Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2416452 A1 2002/01/24

(21) 2 416 452

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2001/07/18
(87) Date publication PCT/PCT Publication Date: 2002/01/24
(85) Entrée phase nationale/National Entry: 2003/01/16
(86) N° demande PCT/PCT Application No.: US 2001/022637
(87) N° publication PCT/PCT Publication No.: 2002/006345
(30) Priorités/Priorities: 2000/07/18 (60/218,903) US;
2000/12/29 (60/259,072) US; 2001/01/08 (60/260,327) US;
2001/01/17 (60/262,158) US; 2001/01/25 (60/264,124) US;
2001/02/02 (60/266,108) US; 2001/02/02 (60/266,084) US;
2001/03/09 (60/274,858) US; 2001/03/26 (60/278,917) US

(51) Cl.Int.$^{7}$/Int.Cl.$^{7}$ C07K 14/705

(71) Demandeur/Applicant:
CURAGEN CORPORATION, US

(72) Inventeurs/Inventors:
ZERHUSEN, BRYAN, US;
PADIGARU, MURALIDHARA, US;
LI, LI, US;
BURGESS, CATHERINE E., US;
CASMAN, STACIE J., US;
SPYTEK, KIMBERLY A., US;
MISHRA, VISHNU, US;
...

(74) Agent: KVAS MILLER EVERITT

(54) Titre : NOUVELLES PROTEINES ET ACIDES NUCLEIQUES LES CODANT
(54) Title: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

(57) Abrégé/Abstract:
Disclosed herein are nucleic acid sequences that encode G-coupled protein-receptor related polypeptides. Also disclosed are polypeptides encoded by these nucleic acid sequences, and antibodies, which immunospeficically-bind to the polypeptide, as well as derivatives, variants, mutants, or fragments of the aforementioned polypeptide, polynucleotide, or antibody. The invention further discloses therapeutic, diagnostic and research methods for diagnosis, treatment, and prevention of disorders involving any one of these novel human acids and proteins.

(21) 2 416 452

(13) A1

(72) **Inventeurs(suite)/Inventors(continued)**: TAYLOR, SARAH, US; SHENOY, SURESH, US; VERNET, CORINE, US; GERLACH, VALERIE, US; ELLERMAN, KAREN, US; MACDOUGALL, JOHN R., US; STONE, DAVID, US; SMITHSON, GLENNDA, US; GROSSE, WILLIAM M., US; ALSOBROOK, JOHN P., II, US; LEPLEY, DENISE M., US; TCHERNEV, VELIZAR T., US; TAILLON, BRUCE, US

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 January 2002 (24.01.2002)

PCT

(10) International Publication Number
WO 02/06345 A2

(51) International Patent Classification[7]: C07K 14/705

(21) International Application Number: PCT/US01/22637

(22) International Filing Date: 18 July 2001 (18.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

| | | |
|---|---|---|
| 60/218,903 | 18 July 2000 (18.07.2000) | US |
| 60/259,072 | 29 December 2000 (29.12.2000) | US |
| 60/260,327 | 8 January 2001 (08.01.2001) | US |
| 60/262,158 | 17 January 2001 (17.01.2001) | US |
| 60/264,124 | 25 January 2001 (25.01.2001) | US |
| 60/266,084 | 2 February 2001 (02.02.2001) | US |
| 60/266,108 | 2 February 2001 (02.02.2001) | US |
| 60/274,858 | 9 March 2001 (09.03.2001) | US |
| 60/278,917 | 26 March 2001 (26.03.2001) | US |

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

| | |
|---|---|
| US | 60/218,903 (CIP) |
| Filed on | 18 July 2000 (18.07.2000) |
| US | 60/259,072 (CIP) |
| Filed on | 29 December 2000 (29.12.2000) |
| US | 60/260,327 (CIP) |
| Filed on | 8 January 2001 (08.01.2001) |
| US | 60/261,030 (CIP) |
| Filed on | 11 January 2001 (11.01.2001) |
| US | 60/262,158 (CIP) |
| Filed on | 17 January 2001 (17.01.2001) |
| US | 60/264,124 (CIP) |
| Filed on | 25 January 2001 (25.01.2001) |
| US | 60/266,084 (CIP) |
| Filed on | 2 February 2001 (02.02.2001) |
| US | 60/266,108 (CIP) |
| Filed on | 2 February 2001 (02.02.2001) |
| US | 60/274,858 (CIP) |
| Filed on | 9 March 2001 (09.03.2001) |
| US | 60/278,917 (CIP) |
| Filed on | 26 March 2001 (26.03.2001) |

