

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

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

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>(51) International Patent Classification <sup>6</sup> :</b><br><b>C12N 15/12, 15/62, C07K 14/705, 16/28,</b><br><b>C12N 5/10, G01N 33/68</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | <b>A1</b> | <b>(11) International Publication Number:</b> <b>WO 99/05276</b><br><br><b>(43) International Publication Date:</b> 4 February 1999 (04.02.99)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>(21) International Application Number:</b> PCT/US98/15493<br><br><b>(22) International Filing Date:</b> 24 July 1998 (24.07.98)<br><br><b>(30) Priority Data:</b><br>60/053,715      25 July 1997 (25.07.97)      US<br><br><b>(71) Applicant:</b> ZYMOGENETICS, INC. [US/US]; 1201 Eastlake Avenue East, Seattle, WA 98102 (US).<br><br><b>(72) Inventors:</b> SHEPPARD, Paul, O.; 20717 N.E. 2nd Street, Redmond, WA 98053 (US). MOORE, Emma, E.; 3507 30th Avenue West, Seattle, WA 98199 (US).<br><br><b>(74) Agent:</b> LINGENFELTER, Susan, E.; ZymoGenetics, Inc., 1201 Eastlake Avenue East, Seattle, WA 98102 (US). |           | <b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).<br><br><b>Published</b><br><i>With international search report.</i> |
| <b>(54) Title:</b> HUMAN CHLORIDE ION CHANNEL ZSIG44<br><br><b>(57) Abstract</b><br><br>The present invention relates to polynucleotide and polypeptide molecules for zsig44, a novel member of the ion channel family. The polypeptides, and polynucleotides encoding them may be used for diagnosis of genetic disease associated with chromosome 10. The present invention also includes antibodies to the zsig44 polypeptides.                                                                                                                                                                                              |           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |                                          |    |                                              |    |                          |
|----|--------------------------|----|------------------------------------------|----|----------------------------------------------|----|--------------------------|
| AL | Albania                  | ES | Spain                                    | LS | Lesotho                                      | SI | Slovenia                 |
| AM | Armenia                  | FI | Finland                                  | LT | Lithuania                                    | SK | Slovakia                 |
| AT | Austria                  | FR | France                                   | LU | Luxembourg                                   | SN | Senegal                  |
| AU | Australia                | GA | Gabon                                    | LV | Latvia                                       | SZ | Swaziland                |
| AZ | Azerbaijan               | GB | United Kingdom                           | MC | Monaco                                       | TD | Chad                     |
| BA | Bosnia and Herzegovina   | GE | Georgia                                  | MD | Republic of Moldova                          | TG | Togo                     |
| BB | Barbados                 | GH | Ghana                                    | MG | Madagascar                                   | TJ | Tajikistan               |
| BE | Belgium                  | GN | Guinea                                   | MK | The former Yugoslav<br>Republic of Macedonia | TM | Turkmenistan             |
| BF | Burkina Faso             | GR | Greece                                   | ML | Mali                                         | TR | Turkey                   |
| BG | Bulgaria                 | HU | Hungary                                  | MN | Mongolia                                     | TT | Trinidad and Tobago      |
| BJ | Benin                    | IE | Ireland                                  | MR | Mauritania                                   | UA | Ukraine                  |
| BR | Brazil                   | IL | Israel                                   | MW | Malawi                                       | UG | Uganda                   |
| BY | Belarus                  | IS | Iceland                                  | MX | Mexico                                       | US | United States of America |
| CA | Canada                   | IT | Italy                                    | NE | Niger                                        | UZ | Uzbekistan               |
| CF | Central African Republic | JP | Japan                                    | NL | Netherlands                                  | VN | Viet Nam                 |
| CG | Congo                    | KE | Kenya                                    | NO | Norway                                       | YU | Yugoslavia               |
| CH | Switzerland              | KG | Kyrgyzstan                               | NZ | New Zealand                                  | ZW | Zimbabwe                 |
| CI | Côte d'Ivoire            | KP | Democratic People's<br>Republic of Korea | PL | Poland                                       |    |                          |
| CM | Cameroon                 | KR | Republic of Korea                        | PT | Portugal                                     |    |                          |
| CN | China                    | KZ | Kazakstan                                | RO | Romania                                      |    |                          |
| CU | Cuba                     | LC | Saint Lucia                              | RU | Russian Federation                           |    |                          |
| CZ | Czech Republic           | LI | Liechtenstein                            | SD | Sudan                                        |    |                          |
| DE | Germany                  | LK | Sri Lanka                                | SE | Sweden                                       |    |                          |
| DK | Denmark                  | LR | Liberia                                  | SG | Singapore                                    |    |                          |
| EE | Estonia                  |    |                                          |    |                                              |    |                          |

5

## HUMAN CHLORIDE ION CHANNEL ZSIG44

Background of the Invention

Chloride channels are membrane proteins that mediate passive transport of chloride and other anions across lipid bilayers. They are found in the plasma membrane and in various organelles, often associated with active cation transport systems. For review, see Pusch, M., Jentsch, T.J., Physiol. Rev., 74:813-827, 1994 and Jentsch, T.J., Gunther, W., BioEssays 19:117-126, 1997.

Several types of chloride channels have been isolated and are subject to diverse regulation. Ligand-gated chloride channels are activated by the external binding of a ligand; for instance,  $\gamma$ -aminobutyric acid (GABA) and glycine-activated channels in neurons and skeletal muscle. Adenosine 3',5'-cyclic monophosphate (cAMP)-dependent chloride channels are activated by a rise in cAMP; such as the cystic fibrosis transmembrane conductance regulator (CFTR) chloride channel. This activation is believed to be mediated by cAMP-dependent phosphorylation by a protein kinase A (PKA). Others include chloride channels activated by elevations in intracellular calcium, and volume-dependent chloride channels which may be regulated by tyrosine phosphorylation.

Plasma-membrane chloride channels, shown to increase voltage-dependent chloride conductance in an appropriate expression system (e.g. *Xenopus laevis* oocytes), are grouped into distinct gene families. The ClC family of voltage-gated chloride channels functions

directly as chloride channels; its members are structurally similar large proteins containing several transmembrane domains but showing great functional diversity. Other ion channels, e.g., K<sup>+</sup> channels, have similar structure (Hille, 5 B., Ionic Channels of Excitable Membranes, Sinauer Assoc., Sunderland, MA, 1992).

Another family of chloride channel proteins has been discovered that increases voltage-dependent chloride conductance but may not function directly as ion channels; 10 instead, these proteins may act as ion channel regulators. These proteins are structurally different from the ClC family, and are small proteins with a single transmembrane domain. This protein family includes rat and human phospholemman (PLM) and human MAT-8 chloride channels 15 (Chen, L.K., Genomics, 41:435-443, 1997; Morrison, B.W. et al., J. Biol. Chem., 270:2176-2182, 1995). Other proteins related to this family are the potassium (K<sup>+</sup>) channel proteins Isk (minK), and rat channel inducing factor (CHIF) (Barhanan, J. et al., Nature, 78-80, 1996; ; Sanguinetti, 20 M.C. et al., Nature, 80-83, 1996; Attali, B et al., Proc. Natl. Acad. Sci., 92:6092-6096, 1995). Recent evidence demonstrates Isk's role as a potassium channel regulator that cannot itself form an ion channel, but enhances voltage-dependent conductance in the presence of a 25 functional K<sup>+</sup> channel (Attali, B., Nature, 384:24-25, 1996; Barhanan, J. et al., ibid.; ; Sanguinetti, M.C. et al., ibid.). Whether these small chloride channel proteins are channel regulators or actual channels is unknown.

Several ion channels are linked to genetic 30 disease in humans (see, Ackerman, M.J., Clapham, D.E., New Eng. J. Med. 336:1575-1586, 1997). The ClC-1 channel is the major chloride channel from skeletal muscle and mutations are associated with hyperexcitable skeletal muscles leading to myotonia (muscle stiffness). Moreover,

the kidney-specific ClC-K1 channel has high expression in Henle's loop; with expression upregulated throughout the kidney in dehydrated rats, suggesting a role in urine concentration. Another rat kidney channel, ClC-K2, has  
5 different cellular localization within the kidney, and presumably a different physiologic role. Human homologues to the rat ClC-K forms are known. The Isk K<sup>+</sup> channel regulator, involved in re-polarization of cardiac cell membranes, may have a role in heart arrhythmias (Attali,  
10 B., ibid.; Barhanan, J. et al., ibid.; Sanguinetti, M.C. et al., ibid.). The CFTR chloride channel is implicated in cystic fibrosis. These discoveries emphasize the diversity of ion channels, their tissue specificity, and their link to human pathological states.

15           Discovery of novel ion channel molecules provides understanding into potential human disease states, offering targets for discovery of novel therapies to correct and act as markers for genetic disease. Thus discovery of novel ion channels is sought. The present invention provides  
20 such polypeptides for these and other uses that should be apparent to those skilled in the art from the teachings herein.

#### Summary of the Invention

25           The present invention addresses this need by providing a novel human ion channel from the new family of ion channel proteins.

          Within one aspect, the present invention provides an isolated polynucleotide encoding a zsig44 polypeptide  
30 comprising a sequence of amino acid residues that is at least 90% identical to an amino acid sequence selected from the group consisting of: (a) the amino acid sequence as shown in SEQ ID NO:2 from residue number 17 (Leu), to residue number 89 (Cys); and (b) the amino acid sequence

as shown in SEQ ID NO:2 from amino acid number 1 (Met) to amino acid number 89 (Cys). Within one embodiment, the isolated polynucleotide disclosed above is selected from the group consisting of: (a) a polynucleotide sequence as shown in SEQ ID NO:1 from nucleotide 52 to nucleotide 270; and (b) a polynucleotide sequence as shown in SEQ ID NO:1 from nucleotide 4 to nucleotide 270. Within another embodiment, the isolated polynucleotide disclosed above comprises nucleotide 1 to nucleotide 267 of SEQ ID NO:8. Within another embodiment, the isolated polynucleotide disclosed above consists essentially of a sequence of amino acid residues that is at least 90% identical to an amino acid sequence as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys). Within another embodiment, the isolated polynucleotide disclosed above consists essentially of a sequence of amino acid residues as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys).

Within a second aspect the present invention provides an expression vector comprising the following operably linked elements: a transcription promoter; a DNA segment encoding a zsig44 polypeptide that is at least 90% identical to an amino acid sequence as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys); and a transcription terminator. Within one embodiment, the expression vector disclosed above further comprises a secretory signal sequence operably linked to the DNA segment.

Within a third aspect, the present invention provides a cultured cell into which has been introduced an expression vector as disclosed above, wherein the cell expresses a polypeptide encoded by the DNA segment.

Within another aspect, the present invention provides a DNA construct encoding a fusion protein, the DNA construct comprising: a first DNA segment encoding a polypeptide that is at least 90% identical to a sequence of

amino acid residues selected from the group consisting of:  
(a) the amino acid sequence of SEQ ID NO:2 from residue  
number 1 (Met), to residue number 16 (Ala); (b) the amino  
acid sequence of SEQ ID NO:2 from residue number 17 (Leu),  
5 to residue number 28 (Asp); (c) the amino acid sequence  
of SEQ ID NO:2 from residue number 29 (Pro), to residue  
number 64 (Lys); (d) the amino acid sequence of SEQ ID  
NO:2 from residue number 65 (Ser), to residue number 89  
(Cys); (e) the amino acid sequence of SEQ ID NO:2 from  
10 residue number 17 (Leu), to residue number 64 (Lys); (f)

the amino acid sequence of SEQ ID NO:2 from residue  
number 29 (Pro), to residue number 89 (Cys); and (g) the  
amino acid sequence of SEQ ID NO:2 from residue number 17  
(Leu), to residue number 89 (Cys); and at least one other  
15 DNA segment encoding an additional polypeptide, wherein the  
first and other DNA segments are connected in-frame; and  
encode the fusion protein. Within an embodiment, the  
fusion protein is produced by a method comprising:  
culturing a host cell into which has been introduced a  
20 vector comprising the following operably linked elements:  
(a) a transcriptional promoter; (b) a DNA construct  
encoding a fusion protein as disclosed above; and (c) a  
transcriptional terminator; and recovering the protein  
encoded by the DNA segment.

25 Within another aspect, the present invention  
provides an isolated polypeptide comprising a sequence of  
amino acid residues that is at least 90% identical to an  
amino acid sequence selected from the group consisting of:  
(a) the amino acid sequence as shown in SEQ ID NO:2 from  
30 residue number 17 (Leu), to residue number 89 (Cys); and  
(b) the amino acid sequence as shown in SEQ ID NO:2 from  
amino acid number 1 (Met) to amino acid number 89 (Cys).  
Within one embodiment, the isolated polypeptide disclosed  
above further contains motifs 1 through 4 and the PLITPGSA  
35 motif spaced apart from N-terminus to C-terminus in a  
configuration M1-{6}-M2-{5}-M3-{1}-M4-{14}-PLITPGSA. Within

another embodiment, the isolated polypeptide disclosed above consists essentially of a sequence of amino acid residues that is at least 90% identical to an amino acid sequence as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys). Within another embodiment, the isolated polypeptide disclosed above is as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys).

Within another aspect, the present invention provides a method of producing a zsig44 polypeptide comprising: culturing a cell according to claim 8; and isolating the zsig44 polypeptide produced by the cell.

Within another aspect, the present invention provides a method of producing an antibody to zsig44 polypeptide comprising: inoculating an animal with a polypeptide selected from the group consisting of: (a) a polypeptide consisting of 9 to 89 amino acids, wherein the polypeptide is at least 90% identical to a contiguous sequence of amino acids in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys); (b) a polypeptide according to claim 11; (c) a polypeptide having an amino acid sequence that is at least 90% identical to residue number 17 (Leu), to residue number 28 (Asp) of SEQ ID NO:2; (d) a polypeptide having an amino acid sequence that is at least 90% identical to residue number 29 (Pro), to residue number 64 (Lys) of SEQ ID NO:2; (e) a polypeptide having an amino acid sequence that is at least 90% identical to residue number 65 (Ser), to residue number 89 (Cys) of SEQ ID NO:2; and wherein the polypeptide elicits an immune response in the animal; and isolating the antibody from the animal. Within one embodiment, the antibody disclosed above binds to a zsig44 polypeptide. Within another embodiment, the antibody disclosed above is a monoclonal antibody.

Within another aspect the present invention provides an antibody which specifically binds to a polypeptide disclosed above.

Within another aspect the present invention provides a method of detecting, in a test sample, the presence of a modulator of zsig44 protein activity, comprising: culturing a cell into which has been introduced an expression vector according to claim 6, wherein the cell expresses the zsig44 protein encoded by the DNA segment in the presence and absence of a test sample; and comparing levels of activity of zsig44 in the presence and absence of a test sample, by a biological or biochemical assay; and determining from the comparison, the presence of modulator of zsig44 activity in the test sample.

These and other aspects of the invention will become evident upon reference to the following detailed description of the invention and attached drawings.

