
| Gln Gln Val Val Ser Lys Met Asp Asn Ile Ser Arg Ala Ala Gly Asn |     |
| 120 125 130 135                                                 |     |
|                                                                 |     |
| act tca gag agt cag agt tgt gcc aga agt ctc tta atg gca gat aga | 544 |
| Thr Ser Glu Ser Gln Ser Cys Ala Arg Ser Leu Leu Met Ala Asp Arg |     |
| 140 145 150                                                     |     |
|                                                                 |     |
| atg caa ata gtt gtg att ctc tca gaa ttt ttt aat acc aca tgg cag | 592 |
| Met Gln Ile Val Val Ile Leu Ser Glu Phe Phe Asn Thr Thr Trp Gln |     |
| 155 160 165                                                     |     |
|                                                                 |     |
| gag gca aat tgt gca aat tgt tta aca aac aac agt gaa gaa tta tca | 640 |
| Glu Ala Asn Cys Ala Asn Cys Leu Thr Asn Asn Ser Glu Glu Leu Ser |     |
| 170 175 180                                                     |     |
|                                                                 |     |
| aac agc aca gta tat ttc ctt aat cta ttt aat cac acc ctg acc tgc | 688 |
| Asn Ser Thr Val Tyr Phe Leu Asn Leu Phe Asn His Thr Leu Thr Cys |     |
| 185 190 195                                                     |     |

ttt gaa cat aac ctt cag ggg aat gca cat agt ctt tta cag aca aaa 736  
 Phe Glu His Asn Leu Gln Gly Asn Ala His Ser Leu Leu Gln Thr Lys  
 200 205 210 215

aat tat tca gaa gta tgc aaa aac tgc cgt gaa gca tac aaa act ctg 784  
 Asn Tyr Ser Glu Val Cys Lys Asn Cys Arg Glu Ala Tyr Lys Thr Leu  
 220 225 230

agt agt ctg tac agt gaa atg caa aaa atg aat gaa ctt gag aat aag 832  
 Ser Ser Leu Tyr Ser Glu Met Gln Lys Met Asn Glu Leu Glu Asn Lys  
 235 240 245

gct gaa cct gga aca cat tta tgc att gat gtg gaa gat gca atg aac 880  
 Ala Glu Pro Gly Thr His Leu Cys Ile Asp Val Glu Asp Ala Met Asn  
 250 255 260

atc act cga aaa cta tgg agt cga act ttc aac tgt tca gtc cct tgc 928  
 Ile Thr Arg Lys Leu Trp Ser Arg Thr Phe Asn Cys Ser Val Pro Cys  
 265 270 275

agt gac aca gtg cct gta att gct gtt tct gtg ttc att ctc ttt cta 976  
 Ser Asp Thr Val Pro Val Ile Ala Val Ser Val Phe Ile Leu Phe Leu  
 280 285 290 295

cct gtt gtc ttc tac ctt agt agc ttt ctt cac tca gag caa aag aaa 1024  
 Pro Val Val Phe Tyr Leu Ser Ser Phe Leu His Ser Glu Gln Lys Lys  
 300 305 310

cgc aaa ctc att ctg ccc aaa cgt ctc aag tcc agt acc agt ttt gca 1072  
Arg Lys Leu Ile Leu Pro Lys Arg Leu Lys Ser Ser Thr Ser Phe Ala  
315 320 325

aat att cag gaa aat tca aac tga gacctacaaa atggagaatt gacatatcac 1126  
Asn Ile Gln Glu Asn Ser Asn  
330

gtgaatgaat ggtggaagac acaacttggt ttcagaaaga agataaactg tgatttgaca 1186

agtcaagctc ttaagaaata caaggacttc agatccatct ttaaataaga attttcgatt 1246

tttctttcct tttccacttc tttctaacag atttggatat ttttaatttc caggcatagc 1306

aatggtatct attttaattgt gtatttggtca caataacaga acatgcaaga acaatcatta 1366

ttttatttta taggcatttg attactattc tagacttctg gtatcttctt actaacataa 1426

atatctcaag tagaaaagtt tttgaaaact aacattttaa aattaatcag ttacagtaaa 1486

gactttgaaa aagaaatgta cttgttagga agtagcttaa ttacccccca ttgcagtatt 1546

attggttatat atatagttaa tatgttggtac atcacaataa tatataattc agtctctagt 1606

ttccctagag tcattttttga aaccactgat tgcaaacctc cctgacaatt tttaaaagta 1666

gtaagccaca ttacatttat ctttgtaaaa agatttatgg taactggttt cttacttgac 1726

ttttataaat agtattttac atcttatttt tgcctttatt tcataagtaa tttaaaaatc 1786

actggattgc tttattatat tcagggcaat atggattatt ttataccaa ggatttgcac 1846

cgtgaattac attaagttat ttggcaattt ataatttatt actacttta atcaaagtga 1906

gcattatcac actgtattta aattgtcatt ttttaaagga atattttctt cttaagatat 1966

atagaggatt ttggagaaga gagacaggag gggtaaaacc agcttaaggt tcagcgagca 2026

gaaagggacc tgagaggatg ctactgtaa gactgttga cagtgggtgtg tattgagggg 2086

atgaatcgga acgatagtct catgcagaaa atagtgagat taagatcatc cttattgttt 2146

ctaaattatt tcaatcagat gaaagtgata cgattgaaat gaaatcacat agttcgtgct 2206

cagaaattct attttgggtat gtttgtatta gccttttagaa aaaacactcc gtttcagaat 2266

tgttcacagt tttatttctt aggttttttag agttcaggat ttcatttatt aatttcttct 2326

tgcttttttg gtggaaatag gctttgttgt aaacattaag aatataaaat ctctctata 2386

tagaaacaag aattttgtta aaaagagaat ttgaatccct tcctatacta taaaatgctc 2446

tatagggaga caaagtgttt cttttttctt ttatgtttac tgtttatgtg gagtgaaata 2506

taaggctctt ggatgtataa cataactcaa agctgttaca ctttctctga tctgctgtga 2566

tccactgaaa atgtgctggg gtttggtctg ctgtcactgt ttatgctgct ggaacttagc 2626

actgtcttga tttgaagcat atgattgaga gccatttgaa gcaatcttca ttaatgcaga 2686

taaaacaagt ttacatgtgc agagttagaa aatgacatgt tcaattctgt aagtgggtgac 2746

tttttgagca cttttcagta ttatgtattt gtaaaaacca ttgttttttg atataaagct 2806

aataagcact ttaaaaagga aaaggcagcc ttactattt tttctggttg agtcattgct 2866

ctttagacct agcatcagca atagatttca aagataagta ttaagcgcta ccctaaagtg 2926

tgtaagtttt tcattttgtc atattgaaaa atgatttgca tagtactgaa tgttgacaca 2986

cagcttatat gtattttacaa gaatatcttt aagtgttttt ttgacacatt aaaataaagg 3046

aaataaggaa attgtaaaaa aaaaaaaaaa aaaaaa 3082

&lt;210&gt; 4

&lt;211&gt; 334

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

Met Glu Pro Gly Pro Thr Ala Ala Gln Arg Arg Cys Ser Leu Pro Pro

1

5

10

15

Trp Leu Pro Leu Gly Leu Leu Leu Trp Ser Gly Leu Ala Leu Gly Ala  
20 25 30

Leu Pro Phe Gly Ser Ser Pro His Arg Val Phe His Asp Leu Leu Ser  
35 40 45

Glu Gln Gln Leu Leu Glu Val Glu Asp Leu Ser Leu Ser Leu Leu Gln  
50 55 60

Gly Gly Gly Leu Gly Pro Leu Ser Leu Pro Pro Asp Leu Pro Asp Leu  
65 70 75 80

Asp Pro Glu Cys Arg Glu Leu Leu Leu Asp Phe Ala Asn Ser Ser Ala  
85 90 95

Glu Leu Thr Gly Cys Leu Val Arg Ser Ala Arg Pro Val Arg Leu Cys  
100 105 110

Gln Thr Cys Tyr Pro Leu Phe Gln Gln Val Val Ser Lys Met Asp Asn  
115 120 125

Ile Ser Arg Ala Ala Gly Asn Thr Ser Glu Ser Gln Ser Cys Ala Arg  
130 135 140

Ser Leu Leu Met Ala Asp Arg Met Gln Ile Val Val Ile Leu Ser Glu  
145 150 155 160