(71) Applicant *(for all designated States except US)*: **CURAGEN CORPORATION** [US/US]; 555 Long Wharf Drive, 11th floor, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ZERHUSEN, Bryan** [US/US]; 45 Jefferson Road, Apartment #22-10, Branford, CT 06405 (US). **PADIGARU, Muralidhara** [IN/US]; 98 Montoya Circle, Branford, CT 06405 (US). **LI, Li** [CN/US]; 487 Oak Avenue #67, Cheshire, CT 06410 (US). **BURGESS, Catherine, E.** [US/US]; 90 Carriage Hill Drive, Wethersfield, CT 06109 (US). **CASMAN, Stacie, J.** [US/US]; 155 Knollwood Drive, Wallingsford, CT 06492 (US). **SPYTEK, Kimberly, A.** [US/US]; 28 Court Street #1, New Haven, CT 06511 (US). **MISHRA, Vishnu** [IN/US]; 309 East Main Street #306, Branford, CT 06405 (US). **TAYLOR, Sarah** [US/US]; 29 Winthop Road, Cheshire, CT 06437 (US). **SHENOY, Suresh** [IN/US]; 15 Millwood Drive, Branford, CT 06405 (US). **VERNET, Corine** [FR/US]; 1739 Foxon Road, Box L, North Branford, CT 06471 (US). **GERLACH, Valerie** [US/US]; 375 East Main Street, Branford, CT 06405 (US). **ELLERMAN, Karen** [US/US]; 87 Montoya Drive, Branford, CT 06405 (US). **MACDOUGALL, John, R.** [CA/US]; 117 Russell Street, Hamden, CT 06517 (US). **STONE, David** [US/US]; 223 Whitethorn Drive, Guilford, CT 06437 (US). **SMITHSON, Glennda** [US/US]; 81 Main Street #32A, Branford, CT 06504 (US). **GROSSE, William, M.** [US/US]; 15 Terrace Road, Apt. C, Branford, CT 06405 (US). **ALSOBROOK, John, P., II** [US/US]; 60 Lake Drive, Madison, CT 06443 (US). **LEPLEY, Denise, M.** [US/US]; 322 East Main Street, Branford, CT 06405 (US). **TCHERNEV, Velizar, T.** [BG/US]; 45 Jefferson Road #3-12, Branford, CT 06405 (US). **TAILLON, Bruce** [US/US]; 181 Elizabeth Lane, Middletown, CT 06457 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

*[Continued on next page]*

(54) Title: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

(57) Abstract: Disclosed herein are nucleic acid sequences that encode G-coupled protein-receptor related polypeptides. Also disclosed are polypeptides encoded by these nucleic acid sequences, and antibodies, which immunospeficically-bind to the polypeptide, as well as derivatives, variants, mutants, or fragments of the aforementioned polypeptide, polynucleotide, or antibody. The invention further discloses therapeutic, diagnostic and research methods for diagnosis, treatment, and prevention of disorders involving any one of these novel human acids and proteins.

WO 02/06345 A2

WO 02/06345 A2

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, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *without international search report and to be republished upon receipt of that report*

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

# NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

## FIELD OF THE INVENTION

The invention generally relates to novel GPCR1, GPCR2, GPCR3, GPCR4, GPCR5, GPCR6, GPCR7, GPCR8 and GPCR9 nucleic acids and polypeptides encoded therefrom. More specifically, the invention relates to nucleic acids encoding novel polypeptides, as well as vectors, host cells, antibodies, and recombinant methods for producing these nucleic acids and polypeptides.

## BACKGROUND OF THE INVENTION

The invention generally relates to nucleic acids and polypeptides. More particularly, the invention relates to nucleic acids encoding novel G-protein coupled receptor (GPCR) polypeptides, as well as vectors, host cells, antibodies, and recombinant methods for producing these nucleic acids and polypeptides.