#### Brief Description of the Drawings

The Figure illustrates a multiple alignment of zsig44 (SEQ ID NO:2), rat CHIF (RATCHI) (SEQ ID NO:3), human MAT-8 (HSMAT8) (SEQ ID NO:4), human PLM (HSU722) (SEQ ID NO:5).

#### Detailed Description of the Invention

Prior to setting forth the invention in detail, it is helpful to the understanding thereof to define the following terms:

The term "affinity tag" is used herein to denote a polypeptide segment that can be attached to a second polypeptide to provide for purification or detection of the second polypeptide or provide sites for attachment of the second polypeptide to a substrate. In principal, any

peptide or protein for which an antibody or other specific binding agent is available can be used as an affinity tag. Affinity tags include a poly-histidine tract, protein A (Nilsson et al., EMBO J. 4:1075, 1985; Nilsson et al.,  
5 Methods Enzymol. 198:3, 1991), glutathione S transferase (Smith and Johnson, Gene 67:31, 1988), Glu-Glu affinity tag (Grussenmeyer, et al., Protein Expression and Purification 2: 95-107, 1991. DNAs encoding affinity tags are available from commercial suppliers (e.g., Pharmacia Biotech,  
10 Piscataway, NJ).

The term "allelic variant" is used herein to denote any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in  
15 phenotypic polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequence. The term allelic variant is also used herein to denote a protein encoded by an allelic variant of a gene.

20 The terms "amino-terminal" and "carboxyl-terminal" are used herein to denote positions within polypeptides. Where the context allows, these terms are used with reference to a particular sequence or portion of a polypeptide to denote proximity or relative position.  
25 For example, a certain sequence positioned carboxyl-terminal to a reference sequence within a polypeptide is located proximal to the carboxyl terminus of the reference sequence, but is not necessarily at the carboxyl terminus of the complete polypeptide.

30 The term "complement/anti-complement pair" denotes non-identical moieties that form a non-covalently associated, stable pair under appropriate conditions. For instance, biotin and avidin (or streptavidin) are prototypical members of a complement/anti-complement pair.

Other exemplary complement/anti-complement pairs include receptor/ligand pairs, antibody/antigen (or hapten or epitope) pairs, sense/antisense polynucleotide pairs, and the like. Where subsequent dissociation of the complement/anti-complement pair is desirable, the complement/anti-complement pair preferably has a binding affinity of  $<10^9 \text{ M}^{-1}$ .

The term "complements of a polynucleotide molecule" is a polynucleotide molecule having a complementary base sequence and reverse orientation as compared to a reference sequence. For example, the sequence 5' ATGCACGGG 3' is complementary to 5' CCCGTGCAT 3'.

The term "contig" denotes a polynucleotide that has a contiguous stretch of identical or complementary sequence to another polynucleotide. Contiguous sequences are said to "overlap" a given stretch of polynucleotide sequence either in their entirety or along a partial stretch of the polynucleotide. For example, representative contigs to the polynucleotide sequence 5'-ATGGCTTAGCTT-3' are 5'-TAGCTTgagtct-3' and 3'-gtcgacTACCGA-5'.

The term "degenerate nucleotide sequence" denotes a sequence of nucleotides that includes one or more degenerate codons (as compared to a reference polynucleotide molecule that encodes a polypeptide). Degenerate codons contain different triplets of nucleotides, but encode the same amino acid residue (i.e., GAU and GAC triplets each encode Asp).

A "DNA segment" is a portion of a larger DNA molecule having specified attributes. For example, a DNA segment encoding a specified polypeptide is a portion of a longer DNA molecule, such as a plasmid or plasmid fragment, that, when read from the 5' to the 3' direction, encodes the sequence of amino acids of the specified polypeptide.

The term "expression vector" is used to denote a DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of interest operably linked to additional segments that provide for its transcription. Such additional segments include promoter and terminator sequences, and may also include one or more origins of replication, one or more selectable markers, an enhancer, a polyadenylation signal, and the like. Expression vectors are generally derived from plasmid or viral DNA, or may contain elements of both.

The term "ion channel," is used generally to denote a polypeptide or protein that has homology to or is in fact an anion, cation, or other channel protein or their putative regulators. The term ion channel is also used herein to denote nucleotides that encode such polypeptides.

The term "isolated", when applied to a polynucleotide, denotes that the polynucleotide has been removed from its natural genetic milieu and is thus free of other extraneous or unwanted coding sequences, and is in a form suitable for use within genetically engineered protein production systems. Such isolated molecules are those that are separated from their natural environment and include cDNA and genomic clones. Isolated DNA molecules of the present invention are free of other genes with which they are ordinarily associated, but may include naturally occurring 5' and 3' untranslated regions such as promoters and terminators. The identification of associated regions will be evident to one of ordinary skill in the art (see for example, Dynan and Tijan, Nature 316:774-78, 1985). An "isolated" polypeptide or protein is a polypeptide or protein that is found in a condition other than its native environment, such as apart from blood and animal tissue. In a preferred form, the isolated polypeptide is substantially free of other polypeptides, particularly

other polypeptides of animal origin. It is preferred to provide the polypeptides in a highly purified form, i.e., greater than 95% pure, more preferably greater than 99% pure. When used in this context, the term "isolated" does  
5 not exclude the presence of the same polypeptide in alternative physical forms, such as dimers or alternatively glycosylated or derivatized forms.

The term "operably linked", when referring to DNA segments, indicates that the segments are arranged so that  
10 they function in concert for their intended purposes, e.g., transcription initiates in the promoter and proceeds through the coding segment to the terminator.

The term "ortholog" denotes a polypeptide or protein obtained from one species that is the functional  
15 counterpart of a polypeptide or protein from a different species. Sequence differences among orthologs are the result of speciation.

"Paralogs" are distinct but structurally related proteins made by an organism. Paralogs are believed to  
20 arise through gene duplication. For example,  $\alpha$ -globin,  $\beta$ -globin, and myoglobin are paralogs of each other.

A "polynucleotide" is a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. Polynucleotides  
25 include RNA and DNA, and may be isolated from natural sources, synthesized *in vitro*, or prepared from a combination of natural and synthetic molecules.

Sizes of polynucleotides are expressed as base pairs (abbreviated "bp"), nucleotides ("nt"), or kilobases  
30 ("kb"). Where the context allows, the latter two terms may describe polynucleotides that are single-stranded or double-stranded. When the term is applied to double-stranded molecules it is used to denote overall length and will be understood to be equivalent to the term "base

pairs". It will be recognized by those skilled in the art that the two strands of a double-stranded polynucleotide may differ slightly in length and that the ends thereof may be staggered as a result of enzymatic cleavage; thus all  
5 nucleotides within a double-stranded polynucleotide molecule may not be paired. Such unpaired ends will in general not exceed 20 nt in length.

A "polypeptide" is a polymer of amino acid residues joined by peptide bonds, whether produced  
10 naturally or synthetically. Polypeptides of less than about 10 amino acid residues are commonly referred to as "peptides".

The term "promoter" is used herein for its art-recognized meaning to denote a portion of a gene containing  
15 DNA sequences that provide for the binding of RNA polymerase and initiation of transcription. Promoter sequences are commonly, but not always, found in the 5' non-coding regions of genes.

A "protein" is a macromolecule comprising one or  
20 more polypeptide chains. A protein may also comprise non-peptidic components, such as carbohydrate groups. Carbohydrates and other non-peptidic substituents may be added to a protein by the cell in which the protein is produced, and will vary with the type of cell. Proteins  
25 are defined herein in terms of their amino acid backbone structures; substituents such as carbohydrate groups are generally not specified, but may be present nonetheless.

The term "receptor" denotes a cell-associated protein that binds to a bioactive molecule (i.e., a ligand)  
30 and mediates the effect of the ligand on the cell. Membrane-bound receptors are generally characterized by a multi-domain structure comprising an extracellular ligand-binding domain and an intracellular effector domain that is typically involved in signal transduction. Binding of

ligand to receptor results in a conformational change in the receptor that causes an interaction between the effector domain and other molecule(s) in the cell. This interaction in turn leads to an alteration in the metabolism of the cell. Metabolic events that are linked to receptor-ligand interactions include gene transcription, phosphorylation, dephosphorylation, increases in cyclic AMP production, mobilization of cellular calcium, mobilization of membrane lipids, cell adhesion, hydrolysis of inositol lipids and hydrolysis of phospholipids. In general, receptors can be membrane bound, cytosolic or nuclear; monomeric (e.g., thyroid stimulating hormone receptor, beta-adrenergic receptor) or multimeric (e.g., PDGF receptor, growth hormone receptor, IL-3 receptor, GM-CSF receptor, G-CSF receptor, erythropoietin receptor and IL-6 receptor).

The term "secretory signal sequence" denotes a DNA sequence that encodes a polypeptide (a "secretory peptide") that, as a component of a larger polypeptide, directs the larger polypeptide through a secretory pathway of a cell in which it is synthesized. The larger polypeptide is commonly cleaved to remove the secretory peptide during transit through the secretory pathway.

The term "splice variant" is used herein to denote alternative forms of RNA transcribed from a gene. Splice variation arises naturally through use of alternative splicing sites within a transcribed RNA molecule, or less commonly between separately transcribed RNA molecules, and may result in several mRNAs transcribed from the same gene. Splice variants may encode polypeptides having altered amino acid sequence. The term splice variant is also used herein to denote a polypeptide or protein encoded by a splice variant of an mRNA transcribed from a gene.

The term "soluble protein" is used herein to denote a protein polypeptide that is not bound to a cell membrane. A soluble protein includes a protein that is naturally intracellular, e.g. cytoplasmic, or is naturally  
5 secreted from the cell. Many cell-surface proteins have naturally occurring, soluble counterparts that are produced by proteolysis or translated from alternatively spliced mRNAs. Receptor and ligand polypeptides are said to be substantially free of transmembrane and intracellular  
10 polypeptide segments when they lack sufficient portions of these segments to provide membrane anchoring or signal transduction, respectively. A soluble protein also includes membrane-bound proteins that have been genetically engineered to be soluble; for instance, by removing a  
15 transmembrane region, or expressing only the soluble portion of the protein in a conventional manner.

Molecular weights and lengths of polymers determined by imprecise analytical methods (e.g., gel electrophoresis) will be understood to be approximate  
20 values. When such a value is expressed as "about" X or "approximately" X, the stated value of X will be understood to be accurate to  $\pm 10\%$ .

25 Teachings of all references cited herein are, in their entirety, incorporated by reference.

The present invention is based in part upon the discovery of a novel DNA sequence that encodes an ion channel or channel regulator. Northern blot analysis of  
30 the tissue distribution of the mRNA corresponding to this novel DNA showed that expression was highest in kidney and bone marrow, with no detectable expression levels in any other tissues or transformed cell lines tested. A putative splice variant was expressed in spinal cord. Dot blot

analysis of several tissues showed a signal in heart. The polypeptide has been designated zsig44.

The novel zsig44 polypeptides of the present invention were initially identified by querying an EST  
5 database for homologous sequences to signal sequences. A single N-terminal EST sequence was discovered and predicted to be related to the ion channel CHIF/MAT-8 family (Attali, B et al., ibid.; Morrison, B.W. et al., ibid.).

The nucleotide sequence of full-length zsig44 is  
10 described in SEQ ID NO:1, and its deduced amino acid sequence is described in SEQ ID NO:2. The multiple alignment revealed that zsig44 is a member of a family of ion channels, characterized by their small size and single transmembrane domain, potentially functioning as ion  
15 channel regulators.

Analysis of the DNA encoding zsig44 polypeptide (SEQ ID NO:1) revealed an open reading frame encoding 89 amino acids (SEQ ID NO:2) comprising a signal peptide of 19 amino acid residues (residue 1 (Met) to residue 16 (Ala) of  
20 SEQ ID NO:2), a transmembrane domain of 22 amino acids (residue 36 (Asn) to residue 58 (Leu) of SEQ ID NO:2), and a mature polypeptide of 73 amino acids (residue 17 (Leu) to residue 89 (Cys) of SEQ ID NO:2). Multiple alignment of zsig44 with other known ion channels revealed a region of  
25 similarity corresponding to amino acid residue 29 (Pro) to amino acid residue 64 (Lys) of SEQ ID NO:2 (see Figure). This region is hereinafter referred to as the "central region" or "central domain." Within this region there are 4 conserved motifs referred to hereinafter as "motif 1"  
30 (SEQ ID NO:6; amino acids 29 to 34 of SEQ ID NO:2), "motif 2" (SEQ ID NO:7; amino acids 41 to 46 of SEQ ID NO:2), "motif 3" (SEQ ID NO:8; amino acids 52 to 57 of SEQ ID NO:2), "motif 4" (SEQ ID NO:9; amino acids 59 to 64 of SEQ ID NO:2). Motifs 3 and 4 are separated by a single

conserved serine residue present in all ion channel family members shown; this serine may serve as a PKC phosphorylation site. The C-terminal "PLITPGSA motif" is also present (SEQ ID NO:10; amino acids 79 to 86 of SEQ ID NO:2). Moreover, zsig44 polypeptide does not have a PKA phosphorylation consensus site, which is characteristic of the CHIF/MAT-8 type ion channels (see, Figure).

Motifs 1 through 4 and the "PLITPGSA motif" are spaced apart from N-terminus to C-terminus in a configuration represented by the following:

M1-{6}-M2-{5}-M3-{1}-M4-{14}-PLITPGSA,

where M# denotes the specific motif disclosed above (e.g., M1 is motif 1, etc.)

PLITPGSA denotes the PLITPGSA motif disclosed above and

{#} denotes the number of amino acids between the motifs.

The presence of conserved motifs generally correlates with or defines important structural regions in proteins. The regions between, or flanking, such motifs may be more variable, but are often functionally significant because they relate to or define important structures and activities such as binding domains, biological and enzymatic activity, signal transduction, tissue localization domains and the like. The regions flanking the central domain that may be functionally significant are from residue 17 (Leu) to residue 28 (Asp), and from residue 65 (Ser) to residue 89 (Cys) of SEQ ID NO:2.

Within preferred embodiments of the invention the isolated polynucleotides will hybridize to similar sized regions of SEQ ID NO:1 or a sequence complementary thereto, under stringent conditions. In general, stringent conditions are selected to be about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a

defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typical stringent conditions are those in which the NaCl concentration is up to about 0.03 M at pH 7 and the temperature is at least about 60°C.

As previously noted, the isolated polynucleotides of the present invention include DNA and RNA. Methods for isolating DNA and RNA are well known in the art. It is generally preferred to isolate RNA from kidney or bone marrow, although DNA can also be prepared using RNA from other tissues, cell lines, or isolated as genomic DNA. Total RNA can be prepared using guanidinium isothiocyanate extraction followed by isolation by centrifugation in a CsCl gradient (Chirgwin et al., Biochemistry 18:52-94, 1979). Poly (A)<sup>+</sup> RNA is prepared from total RNA using the method of Aviv and Leder (Proc. Natl. Acad. Sci. USA 69:1408-1412, 1972). Complementary DNA (cDNA) is prepared from poly(A)<sup>+</sup> RNA using known methods. Polynucleotides encoding zsig44 polypeptides are then identified and isolated by, for example, hybridization or PCR.