Phe Phe Asn Thr Thr Trp Gln Glu Ala Asn Cys Ala Asn Cys Leu Thr  
165 170 175

Asn Asn Ser Glu Glu Leu Ser Asn Ser Thr Val Tyr Phe Leu Asn Leu  
180 185 190

Phe Asn His Thr Leu Thr Cys Phe Glu His Asn Leu Gln Gly Asn Ala  
195 200 205

His Ser Leu Leu Gln Thr Lys Asn Tyr Ser Glu Val Cys Lys Asn Cys  
210 215 220

Arg Glu Ala Tyr Lys Thr Leu Ser Ser Leu Tyr Ser Glu Met Gln Lys  
225 230 235 240

Met Asn Glu Leu Glu Asn Lys Ala Glu Pro Gly Thr His Leu Cys Ile  
245 250 255

Asp Val Glu Asp Ala Met Asn Ile Thr Arg Lys Leu Trp Ser Arg Thr  
260 265 270

Phe Asn Cys Ser Val Pro Cys Ser Asp Thr Val Pro Val Ile Ala Val  
275 280 285

Ser Val Phe Ile Leu Phe Leu Pro Val Val Phe Tyr Leu Ser Ser Phe  
290 295 300

Leu His Ser Glu Gln Lys Lys Arg Lys Leu Ile Leu Pro Lys Arg Leu  
305 310 315 320

Lys Ser Ser Thr Ser Phe Ala Asn Ile Gln Glu Asn Ser Asn  
325 330

## CLAIMS

1. A method for preventing or treating a bone resorption-related disease in a mammal subject, comprising modulating in said mammal subject:
  - the expression of a nucleic acid sequence at least 90% identical to a sequence selected from the group consisting of SEQ ID NO: 1 and SEQ ID NO: 2, and/or
  - the concentration of an GI polypeptide comprising an amino acid sequence that is encoded by said nucleic acid sequence.
2. The method according to claim 1, wherein said nucleic acid sequence is 100% identical to a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, and fragments thereof.
3. The method according to claim 2, wherein said nucleic acid sequence is SEQ ID NO: 2.
4. The method according to any one of claims 1 to 3, wherein said amino acid sequence is at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO: 3 and SEQ ID NO: 4.
5. The method according to claim 4, wherein said amino acid sequence is 100% identical to an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 4, and fragments thereof.
6. The method according to claim 5, wherein said amino acid sequence is SEQ ID NO: 4.

7. The method according to any one of claims 1 to 6, wherein said GI polypeptide comprises a transmembrane domain.
8. The method according to claim 7, wherein said GI polypeptide is a receptor.
9. The method according to any one of claims 1 to 8, wherein said expression of said nucleic acid sequence and/or said concentration of said GI polypeptide is increased, thereby preventing or treating a lack of bone resorption in said mammal subject.
10. The method according to claim 9, wherein said expression of said nucleic acid sequence and/or said concentration of said GI polypeptide is increased by administering to said mammal subject at least one of the following:
- a functional nucleic acid sequence at least 90% identical to a sequence selected from the group consisting of SEQ ID NO: 1 and SEQ ID NO: 2,
  - an expression or cloning vector having said nucleic acid sequence,
  - a molecule for activating, in said mammal subject, said expression of said nucleic acid sequence,
  - said GI polypeptide,
  - a molecule for activating, in said mammal subject, the production of said GI polypeptide,

- a molecule for increasing, in said mammal subject, said concentration of said GI polypeptide,
- a host cell transformed or transfected with said expression or cloning vector, and
- a viral vector having said nucleic acid sequence.

11. The method according to any one of claims 1 to 10, wherein said bone resorption-related disease is osteopetrosis.

12. The method according to claim 1, wherein said expression of said nucleic acid sequence and/or said concentration of said GI polypeptide is reduced, thereby preventing or treating an excess of bone resorption in said mammal subject.

13. The method according to claim 12, wherein said expression of said nucleic acid sequence and/or said concentration of said GI polypeptide is reduced by administering to said mammal subject at least one of the following:

- a molecule having the function of inhibiting said expression of said nucleic acid sequence,
- a molecule having the function of inhibiting said production of said GI polypeptide, and/or
- a molecule having the function of reducing said concentration of said GI polypeptide.

14. The method according to claim 13, wherein said molecule having the function of inhibiting said expression of said nucleic acid sequence is a molecule that binds therewith.

15. The method according to claim 13, wherein said molecule having the function of inhibiting said production of said GI polypeptide is a molecule that blocks the translation of said GI polypeptide.

16. The method according to any one of claims 1 to 8, and 12 to 15, wherein said bone resorption-related disease is osteoporosis.

17. Use of a pharmaceutically effective amount of at least one of the following:

- a functional nucleic acid sequence at least 90% identical to a sequence selected from the group consisting of SEQ ID NO: 1 and SEQ ID NO: 2,

- an expression or cloning vector having said nucleic acid sequence,

- a molecule for activating the expression of said nucleic sequence,

- an GI polypeptide comprising an amino acid sequence that is encoded by said nucleic acid sequence,

- a molecule for activating the production of said GI polypeptide,

- a molecule for increasing the concentration of said GI polypeptide,

- a host cell transformed or transfected with said expression or cloning vector, and

- a viral vector having said nucleic acid sequence,

for the manufacture of a pharmaceutical composition for treating or preventing a lack of bone resorption in a mammal subject.

18. The use according to claim 17, wherein said lack of bone resorption is osteopetrosis.

19. Use of a pharmaceutically effective amount of at least one of the following:

- a molecule having the function of inhibiting the expression of a nucleic acid sequence at least 90% identical to a sequence selected from the group consisting of SEQ ID NO: 1 and SEQ ID NO: 2,

- a molecule having the function of inhibiting the production of an GI polypeptide comprising an amino acid sequence that is encoded by said nucleic acid sequence, and/or

- a molecule having the function of reducing the concentration of said GI polypeptide,

for the manufacture of a pharmaceutical composition for treating or preventing an excess of bone resorption in a mammal subject.

20. The use according to claim 19, wherein said excess of bone resorption is osteoporosis.

21. The use according to any one of claims 17 to 20, wherein said nucleic acid sequence is 100% identical to a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO: 2, and fragments thereof.

22. The use according to claim 21, wherein said nucleic acid sequence is SEQ ID NO: 2.

23. The use according to any one of claims 17 to 22, wherein said amino acid sequence is at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO: 3 and SEQ ID NO: 4.

24. The use according to claim 23, wherein said amino acid sequence is 100% identical to an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 4, and fragments thereof.