## SUMMARY OF THE INVENTION

The invention is based in part upon the discovery of nucleic acid sequences encoding novel polypeptides. The novel nucleic acids and polypeptides are referred to herein as GPCRX, or GPCR1, GPCR2, GPCR3, GPCR4, GPCR5, GPCR6, GPCR7, GPCR8 and GPCR9 nucleic acids and polypeptides. These nucleic acids and polypeptides, as well as derivatives, homologs, analogs and fragments thereof, will hereinafter be collectively designated as "GPCRX" nucleic acid or polypeptide sequences.

In one aspect, the invention provides an isolated GPCRX nucleic acid molecule encoding a GPCRX polypeptide that includes a nucleic acid sequence that has identity to the nucleic acids disclosed in SEQ ID NOS1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37. In some embodiments, the GPCRX nucleic acid molecule will hybridize under stringent conditions to a nucleic acid sequence complementary to a nucleic acid molecule that includes a protein-coding sequence of a GPCRX nucleic acid sequence. The invention also includes an isolated nucleic acid that encodes a GPCRX polypeptide, or a fragment, homolog, analog or derivative thereof. For example, the nucleic acid can encode a polypeptide at least

80% identical to a polypeptide comprising the amino acid sequences of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37. The nucleic acid can be, for example, a genomic DNA fragment or a cDNA molecule that includes the nucleic acid sequence of any of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37.

Also included in the invention is an oligonucleotide, *e.g.*, an oligonucleotide which includes at least 6 contiguous nucleotides of a GPCRX nucleic acid (*e.g.*, SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37) or a complement of said oligonucleotide.

Also included in the invention are substantially purified GPCRX polypeptides (SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37). In certain embodiments, the GPCRX polypeptides include an amino acid sequence that is substantially identical to the amino acid sequence of a human GPCRX polypeptide.

The invention also features antibodies that immunoselectively bind to GPCRX polypeptides, or fragments, homologs, analogs or derivatives thereof.

In another aspect, the invention includes pharmaceutical compositions that include therapeutically- or prophylactically-effective amounts of a therapeutic and a pharmaceutically-acceptable carrier. The therapeutic can be, *e.g.*, a GPCRX nucleic acid, a GPCRX polypeptide, or an antibody specific for a GPCRX polypeptide. In a further aspect, the invention includes, in one or more containers, a therapeutically- or prophylactically-effective amount of this pharmaceutical composition.

In a further aspect, the invention includes a method of producing a polypeptide by culturing a cell that includes a GPCRX nucleic acid, under conditions allowing for expression of the GPCRX polypeptide encoded by the DNA. If desired, the GPCRX polypeptide can then be recovered.

In another aspect, the invention includes a method of detecting the presence of a GPCRX polypeptide in a sample. In the method, a sample is contacted with a compound that selectively binds to the polypeptide under conditions allowing for formation of a complex between the polypeptide and the compound. The complex is detected, if present, thereby identifying the GPCRX polypeptide within the sample.

The invention also includes methods to identify specific cell or tissue types based on their expression of a GPCRX.

Also included in the invention is a method of detecting the presence of a GPCRX nucleic acid molecule in a sample by contacting the sample with a GPCRX nucleic acid probe or primer, and detecting whether the nucleic acid probe or primer bound to a GPCRX nucleic acid molecule in the sample.

In a further aspect, the invention provides a method for modulating the activity of a GPCRX polypeptide by contacting a cell sample that includes the GPCRX polypeptide with a compound that binds to the GPCRX polypeptide in an amount sufficient to modulate the activity of said polypeptide. The compound can be, *e.g.*, a small molecule, such as a nucleic acid, peptide, polypeptide, peptidomimetic, carbohydrate, lipid or other organic (carbon containing) or inorganic molecule, as further described herein.

Also within the scope of the invention is the use of a therapeutic in the manufacture of a medicament for treating or preventing disorders or syndromes including, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, or other disorders related to cell signal processing and metabolic pathway modulation. The therapeutic can be, *e.g.*, a GPCRX nucleic acid, a GPCRX polypeptide, or a GPCRX-specific antibody, or biologically-active derivatives or fragments thereof.