The present invention also provides polynucleotide molecules, including DNA and RNA molecules, that encode the zsig44 polypeptides disclosed herein. Those skilled in the art will readily recognize that, in view of the degeneracy of the genetic code, considerable sequence variation is possible among these polynucleotide molecules. SEQ ID NO:11 is a degenerate DNA sequence that encompasses all DNAs that encode the zsig44 polypeptide of SEQ ID NO:2. Those skilled in the art will recognize that the degenerate sequence of SEQ ID NO:11 also provides all RNA sequences encoding SEQ ID NO:2 by substituting U for T. Thus, zsig44 polypeptide-encoding polynucleotides comprising nucleotide 1 to nucleotide 267 of SEQ ID NO:11

and their RNA equivalents are contemplated by the present invention. Table 1 sets forth the one-letter codes used within SEQ ID NO:11 to denote degenerate nucleotide positions. "Resolutions" are the nucleotides denoted by a  
5 code letter. "Complement" indicates the code for the complementary nucleotide(s). For example, the code Y denotes either C or T, and its complement R denotes A or G, A being complementary to T, and G being complementary to C.

TABLE 1

| Nucleotide | Resolution | Nucleotide | Complement |
|------------|------------|------------|------------|
| A          | A          | T          | T          |
| C          | C          | G          | G          |
| G          | G          | C          | C          |
| T          | T          | A          | A          |
| R          | A G        | Y          | C T        |
| Y          | C T        | R          | A G        |
| M          | A C        | K          | G T        |
| K          | G T        | M          | A C        |
| S          | C G        | S          | C G        |
| W          | A T        | W          | A T        |
| H          | A C T      | D          | A G T      |
| B          | C G T      | V          | A C G      |
| V          | A C G      | B          | C G T      |
| D          | A G T      | H          | A C T      |
| N          | A C G T    | N          | A C G T    |

The degenerate codons used in SEQ ID NO:11,  
 5 encompassing all possible codons for a given amino acid,  
 are set forth in Table 2.

TABLE 2

| Amino Acid | One Letter Code | Codons                  | Degenerate Codon |
|------------|-----------------|-------------------------|------------------|
| Cys        | C               | TGC TGT                 | TGY              |
| Ser        | S               | AGC AGT TCA TCC TCG TCT | WSN              |
| Thr        | T               | ACA ACC ACG ACT         | ACN              |
| Pro        | P               | CCA CCC CCG CCT         | CCN              |
| Ala        | A               | GCA GCC GCG GCT         | GCN              |
| Gly        | G               | GGA GGC GGG GGT         | GGN              |
| Asn        | N               | AAC AAT                 | AAY              |
| Asp        | D               | GAC GAT                 | GAY              |
| Glu        | E               | GAA GAG                 | GAR              |
| Gln        | Q               | CAA CAG                 | CAR              |
| His        | H               | CAC CAT                 | CAY              |
| Arg        | R               | AGA AGG CGA CGC CGG CGT | MGN              |
| Lys        | K               | AAA AAG                 | AAR              |
| Met        | M               | ATG                     | ATG              |
| Ile        | I               | ATA ATC ATT             | ATH              |
| Leu        | L               | CTA CTC CTG CTT TTA TTG | YTN              |
| Val        | V               | GTA GTC GTG GTT         | GTN              |
| Phe        | F               | TTC TTT                 | TTY              |
| Tyr        | Y               | TAC TAT                 | TAY              |
| Trp        | W               | TGG                     | TGG              |
| Ter        | .               | TAA TAG TGA             | TRR              |
| Asn Asp    | B               |                         | RAY              |
| Glu Gln    | Z               |                         | SAR              |
| Any        | X               |                         | NNN              |

One of ordinary skill in the art will appreciate that some ambiguity is introduced in determining a degenerate codon, representative of all possible codons encoding each amino acid. For example, the degenerate codon for serine (WSN) can, in some circumstances, encode

arginine (AGR), and the degenerate codon for arginine (MGN) can, in some circumstances, encode serine (AGY). A similar relationship exists between codons encoding phenylalanine and leucine. Thus, some polynucleotides encompassed by the  
5 degenerate sequence may encode variant amino acid sequences, but one of ordinary skill in the art can easily identify such variant sequences by reference to the amino acid sequence of SEQ ID NO:2. Variant sequences can be readily tested for functionality as described herein.

10 One of ordinary skill in the art will also appreciate that different species can exhibit "preferential codon usage." In general, see, Grantham, et al., Nuc. Acids Res. 8:1893-912, 1980; Haas, et al. Curr. Biol. 6:315-24, 1996; Wain-Hobson, et al., Gene 13:355-64, 1981;  
15 Grosjean and Fiers, Gene 18:199-209, 1982; Holm, Nuc. Acids Res. 14:3075-87, 1986; Ikemura, J. Mol. Biol. 158:573-97, 1982. As used herein, the term "preferential codon usage" or "preferential codons" is a term of art referring to protein translation codons that are most frequently used in  
20 cells of a certain species, thus favoring one or a few representatives of the possible codons encoding each amino acid (See Table 2). For example, the amino acid Threonine (Thr) may be encoded by ACA, ACC, ACG, or ACT, but in mammalian cells ACC is the most commonly used codon; in  
25 other species, for example, insect cells, yeast, viruses or bacteria, different Thr codons may be preferential. Preferential codons for a particular species can be introduced into the polynucleotides of the present invention by a variety of methods known in the art.  
30 Introduction of preferential codon sequences into recombinant DNA can, for example, enhance production of the protein by making protein translation more efficient within a particular cell type or species. Therefore, the degenerate codon sequence disclosed in SEQ ID NO:11 serves

as a template for optimizing expression of polynucleotides in various cell types and species commonly used in the art and disclosed herein. Sequences containing preferential codons can be tested and optimized for expression in  
5 various species, and tested for functionality as disclosed herein.

The present invention further provides counterpart polypeptides and polynucleotides from other species (orthologs). These species include, but are not  
10 limited to mammalian, avian, amphibian, reptile, fish, insect and other vertebrate and invertebrate species. Of particular interest are zsig44 polypeptides from other mammalian species, including murine, rat, porcine, ovine, bovine, canine, feline, equine and other primate  
15 polypeptides. Orthologs of human zsig44 can be cloned using information and compositions provided by the present invention in combination with conventional cloning techniques. For example, a cDNA can be cloned using mRNA obtained from a tissue or cell type that expresses zsig44  
20 as disclosed herein. Suitable sources of mRNA can be identified by probing Northern blots with probes designed from the sequences disclosed herein. A library is then prepared from mRNA of a positive tissue or cell line. A zsig44-encoding cDNA can then be isolated by a variety of  
25 methods, such as by probing with a complete or partial human cDNA or with one or more sets of degenerate probes based on the disclosed sequences. The highly conserved amino acid sequence in the central domain of zsig44 can be used as a tool to identify novel ion channel family  
30 members. For instance, reverse transcription-polymerase chain reaction (RT-PCR) can be used to amplify sequences encoding the conserved motifs 1-4 and the PLITPGSA motif from RNA obtained from a variety of tissue sources. In particular, highly degenerate primers designed from the 5'

signal sequences are useful for this purpose. A cDNA can also be cloned using the polymerase chain reaction, or PCR (Mullis, U.S. Patent No. 4,683,202), using primers designed from the representative human zsig44 sequences disclosed  
5 herein. Within an additional method, the cDNA library can be used to transform or transfect host cells, and expression of the cDNA of interest can be detected with an antibody to zsig44 polypeptide. Similar techniques can also be applied to the isolation of genomic clones.

10 Those skilled in the art will recognize that the sequences disclosed in SEQ ID NO:1 represents a single allele of human zsig44, and that allelic variation and alternative splicing are expected to occur. Allelic variants of this sequence can be cloned by probing cDNA or  
15 genomic libraries from different individuals according to standard procedures. Splice variants within the same or different cell types or tissues can be cloned using a variety of molecular biological techniques known in the art. For example, 5' and 3' rapid amplification of cDNA  
20 ends (RACE), PCR using degenerate oligo primers and traditional hybridization cloning techniques among others can be used to identify and clone both allelic and splice variant cDNAs (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory  
25 Press, Cold Spring Harbor, NY, 1989, and Ausubel et al. (eds.), Current Protocols in Molecular Biology, John Wiley and Sons, Inc., NY, 1987). Allelic variants of the DNA sequence shown in SEQ ID NO:1, including those containing silent mutations and those in which mutations result in  
30 amino acid sequence changes, are within the scope of the present invention, as are proteins which are allelic variants of SEQ ID NO:2. CDNAs generated from alternatively spliced mRNAs, which retain the properties of the zsig44 polypeptide are included within the scope of the

present invention, as are polypeptides encoded by such cDNAs and mRNAs. Allelic variants and splice variants of these sequences can be cloned by probing cDNA or genomic libraries from different individuals or tissues according to standard procedures known in the art.

The present invention also provides isolated zsig44 polypeptides that are substantially homologous to the polypeptide of SEQ ID NO:2 and its orthologs. The term "substantially homologous" is used herein to denote polypeptides having 50%, preferably 60%, more preferably at least 80%, sequence identity to the sequence shown in SEQ ID NO:2 or its orthologs. Such polypeptides will more preferably be at least 90% identical, and most preferably 95% or more identical to SEQ ID NO:2 or its orthologs or paralogues. Percent sequence identity is determined by conventional methods. See, for example, Altschul et al., Bull. Math. Bio. 48: 603-616; 1986 and Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA 89:10915-10919, 1992. Briefly, two amino acid sequences are aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "blosum 62" scoring matrix of Henikoff and Henikoff (ibid.) as shown in Table 3 (amino acids are indicated by the standard one-letter codes). The percent identity is then calculated as:

$$\frac{\text{Total number of identical matches}}{[\text{length of the longer sequence plus the number of gaps introduced into the longer sequence in order to align the two sequences}]} \times 100$$

Table 3

|   | A  | R  | N  | D  | C  | Q  | E  | G  | H  | I  | L  | K  | M  | F  | P  | S  | T  | W  | Y  | V |
|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|
| A | 4  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |
| R | -1 | 5  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |
| N | -2 | 0  | 6  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |
| D | -2 | -2 | 1  | 6  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |
| C | 0  | -3 | -3 | -3 | 9  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |
| Q | -1 | 1  | 0  | 0  | -3 | 5  |    |    |    |    |    |    |    |    |    |    |    |    |    |   |
| E | -1 | 0  | 0  | 2  | -4 | 2  | 5  |    |    |    |    |    |    |    |    |    |    |    |    |   |
| G | 0  | -2 | 0  | -1 | -3 | -2 | -2 | 6  |    |    |    |    |    |    |    |    |    |    |    |   |
| H | -2 | 0  | 1  | -1 | -3 | 0  | 0  | -2 | 8  |    |    |    |    |    |    |    |    |    |    |   |
| I | -1 | -3 | -3 | -3 | -1 | -3 | -3 | -4 | -3 | 4  |    |    |    |    |    |    |    |    |    |   |
| L | -1 | -2 | -3 | -4 | -1 | -2 | -3 | -4 | -3 | 2  | 4  |    |    |    |    |    |    |    |    |   |
| K | -1 | 2  | 0  | -1 | -3 | 1  | 1  | -2 | -1 | -3 | -2 | 5  |    |    |    |    |    |    |    |   |
| M | -1 | -1 | -2 | -3 | -1 | 0  | -2 | -3 | -2 | 1  | 2  | -1 | 5  |    |    |    |    |    |    |   |
| F | -2 | -3 | -3 | -3 | -2 | -3 | -3 | -3 | -1 | 0  | 0  | -3 | 0  | 6  |    |    |    |    |    |   |
| P | -1 | -2 | -2 | -1 | -3 | -1 | -1 | -2 | -2 | -3 | -3 | -1 | -2 | -4 | 7  |    |    |    |    |   |
| S | 1  | -1 | 1  | 0  | -1 | 0  | 0  | 0  | -1 | -2 | -2 | 0  | -1 | -2 | -1 | 4  |    |    |    |   |
| T | 0  | -1 | 0  | -1 | -1 | -1 | -1 | -2 | -2 | -1 | -1 | -1 | -1 | -2 | -1 | 1  | 5  |    |    |   |
| W | -3 | -3 | -4 | -4 | -2 | -2 | -3 | -2 | -2 | -3 | -2 | -3 | -1 | 1  | -4 | -3 | -2 | 11 |    |   |
| Y | -2 | -2 | -2 | -3 | -2 | -1 | -2 | -3 | 2  | -1 | -1 | -2 | -1 | 3  | -3 | -2 | -2 | 2  | 7  |   |
| V | 0  | -3 | -3 | -3 | -1 | -2 | -2 | -3 | -3 | 3  | 1  | -2 | 1  | -1 | -2 | -2 | 0  | -3 | -1 | 4 |

5

10

15

20

Sequence identity of polynucleotide molecules is determined by similar methods using a ratio as disclosed above.

Variant zsig44 polypeptides or substantially homologous zsig44 polypeptides are characterized as having one or more amino acid substitutions, deletions or additions. These changes are preferably of a minor nature, that is conservative amino acid substitutions (see Table 4) and other substitutions that do not significantly affect the folding or activity of the protein or polypeptide, e.g., small deletions, typically of one to about 30 amino acids; and small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue, a small linker peptide of up to about 20-25 residues, or a small extension that facilitates purification, such as an affinity tag. The present invention thus includes zsig44 polypeptides containing tags, such as a poly-histidine tract, protein A (Nilsson et al., EMBO J. 4:1075, 1985; Nilsson et al., Methods Enzymol. 198:3, 1991), glutathione S transferase (Smith and Johnson, Gene 67:31, 1988), maltose binding protein (Kellerman and Ferenci, Methods Enzymol. 90:459-463, 1982; Guan et al., Gene 67:21-30, 1987), thioredoxin, ubiquitin, cellulose binding protein, T7 polymerase, or other antigenic epitope or binding domain. See, in general Ford et al., Protein Expression and Purification 2: 95-107, 1991. DNAs encoding affinity tags are available from commercial suppliers (e.g., Pharmacia Biotech, Piscataway, NJ; New England Biolabs, Beverly, MA). Polypeptides comprising affinity tags can further comprise a proteolytic cleavage site between the zsig44 polypeptide and the affinity tag. Preferred sites include thrombin cleavage sites and factor Xa cleavage sites. In addition amino acid residues of zsig44 can be photoaffinity labeled (Brunner et al., Ann. Rev. Biochem.

62:483-514, 1993 and Fedan et al., Biochem. Pharmacol. 33:1167-1180, 1984).