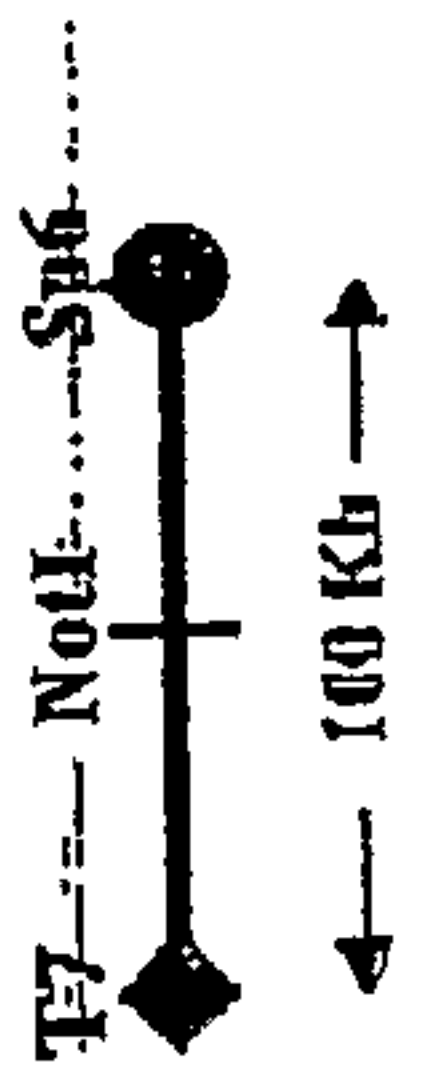
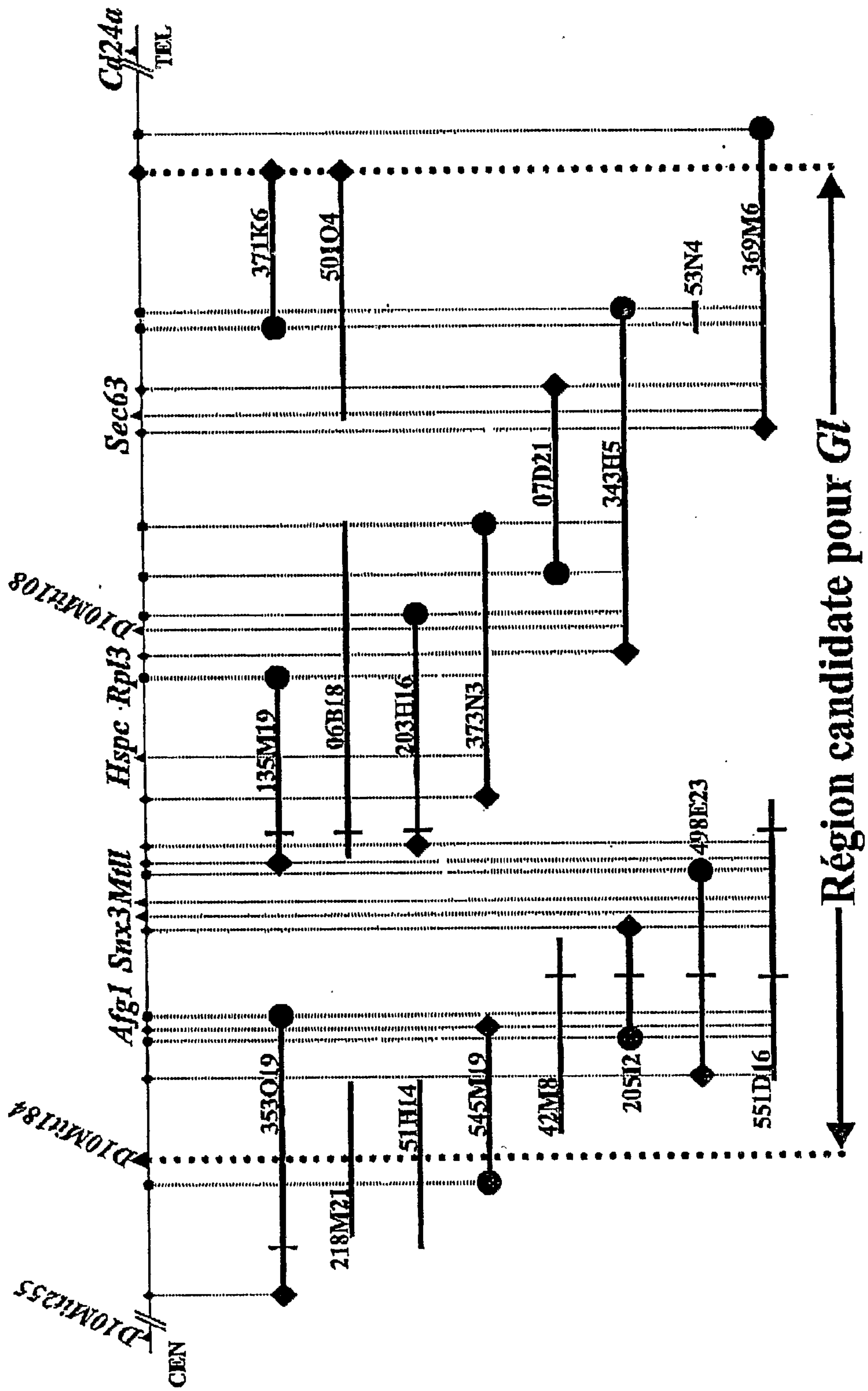
25. The use according to claim 24, wherein said amino acid sequence is SEQ ID NO: 4.

26. The use according to any one of claims 17 to 25, wherein said GI polypeptide comprises a transmembrane domain.

27. The use according to claim 26, wherein said GI polypeptide is a receptor.

28. The use according to any one of claims 17 to 27, wherein said pharmaceutical composition further comprises a pharmaceutically acceptable carrier.

FIGURE 1



## FIGURE 2

| BACs   | Kb  | D10M1255 | M19-Sp6 | D10M1254 | D10M1184 | C1-89AluI | E23-T7 | 12-Sp6 | M19-T7 | O19-Sp6 | D10-89-RsaI | 12-T7 | Sax3 | MuI | B5-89-HindIII | E23-Sp6 | M19-T7 | H16-T7 | N3-T7 | M19-Sp6 | H5-T7 | D10M1108 | H16-Sp6 | D21-Sp6 | N3-Sp6 | A5-89-HindIII | M6-T7 | Sec63 | D21-T7 | K6-Sp6 | H5-Sp6 | K6-T7=O4-T7 | M6-Sp6 | Cd24 |
|--------|-----|----------|---------|----------|----------|-----------|--------|--------|--------|---------|-------------|-------|------|-----|---------------|---------|--------|--------|-------|---------|-------|----------|---------|---------|--------|---------------|-------|-------|--------|--------|--------|-------------|--------|------|
|        |     | MOL/GL   | NP      | MOL/GL   | MOL/GL   | NP        | NP     | ?      | NP     | ?       | MOL/GL      | NP    | +/GL | ?   | NP            | NP      | NP     | NP     | NP    | +/GL    | +/GL  | NP       | +/GL    | ?       | ?      | +/GL          | +/GL  | NP    | ?      | NP     | NP     | +/GL        | ?      | NP   |
| 218M21 | 100 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 51H14  | 120 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 545M19 | 105 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 353O19 | 180 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 42M8   | 125 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 205I2  | 75  |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 498E23 | 130 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 551D16 | 180 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 135M19 | 120 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 06B18  | 200 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 203H16 | 150 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 173N3  | 180 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 07D21  | 120 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 343H5  | 220 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 53N4   | 20  |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 369M6  | 190 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 371K6  | 100 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |
| 501O4  | 160 |          |         |          |          |           |        |        |        |         |             |       |      |     |               |         |        |        |       |         |       |          |         |         |        |               |       |       |        |        |        |             |        |      |