For example, the compositions of the present invention will have efficacy for treatment of patients suffering from: developmental diseases, MHCII and III diseases (immune diseases), taste and scent detectability Disorders, Burkitt's lymphoma, corticoneurogenic disease, signal transduction pathway disorders, Retinal diseases including those involving photoreception, Cell growth rate disorders; cell shape disorders, feeding disorders; control of feeding; potential obesity due to over-eating; potential disorders due to starvation (lack of appetite), noninsulin-dependent diabetes mellitus (NIDDM1), bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer), anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, Crohn's disease; multiple sclerosis; Albright Hereditary Ostoeodystrophy, angina pectoris, myocardial infarction, ulcers, allergies, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation. Dentatorubro-pallidoluysian

atrophy (DRPLA) Hypophosphatemic rickets, autosomal dominant (2) Acrocallosal syndrome and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders of the like.

The polypeptides can be used as immunogens to produce antibodies specific for the invention, and as vaccines. They can also be used to screen for potential agonist and antagonist compounds. For example, a cDNA encoding GPCRX may be useful in gene therapy, and GPCRX may be useful when administered to a subject in need thereof. By way of nonlimiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to Neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer), anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, Crohn's disease; multiple sclerosis; and Treatment of Albright Hereditary Ostoeodystrophy, angina pectoris, myocardial infarction, ulcers, allergies, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders.

The invention further includes a method for screening for a modulator of disorders or syndromes including, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders or other disorders related to cell signal processing and metabolic pathway modulation. The method includes contacting a test compound with a GPCRX polypeptide and determining if the test compound binds to said

Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders or other disorders related to cell signal processing and metabolic pathway modulation by administering a test compound to a test animal at increased risk for the aforementioned disorders or syndromes. The test animal expresses a recombinant polypeptide encoded by a GPCRX nucleic acid. Expression or activity of GPCRX polypeptide is then measured in the test animal, as is expression or activity of the protein in a control animal which recombinantly-expresses GPCRX polypeptide and is not at increased risk for the disorder or syndrome. Next, the expression of GPCRX polypeptide in both the test animal and the control animal is compared. A change in the activity of GPCRX polypeptide in the test animal relative to the control animal indicates the test compound is a modulator of latency of the disorder or syndrome.

In yet another aspect, the invention includes a method for determining the presence of or predisposition to a disease associated with altered levels of a GPCRX polypeptide, a GPCRX nucleic acid, or both, in a subject (*e.g.*, a human subject). The method includes measuring the amount of the GPCRX polypeptide in a test sample from the subject and comparing the amount of the polypeptide in the test sample to the amount of the GPCRX polypeptide present in a control sample. An alteration in the level of the GPCRX polypeptide in the test sample as compared to the control sample indicates the presence of or predisposition to a disease in the subject. Preferably, the predisposition includes, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders. Also, the expression levels of the new polypeptides of the invention can be used in a method to screen for various cancers as well as to determine the stage of cancers.

In a further aspect, the invention includes a method of treating or preventing a pathological condition associated with a disorder in a mammal by administering to the subject a GPCRX polypeptide, a GPCRX nucleic acid, or a GPCRX-specific antibody to a subject (*e.g.*, a human subject), in an amount sufficient to alleviate or prevent the pathological condition. In preferred embodiments, the disorder, includes, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease,

anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders.

In yet another aspect, the invention can be used in a method to identity the cellular receptors and downstream effectors of the invention by any one of a number of techniques commonly employed in the art. These include but are not limited to the two-hybrid system, affinity purification, co-precipitation with antibodies or other specific-interacting molecules.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

## DETAILED DESCRIPTION OF THE INVENTION

The invention is based, in part, upon the discovery of novel nucleic acid sequences that encode novel polypeptides. The novel nucleic acids and their encoded polypeptides are referred to individually as GPCR1, GPCR2, GPCR3, GPCR4, GPCR5, GPCR6, GPCR7, GPCR8 and GPCR9. The nucleic acids, and their encoded polypeptides, are collectively designated herein as "GPCRX".

The novel GPCRX nucleic acids of the invention include the nucleic acids whose sequences are provided in Tables 1A, 2A, 2C, 3A, 4A, 4C, 4E, 5A, 5C, 6A, 6C, 6E, 6G, 7A, 7C, 8A, 8C, 8E, 8G and 9A, inclusive, or a fragment, derivative, analog or homolog thereof. The novel GPCRX proteins of the invention include the protein fragments whose sequences are provided in Tables 1B, 2B, 2D, 3B, 4B, 4D, 4F, 5B, 5D, 6B, 6D, 6F, 6H, 7B, 7D, 8B, 8D, 8F, 8H and 9B, inclusive. The individual GPCRX nucleic acids and proteins are described below. Within the scope of this invention is a method of using these nucleic acids and peptides in the treatment or prevention of a disorder related to cell signaling or metabolic pathway modulation.