Table 4

Conservative amino acid substitutions

|    |              |               |
|----|--------------|---------------|
| 5  | Basic:       | arginine      |
|    |              | lysine        |
|    |              | histidine     |
| 10 | Acidic:      | glutamic acid |
|    |              | aspartic acid |
|    |              | glutamine     |
| 15 | Polar:       | asparagine    |
|    |              | leucine       |
|    |              | isoleucine    |
| 20 | Hydrophobic: | valine        |
|    |              | phenylalanine |
|    |              | tryptophan    |
| 25 | Aromatic:    | tyrosine      |
|    |              | glycine       |
|    |              | alanine       |
| 30 | Small:       | serine        |
|    |              | threonine     |
|    |              | methionine    |

The proteins of the present invention can also  
 25 comprise non-naturally occurring amino acid residues. Non-  
 naturally occurring amino acids include, without  
 limitation, *trans*-3-methylproline, 2,4-methanoproline, *cis*-  
 4-hydroxyproline, *trans*-4-hydroxyproline, *N*-methylglycine,  
*allo*-threonine, methylthreonine, hydroxyethylcysteine,  
 30 hydroxyethylhomocysteine, nitroglutamine, homoglutamine,  
 pipecolic acid, thiazolidine carboxylic acid,  
 dehydroproline, 3- and 4-methylproline, 3,3-  
 dimethylproline, *tert*-leucine, norvaline, 2-  
 azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine,

and 4-fluorophenylalanine. Several methods are known in the art for incorporating non-naturally occurring amino acid residues into proteins. For example, an *in vitro* system can be employed wherein nonsense mutations are suppressed using chemically aminoacylated suppressor tRNAs. Methods for synthesizing amino acids and aminoacylating tRNA are known in the art. Transcription and translation of plasmids containing nonsense mutations is carried out in a cell-free system comprising an *E. coli* S30 extract and commercially available enzymes and other reagents. Proteins are purified by chromatography. See, for example, Robertson et al., J. Am. Chem. Soc. 113:2722, 1991; Ellman et al., Methods Enzymol. 202:301, 1991; Chung et al., Science 259:806-9, 1993; and Chung et al., Proc. Natl. Acad. Sci. USA 90:10145-9, 1993). In a second method, translation is carried out in *Xenopus* oocytes by microinjection of mutated mRNA and chemically aminoacylated suppressor tRNAs (Turcatti et al., J. Biol. Chem. 271:19991-8, 1996). Within a third method, *E. coli* cells are cultured in the absence of a natural amino acid that is to be replaced (e.g., phenylalanine) and in the presence of the desired non-naturally occurring amino acid(s) (e.g., 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, or 4-fluorophenylalanine). The non-naturally occurring amino acid is incorporated into the protein in place of its natural counterpart. See, Koide et al., Biochem. 33:7470-6, 1994. Naturally occurring amino acid residues can be converted to non-naturally occurring species by *in vitro* chemical modification. Chemical modification can be combined with site-directed mutagenesis to further expand the range of substitutions (Wynn and Richards, Protein Sci. 2:395-403, 1993).

A limited number of non-conservative amino acids, amino acids that are not encoded by the genetic code, non-

naturally occurring amino acids, and unnatural amino acids may be substituted for zsig44 amino acid residues.

Essential amino acids in the zsig44 polypeptides of the present invention can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, Science 244: 1081-1085, 1989; Bass et al., Proc. Natl. Acad. Sci. USA 88:4498-502, 1991). In the latter technique, single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for biological activity (e.g., increased voltage-dependent conductance), as disclosed below, to identify amino acid residues that are critical to the activity of the molecule. See also, Hilton et al., J. Biol. Chem. 271:4699-4708, 1996. Sites of protein-protein and intramolecular amino acid interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction or photoaffinity labeling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos et al., Science 255:306-312, 1992; Smith et al., J. Mol. Biol. 224:899-904, 1992; Wlodaver et al., FEBS Lett. 309:59-64, 1992. The identities of essential amino acids of the zsig44 polypeptide can also be inferred from analysis of homologies with related proteins, such as known ion channels.

In addition a hydrophobicity plot of zsig44, which identifies putative highly antigenic sites, is useful for predicting allowable amino acid substitutions and antibody epitopes, as discussed herein. Generation of hydrophobicity plots, and determination of antigenic sites, allowable amino acid substitutions and epitopes, is well within the skill of one in the art.

Multiple amino acid substitutions can be made and tested using known methods of mutagenesis and screening. For example, see Reidhaar-Olson and Sauer Science 241:53-57, 1988; or Bowie and Sauer, Proc. Natl. Acad. Sci. USA 86:2152-2156, 1989. Briefly, these authors disclose methods for simultaneously randomizing two or more positions in a polypeptide, selecting for functional polypeptide, and then sequencing the mutagenized polypeptides to determine the spectrum of allowable substitutions at each position. Other methods that can be used include phage display (e.g., Lowman et al., Biochem. 30:10832-10837, 1991; Ladner et al., U.S. Patent No. 5,223,409; Huse, WIPO Publication WO 92/06204) and region-directed mutagenesis (Derbyshire et al., Gene 46:145, 1986; Ner et al., DNA 7:127, 1988).

Variants of the disclosed zsig44 DNA and polypeptide sequences can be generated through DNA shuffling as disclosed by Stemmer, Nature 370:389-91, 1994; Stemmer, Proc. Natl. Acad. Sci. USA 91:10747-51, 1994; and WIPO Publication WO 97/20078. Briefly, variant DNAs are generated by *in vitro* homologous recombination by random fragmentation of a parent DNA followed by reassembly using PCR, resulting in randomly introduced point mutations. This technique can be modified by using a family of parent DNAs, such as allelic variants or DNAs from different species, to introduce additional variability into the process. Selection or screening for the desired activity, followed by additional iterations of mutagenesis and assay provides for rapid "evolution" of sequences by selecting for desirable mutations while simultaneously selecting against detrimental changes.

Mutagenesis methods as disclosed herein can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides in

host cells. Mutagenized DNA molecules that encode active polypeptides (e.g., increase voltage-dependent conductance in *Xenopus laevis* oocytes or mammalian cells, or detected as expressed on the surface of cells by an antibody raised to zsig44) can be recovered from the host cells and rapidly sequenced using modern equipment. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide of interest, and can be applied to polypeptides of unknown structure.

Using the methods discussed herein, one of ordinary skill in the art can identify and/or prepare a variety of polypeptide fragments or variants of SEQ ID NO:2 that retain the ion channel or regulatory properties of the wild-type protein. Activity can be assessed by techniques discussed herein. Such polypeptides may include additional amino acids from, for example, part or all of the transmembrane, central, and intracellular domains; other domains; affinity tags; and the like. Such polypeptides may also include additional polypeptide segments as generally disclosed above.

For any zsig44 polypeptide, including variants and fusion proteins, one of ordinary skill in the art can readily generate a fully degenerate polynucleotide sequence encoding that variant using the information set forth in Tables 1 and 2 above.

The polypeptides of the present invention, including full-length proteins, fragments thereof, biologically active fragments, and fusion proteins, can be produced in genetically engineered host cells according to conventional techniques. Suitable host cells are those cell types that can be transformed or transfected with exogenous DNA and grown in culture, and include bacteria, fungal cells, and cultured higher eukaryotic cells.

Eukaryotic cells, particularly cultured cells of multicellular organisms, are preferred. Techniques for manipulating cloned DNA molecules and introducing exogenous DNA into a variety of host cells are disclosed by Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, and Ausubel et al. eds., Current Protocols in Molecular Biology, John Wiley and Sons, Inc., NY, 1987.

In general, a DNA sequence encoding a zsig44 polypeptide is operably linked to other genetic elements required for its expression, generally including a transcription promoter and terminator, within an expression vector. The vector will also commonly contain one or more selectable markers and one or more origins of replication, although those skilled in the art will recognize that within certain systems selectable markers may be provided on separate vectors, and replication of the exogenous DNA may be provided by integration into the host cell genome. Selection of promoters, terminators, selectable markers, vectors and other elements is a matter of routine design within the level of ordinary skill in the art. Many such elements are described in the literature and are available through commercial suppliers.

To direct a zsig44 polypeptide into the secretory pathway of a host cell, a secretory signal sequence (also known as a leader sequence, prepro sequence or pre sequence) is provided in the expression vector. The secretory signal sequence is operably linked to the zsig44 DNA sequence, i.e., the two sequences are joined in the correct reading frame and positioned to direct the newly synthesized polypeptide into the secretory pathway of the host cell. Secretory signal sequences are commonly positioned 5' to the DNA sequence encoding the polypeptide of interest, although certain secretory signal sequences

may be positioned elsewhere in the DNA sequence of interest (see, e.g., Welch et al., U.S. Patent No. 5,037,743; Holland et al., U.S. Patent No. 5,143,830).

Alternatively, the secretory signal sequence  
5 contained in the polypeptides of the present invention, and disclosed herein, is used to direct other polypeptides into the secretory pathway. The present invention provides for such fusion polypeptides. A signal fusion polypeptide can be made wherein a secretory signal sequence derived from  
10 residue 1 (Met) to residue 16 (Ala) of SEQ ID NO:2 is operably linked to another polypeptide using methods known in the art and disclosed herein. The secretory signal sequence contained in the fusion polypeptides of the present invention is preferably fused amino-terminally to  
15 an additional peptide to direct the additional peptide into the secretory pathway. Such constructs have numerous applications known in the art. For example, these novel secretory signal sequence fusion constructs can direct the secretion of an active component of a normally non-secreted  
20 protein. Such fusions may be used *in vivo* or *in vitro* to direct peptides through the secretory pathway.

Cultured mammalian cells are suitable hosts within the present invention. Methods for introducing  
25 exogenous DNA into mammalian host cells include calcium phosphate-mediated transfection (Wigler et al., Cell 14:725, 1978; Corsaro and Pearson, Somatic Cell Genetics 7:603, 1981; Graham and Van der Eb, Virology 52:456, 1973), electroporation (Neumann et al., EMBO J. 1:841-845, 1982),  
30 DEAE-dextran mediated transfection (Ausubel et al., ibid.), liposome-mediated transfection (Hawley-Nelson et al., Focus 15:73, 1993; Ciccarone et al., Focus 15:80, 1993), and viral vectors (Miller, A., Rosman, G., BioTechniques 7:980-90, 1989; Wang, Q., Finer, M., Nature Med. 2:714-16, 1996).

The production of recombinant polypeptides in cultured mammalian cells is disclosed, for example, by Levinson et al., U.S. Patent No. 4,713,339; Hagen et al., U.S. Patent No. 4,784,950; Palmiter et al., U.S. Patent No. 4,579,821; 5 and Ringold, U.S. Patent No. 4,656,134. Suitable cultured mammalian cells include the COS-1 (ATCC No. CRL 1650), COS-7 (ATCC No. CRL 1651), BHK (ATCC No. CRL 1632), BHK 570 (ATCC No. CRL 10314), 293 (ATCC No. CRL 1573; Graham et al., J. Gen. Virol. 36:59-72, 1977) and Chinese hamster 10 ovary (e.g. CHO-K1; ATCC No. CCL 61) cell lines. Additional suitable cell lines are known in the art and available from public depositories such as the American Type Culture Collection, Rockville, Maryland. In general, strong transcription promoters are preferred, such as 15 promoters from SV-40 or cytomegalovirus. See, e.g., U.S. Patent No. 4,956,288. Other suitable promoters include those from metallothionein genes (U.S. Patent Nos. 4,579,821 and 4,601,978) and the adenovirus major late promoter.

20 Drug selection is generally used to select for cultured mammalian cells into which foreign DNA has been inserted. Such cells are commonly referred to as "transfectants". Cells that have been cultured in the presence of the selective agent and are able to pass the 25 gene of interest to their progeny are referred to as "stable transfectants." A preferred selectable marker is a gene encoding resistance to the antibiotic neomycin. Selection is carried out in the presence of a neomycin-type drug, such as G-418 or the like. Selection systems can 30 also be used to increase the expression level of the gene of interest, a process referred to as "amplification." Amplification is carried out by culturing transfectants in the presence of a low level of the selective agent and then increasing the amount of selective agent to select for

cells that produce high levels of the products of the introduced genes. A preferred amplifiable selectable marker is dihydrofolate reductase, which confers resistance to methotrexate. Other drug resistance genes (e.g.,  
5 hygromycin resistance, multi-drug resistance, puromycin acetyltransferase) can also be used. Alternative markers that introduce an altered phenotype, such as green fluorescent protein, or cell surface proteins such as CD4, CD8, Class I MHC, placental alkaline phosphatase may be  
10 used to sort transfected cells from untransfected cells by such means as FACS sorting or magnetic bead separation technology.

Other higher eukaryotic cells can also be used as hosts, including insect cells, amphibian cells (e.g.  
15 *Xenopus laevis* oocytes), plant cells and avian cells. Transformation of insect cells and production of foreign polypeptides therein is disclosed by Guarino et al., U.S. Patent No. 5,162,222; Bang et al., U.S. Patent No. 4,775,624; and WIPO publication WO 94/06463. The use of  
20 *Agrobacterium rhizogenes* as a vector for expressing genes in plant cells has been reviewed by Sinkar et al., J. Biosci. (Bangalore) 11:47-58, 1987.

Insect cells can be infected with recombinant baculovirus, commonly derived from *Autographa californica*  
25 nuclear polyhedrosis virus (AcNPV). DNA encoding the zsig44 polypeptide is inserted into the baculoviral genome in place of the AcNPV polyhedrin gene coding sequence by one of two methods. The first is the traditional method of homologous DNA recombination between wild-type AcNPV and a  
30 transfer vector containing the zsig44 flanked by AcNPV sequences. Suitable insect cells, e.g. SF9 cells, are infected with wild-type AcNPV and transfected with a transfer vector comprising a zsig44 polynucleotide operably linked to an AcNPV polyhedrin gene promoter, terminator,

and flanking sequences. See, King, L.A. and Possee, R.D., The Baculovirus Expression System: A Laboratory Guide, London, Chapman & Hall; O'Reilly, D.R. et al., Baculovirus Expression Vectors: A Laboratory Manual, New York, Oxford University Press., 1994; and, Richardson, C. D., Ed., Baculovirus Expression Protocols. Methods in Molecular Biology, Totowa, NJ, Humana Press, 1995. Natural recombination within an insect cell will result in a recombinant baculovirus which contains zsig44 driven by the polyhedrin promoter. Recombinant viral stocks are made by methods commonly used in the art.

The second method of making recombinant baculovirus utilizes a transposon-based system described by Luckow (Luckow, V.A, et al., J Virol 67:4566-79, 1993). This system is sold in the Bac-to-Bac™ kit (Life Technologies, Rockville, MD). This system utilizes a transfer vector, pFastBac1™ (Life Technologies) containing a Tn7 transposon to move the DNA encoding the zsig44 polypeptide into a baculovirus genome maintained in E. coli as a large plasmid called a "bacmid." The pFastBac1™ transfer vector utilizes the AcNPV polyhedrin promoter to drive the expression of the gene of interest, in this case zsig44. However, pFastBac1™ can be modified to a considerable degree. See, Hill-Perkins, M.S. and Possee, R.D., J. Gen. Virol. 71:971-6, 1990; Bonning, B.C. et al., J. Gen. Virol. 75:1551-6, 1994; and, Chazenbalk, G.D., and Rapoport, B., J. Biol. Chem. 270:1543-9, 1995. In addition, transfer vectors can include an in-frame fusion with DNA encoding an epitope tag at the C- or N-terminus of the expressed zsig44 polypeptide, for example, a Glu-Glu epitope tag (Grussenmeyer, T. et al., Proc. Natl. Acad. Sci. 82:7952-4, 1985). Using a technique known in the art, a transfer vector containing zsig44 is transformed into E. Coli, and screened for bacmids which contain an interrupted

lacZ gene indicative of recombinant baculovirus. The bacmid DNA containing the recombinant baculovirus genome is isolated, using common techniques, and used to transfect *Spodoptera frugiperda* cells, e.g. Sf9 cells. Recombinant virus that expresses zsig44 is subsequently produced. Recombinant viral stocks are made by methods commonly used the art.