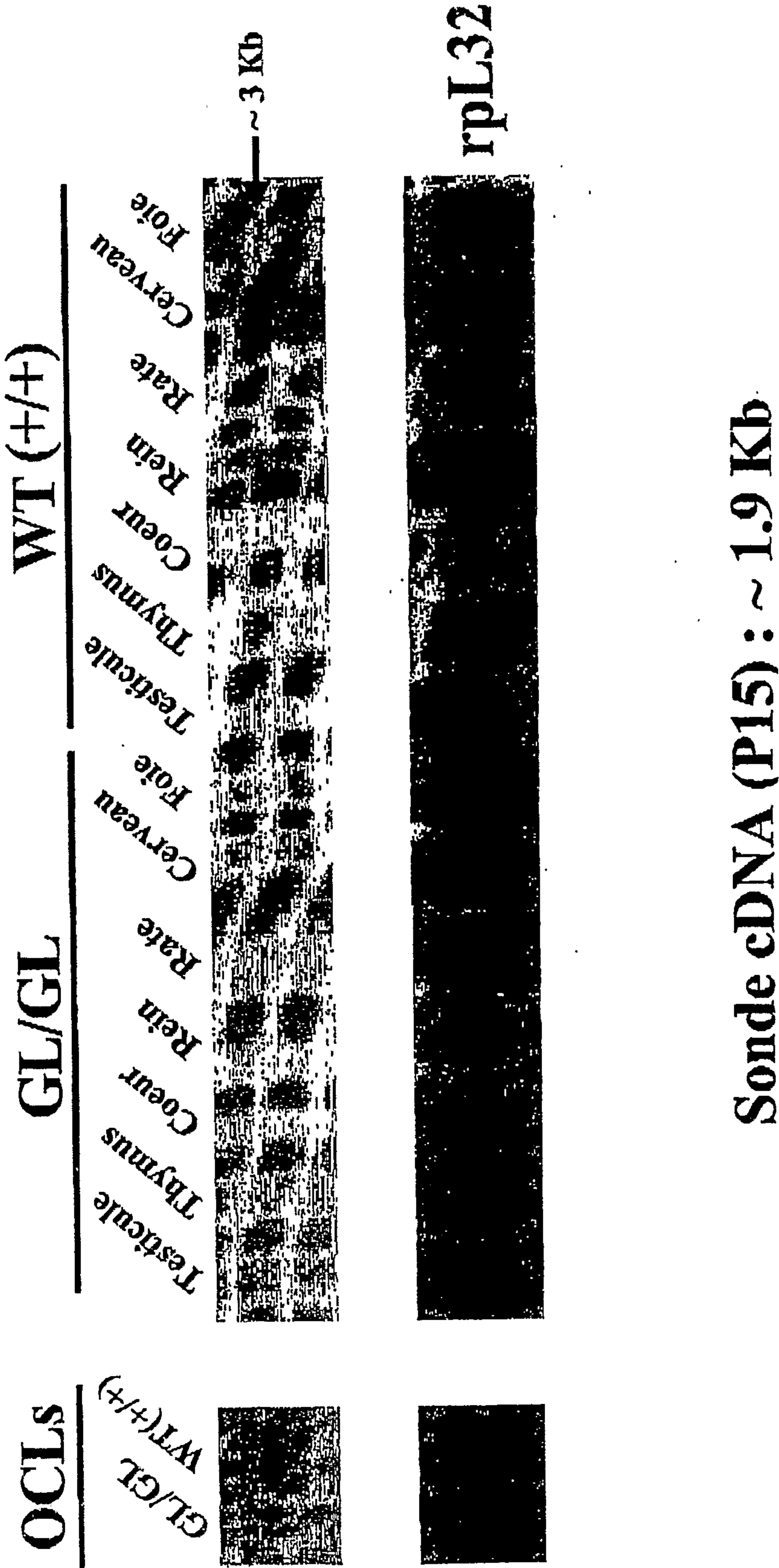
P : Polymorpheque

NP : Non-Polymorpheque

Région candidate pour Gl (~ 500 Kb)