The GPCRX proteins of the invention have a high homology to the 7tm_1 domain (PFam Acc. No. pfam00001). The 7tm_1 domain is from the 7 transmembrane receptor family, which includes a number of different proteins, including, for example, serotonin receptors, dopamine receptors, histamine receptors, andrenergic receptors, cannabinoid receptors, angiotensin II receptors, chemokine receptors, opioid receptors, G-protein coupled receptor (GPCR) proteins, olfactory receptors (OR), and the like. Some proteins and the Protein Data Base Ids/gene indexes include, for example: rhodopsin (129209); 5-hydroxytryptamine receptors; (112821, 8488960, 112805, 231454, 1168221, 398971, 112806); G protein-coupled receptors (119130, 543823, 1730143, 132206, 137159, 6136153, 416926, 1169881, 136882, 134079); gustatory receptors (544463, 462208); c-x-c chemokine receptors (416718, 128999, 416802, 548703, 1352335); opsins (129193, 129197, 129203); and olfactory receptor-like proteins (129091, 1171893, 400672, 548417).

Because of the close homology among the members of the GPCRX family, proteins that are homologous to any one member of the family are also largely homologous to the other members, except where the sequences are different as shown below.

The similarity information for the GPCRX proteins and nucleic acids disclosed herein suggest that GPCR1-GPCR9 may have important structural and/or physiological functions characteristic of the Olfactory Receptor family and the GPCR family. Therefore, the nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications and as a research tool. These include serving as a specific or selective nucleic acid or protein diagnostic and/or prognostic marker, wherein the presence or amount of the nucleic acid or the protein are to be assessed, as well as potential therapeutic applications such as the following: (i) a protein therapeutic, (ii) a small molecule drug target, (iii) an antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) a nucleic acid useful in gene therapy (gene delivery/gene ablation), and (v) a composition promoting tissue regeneration in vitro and in vivo (vi) biological defense weapon.

G-Protein Coupled Receptor proteins (GPCRs) have been identified as a large family of G protein-coupled receptors in a number of species. These receptors share a seven transmembrane domain structure with many neurotransmitter and hormone receptors, and are likely to underlie the recognition and G-protein-mediated transduction of various signals. Human GPCR generally do not contain introns and belong to four different gene subfamilies, displaying great sequence variability. These genes are dominantly expressed in olfactory epithelium. See, *e.g.*, Ben-Arie et al., *Hum. Mol. Genet*. 1994 3:229-235; and, Online

Mendelian Inheritance in Man (OMIM) entry # 164342 (http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?).

The olfactory receptor (OR) gene family constitutes one of the largest GPCR multigene families and is distributed among many chromosomal sites in the human genome. See Rouquier et al., *Hum. Mol. Genet.* 7(9):1337-45 (1998); Malnic et al., *Cell* 96:713-23 (1999). Olfactory receptors constitute the largest family among G protein-coupled receptors, with up to 1000 members expected. See Vanderhaeghen et al., *Genomics* 39(3):239-46 (1997); Xie et al., *Mamm. Genome* 11(12):1070-78 (2000); Issel-Tarver et al., *Proc. Natl. Acad. Sci. USA* 93(20):10897-902 (1996). The recognition of odorants by olfactory receptors is the first stage in odor discrimination. See Krautwurst et al., *Cell* 95(7):917-26 (1998); Buck et al., *Cell* 65(1):175-87 (1991). Many ORs share some characteristic sequence motifs and have a central variable region corresponding to a putative ligand binding site. See Issel-Tarver et al., *Proc. Natl. Acad. Sci. USA* 93:10897-902 (1996).

Other examples of seven membrane spanning proteins that are related to GPCRs are chemoreceptors. See Thomas et al., Gene 178(1-2):1-5 (1996). Chemoreceptors have been identified in taste, olfactory, and male reproductive tissues. See *id.*; Walensky et al., *J. Biol. Chem.* 273(16):9378-87 (1998); Parmentier et al., *Nature* 355(6359):453-55 (1992); Asai et al., *Biochem. Biophys. Res. Commun.* 221(2):240-47 (1996).