The recombinant virus is used to infect host cells, typically a cell line derived from the fall armyworm, *Spodoptera frugiperda*. See, in general, Glick and Pasternak, Molecular Biotechnology: Principles and Applications of Recombinant DNA, ASM Press, Washington, D.C., 1994. Another suitable cell line is the High Five<sup>TM</sup> cell line (Invitrogen) derived from *Trichoplusia ni* (U.S. Patent No. 5,300,435). Commercially available serum-free media are used to grow and maintain the cells. Suitable media are Sf900 II<sup>TM</sup> (Life Technologies) or ESF 921<sup>TM</sup> (Expression Systems) for the Sf9 cells; and Ex-cello405<sup>TM</sup> (JRH Biosciences, Lenexa, KS) or Express Five<sup>TM</sup> (Life Technologies) for the *T. ni* cells. The cells are grown up from an inoculation density of approximately  $2-5 \times 10^5$  cells to a density of  $1-2 \times 10^6$  cells at which time a recombinant viral stock is added at a multiplicity of infection (MOI) of 0.1 to 10, more typically near 3. Procedures used are generally described in available laboratory manuals (King, L. A. and Possee, R.D., ibid.; O'Reilly, D.R. et al., ibid.; Richardson, C. D., ibid.). Subsequent purification of the zsig44 polypeptide from the supernatant can be achieved using methods described herein.

Fungal cells, including yeast cells, can also be used within the present invention. Yeast species of particular interest in this regard include *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Pichia methanolica*. Methods for transforming *S. cerevisiae* cells with exogenous

DNA and producing recombinant polypeptides therefrom are disclosed by, for example, Kawasaki, U.S. Patent No. 4,599,311; Kawasaki et al., U.S. Patent No. 4,931,373; Brake, U.S. Patent No. 4,870,008; Welch et al., U.S. Patent No. 5,037,743; and Murray et al., U.S. Patent No. 4,845,075. Transformed cells are selected by phenotype determined by the selectable marker, commonly drug resistance or the ability to grow in the absence of a particular nutrient (e.g., leucine). A preferred vector system for use in *Saccharomyces cerevisiae* is the POT1 vector system disclosed by Kawasaki et al. (U.S. Patent No. 4,931,373), which allows transformed cells to be selected by growth in glucose-containing media. Suitable promoters and terminators for use in yeast include those from glycolytic enzyme genes (see, e.g., Kawasaki, U.S. Patent No. 4,599,311; Kingsman et al., U.S. Patent No. 4,615,974; and Bitter, U.S. Patent No. 4,977,092) and alcohol dehydrogenase genes. See also U.S. Patents Nos. 4,990,446; 5,063,154; 5,139,936 and 4,661,454. Transformation systems for other yeasts, including *Hansenula polymorpha*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces fragilis*, *Ustilago maydis*, *Pichia pastoris*, *Pichia guilliermondii*, *Pichia methanolica* and *Candida maltosa* are known in the art. See, for example, Gleeson et al., J. Gen. Microbiol. 132:3459-3465, 1986 and Cregg, U.S. Patent No. 4,882,279. *Aspergillus* cells may be utilized according to the methods of McKnight et al., U.S. Patent No. 4,935,349. Methods for transforming *Acremonium chrysogenum* are disclosed by Sumino et al., U.S. Patent No. 5,162,228. Methods for transforming *Neurospora* are disclosed by Lambowitz, U.S. Patent No. 4,486,533.

Transformed or transfected host cells are cultured according to conventional procedures in a culture medium containing nutrients and other components required

for the growth of the chosen host cells. A variety of suitable media, including defined media and complex media, are known in the art and generally include a carbon source, a nitrogen source, essential amino acids, vitamins and minerals. Media may also contain such components as growth factors or serum, as required. The growth medium will generally select for cells containing the exogenously added DNA by, for example, drug selection or deficiency in an essential nutrient which is complemented by the selectable marker carried on the expression vector or co-transfected into the host cell.

The use of *Pichia methanolica* as host for the production of recombinant proteins is disclosed in WIPO Publications WO 97/17450, WO 97/17451, WO 98/02536, and WO 98/02565. DNA molecules for use in transforming *P. methanolica* will commonly be prepared as double-stranded, circular plasmids, which are preferably linearized prior to transformation. For polypeptide production in *P. methanolica*, it is preferred that the promoter and terminator in the plasmid be that of a *P. methanolica* gene, such as a *P. methanolica* alcohol utilization gene (*AUG1* or *AUG2*). Other useful promoters include those of the dihydroxyacetone synthase (*DHAS*), formate dehydrogenase (*FMD*), and catalase (*CAT*) genes. To facilitate integration of the DNA into the host chromosome, it is preferred to have the entire expression segment of the plasmid flanked at both ends by host DNA sequences. A preferred selectable marker for use in *Pichia methanolica* is a *P. methanolica* *ADE2* gene, which encodes phosphoribosyl-5-aminoimidazole carboxylase (*AIRC*; EC 4.1.1.21), which allows *ade2* host cells to grow in the absence of adenine. For large-scale, industrial processes where it is desirable to minimize the use of methanol, it is preferred to use host cells in which both methanol utilization genes (*AUG1* and *AUG2*) are

deleted. For production of secreted proteins, host cells deficient in vacuolar protease genes (*PEP4* and *PRB1*) are preferred. Electroporation is used to facilitate the introduction of a plasmid containing DNA encoding a polypeptide of interest into *P. methanolica* cells. It is preferred to transform *P. methanolica* cells by electroporation using an exponentially decaying, pulsed electric field having a field strength of from 2.5 to 4.5 kV/cm, preferably about 3.75 kV/cm, and a time constant (t) of from 1 to 40 milliseconds, most preferably about 20 milliseconds.

Prokaryotic host cells, including strains of the bacteria *Escherichia coli*, *Bacillus* and other genera are also useful host cells within the present invention. Techniques for transforming these hosts and expressing foreign DNA sequences cloned therein are well known in the art (see, e.g., Sambrook et al., *ibid.*). When expressing a *zsig44* polypeptide in bacteria such as *E. coli*, the polypeptide may be retained in the cytoplasm, typically as insoluble granules, or may be directed to the periplasmic space by a bacterial secretion sequence. In the former case, the cells are lysed, and the granules are recovered and denatured using, for example, guanidine isothiocyanate or urea. The denatured polypeptide can then be refolded and dimerized by diluting the denaturant, such as by dialysis against a solution of urea and a combination of reduced and oxidized glutathione, followed by dialysis against a buffered saline solution. In the latter case, the polypeptide can be recovered from the periplasmic space in a soluble and functional form by disrupting the cells (by, for example, sonication or osmotic shock) to release the contents of the periplasmic space and recovering the protein, thereby obviating the need for denaturation and refolding.

*P. methanolica* cells are cultured in a medium comprising adequate sources of carbon, nitrogen and trace nutrients at a temperature of about 25°C to 35°C. Liquid cultures are provided with sufficient aeration by conventional means, such as shaking of small flasks or sparging of fermentors. A preferred culture medium for *P. methanolica* is YEPD (2% D-glucose, 2% Bacto™ Peptone (Difco Laboratories, Detroit, MI), 1% Bacto™ yeast extract (Difco Laboratories), 0.004% adenine and 0.006% L-leucine).

It is preferred to purify the protein to >80% purity, more preferably to >90% purity, even more preferably >95%, and particularly preferred is a pharmaceutically pure state, that is greater than 99.9% pure with respect to contaminating macromolecules, particularly other proteins and nucleic acids, and free of infectious and pyrogenic agents. Preferably, a purified protein is substantially free of other proteins, particularly other proteins of animal origin.

Expressed recombinant zsig44 polypeptides (or chimeric zsig44 polypeptides) can be purified using fractionation and/or conventional purification methods and media. Ammonium sulfate precipitation and acid or chaotrope extraction may be used for fractionation of samples. Exemplary purification steps include hydroxyapatite, size exclusion, FPLC and reverse-phase high performance liquid chromatography. Suitable chromatographic media include derivatized dextrans, agarose, cellulose, polyacrylamide, specialty silicas, and the like. PEI, DEAE, QAE and Q derivatives are preferred. Exemplary chromatographic media include those media derivatized with phenyl, butyl, or octyl groups, such as Phenyl-Sepharose FF (Pharmacia), Toyopearl butyl 650 (Toso Haas, Montgomeryville, PA), Octyl-Sepharose (Pharmacia) and the like; or polyacrylic resins, such as Amberchrom CG 71

(Toso Haas) and the like. Suitable solid supports include glass beads, silica-based resins, cellulosic resins, agarose beads, cross-linked agarose beads, polystyrene beads, cross-linked polyacrylamide resins, and the like, that are insoluble under the conditions in which they are to be used. These supports may be modified with reactive groups that allow attachment of proteins by amino groups, carboxyl groups, sulfhydryl groups, hydroxyl groups and/or carbohydrate moieties. Examples of coupling chemistries include cyanogen bromide activation, N-hydroxysuccinimide activation, epoxide activation, sulfhydryl activation, hydrazide activation, and carboxyl and amino derivatives for carbodiimide coupling chemistries. These and other solid media are well known and widely used in the art, and are available from commercial suppliers. Methods for binding receptor polypeptides to support media are well known in the art. Selection of a particular method is a matter of routine design and is determined in part by the properties of the chosen support. See, for example, Affinity Chromatography: Principles & Methods, Pharmacia LKB Biotechnology, Uppsala, Sweden, 1988.

The polypeptides of the present invention can be isolated by exploitation of their structural and physical properties. For example, immobilized metal ion adsorption (IMAC) chromatography can be used to purify histidine-rich proteins, including those comprising polyhistidine tags. Briefly, a gel is first charged with divalent metal ions to form a chelate (E. Sulkowski, Trends in Biochem. 3:1-7, 1985). Histidine-rich proteins will be adsorbed to this matrix with differing affinities, depending upon the metal ion used, and will be eluted by competitive elution, lowering the pH, or use of strong chelating agents. Other methods of purification include purification of glycosylated proteins by lectin affinity chromatography and

ion exchange chromatography (Methods in Enzymol., Vol. 182, "Guide to Protein Purification", M. Deutscher, (ed.), Acad. Press, San Diego, 1990, pp. 529-39). Within additional embodiments of the invention, a fusion of the polypeptide of interest and an affinity tag (e.g., polyhistidine, maltose-binding protein, an immunoglobulin domain) may be constructed to facilitate purification.

Moreover, using methods described in the art, polypeptide fusions, or hybrid zsig44 ion channel proteins, are constructed using regions or domains of the inventive zsig44 in combination with those of other known ion channel proteins (e.g. MAT-8, PLM, and IsK), or heterologous proteins (Sambrook et al., ibid.; Altschul et al., ibid.; Picard. D. Cur. Opin. Biology, 5:511-515, 1994, and references therein). These methods allow the determination of the biological importance of larger domains or regions in a polypeptide of interest. Such hybrids may alter reaction kinetics, binding, constrict or expand the substrate specificity, or alter tissue and cellular localization of a polypeptide, and can be applied to polypeptides of unknown structure.

Fusion proteins can be prepared by methods known to those skilled in the art by preparing each component of the fusion protein and chemically conjugating them. Alternatively, a polynucleotide encoding both components of the fusion protein in the proper reading frame can be generated using known techniques and expressed by the methods described herein. For example, part or all of a domain(s) conferring a biological function may be swapped between zsig44 of the present invention with the functionally equivalent domain(s) from another ion channel, such as MAT-8 or IsK. Such domains include, but are not limited to the secretory signal sequence, transmembrane domain, central domain, and regions flanking the central

domain, as described herein. Such fusion proteins would be expected to have a biological functional profile that is the same or similar to polypeptides of the present invention or other known ion channel proteins (e.g. MAT-8, CHIF or IsK), depending on the fusion constructed. Moreover, such fusion proteins may exhibit other properties as disclosed herein.

Zsig44 polypeptides or fragments thereof can also be prepared through chemical synthesis. Zsig44 polypeptides may be monomers or multimers; glycosylated or non-glycosylated; pegylated or non-pegylated; and may or may not include an initial methionine amino acid residue.

Polypeptides of the present invention can also be synthesized by exclusive solid phase synthesis, partial solid phase methods, fragment condensation or classical solution synthesis. Methods for synthesizing polypeptides are well known in the art. See, for example, Merrifield, J. Am. Chem. Soc. 85:2149, 1963; Kaiser et al., Anal. Biochem. 34:595, 1970. After the entire synthesis of the desired peptide on a solid support, the peptide-resin is with a reagent which cleaves the polypeptide from the resin and removes most of the side-chain protecting groups. Such methods are well established in the art.

The activity of molecules of the present invention can be measured using a variety of assays that measure ion channel activity. Of particular interest is measuring ion transfer cross cell membranes. Such assays are well known in the art. Specific assays to assess the activity of novel ion channels or their regulators include, but are not limited to, bioassays measuring voltage-dependent conductance in *Xenopus laevis* oocytes (see, Rudy, B., and Iverson, L.E., eds., Meth. Enzymol., vol. 207, Academic Press, San Diego, CA, 1992; Hamill, O.P et al., Pfluegers Arch. 391:85-100, 1981; Moorman, J.R. et al., J.

Biol. Chem. 267:14551-14554, 1992; Durieux, M.E., et al.,  
Am. J. Physiol. 263:C896-C900, 1992). This method involves  
injecting *in vitro* expressed mRNAs into isolated oocytes  
and assessing voltage-dependent conductance using a patch-  
5 clamp technique. An ion channel or its regulator may  
increase voltage-dependent conductance in this assay  
system. This system may be applied to other cell types,  
such as insect and mammalian cells (see, Rudy, B., Iverson,  
L.E., eds., ibid.). Other assays involve measuring ion  
10 channel activity indirectly in mammalian or other cell  
types, through the use of a chelator dye, such as Fura2  
(See, for example, James-Kracke M.R., J. Gen. Physiol.  
99:41-62, 1992; Raghu, P. et al., Gene 190:151-156, 1997).  
Ion channel activity can also be monitored by using a  
15 radiolabeled ion, such as a <sup>125</sup>I efflux assay (Xia, Y. et  
al., J. Membr. Biol. 151:269-278, 1996). Other assays  
involve measuring changes in gene expression in mammalian  
cells signaled by ion flux or ion channel phosphorylation;  
for example, by driving expression of a measurable reporter  
20 gene, e.g. luciferase, under a suitable promoter, for  
example, as described below.