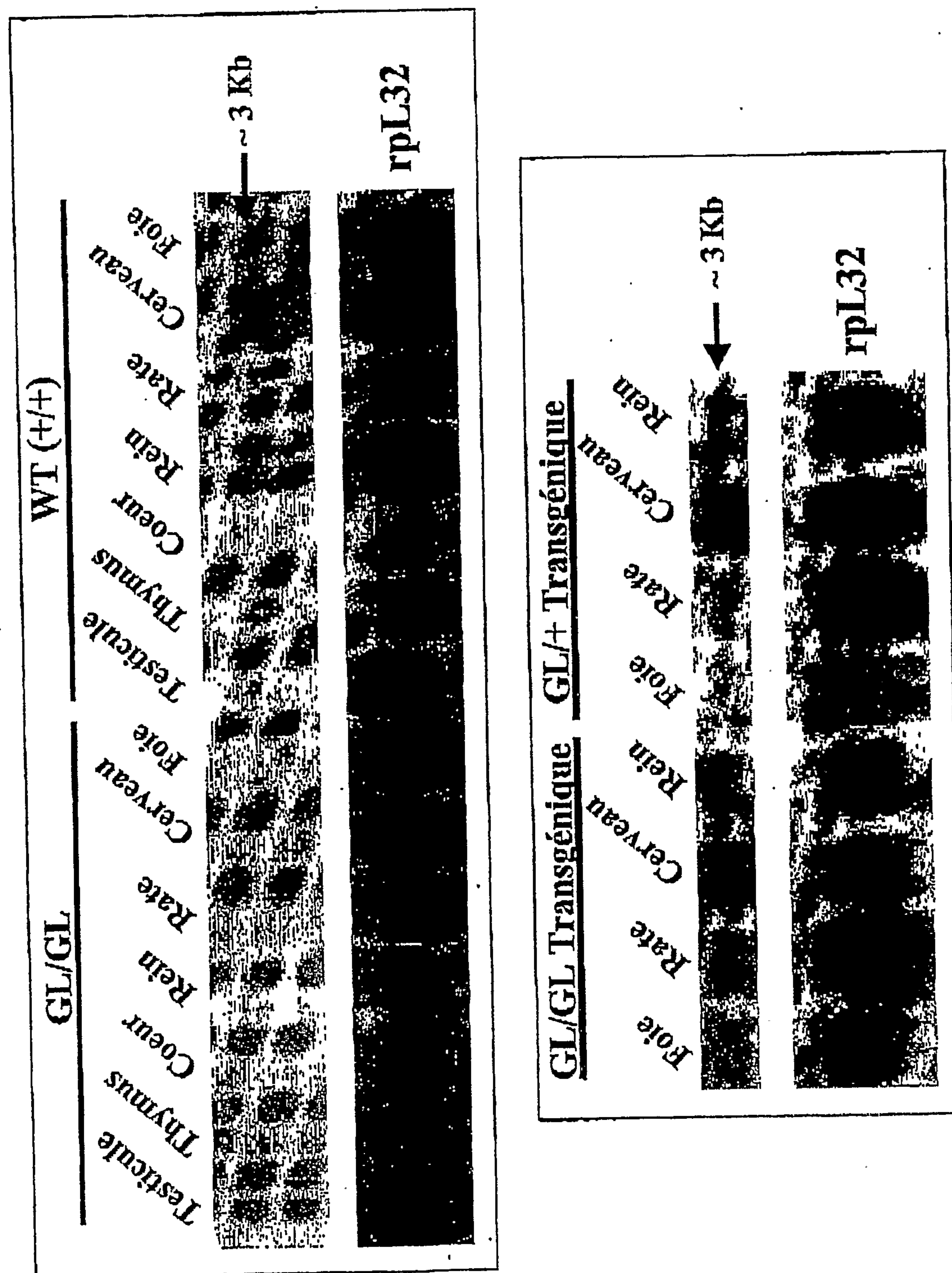
←

FIGURE 3



4/14

FIGURE 4



Western blot analysis of Gl protein in Wild-type and *gl* osteoclasts

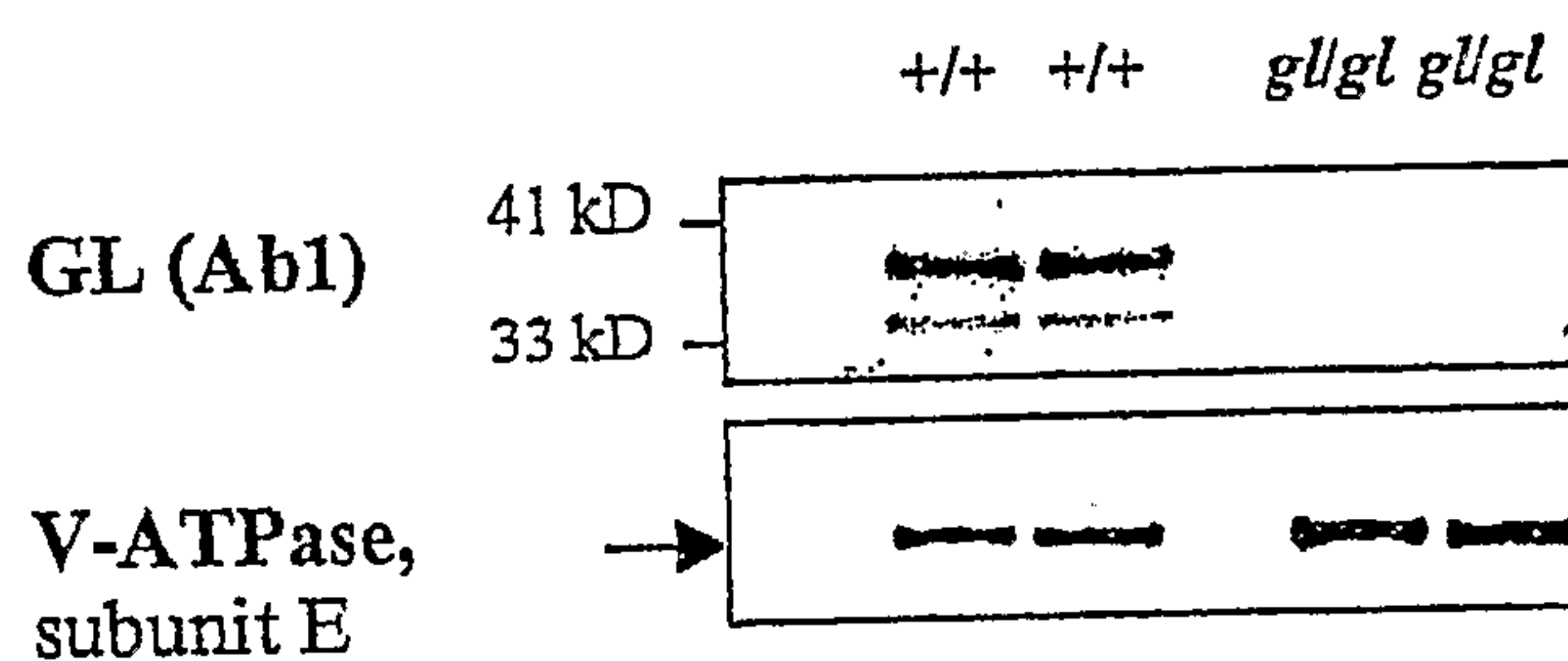
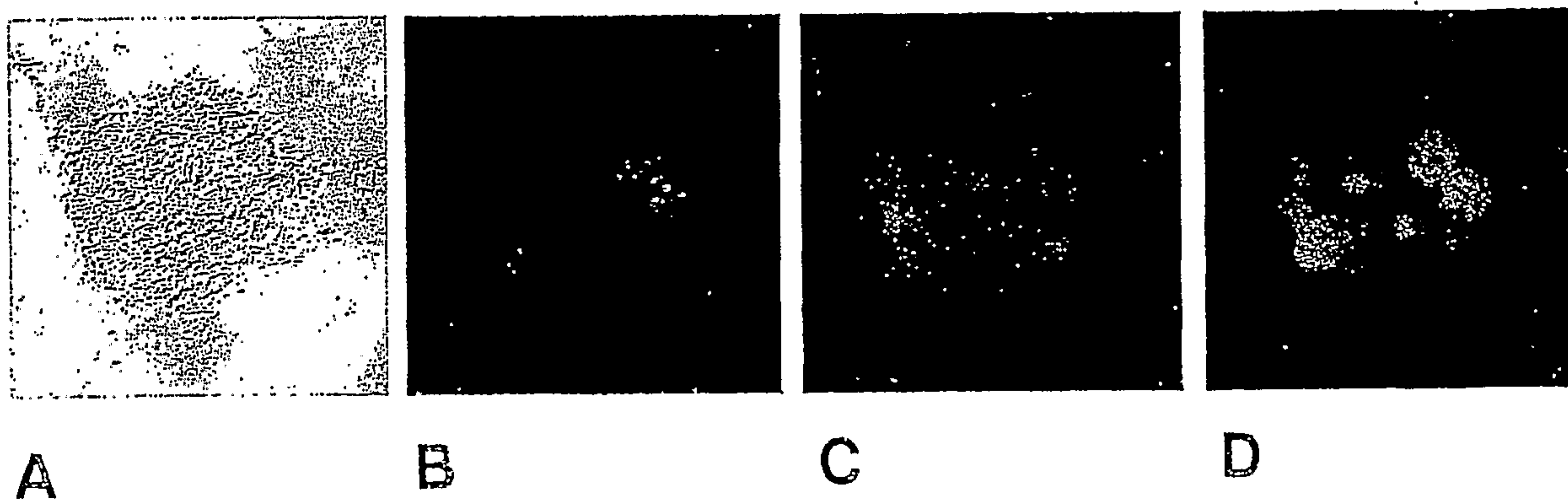


FIGURE 5A

Cytoplasmic localization of the Gl protein

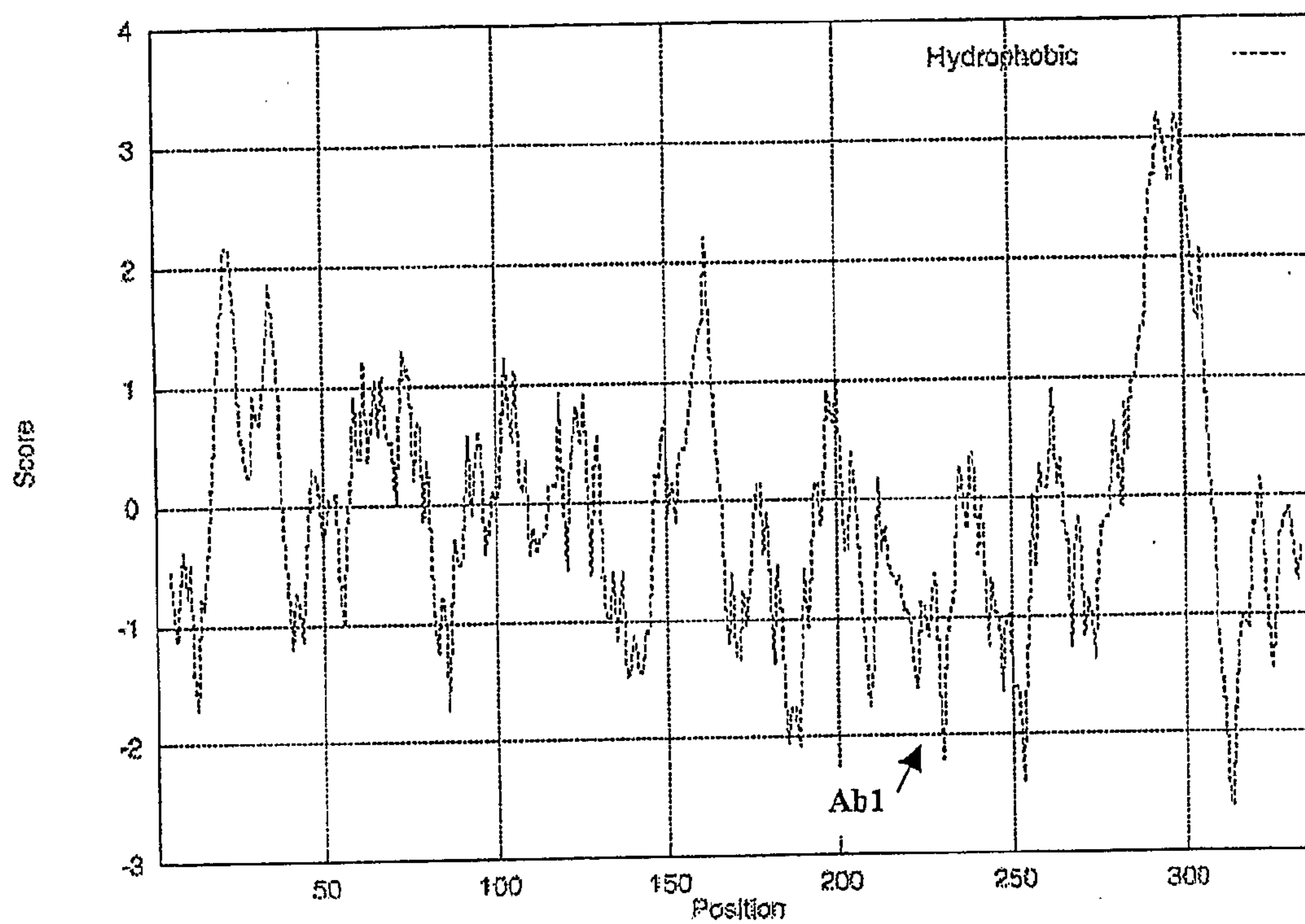
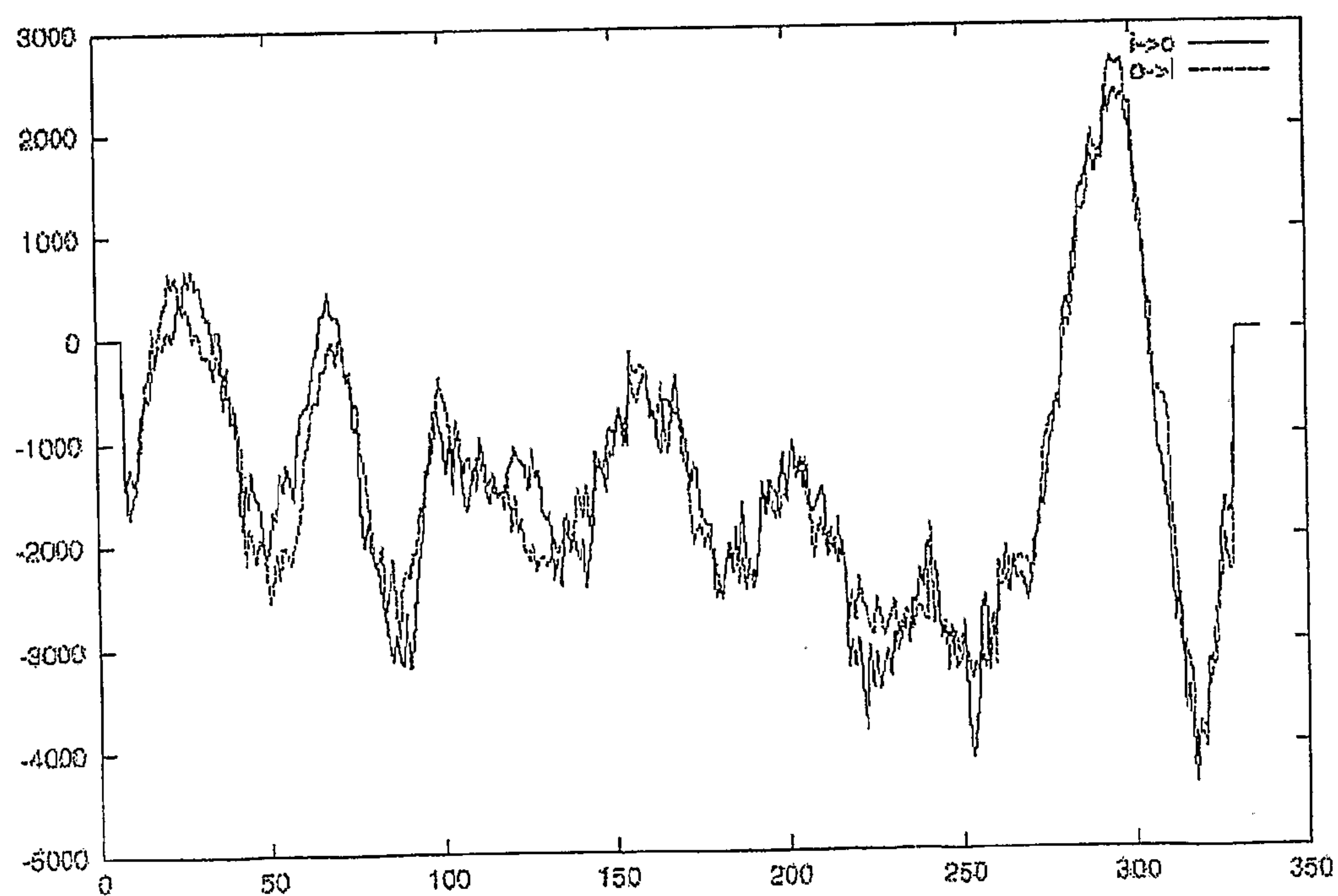