The GPCRX nucleic acids and proteins are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further below. For example, a cDNA encoding the GPCR (or olfactory-receptor) like protein may be useful in gene therapy, and the receptor -like protein may be useful when administered to a subject in need thereof. The nucleic acids and proteins of the invention are also useful in potential therapeutic applications used in the treatment of developmental diseases, MHCII and III diseases (immune diseases), taste and scent detectability disorders, Burkitt's lymphoma, corticoneurogenic disease, signal transduction pathway disorders, retinal diseases including those involving photoreception, cell growth rate disorders, cell shape disorders, feeding disorders, potential obesity due to over-eating, potential disorders due to starvation (lack of appetite), noninsulin-dependent diabetes mellitus (NIDDM1), bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer), anorexia, bulimia, asthma, allergies, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, Crohn's

disease, multiple sclerosis, Albright hereditary ostoeodystrophy, angina pectoris, myocardial infarction, ulcers, benign prostatic hypertrophy, psychotic and neurological disorders (including anxiety, schizophrenia, manic depression, delirium, dementia, and severe mental retardation), dentatorubro-pallidoluysian atrophy (DRPLA), hypophosphatemic rickets, autosomal dominant (2) acrocallosal syndrome and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders. Other GPCR-related diseases and disorders are contemplated.

The polypeptides can be used as immunogens to produce antibodies specific for the invention, and as vaccines. They can also be used to screen for potential agonist and antagonist compounds. For example, a cDNA encoding the GPCR-like protein may be useful in gene therapy, and the GPCR-like protein may be useful when administered to a subject in need thereof. By way of nonlimiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from developmental diseases, MHCII and III diseases (immune diseases), taste and scent detectability disorders, Burkitt's lymphoma, corticoneurogenic disease, signal transduction pathway disorders, retinal diseases including those involving photoreception, cell growth rate disorders, cell shape disorders, feeding disorders, potential obesity due to over-eating, potential disorders due to starvation (lack of appetite), noninsulin-dependent diabetes mellitus (NIDDM1), bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer), anorexia, bulimia, asthma, allergies, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, Crohn's disease, multiple sclerosis, Albright hereditary ostoeodystrophy, angina pectoris, myocardial infarction, ulcers, benign prostatic hypertrophy, psychotic and neurological disorders (including anxiety, schizophrenia, manic depression, delirium, dementia, and severe mental retardation), dentatorubro-pallidoluysian atrophy (DRPLA), hypophosphatemic rickets, autosomal dominant (2) acrocallosal syndrome and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders. The novel nucleic acid encoding GPCR-like protein, and the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods.

**GPCR1**

The disclosed novel GPCR1 (alternatively referred to herein as GMba64p14_A) includes the 968 nucleotide sequence (SEQ ID NO:1) shown in Table 1A. A GPCR1 ORF begins with a Kozak consensus ATG initiation codon at nucleotides 3-5 and ends with a TGA codon at nucleotides 951-953. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 1A, and the start and stop codons are in bold letters.