An alternative *in vivo* approach for assaying  
proteins of the present invention involves viral delivery  
systems. Exemplary viruses for this purpose include  
25 adenovirus, herpesvirus, vaccinia virus and adeno-  
associated virus (AAV). Adenovirus, a double-stranded DNA  
virus, is currently the best studied gene transfer vector  
for delivery of heterologous nucleic acid (for a review,  
see T.C. Becker et al., Meth. Cell Biol. 43:161-89, 1994;  
30 and J.T. Douglas and D.T. Curiel, Science & Medicine 4:44-  
53, 1997). The adenovirus system offers several  
advantages: adenovirus can (i) accommodate relatively  
large DNA inserts; (ii) be grown to high-titer; (iii)  
infect a broad range of mammalian cell types; and (iv) be

used with a large number of available vectors containing different promoters. Also, because adenoviruses are stable in the bloodstream, they can be administered by intravenous injection.

5           By deleting portions of the adenovirus genome, larger inserts (up to 7 kb) of heterologous DNA can be accommodated. These inserts can be incorporated into the viral DNA by direct ligation or by homologous recombination with a co-transfected plasmid. In an exemplary system, the  
10   essential E1 gene has been deleted from the viral vector, and the virus will not replicate unless the E1 gene is provided by the host cell (the human 293 cell line is exemplary). When intravenously administered to intact animals, adenovirus primarily targets the liver. If the  
15   adenoviral delivery system has an E1 gene deletion, the virus cannot replicate in the host cells. However, the host's tissue (e.g., liver) will express and process (and, if a secretory signal sequence is present, secrete) the heterologous protein. Secreted proteins will enter the  
20   circulation in the highly vascularized liver, and effects on the infected animal can be determined.

          The adenovirus system can also be used for protein production *in vitro*. By culturing adenovirus-infected non-293 cells under conditions where the cells are  
25   not rapidly dividing, the cells can produce proteins for extended periods of time. For instance, BHK cells are grown to confluence in cell factories, then exposed to the adenoviral vector encoding the secreted protein of interest. The cells are then grown under serum-free  
30   conditions, which allows infected cells to survive for several weeks without significant cell division. Alternatively, adenovirus vector infected 293S cells can be grown in suspension culture at relatively high cell density to produce significant amounts of protein (see Garnier et

al., Cytotechnol. 15:145-55, 1994). With either protocol, an expressed, secreted heterologous protein can be repeatedly isolated from the cell culture supernatant. Within the infected 293S cell production protocol, non-  
5 secreted proteins may also be effectively obtained.

The tissue distribution of zsig44 suggests a role in kidney or bone marrow function. Zsig44 may act as a novel ion channel or regulate existing or unknown ion channels in the bone marrow, kidney or other tissues, such  
10 as heart and spinal cord. For example, several ClC chloride channels ClC-1 and ClC-2 human homologs, disclosed above, are known to be kidney-specific and may be regulated by zsig44. Moreover, zsig44 may play a role in kidney, bone marrow, heart and or neural pathologies associated  
15 with genetic and other human disease states, such as diabetes, kidney stones, bone disease, hematopoietic disorders, immune disorders, leukemias, hypertension, cardiac disorders and neural diseases. Discovery of antagonists and or agonists of zsig44 activity would offer  
20 therapeutic uses for zsig44. In view of the tissue distribution observed for this zsig44, agonists (including the natural ligand, substrate, cofactor, etc.) and antagonists have potential in both *in vitro* and *in vivo* applications.

25 Proteins of the present invention are also used in a cell-based screen for modulators (e.g., antagonists and agonists) to zsig44. Antagonists and agonists can affect zsig44 in several ways, e.g. regulatory activity, gene expression, binding or interaction with other ion  
30 channel polypeptides or ion channel subunits (Kim, J.W. et al., Biochim. Biophys. ACTA 350:133-135, 1997). Such antagonists and agonists could affect zsig44 polypeptide by respectively decreasing or increasing transport of ions through an ion channel; either by affecting the putative

regulatory function of zsig44 activity on other ion channels or affecting zsig44 activity as a direct ion channel or subunit thereof. This increase or decrease is measured by assessing voltage-dependent conductance or in  
5 another appropriate assay system known in the art. In such application, zsig44 is expressed alone, or co-expressed with an indicator ion channel regulated by polypeptides of the present invention. Methods to construct such a cell are known in the art and disclosed herein. Preferred  
10 indicator ion channels have a readily measurable biological activity such as measurable voltage conductance or calcium flux as indicated by a fluorescent chelator dye, e.g. Fura2. Examples of such activity assays are known in the art. Antagonists and agonists are identified by screening  
15 the voltage conductance or calcium flux of the cells after exposure to the presence of various agents, discussed below. Changes in the voltage conductance or the indicator substrate reflect activities that the agents exert on zsig44 by either enhancing or inhibiting zsig44 activity,  
20 relative to control cells not subjected to the agent. For example, relative to the control, an agonist increasing the activity of zsig44, would result in increased conductance. Conversely, relative to the control, an antagonist decreasing the activity of zsig44, would result in  
25 decreased conductance. Sources for agents include any natural or chemical source including but not limited to plant, microbial and fungal extracts, chemical libraries, and combinatorial chemical libraries, and the like. Methods of establishing and employing this type of cell-  
30 based screening assay are known in the art.

Zsig44 can also be used to identify modulators (e.g, antagonists) of its activity. Test compounds are added to the assays disclosed herein to identify compounds that inhibit the activity of zsig44. In addition to those

assays disclosed herein, samples can be tested for inhibition of zsig44 activity within a variety of assays designed to measure zsig44 binding, oligomerization, or the stimulation/inhibition of zsig44-dependent cellular responses. For example, zsig44-responsive cell lines can be transfected with a reporter gene construct that is responsive to a zsig44-stimulated cellular pathway. Reporter gene constructs of this type are known in the art, and will generally comprise a zsig44-DNA response element operably linked to a gene encoding an assay detectable protein, such as luciferase. DNA response elements can include, but are not limited to, cyclic AMP response elements (CRE), hormone response elements (HRE) insulin response element (IRE) (Nasrin et al., Proc. Natl. Acad. Sci. USA 87:5273-7, 1990) and serum response elements (SRE) (Shaw et al. Cell 56: 563-72, 1989). Cyclic AMP response elements are reviewed in Roestler et al., J. Biol. Chem. 263 (19):9063-6; 1988 and Habener, Molec. Endocrinol. 4 (8):1087-94; 1990. Hormone response elements are reviewed in Beato, Cell 56:335-44; 1989. Candidate compounds, solutions, mixtures or extracts are tested for the ability to inhibit the activity of zsig44 on the target cells as evidenced by a decrease in zsig44 stimulation of reporter gene expression. Assays of this type will detect compounds that directly block zsig44 binding to cell-surface receptors, e.g., through dimerization, as well as compounds that block processes in the cellular pathway subsequent to such binding. Alternatively, compounds or other samples can be tested for direct blocking of zsig44, or blocking of zsig44 binding to other cell surface molecules, using zsig44 tagged with a detectable label (e.g., <sup>125</sup>I, biotin, horseradish peroxidase, FITC, and the like). Within assays of this type, the ability of a test sample to inhibit labeled zsig44 binding to another protein can be indicative of inhibitory activity, which can be confirmed through secondary assays.

Zsig44 polypeptides can also be used to prepare antibodies that specifically bind to zsig44 epitopes, peptides or polypeptides. The zsig44 polypeptide or a fragment thereof serves as an antigen (immunogen) to inoculate an animal and elicit an immune response. Suitable antigens would be the mature zsig44 polypeptide encoded by SEQ ID NO:2 from amino acid number residue 17 (Leu) to residue 89 (Cys) or a contiguous 9 to 89 amino acid fragment thereof. In addition, suitable antigens include the central domain from amino acid residue 30 (Pro) to amino acid residue 64 (Lys) of SEQ ID NO:2, and the flanking regions from residue 17 (Leu) to residue 28 (Asp), and from residue 65 (Ser) to residue 89 (Cys) of SEQ ID NO:2. Moreover, predicted antigenic stretches, based on a hydrophobicity plot of zsig44 polypeptide, can also serve as suitable antigens for the generation of antibodies. Antibodies generated from this immune response can be isolated and purified as described herein. Methods for preparing and isolating polyclonal and monoclonal antibodies are well known in the art. See, for example, Current Protocols in Immunology, Cooligan, et al. (eds.), National Institutes of Health, John Wiley and Sons, Inc., 1995; Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, NY, 1989; and Hurrell, J. G. R., Ed., Monoclonal Hybridoma Antibodies: Techniques and Applications, CRC Press, Inc., Boca Raton, FL, 1982.

As would be evident to one of ordinary skill in the art, polyclonal antibodies can be generated from inoculating a variety of warm-blooded animals, such as horses, cows, goats, sheep, dogs, chickens, rabbits, mice, and rats with a zsig44 polypeptide or a fragment thereof.

The immunogenicity of a zsig44 polypeptide may be increased through the use of an adjuvant, such as alum (aluminum hydroxide) or Freund's complete or incomplete adjuvant. Polypeptides useful for immunization also include fusion polypeptides, such as fusions of zsig44 or a

portion thereof with an immunoglobulin polypeptide or with maltose binding protein. The polypeptide immunogen may be a full-length molecule or a portion thereof, for example, a peptide or soluble zsig44 protein. If the polypeptide  
5 portion is "hapten-like", such portion may be advantageously joined or linked to a macromolecular carrier (such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA) or tetanus toxoid) for immunization.

As used herein, the term "antibodies" includes  
10 polyclonal antibodies, affinity-purified polyclonal antibodies, monoclonal antibodies, and antigen-binding fragments, such as F(ab')<sub>2</sub> and Fab proteolytic fragments. Genetically engineered intact antibodies or fragments, such as chimeric antibodies, Fv fragments, single chain  
15 antibodies and the like, as well as synthetic antigen-binding peptides and polypeptides, are also included. Non-human antibodies may be humanized by grafting non-human CDRs onto human framework and constant regions, or by incorporating the entire non-human variable domains  
20 (optionally "cloaking" them with a human-like surface by replacement of exposed residues, wherein the result is a "veneered" antibody). In some instances, humanized antibodies may retain non-human residues within the human variable region framework domains to enhance proper binding  
25 characteristics. Through humanizing antibodies, biological half-life may be increased, and the potential for adverse immune reactions upon administration to humans is reduced.

Alternative techniques for generating or selecting antibodies useful herein include *in vitro*  
30 exposure of lymphocytes to zsig44 protein or peptide, and selection of antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled zsig44 polypeptide). Genes encoding polypeptides having potential zsig44 polypeptide binding domains can be

obtained by screening random peptide libraries displayed on phage (phage display) or on bacteria, such as *E. coli*. Nucleotide sequences encoding the polypeptides can be obtained in a number of ways, such as through random  
5 mutagenesis and random polynucleotide synthesis. In general, these random peptide display libraries can be used to screen for peptides which interact with a known target which can be a protein or polypeptide, such as a ligand or receptor, a biological or synthetic macromolecule, or  
10 organic or inorganic substances. Techniques for creating and screening such random peptide display libraries are known in the art (Ladner et al., US Patent NO. 5,223,409; Ladner et al., US Patent NO. 4,946,778; Ladner et al., US Patent NO. 5,403,484 and Ladner et al., US Patent NO.  
15 5,571,698) and random peptide display libraries and kits for screening such libraries are available commercially, for instance from Clontech (Palo Alto, CA), Invitrogen Inc. (San Diego, CA), New England Biolabs, Inc. (Beverly, MA) and Pharmacia LKB Biotechnology Inc. (Piscataway, NJ).  
20 Random peptide display libraries can be screened using the zsig44 sequences disclosed herein to identify proteins which bind to zsig44. These "binding proteins" which interact with zsig44 polypeptides can be used for tagging cells; for isolating homolog polypeptides by affinity  
25 purification; they can be directly or indirectly conjugated to drugs, toxins, radionuclides and the like. These binding proteins can also be used in analytical methods such as for screening expression libraries and neutralizing activity. The binding proteins can also be used for  
30 diagnostic assays for determining circulating levels of polypeptides; for detecting or quantitating soluble polypeptides as marker of underlying pathology or disease. These binding proteins can also act as zsig44 "antagonists" to block zsig44 binding, multimerization, or zsig44-

mediated cell-cell interactions, and signal transduction *in vitro* and *in vivo*.

Antibodies are determined to be specifically binding if: 1) they exhibit a threshold level of binding activity, and/or 2) they do not significantly cross-react with related polypeptide molecules. First, antibodies herein specifically bind if they bind to a zsig44 polypeptide, peptide or epitope with an affinity at least 10-fold greater than the binding affinity to control (non-zsig44) polypeptide. It is preferred that the antibodies exhibit a binding affinity ( $K_a$ ) of  $10^6 \text{ M}^{-1}$  or greater, preferably  $10^7 \text{ M}^{-1}$  or greater, more preferably  $10^8 \text{ M}^{-1}$  or greater, and most preferably  $10^9 \text{ M}^{-1}$  or greater. The binding affinity of an antibody can be readily determined by one of ordinary skill in the art, for example, by Scatchard analysis (Scatchard, G., Ann. NY Acad. Sci. 51: 660-672, 1949).

Second, antibodies are determined to specifically bind if they do not significantly cross-react with related polypeptides. Antibodies do not significantly cross-react with related polypeptide molecules, for example, if they detect zsig44 but not known related polypeptides using a standard Western blot analysis (Ausubel et al., ibid.). Examples of known related polypeptides are orthologs, e.g., CHIF (SEQ ID NO:3); paralogs, such as other known human ion channels, e.g., MAT-8 (SEQ ID NO:4); mutant zsig44 polypeptides; and other related ion channels or their regulators. Moreover, antibodies may be "screened against" known related polypeptides to isolate a population that specifically binds to the inventive polypeptides. For example, antibodies raised to zsig44 are adsorbed to related polypeptides adhered to insoluble matrix; antibodies specific to zsig44 will flow through the matrix under the proper buffer conditions. Such screening allows

isolation of polyclonal and monoclonal antibodies non-crossreactive to closely related polypeptides (Antibodies: A Laboratory Manual, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988; Current Protocols in Immunology, Cooligan, et al. (eds.), National Institutes of Health, John Wiley and Sons, Inc., 1995). Screening and isolation of specific antibodies is well known in the art. See, Fundamental Immunology, Paul (eds.), Raven Press, 1993; Getzoff et al., Adv. in Immunol. 43: 1-98, 1988; 5 Monoclonal Antibodies: Principles and Practice, Goding, J.W. (eds.), Academic Press Ltd., 1996; Benjamin et al., Ann. Rev. Immunol. 2: 67-101, 1984. 10

A variety of assays known to those skilled in the art can be utilized to detect antibodies which specifically bind to zsig44 proteins or peptides. Exemplary assays are described in detail in Antibodies: A Laboratory Manual, Harlow and Lane (Eds.), Cold Spring Harbor Laboratory Press, 1988. Representative examples of such assays include: concurrent immunoelectrophoresis, 15 radioimmunoassay, radioimmuno-precipitation, enzyme-linked immunosorbent assay (ELISA), dot blot or Western blot assay, inhibition or competition assay, and sandwich assay. In addition, antibodies can be screened for binding to wild-type versus mutant zsig44 protein or polypeptide. 20

Antibodies to zsig44 may be used for tagging cells that express zsig44; for isolating zsig44 by affinity purification; for diagnostic assays for determining circulating levels of zsig44 polypeptides; for detecting or quantitating zsig44 polynucleotide or polypeptide as marker 25 of underlying pathology or disease; in analytical methods employing FACS; for screening expression libraries; for generating anti-idiotypic antibodies; and as neutralizing antibodies or as antagonists to block zsig44 activity *in vitro* and *in vivo*. Suitable direct tags or labels include 30

radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent markers, chemiluminescent markers, magnetic particles and the like; indirect tags or labels may feature use of biotin-avidin or other complement/anti-complement pairs as intermediates. Antibodies herein may also be directly or indirectly conjugated to drugs, toxins, radionuclides and the like, and these conjugates used for *in vivo* diagnostic or therapeutic applications. Moreover, antibodies to zsig44 or fragments thereof may be used *in vitro* to detect denatured zsig44 or fragments thereof in assays, for example, Western Blots or other assays known in the art.