Osteoclasts were processed for immunofluorescence microscopy, by using rabbit anti-Gl antibodies.

(A) Normarski; (B) Hoechst; (C) Gl antibodies; (D) Merge pictures.

FIGURE 5B

6/14

**Kyte-Doolittle Hydropathy plot  
for Mouse G1 protein****FIGURE 6****TMpred-Prediction of transmembrane Regions  
and Orientation for Mouse G1 protein****FIGURE 7**

7/14

cDNA sequence of the mouse *gl* gene (SEQ ID 1)

gtcggaagcaccgggagctatctgttacagtccggccccggg ATG GCT CGG GAC GCG GAG CTG GCG 68  
 CGC AGT AGC GGG TGG CCG TGG CGG TGG CTG CCG GCG CTG CTG CTG CTG CAG CTG CTG 125  
 CGG TGG AGG TGC GCC CTG TGC GCG CTC CCC TTC ACC AGC AGT CGG CAC CCA GGC TTT 182  
 GCG GAC CTG CTG TCG GAG CAG CAG CTG TTG GAG GTG CAG GAC TTG ACC CTG TCT TTG 239  
 CTG CAG GGC GGA GGT CTA GGG CCG CTG TCA CTG CTA CCT CCG GAC CTG CCG GAT CTG 296  
 GAG CCT GAG TGC CGG GAG CTG CTG ATG GAC TTC GCC AAT AGC AGC GCC GAG CTG ACC 353  
 GCC TGT ATG GTG CGC AGC GCT CGG CCC GTG CGC CTC TGC CAG ACC TGC TAC CCG CTC 410  
 TTC CAA CAG GTC GCA ATC AAG ATG GAC AAC ATC AGC CGA AAC ATC GGG AAT ACC TCC 467  
 GAG GGC CCG CGC TGT GGC GGA AGT CTC CTG ACG GCA GAC AGA ATG CAG ATA GTT CTC 524  
 ATG GTC TCT GAG TTT TTC AAC AGC ACG TGG CAG GAG GCG AAC TGC GCA AAT TGC CTA 581  
 ACA AAC AAT GGT GAG GAT TTG TCA AAC AAC ACA GAG GAC TTC CTC AGT CTG TTT AAC 638  
 AAG ACT TTG GCC TGC TTT GAG CAT AAC CTG CAG GGG CAC ACA TAC AGT CTC CTC CCA 695  
 CCA AAA AAT TAC TCC GAA GTG TGC AGA AAC TGT AAA GAG GCA TAT AAA AAC CTG AGC 752  
 CTC CTG TAC AGT CAA ATG CAG AAA CTG AAC GGG CTT GAG AAC AAG GCT GAG CCT GAG 809  
 ACG CAC TTG TGC ATC GAT GTG GAG GAT GCA ATG AAC ATT ACT CGG AAG CTT TGG AGT 866  
 CGA ACC TTC AAC TGT TCG GTC ACC TGC AGC GAC ACG GTG TCC GTG GTT GCT GTG TCT 923  
 GTG TTC ATT CTC TTC CTG CCT GTC GTC TTC TAC CTC AGT AGC TTC CTT CAC TCA GAG 980  
 CAA AAG AAA CGC AAA CTC ATT CTA CCC AAA CGT CTC AAG TCG AGC ACC AGT TTT GCC 1037  
 AAC ATT CAA GAA AAT GCC ACC TGA agcctgcaaaacggggactcgacctcacgttggtgccagtaagt 1104

FIGURE 8

8/14

gttagaccacagcacagtcgagaagaagatgagccaaggtcggacaagttgcattctcacgaaatggtgggactgc 1180  
agacctataattttattctgaataaggggttctcaaattcccttttctgagcaccctttttttttttttgaagatt 1256  
tctgtatttttagttttcaaacatagcaatgttacatattttaaggtatatctgttacaataacaagtgagggcctt 1332  
ttttctcaggcatatgaatgactactggacacttctgatttatcctcgtttagcagaagtacacaaagcagaaaagg 1408  
ctgaggtctgctattttacacattagtcactgggagcccactctgaaaaagaaacatacttgccaaatggtagcagg 1484  
ctcagtgatttaacttaagtgaattcccattgtagtattggtgtatgtatatacatatgcacacacacacacata 1560  
tatatatacacgtatacatagatgtatatatgtaatgtatacttaatatatcatatacattaaaataatgttctctag 1636  
ttccctgaagtcctttttgaaaccactagtttgattataaacctccttaacagttttcagagagtgattccacatta 1712  
tgcatttatccttggttaaagggtttacagtaactgaggttctaataatgactttttataaatactattttacatcttat 1788  
ttttgtctttatttagtaagtaatttataatcactggactgcttaattacctttgaggacaagatggattcatctt 1864  
atgccagggtatttgcatcatgaatttcattaagttatttggcaacctgttaacttggttagtagttcaagtcgaatgt 1940  
caccaagtgtgtcatactgtgttttaaatttgtgattttttttaatgaaaattttatccttggaatatgttgagatt 2016  
tggggagaaaacaaggacaaacacaagagcttaaatttcagaaaaatagacagggacctgagggatgctcacgggtgag 2092  
acagctgcgtgggtttacactggagatgactcgggttgacaggctcgcacaggaagcctcccagttacgggaaagatg 2168  
aaagtcacatgactgaaacgaaattacccatctcactgtcaggaaactagttcttctttggcatatttctagcaac 2244  
ctttaaaaccatgcttggtttcagtgtcactcagttgtatttctcaagatgtagaagttgatgggtttgttggttaa 2320  
tccggtggaacgggctttgttgtaaaggtaatgaataggaaactcctcagattcaatggttaagaaaatgtgact 2396  
cccttcacaacctgtaattgccctacaggaaggcaggagtgtttgggtattttttgtatgtttcccatatatgcag 2472  
agtgtgagagcaggctagtcttagtcccagagtgtgtcacaccgggtatgtgacaatcagacgacgctgtgatcca 2548  
ctagatgtgccgggggttcattgtgtgtcattgttcctgtcttgatttgaagcacatgggttgaggggtcattggaag 2624  
ccatcttcacagtacatgtaaagcttattttacatgtgcaaagtgagtgaagtgacatatttaaaactgtgagtagc 2700  
gactcctcgggtacctttcagtactgtgtgtacaaaccactgcttttggctaagaagctggagagcactttaacaa 2776  
gccagccatctctgttctgatcagggtctggctctctagaggttgcattagaaatatatttgaaaatgtgccaaa 2852  
gaatttcaccttgtgggtcatattaaaaaaatgtacatagttctgaatcctgaggcacatagggttatgtgtgtgca 2928  
caagaaaacctgttttttctcttatgctttacaataaaggaaataacaaggaaaaaaaaaaaaaaaaaaaaa 2978

cDNA sequence of the mouse *gl* gene (SEQ ID 3)

[illegible]

FIGURE 9

|                                                                                   |      |
|-----------------------------------------------------------------------------------|------|
| gttagaccacagcacagtgcgagaagaagatgagccaaggctcggaacaagttgcatttctcacgaaatgttgggactgc  | 1180 |
| agacctataattttattctgaataagggttctcaaattcccttttcctgagcaccctttttttttttttgaagatt      | 1256 |
| tctgtatttttagttttcaaacatagcaatgttacatattttaagggtatatctgttacaataacaagtgagggctt     | 1332 |
| ttttctcaggcatatgaatgactactggacacttctgatttatcctcgttagcagaagtacacaaagcagaaaagg      | 1408 |
| ctgaggtctgctattttacacatttagtcactgggagcccactctgaaaaagaaacatacttgccaaatggtagcagg    | 1484 |
| ctcagtgattaacttaagtgaattcccattgtagtattggttgatatgtatatacatatgcacacacacacacata      | 1560 |
| tatatatacacgtatacatagatgtatataatgtataacttaatatatcatacattaaaaataatgttctctag        | 1636 |
| ttccctgaagtcctttttgaaaccactagttgattataaacctccttaacagttttcagagagtgattccacatta      | 1712 |
| tgcattttatccttggttaaagggttacagtaactgaggttctaatatgactttttataaatactattttacatcttat   | 1788 |
| ttttgtctttatttagtaagtaattttataatcactggactgcttaattacctttgaggacaagatggattcatctt     | 1864 |
| atgccagggatgtgcatcatgaatttcattaaagttattttggcaacctgtaacttggttagtagttcaagtcgaatgt   | 1940 |
| caccaagtgtgtcatactgtgttttaaattttgtgattttttttaatgaaaatttttatcttggaatatttgagatt     | 2016 |
| tggggagaaacaaggacaaacacaagagcttaaatttcagaaaatagacagggacctgagggatgctcacggtgag      | 2092 |
| acagctgcgtgggtttacactggagatgactcgggttgacaggctcgcacaggaagcctccagttacgggaaagatg     | 2168 |
| aaagtcacatgactgaaacgaaattacccatctcactgtcaggaaactagttcttctttggcatattttctagcaac     | 2244 |
| ctttaaaaccatgcttggttcagtgctcactcagttgtattttctcaagatgtagaagttgatgggttttgttggttaa   | 2320 |
| tccggttgaaacgggctttgttgtaaaggtaatgaataggaaactcctcagattcaatggttaagaaaatgtgact      | 2396 |
| cccttcacaacctgtaattgccctacaggaaggcaggagtggttggttattttttgtatgtttcccatatgacag       | 2472 |
| agtgtgagagcaggctagtccttagtcccagagtggtgtcacaccgggtatgtgacaatcagacgacgctgtgatcca    | 2548 |
| ctagatgtgccgggggttcattgtgctgtcattgttctgtcttgatttgaaacacatgggttgaggggtcattggaag    | 2624 |
| ccatcttcacatcagtacatgtaaagcttattttacatgtgcaaagtgagtgaagtgcacataatttaaactgtgagtagc | 2700 |
| gactcctcgggtacctttcagtagtctgtgtgtacaaaccactgcttttggttaagaagctggagagcactttaacaa    | 2776 |
| gccagccatctctgttccctgatcagggtctggctctctagagggtgcattagaaatatatttgaaaatgtgccaaa     | 2852 |
| gaattttcatcttggtggtcatattaaaaaaaaatgtacatagttctgaatcctgaggcacatagggttatgtgtgtgca  | 2928 |
| caagaaaacctgttttttcccttatgctttacaataaaggaaataacaagggaiaaaaaaaaaaaaaaaaaaaaaa      | 2978 |

11/14

cDNA sequence of the human homologue of the mouse *gl* gene (SEQ ID 2)

cgctcgcggaaccggaagcggcggtgtccgcggtgccggctgggggaggagaggcggcggtgggctccctgggg 76  
 tgtgtgagcccggtg ATG GAG CCG GGC CCG ACA GCC GCG CAG CGG AGG TGT TCG TTG CCG 136  
 CCG TGG CTG CCG CTG GGG CTG CTG CTG TGG TCG GGG CTG GCC CTG GGC GCG CTC CCC 193  
 TTC GGC AGC AGT CCG CAC AGG GTC TTC CAC GAC CTC CTG TCG GAG CAG CAG TTG CTG 250  
 GAG GTG GAG GAC TTG TCC CTG TCC CTC CTG CAG GGT GGA GGG CTG GGG CCT CTG TCG 307  
 CTG CCC CCG GAC CTG CCG GAT CTG GAT CCT GAG TGC CGG GAG CTC CTG CTG GAC TTC 364  
 GCC AAC AGC AGC GCA GAG CTG ACA GGG TGT CTG GTG CGC AGC GCC CGG CCC GTG CGC 421  
 CTC TGT CAG ACC TGC TAC CCC CTC TTC CAA CAG GTC GTC AGC AAG ATG GAC AAC ATC 478  
 AGC CGA GCC GCG GGG AAT ACT TCA GAG AGT CAG AGT TGT GCC AGA AGT CTC TTA ATG 535  
 GCA GAT AGA ATG CAA ATA GTT GTG ATT CTC TCA GAA TTT TTT AAT ACC ACA TGG CAG 592  
 GAG GCA AAT TGT GCA AAT TGT TTA ACA AAC AAC AGT GAA GAA TTA TCA AAC AGC ACA 649  
 GTA TAT TTC CTT AAT CTA TTT AAT CAC ACC CTG ACC TGC TTT GAA CAT AAC CTT CAG 706  
 GGG AAT GCA CAT AGT CTT TTA CAG ACA AAA AAT TAT TCA GAA GTA TGC AAA AAC TGC 763  
 CGT GAA GCA TAC AAA ACT CTG AGT AGT CTG TAC AGT GAA ATG CAA AAA ATG AAT GAA 820  
 CTT GAG AAT AAG GCT GAA CCT GGA ACA CAT TTA TGC ATT GAT GTG GAA GAT GCA ATG 877  
 AAC ATC ACT CGA AAA CTA TGG AGT CGA ACT TTC AAC TGT TCA GTC CCT TGC AGT GAC 934  
 ACA GTG CCT GTA ATT GCT GTT TCT GTG TTC ATT CTC TTT CTA CCT GTT GTC TTC TAC 991  
 CTT AGT AGC TTT CTT CAC TCA GAG CAA AAG AAA CGC AAA CTC ATT CTG CCC AAA CGT 1048  
 CTC AAG TCC AGT ACC AGT TTT GCA AAT ATT CAG GAA AAT TCA AAC TGA gacctacaaaa 1107  
 tggagaattgacatatcacgtgaatgaatggtggaagacacaacttggtttcagaaagaagataaactgtgatttg 1183