Polynucleotides of the present invention are also used to detect abnormalities on human chromosome 10 associated with disease or other human traits. The polynucleotides of the present invention map to the 10q11.1 region on human chromosome 10. The zsig44 gene maps 308.19 cR\_3000 from the top of the human chromosome 10 linkage group on the WICGR radiation hybrid map. Proximal and distal framework markers were NIB353 and AFMB032YH1 respectively. The use of surrounding markers positions zsig44 in the 10q11.1 region on the integrated LDB chromosome 10 map.

The present invention also provides reagents which will find use in diagnostic applications. For example, the zsig44 gene, a probe comprising zsig44 DNA or RNA or a sub-sequence thereof can be used to determine if the zsig44 gene is present or absent on chromosome 10 or if a mutation has occurred. Detectable chromosomal aberrations at the zsig44 gene locus include but are not limited to aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements. Such aberrations can be detected using

polynucleotides of the present invention by employing molecular genetic techniques, such as restriction fragment length polymorphism (RFLP) analysis, short tandem repeat (STR) analysis employing PCR techniques, and other genetic linkage analysis techniques known in the art (Sambrook et al., ibid.; Ausubel, et. al., ibid.; Marian, A.J., Chest, 108: 255-265, 1995).

The molecules of the present invention will be useful in diagnosing genetic chromosomal abnormalities. The polypeptides, nucleic acid and/or antibodies of the present invention may be used in treatment of disorders associated with diabetes, bone diseases or leukemia. The molecules of the present invention may used to modulate other ion channels or to treat or prevent development of pathological conditions in such diverse tissues as kidney, bone marrow, and heart. In particular, certain genetic syndromes and other human diseases may be amenable to such diagnosis, treatment or prevention.

Transgenic mice, engineered to express the zsig44 gene, and mice that exhibit a complete absence of zsig44 gene function, referred to as "knockout mice" (Snouwaert et al., Science 257:1083, 1992), may also be generated (Lowell et al., Nature 366:740-42, 1993). These mice may be employed to study the zsig44 gene and the protein encoded thereby in an *in vivo* system.

Polynucleotides encoding zsig44 polypeptides are useful within gene therapy applications where it is desired to increase or inhibit zsig44 activity. If a mammal has a mutated or lacks a zsig44 gene, the zsig44 gene can be introduced into the cells of the mammal. In one embodiment, a gene encoding a zsig44 polypeptide is introduced *in vivo* in a viral vector. Such vectors include an attenuated or defective DNA virus, such as but not limited to herpes simplex virus (HSV), papillomavirus, Epstein Barr virus

(EBV), adenovirus, adeno-associated virus (AAV), and the like. Defective viruses, which entirely or almost entirely lack viral genes, are preferred. A defective virus does not produce viable infective viruses and cannot re-infect after introduction into a cell. Use of defective viral vectors allows for administration to cells in a specific, localized area, without concern that the vector can infect other cells. Examples of particular vectors include, but are not limited to, a defective herpes simplex virus 1 (HSV1) vector (Kaplitt et al., Molec. Cell. Neurosci., 2:320-330, 1991), an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet et al., J. Clin. Invest., 90:626-630 (1992), and a defective adeno-associated virus vector (Samulski et al., J. Virol., 61:3096-3101, 1987; Samulski et al., J. Virol., 63:3822-3828, 1989).

In another embodiment, a zsig44 gene can be introduced in a retroviral vector, e.g., as described in Anderson et al., U.S. Patent No. 5,399,346; Mann et al., Cell, 33:153, 1983; Temin et al., U.S. Patent No. 4,650,764; Temin et al., U.S. Patent No. 4,980,289; Markowitz et al., J. Virol., 62:1120, 1988; Temin et al., U.S. Patent No. 5,124,263; International Patent Publication No. WO 95/07358, published March 16, 1995 by Dougherty et al.; and Kuo et al., Blood, 82:845, 1993.

Alternatively, the vector can be introduced by lipofection *in vivo* using liposomes. Synthetic cationic lipids can be used to prepare liposomes for *in vivo* transfection of a gene encoding a marker (Felgner et al., Proc. Natl. Acad. Sci. USA, 84:7413-7417, 1987; see Mackey et al., Proc. Natl. Acad. Sci. USA, 85:8027-8031, 1988). The use of lipofection to introduce exogenous genes into specific organs *in vivo* has certain practical advantages. Molecular targeting of liposomes to specific cells

represents one area of benefit. More particularly, directing transfection to particular cells represents one area of benefit. For instance, directing transfection to particular cell types would be particularly advantageous in a tissue with cellular heterogeneity, such as the pancreas, liver, kidney, and brain. Lipids may be chemically coupled to other molecules for the purpose of targeting. Targeted peptides, e.g., hormones or neurotransmitters, and proteins such as antibodies, or non-peptide molecules can be coupled to liposomes chemically.

It is possible to remove the target cells from the body; to introduce the vector as a naked DNA plasmid; and then to re-implant the transformed cells into the body. Naked DNA vectors for gene therapy can be introduced into the desired host cells by methods known in the art, e.g., transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun or use of a DNA vector transporter. See, e.g., Wu et al., J. Biol. Chem., 267:963-967, 1992; Wu et al., J. Biol. Chem., 263:14621-14624, 1988; and Johnston and Tang, Methods in Cell Biology, 43: 353-65 (1994).

Another aspect of the present invention involves antisense polynucleotide compositions that are complementary to a segment of the polynucleotides set forth in SEQ ID NO:1. Such synthetic antisense oligonucleotides are designed to bind to mRNA encoding zsig44 polypeptides and inhibit zsig44 gene transcription and translation of such mRNA. Such antisense oligonucleotides are useful to inhibit expression of zsig44 polypeptide-encoding genes in cell culture or in a subject.

Molecules of the present invention can be used to identify and isolate zsig44 associated or bound ion channel proteins. For example, proteins and peptides of the

present invention can be immobilized on a column and membrane preparations run over the column (Immobilized Affinity Ligand Techniques, Hermanson et al., eds., Academic Press, San Diego, CA, 1992, pp. 195-202).

5 Proteins and peptides can also be radiolabeled (Methods in Enzymol., vol. 182, "Guide to Protein Purification", M. Deutscher, ed., Acad. Press, San Diego, 1990, 721-737) or photoaffinity labeled (Brunner et al., Ann. Rev. Biochem. 62:483-514, 1993 and Fedan et al., Biochem. Pharmacol. 10 33:1167-1180, 1984) and specific cell-surface proteins can be identified.

The invention is further illustrated by the following non-limiting examples.

15

## EXAMPLES

### Example 1

20

#### Identification of zsig44

##### Using an EST Sequence to Obtain full-length zsig44

Scanning of a translated DNA database using a signal trap as a query resulted in identification of an expressed sequence tag (EST) sequence found to be  
25 homologous to a human secretory signal sequence. Oligonucleotide primers ZC 13,652 (SEQ ID NO:12) and ZC 13,653 (SEQ ID NO:13) were designed from the sequence of the identified EST. The primers were used for priming internally within the EST.

30

##### Identification of tissue source:

To identify tissue sources in which the full-length mRNA containing the EST was expressed, oligos specific for the EST polynucleotide sequence were used to

amplify cDNA from several human tissues (fetal brain, bone marrow, HUVEC, fetal lung, lymph node, pancreas, small intestine, stomach). Approximately 1 ng of Marathon™ ready cDNA (Clontech Laboratories, Inc., Palo Alto, CA) from various tissue sources was used as a template in polymerase chain reaction (PCR) using the oligos designed from the EST.

The conditions used for PCR were 1 cycle at 94°C for 5 minutes; 30 cycles at 94°C for 30 seconds, then 68°C for 4 minutes; 4°C soak. PCR product was analyzed on a 1.5% agarose gel and the expected 200 bp PCR product was seen only in fetal brain. This tissue source, fetal brain, was identified as having a high probability of containing a cDNA for the EST sequence. Other cDNAs tested could not be amplified with the oligonucleotide primers. Sequence analysis confirmed that the fetal brain PCR product sequence was that of the EST.

Isolation of full length zsig44 cDNA:

To obtain a full-length cDNA, 3' RACE was employed. A 3' RACE product was generated using a "marathon ready" human fetal brain cDNA as template and AP-1 (Clontech) and oligonucleotide ZC 13,653 (SEQ ID NO:13) as primers. This first-round 3' RACE PCR reaction was run as follows: 1 cycle at 94°C for 5 minutes; 30 cycles at 94°C for 30 seconds, then 68°C for 4 minutes; 4°C soak. An aliquot of 3' RACE PCR product was removed and analyzed on a 1.5% agarose gel. Multiple bands were seen on the gel.

The remaining DNA was diluted 1:100. A second-round nested 3' RACE PCR reaction was run to amplify template cDNA sequence. This PCR reaction used AP-2 (Clontech) and oligonucleotide ZC 13,611 (SEQ ID NO:14), which was designed to anneal to sequence internal of ZC 13,653 (SEQ ID NO:13). This nested PCR reaction was run as

per the first-round 3' RACE reaction disclosed above. The resulting DNA products were electrophoresed on a 1.5% agarose gel and a prominent band at approximately 453 bp, was seen. The DNA band was gel purified and sequenced. Sequence analyses revealed that the DNA products included the EST DNA sequence. Moreover, this 453 bp sequence generated from the 3' RACE product was a full-length cDNA encoding zsig44 protein.

## Example 2

### Tissue Distribution

Northern blot analysis was performed using Human Multiple Tissue Blots (MTN I, MTN II, and MTN III) from Clontech (Palo Alto, CA). The 453 bp PCR product described in Example 1 was electrophoresed on a 1% agarose gel, the fragment was electroeluted, and then radioactively labeled with  $^{32}\text{P}$ -dCTP using Prime-It II, a random prime labeling system (Stratagene Cloning Systems, La Jolla, CA), according to the manufacturer's specifications. The probe was then purified using Chroma Spin + TE-30 LC columns (Clontech) according to the manufacturer's instructions. ExpressHyb<sup>TM</sup> (Clontech) solution was used for prehybridization and as a hybridizing solution for the Northern blots. Hybridization took place overnight at 65°C using  $5 \times 10^6$  cpm/ml of labeled probe. The blots were then washed in 2X SSC/1% SDS at 65 °C, followed by a wash in 0.1X SSC/0.1% SDS at 55°C. One transcript size was detected at approximately 1.0 kb. Signal intensity was highest for kidney and bone marrow. No signals at 1.0 kb were present in any other tissues represented on the blots. A 2.4 kb transcript was noted only in spinal cord, and may be a splice variant of zsig44. Using the same probe and conditions for the Northern blots, dot blot analysis of

mRNAs from various tissues was performed. A strong signal was noted in heart tissue.

Northern blot analysis was also performed using human Tumor panel Blots I-VI (Invitrogen, San Diego, CA).  
5 These blots were probed using the 453 bp probe as disclosed above. No transcript was detected in any of the tumor samples.

### Example 3

#### 10 PCR-Based Chromosomal Mapping of the zsig44 Gene

Zsig44 was mapped to chromosome 10 using the commercially available GeneBridge 4 Radiation Hybrid Panel (Research Genetics, Inc., Huntsville, AL). The GeneBridge 4  
15 Radiation Hybrid Panel contains DNAs from each of 93 radiation hybrid clones, plus two control DNAs (the HFL donor and the A23 recipient). A publicly available WWW server (<http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>) allows mapping relative to the  
20 Whitehead Institute/MIT Center for Genome Research's radiation hybrid map of the human genome (the "WICGR" radiation hybrid map) which was constructed with the GeneBridge 4 Radiation Hybrid Panel.

25 For the mapping of zsig44 with the GeneBridge 4 RH Panel, 20  $\mu$ l PCR reactions were set up in a 96-well microtiter plate (Stratagene, La Jolla, CA) and used in a RoboCycler Gradient 96 thermal cycler (Stratagene). Each of the 95 PCR reactions consisted of 2  $\mu$ l 10X KlenTaq PCR  
30 reaction buffer (Clontech) 1.6  $\mu$ l dNTPs mix (2.5 mM each, PERKIN-ELMER, Foster City, CA), 1  $\mu$ l sense primer, ZC 14,842 (SEQ ID NO:15), 1  $\mu$ l antisense primer, ZC 14,838 (SEQ ID NO:16), 2  $\mu$ l RediLoad (Research Genetics, Inc.), 0.4  $\mu$ l 50X Advantage KlenTaq Polymerase Mix (Clontech), 25

ng of DNA from an individual hybrid clone or control and ddH<sub>2</sub>O for a total volume of 20  $\mu$ l. The reactions were overlaid with an equal amount of mineral oil and sealed. The PCR cycler conditions were as follows: an initial 1 cycle 5 minute denaturation at 95°C; 35 cycles of a 1 minute denaturation at 95°C, 1 minute annealing at 62°C and 1.5 minute extension at 72°C; followed by a final 1 cycle extension of 7 minutes at 72°C. The reactions were separated by electrophoresis on a 2% agarose gel (Life Technologies, Gaithersburg, MD).

The results showed that zsig44 maps 308.49 cR\_3000 from the top of the human chromosome 10 linkage group on the WICGR radiation hybrid map. Proximal and distal framework markers were NIB353 and AFMB032YH1, respectively. The use of surrounding markers positions zsig44 in the 10q11.1 region on the integrated LDB chromosome 10 map (The Genetic Location Database, University of Southampton, WWW server: [http://cedar.genetics.soton.ac.uk/public\\_html/](http://cedar.genetics.soton.ac.uk/public_html/)).

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

## WHAT IS CLAIMED IS:

1. An isolated polynucleotide encoding a zsig44 polypeptide comprising a sequence of amino acid residues that is at least 90% identical to an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as shown in SEQ ID NO:2 from residue number 17 (Leu), to residue number 89 (Cys); and

(b) the amino acid sequence as shown in SEQ ID NO:2 from amino acid number 1 (Met) to amino acid number 89 (Cys).

2. An isolated polynucleotide according to claim 1, wherein the polynucleotide is selected from the group consisting of:

(a) a polynucleotide sequence as shown in SEQ ID NO:1 from nucleotide 52 to nucleotide 270; and

(b) a polynucleotide sequence as shown in SEQ ID NO:1 from nucleotide 4 to nucleotide 270.