FIGURE 10

12/14

acaagtcaagctcttaagaaatacaaggacttcagatccatttttaataagaattttcgatttttctttcctttt 1259  
ccacttctttctaacagattttggatatttttaatttccaggcatagcaatgttatctattttaatgtgtatttgtc 1335  
acaataacagaacatgcaagaacaatcattattttatattttatagggcatttgattactattctagacttctgggtatc 1411  
ttcttactaacataaatatctcaagtagaaaaagtttttgaaaactaacattttaaaaattaatcagttacagtaaag 1487  
actttgaaaaagaaatgtacttgtaggaagtagcttaattacccccattgcagttatttattgttatatatatagt 1563  
taatattgttgtagacatcacataatatataattcagtcctctagtttccctagagtcatttttgaaaccactgattgc 1639  
aaacctccctgacaatttttaaaagtagtaagccacattacatttatctttgtaaaaagatttatggtaactgggtt 1715  
tcttacttgactttttataaatagtttttacatcttatttttgccctttatttcataagtaatttaaaaatcactgg 1791  
attgctttatttatattcagggcaatatggattatttttataccaaggatttgcacgtgaattacattaagttatt 1867  
tggcaattttataattttattactacttttaaatcaaatgtagcattatcacactgtattttaaatgtcatttttttaa 1943  
ggaatattttcttcttaagatatatagaggattttggagaagagagacaggaggggttaaaaccagcttaagggttca 2019  
gagagcagaaaagggacctgagaggatgctcactgttaagactgttggacagtggtgtgtattgaggggatgaatcgg 2095  
aacgatagtcctcatgcagaaaatagtgagattaagatcattccttattgtttctaaattatttcaatcagatgaaag 2171  
tgatacgattgaaatgaaatcacatagttcgtgctcagaaattctatttttggtatgtttgtattagccttttagaaa 2247  
aaacactccgttttcagaattgtttcacagttttattttcttaggttttttagagttcaggattttcattttattaatttct 2323  
tcttgctttttttggtggaaataggctttgttgtaaacatttaagaatatataaatctcctctatatagaaacaagaat 2399  
tttggttaaaaagagaatttgaatcccttcctataactataaaaatgctctatagggagacaaaagtgtttcttttttct 2475  
tttatgttttactgtttatgtggagtgaatatataaggctccttggtatgtataacataactcaaaaagctgttacactttc 2551  
tctgatctgctgtgatccactgaaaatgtgctggggtttgttctgctgtcactgtttatgctgctggaacttagca 2627  
ctgtcttgatttgaagcatatgattgagagccatttgaagcaatcttcattaatgcagataaaacaagtttacatg 2703  
tgcagagtttagaaaatgacatgttcaattctgtaagtggtagctttttgagcacctttcagttattatgtatttgta 2779  
aaaaccattgtttttggatataaagctataaagcacttttaaaaaggaaaaggcagcctttactattttttctgggtt 2855  
gagtcattgctcttttagacctagcatcagcaatagatttcaaagataagtattaagcgctaccctaaagtgtgtaa 2931  
gtttttcattttgtcatattgaaaaatgatttgcatagtactgaatgttgacacacagcttatatgtattttacaag 3007  
aatactctttaagtgttttttttgacacattaaaataaaggaaataaggaaattgttaaaaaaaaaaaaaaaaaaaaaa 3063

13/14

cDNA sequence of the human homologue of the mouse *gl* gene (SEQ ID 4)

|                                                                                |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |      |   |   |     |
|--------------------------------------------------------------------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|------|---|---|-----|
|                                                                                | M | E | P | G | P | T | A | A | Q | R | R | C | S | L | P | 15   |   |   |     |
| P                                                                              | W | L | P | L | G | L | L | L | W | S | G | L | A | L | G | A    | L | P | 34  |
| F                                                                              | G | S | S | P | H | R | V | F | H | D | L | L | S | E | Q | Q    | L | L | 53  |
| E                                                                              | V | E | D | L | S | L | S | L | L | Q | G | G | G | L | G | P    | L | S | 72  |
| L                                                                              | P | P | D | L | P | D | L | D | P | E | C | R | E | L | L | L    | D | F | 91  |
| A                                                                              | N | S | S | A | E | L | T | G | C | L | V | R | S | A | R | P    | V | R | 110 |
| L                                                                              | C | Q | T | C | Y | P | L | F | Q | Q | V | V | S | K | M | D    | N | I | 129 |
| S                                                                              | R | A | A | G | N | T | S | E | S | Q | S | C | A | R | S | L    | L | M | 148 |
| A                                                                              | D | R | M | Q | I | V | V | I | L | S | E | F | F | N | T | T    | W | Q | 167 |
| E                                                                              | A | N | C | A | N | C | L | T | N | N | S | E | E | L | S | N    | S | T | 186 |
| V                                                                              | Y | F | L | N | L | F | N | H | T | L | T | C | F | E | H | N    | L | Q | 205 |
| G                                                                              | N | A | H | S | L | L | Q | T | K | N | Y | S | E | V | C | K    | N | C | 224 |
| R                                                                              | E | A | Y | K | T | L | S | S | L | Y | S | E | M | Q | K | M    | N | E | 243 |
| L                                                                              | E | N | K | A | E | P | G | T | H | L | C | I | D | V | E | D    | A | M | 262 |
| N                                                                              | I | T | R | K | L | W | S | R | T | F | N | C | S | V | P | C    | S | D | 281 |
| T                                                                              | V | P | V | I | A | V | S | V | F | I | L | F | L | P | V | V    | F | Y | 300 |
| L                                                                              | S | S | F | L | H | S | E | Q | K | K | R | K | L | I | L | P    | K | R | 319 |
| L                                                                              | K | S | S | T | S | F | A | N | I | Q | E | N | S | N | * |      |   |   | 335 |
| tggagaattgacatatcacgtgaatgaatggtggaagacacaacttggttttcagaaagaagataaaactgtgatttg |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   | 1183 |   |   |     |

FIGURE 11

acaagtcaagctcttaagaaatacaaggacttcagatccatTTTTtaaataagaattttcgatttttctttcctttt 1259  
ccacttctttctaacagatttggatatttttaatttccaggcatagcaatgttatctatttttaattgtgtatttgtc 1335  
acaataacagaacatgcaagaacaatcattattttatattttatagggcatttgattactattctagacttctgggtatc 1411  
ttcttactaacataaatatctcaagtagaaaaagtttttgaaaactaacatttaaaaattaatcagttacagtaaag 1487  
actttgaaaaagaaatgtacttgttaggaagtagcttaattacccccattgcagtattattgttatatatatagt 1563  
taatatgttgtagacatcacataatatataattcagtcctctagtttccctagagtcatttttgaaaccactgattgc 1639  
aaacctccctgacaattttttaaagtagtaagccacattacatttatctttgtaaaaagatttatggtaactgggt 1715  
tcttacttgacttttataaatagtttttacatcttatttttgcctttatttcataagtaatttaaaaatcactgg 1791  
attgctttatttatattcagggcaatatggattatttttataccaaggatttgcctcgtgaattacattaagttatt 1867  
tggaattttataattttattactacttttaaatcaaatgtagcattatcacactgtatttaaatgtcatttttttaa 1943  
ggaatattttcttcttaagatatatagaggatttttgagaagagagacaggaggggtaaaaccagcttaagggttca 2019  
gagagcagaaagggacctgagaggatgctcactgtaagactgttgagacagtgggtgtgtattgaggggatgaatcgg 2095  
aacgatagtctcatgcagaaaatagtgagattaagatcatccttattgtttctaaattatttcaatcagatgaaag 2171  
tgatacgattgaaatgaaatcacatagttcgtgctcagaaattctatttttggtatgtttgtattagcctttagaaa 2247  
aaacactccgtttcagaattgttcacagttttattttcttaggttttttagagttcaggatttcatttattaatttct 2323  
tcttgcttttttggtggaatataggctttgttgtaaacattaagaatataaaatctcctctatatagaaacaagaat 2399  
tttggtaaaaagagaatttgaatcccttctataactataaaatgctctatagggagacaaagtgtttcttttttct 2475  
tttatgtttactgtttatgtggagtgaatataggctcttggtatgtataacatactcaaaagctgttacactttc 2551  
tctgatctgctgtgatccactgaaaatgtgctgggtttgttctgctgtcactgtttatgctgctggaacttagca 2627  
ctgtcttgatttgaagcatatgattgagagccatttgaagcaatcttcattaatgcagataaaacaagtttacatg 2703  
tgcagagttagaaaatgacatgttcaattctgttaagtgggtgactttttgagcacctttcagtattatgtatttcta 2779  
aaaaccattgtttttggatatataagctaataagcactttaaaaaggaaaaggcagcctttactattttttctgggt 2855  
gagtcattgctcttttagacctagcatcagcaatagatttcaaagataagtattaagcgctaccctaaagtgtgtaa 2931  
gtttttcattttgtcatattgaaaaatgatttgcatagtactgaatgttgacacacagcttatatgtatttacaag 3007  
aatatctttaagtgtttttttgacacattaaataaaaggaaataaggaaattgtaaaaa 3063