3. An isolated polynucleotide sequence according to claim 1, wherein the polynucleotide comprises nucleotide 1 to nucleotide 267 of SEQ ID NO:8.

4. An isolated polynucleotide according to claim 1, wherein the zsig44 polypeptide consists essentially of a sequence of amino acid residues that is at least 90% identical to an amino acid sequence as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys).

5. An isolated polynucleotide according to claim 4, wherein the zsig44 polypeptide consists essentially of a sequence of amino acid residues as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys).

6. An expression vector comprising the following operably linked elements:

a transcription promoter;

a DNA segment encoding a zsig44 polypeptide that is at least 90% identical to an amino acid sequence as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys); and

a transcription terminator.

7. An expression vector according to claim 6, further comprising a secretory signal sequence operably linked to the DNA segment.

8. A cultured cell into which has been introduced an expression vector according to claim 6, wherein the cell expresses a polypeptide encoded by the DNA segment.

9. A DNA construct encoding a fusion protein, the DNA construct comprising:

a first DNA segment encoding a polypeptide that is at least 90% identical to a sequence of amino acid residues selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2 from residue number 1 (Met), to residue number 16 (Ala);

(b) the amino acid sequence of SEQ ID NO:2 from residue number 17 (Leu), to residue number 28 (Asp);

(c) the amino acid sequence of SEQ ID NO:2 from residue number 29 (Pro), to residue number 64 (Lys);

(d) the amino acid sequence of SEQ ID NO:2 from residue number 65 (Ser), to residue number 89 (Cys);

(e) the amino acid sequence of SEQ ID NO:2 from residue number 17 (Leu), to residue number 64 (Lys);

(f) the amino acid sequence of SEQ ID NO:2 from residue number 29 (Pro), to residue number 89 (Cys); and

(g) the amino acid sequence of SEQ ID NO:2 from residue number 17 (Leu), to residue number 89 (Cys); and

at least one other DNA segment encoding an additional polypeptide,

wherein the first and other DNA segments are connected in-frame; and  
encode the fusion protein.

10. A fusion protein produced by a method comprising:

culturing a host cell into which has been introduced a vector comprising the following operably linked elements:

- (a) a transcriptional promoter;
- (b) a DNA construct encoding a fusion protein according to claim 9; and
- (c) a transcriptional terminator; and  
recovering the protein encoded by the DNA segment.

11. An isolated polypeptide comprising a sequence of amino acid residues that is at least 90% identical to an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence as shown in SEQ ID NO:2 from residue number 17 (Leu), to residue number 89 (Cys); and
- (b) the amino acid sequence as shown in SEQ ID NO:2 from amino acid number 1 (Met) to amino acid number 89 (Cys).

12. An isolated polypeptide according to claim 11, wherein the polypeptide further contains motifs 1 through 4 and the PLITPGSA motif spaced apart from N-terminus to C-terminus in a configuration M1-{6}-M2-{5}-M3-{1}-M4-{14}-PLITPGSA.

13. An isolated polypeptide according to claim 11, wherein the polypeptide consists essentially of a sequence of amino acid residues that is at least 90% identical to an amino acid sequence as shown in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys).

14. An isolated polypeptide according to claim 13, wherein the sequence of amino acid residues is as shown in SEQ

ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys).

15. A method of producing a zsig44 polypeptide comprising:

culturing a cell according to claim 8; and  
isolating the zsig44 polypeptide produced by the cell.

16. A method of producing an antibody to zsig44 polypeptide comprising:

inoculating an animal with a polypeptide selected from the group consisting of:

(a) a polypeptide consisting of 9 to 89 amino acids, wherein the polypeptide is at least 90% identical to a contiguous sequence of amino acids in SEQ ID NO:2 from amino acid number 17 (Leu) to amino acid number 89 (Cys);

(b) a polypeptide according to claim 11;

(c) a polypeptide having an amino acid sequence that is at least 90% identical to residue number 17 (Leu), to residue number 28 (Asp) of SEQ ID NO:2;

(d) a polypeptide having an amino acid sequence that is at least 90% identical to residue number 29 (Pro), to residue number 64 (Lys) of SEQ ID NO:2;

(e) a polypeptide having an amino acid sequence that is at least 90% identical to residue number 65 (Ser), to residue number 89 (Cys) of SEQ ID NO:2; and

wherein the polypeptide elicits an immune response in the animal; and

isolating the antibody from the animal.

17. An antibody produced by the method of claim 16, which binds to a zsig44 polypeptide.

18. The antibody of claim 16, wherein the antibody is a monoclonal antibody.

19. An antibody which specifically binds to a polypeptide of claim 11.

20. A method of detecting, in a test sample, the presence of a modulator of zsig44 protein activity, comprising:

culturing a cell into which has been introduced an expression vector according to claim 6, wherein the cell expresses the zsig44 protein encoded by the DNA segment in the presence and absence of a test sample; and

comparing levels of activity of zsig44 in the presence and absence of a test sample, by a biological or biochemical assay; and

determining from the comparison, the presence of modulator of zsig44 activity in the test sample.

1/1

TMTMTMTMTMTMTM  
zsig44 -MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLA ; 49  
RATCHI MEGITCAFLLVLAGLPVLEANGP-VDKGSPFYYDWESLQLGGMIFGGLLC ; 49  
HSMAT8 MQKVTLGLLVFLAGFPVLDAND-LEDKNSPFYYDWHSLOVGGGLICAGVLC ; 49  
HSU722 MAPLHHILVFCVGLLTMAKAES--PKEHDPFTYDYQSLQIGGLVIAGILF ; 48  
111111 222222

TMTMTMTM  
zsig44 IAGIAAVLSGKCKCKSSQKQHSPVPEKAIPLITPGSATTC---- ; 89  
RATCHI IAGIAMALSGKCKCRRNHT-PSSLPEKVTPLITPGSAST---- ; 87  
HSMAT8 AMGIIIVMSAKCKCKFGQK-SGHPGETPPLITPGSAQS----- ; 87  
HSU722 ILGILIVLSRRRCRCKFNQQQRTGEPDEEEGTFRSSIRRLSTRRR ; 92  
333333 444444

Figure

## SEQUENCE LISTING

<110> ZymoGenetics, Inc.

<120> A NOVEL HUMAN ION CHANNEL ZSIG44

<130> 97-45PC

<150> 60/053,715

<151> 1997-07-25

<160> 16

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 451

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (4)...(270)

<400> 1

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| gac atg gag aga gtg acc ctg gcc ctt ctc cta ctg gca ggc ctg act | 48  |
| Met Glu Arg Val Thr Leu Ala Leu Leu Leu Leu Ala Gly Leu Thr     |     |
| 1 5 10 15                                                       |     |
| gcc ttg gaa gcc aat gac cca ttt gcc aat aaa gac gat ccc ttc tac | 96  |
| Ala Leu Glu Ala Asn Asp Pro Phe Ala Asn Lys Asp Asp Pro Phe Tyr |     |
| 20 25 30                                                        |     |
| tat gac tgg aaa aac ctg cag ctg agc gga ctg atc tgc gga ggg ctc | 144 |
| Tyr Asp Trp Lys Asn Leu Gln Leu Ser Gly Leu Ile Cys Gly Gly Leu |     |
| 35 40 45                                                        |     |
| ctg gcc att gct ggg atc gcg gca gtt ctg agt ggc aaa tgc aaa tgc | 192 |
| Leu Ala Ile Ala Gly Ile Ala Ala Val Leu Ser Gly Lys Cys Lys Cys |     |
| 50 55 60                                                        |     |
| aag agc agc cag aag cag cac agt cct gta cct gag aag gcc atc cca | 240 |
| Lys Ser Ser Gln Lys Gln His Ser Pro Val Pro Glu Lys Ala Ile Pro |     |
| 65 70 75                                                        |     |

ctc atc act cca ggc tct gcc act act tgc tgagcacagg actggcctcc 290  
 Leu Ile Thr Pro Gly Ser Ala Thr Thr Cys  
 80 85

agggatggcc tgaagcctaa cactggcccc cagcacctcc tcccctggga ggccttatcc 350  
 tcaaggaagg acttctctcc aagggcaggc tgtaggcccc ctttctgatc aggaggcttc 410  
 tttatgaatt aaactcgccc caccaccccc tcaaaaaaaaa a 451

<210> 2

<211> 89

<212> PRT

<213> Homo sapiens

<400> 2

Met Glu Arg Val Thr Leu Ala Leu Leu Leu Leu Ala Gly Leu Thr Ala  
 1 5 10 15  
 Leu Glu Ala Asn Asp Pro Phe Ala Asn Lys Asp Asp Pro Phe Tyr Tyr  
 20 25 30  
 Asp Trp Lys Asn Leu Gln Leu Ser Gly Leu Ile Cys Gly Gly Leu Leu  
 35 40 45  
 Ala Ile Ala Gly Ile Ala Ala Val Leu Ser Gly Lys Cys Lys Cys Lys  
 50 55 60  
 Ser Ser Gln Lys Gln His Ser Pro Val Pro Glu Lys Ala Ile Pro Leu  
 65 70 75 80  
 Ile Thr Pro Gly Ser Ala Thr Thr Cys  
 85

<210> 3

<211> 87

<212> PRT

<213> Rattus norvegicus

<400> 3

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

Thr Pro Gly Ser Ala Ser Thr  
85

<210> 4

<211> 87

<212> PRT

<213> Homo sapiens

<400> 4

Met Gln Lys Val Thr Leu Gly Leu Leu Val Phe Leu Ala Gly Phe Pro  
1 5 10 15  
Val Leu Asp Ala Asn Asp Leu Glu Asp Lys Asn Ser Pro Phe Tyr Tyr  
20 25 30  
Asp Trp His Ser Leu Gln Val Gly Gly Leu Ile Cys Ala Gly Val Leu  
35 40 45  
Cys Ala Met Gly Ile Ile Ile Val Met Ser Ala Lys Cys Lys Cys Lys  
50 55 60  
Phe Gly Gln Lys Ser Gly His His Pro Gly Glu Thr Pro Pro Leu Ile  
65 70 75 80  
Thr Pro Gly Ser Ala Gln Ser  
85

<210> 5

<211> 92

<212> PRT

<213> Homo Sapiens

<400> 5

Met Ala Pro Leu His His Ile Leu Val Phe Cys Val Gly Leu Leu Thr  
1 5 10 15  
Met Ala Lys Ala Glu Ser Pro Lys Glu His Asp Pro Phe Thr Tyr Asp  
20 25 30  
Tyr Gln Ser Leu Gln Ile Gly Gly Leu Val Ile Ala Gly Ile Leu Phe  
35 40 45  
Ile Leu Gly Ile Leu Ile Val Leu Ser Arg Arg Cys Arg Cys Lys Phe  
50 55 60  
Asn Gln Gln Gln Arg Thr Gly Glu Pro Asp Glu Glu Gly Thr Phe  
65 70 75 80  
Arg Ser Ser Ile Arg Arg Leu Ser Thr Arg Arg Arg  
85 90

<210> 6

<211> 6

<212> PRT

<213> Homo sapiens

&lt;400&gt; 6

Pro Phe Tyr Tyr Asp Trp  
1 5

&lt;210&gt; 7

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 7

Gly Leu Ile Cys Gly Gly  
1 5

&lt;210&gt; 8

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 8

Gly Ile Ala Ala Val Leu  
1 5

&lt;210&gt; 9

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 9

Gly Lys Cys Lys Cys Lys  
1 5

&lt;210&gt; 10

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

Pro Leu Ile Thr Pro Gly Ser Ala  
1 5

&lt;210&gt; 11

&lt;211&gt; 267

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; degenerate sequence derived from zsig44 amino acid sequence

&lt;400&gt; 11

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atggarmgng tnacnytngc nytnytnytn ytngcnggny tnacngcnyt ngargcnaay | 60  |
| gayccnttyg cnaayaarga ygayccntty taytaygayt ggaaraayt ncarytnwsn  | 120 |
| ggnytnatht gyggnggnyt nytngcnath gcnggnathg cngcngtnyt nwsnggnaar | 180 |
| tgyaartgya arwsnwsnca raarcarcay wsncngtnc cngaraargc nathccnytn  | 240 |
| athacnccng gnwsngcnac nacntgy                                     | 267 |

&lt;210&gt; 12

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; oligonucleotide primer: ZC13652

&lt;400&gt; 12

|                            |    |
|----------------------------|----|
| cagtcaggcc tgccagtagg agaa | 24 |
|----------------------------|----|

&lt;210&gt; 13

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; oligonucleotide primer: ZC13653

&lt;400&gt; 13

|                         |    |
|-------------------------|----|
| cgcaggacac tggtaggga gc | 22 |
|-------------------------|----|

&lt;210&gt; 14

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; oligonucleotide primer: ZC13611

&lt;400&gt; 14

|                            |    |
|----------------------------|----|
| gtgacatgga gagagtgacc ctgg | 24 |
|----------------------------|----|

<210> 15  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide primer: ZC14842

<400> 15  
ctgccactac ttgctgag 18

<210> 16  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide primer: ZC14838

<400> 16  
gcctgccctt ggagagaa 18

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 98/15493

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/62 C07K14/705 C07K16/28 C12N5/10  
G01N33/68

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

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K

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

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

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages                                                                         | Relevant to claim No. |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Y          | WO 96 05322 A (GEN HOSPITAL CORP ; HARVARD COLLEGE (US)) 22 February 1996<br>see page 1 - page 8; figure 3<br>see figures SEQ.1,2<br>---                   | 1-11,<br>13-20        |
| Y          | US 5 487 976 A (SODERLUND DAVID M ET AL)<br>30 January 1996<br>see abstract; example 5<br>---                                                              | 1-11,<br>13-20        |
| A          | WO 95 25796 A (UNIV IOWA RES FOUND ; WELSH MICHAEL J (US); SHEPPARD DAVID N (US))<br>28 September 1995<br>see page 1 - page 6; figure SEQ.1<br>---<br>-/-- | 1-15                  |

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

6 November 1998

Date of mailing of the international search report

18/11/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Gurdjian, D

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/15493

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                               | Relevant to claim No. |
|----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| P,X      | <p>EMEST DATABASE<br/> Accession number AI003791<br/> 15-JUN-1998 (Rel. 56, Last updated,<br/> Version 1)<br/> TRANSMEMBRANE PROTEIN<br/> XP002083511<br/> see the whole document<br/> -----</p> | 1-9                   |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/15493

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|-------------------------------------------|---------------------|----------------------------|---------------------|
| WO 9605322 A                              | 22-02-1996          | US 5728579 A               | 17-03-1998          |
| US 5487976 A                              | 30-01-1996          | NONE                       |                     |
| WO 9525796 A                              | 28-09-1995          | US 5639661 A               | 17-06-1997          |
|                                           |                     | AU 2193595 A               | 09-10-1995          |
|                                           |                     | CA 2186122 A               | 28-09-1995          |
|                                           |                     | EP 0751994 A               | 08-01-1997          |