**Table 1A. GPCR1 Nucleotide Sequence (SEQ ID NO:1)**

GG**ATG**GGAAAACCAGGCAGAGTGAACCAAACCACTGTTTCAGACTTCCTCCTTCTAGGACTCTCTGA
GTGGCCAGAGGAGCAGCCTCTTCTGTTTGGCATCTTCCTTGGCATGTACCTGGTCACCATGGTGGGG
AACCTGCTCATTATCCTGGCCATCAGCTCTGACCCACACCTCCATACTCCCATGTACTTCTTTCTGG
CCAACCTGTCATTAACTGATGCCTGTTTCACTTCTGCCTCCATCCCCAAAATGCTGGCCAACATTCA
TACCCAGAGTCAGATCATCTCGTATTCTGGGTGTCTTGCACAGCTATATTTCCTCCTTATGTTTGGT
GGCCTTGACAACTGCCTGCTGGCTGTGATGGCATATGACCGCTATGTGGCCATCTGCCAACCACTCC
ATTACAGCACATCTATGAGTCCCCAGCTCTGTGCACTAATGCTGGGTGTGTGCTGGGTGCTAACCAA
CTGTCCTGCCCTGATGCACACACTGTTGCTGACCCGCGTGGCTTTCTGTGCCCAGAAAGCCATCCCT
CATTTCTATTGTGATCCTAGTGCTCTCCTGAAGCTTGCCTGCTCAGATACCCATGTAAACGAGCTGA
TGATCATCACCATGGGCTTGCTGTTCCTCACTGTTCCCCTCCTGCTGATCGTCTTCTCCTATGTCCG
CATTTTCTGGGCTGTGTTTGTCATCTCATCTCCTGGAGGGAGATGGAAGGCCTTCTCTACCTGTGGT
TCTCATCTCACGGTGGTTCTGCTCTTCTATGGGTCTCTTATGGGTGTGTATTTACTTCCTCCATCAA
CTTACTCTACAGAGAGGGAAAGTAGGGCTGCTGTTCTCTATATGGTGATTATTCCCACGCTAAACCC
ATTCATTTATAGCTTGAGGAACAGAGACATGAAGGAGGCTTTGGGTAAACTTTTTGTCAGTGGAAAA
ACATTCTTTTTA**TGA**TTAGACATCTAGACG

A GPCR-like protein of the invention, referred to herein as GPCR1, is an Olfactory Receptor ("OR")-like protein. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR1 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

The GPCR1 polypeptide (SEQ ID NO:2) encoded by SEQ ID NO:1 is 316 aa in length, has a molecular weight of 35183.4 Daltons, and is presented using the one-letter amino acid code in Table 1B. The Psort profile for GPCR1 predicts that these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR1 polypeptide is located to the Golgi body with a certainty of 0.400, the endoplasmic reticulum (membrane) with a certainty of 0.300, or a microbody (peroxisome) with a certainty of 0.300. The Signal P predicts a likely cleavage site for a GPCR1 peptide is between positions 54 and 55, *i.e.*, at the dash in the sequence ISS-DP.

**Table 1B. GPCR1 protein sequence (SEQ ID NO:2)**

MGKPGRVNQTTVSDFLLLGLSEWPEEQPLLFGIFLGMYLVTMVGNLLIILAISSDPHLHTPMYFFLA
NLSLTDACFTSASIPKMLANIHTQSQIISYSGCLAQLYFLLMFGGLDNCLLAVMAYDRYVAICQPLH
YSTSMSPQLCALMLGVCWVLTNCPALMHTLLLTRVAFCAQKAIPHFYCDPSALLKLACSDTHVNELM
IITMGLLFLTVPLLLIVFSYVRIFWAVFVISSPGGRWKAFSTCGSHLTVVLLFYGSLMGVYLLPPST
YSTERESRAAVLYMVIIPTLNPFIYSLRNRDMKEALGKLFVSGKTFFL

The amino acid sequence of GPCR1 had high homology to other proteins as shown in Table 1C.

**Table 1C. BLASTX results for GPCR1**

| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
|---|---|---|---|
| patp:AAR27868 Odorant receptor clone F5 - Rattus rattus, 313 aa | +3 | 952 | 6.7e-95 |
| patp:AAR27874 Odorant receptor clone I9 - Rattus rattus, 314 aa | +3 | 888 | 4.1e-88 |

Additional BLASTP results are shown in Table 1D.

**Table 1D. GPCR1 BLASTP results**

| Gene Index / Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |
|---|---|---|---|---|---|
| AB038167 | Q9JHE2 rattus norvegicus (rat). gustatory receptor 43. 3/2001 | 311 | 179/300 (60%) | 226/300, (75%) | 1e-104 |
| M64377; P23266 | OLF5_RAT rattus norvegicus (rat). olfactory receptor-like protein f5. 7/1993 | 313 | 178/300 (59%) | 228/300, (76%) | 1e-103 |
| U50947 | Q62942 rattus norvegicus (rat). taste bud receptor protein tb 334. 3/2001 | 311 | 179/300 (60%) | 225/300, (75%) | 1e-102 |
| AF101730 | Q9TUA9 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 176/308 (57%) | 224/308, (73%) | 1e-100 |
| AF101760; AF101761 | Q9TQX4 gorilla gorilla (gorilla). olfactory receptor. 3/2001 | 314 | 176/309 (57%) | 224/309, (72%) | 1e-100 |
| AF101741 | Q9TUA1 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 176/309 (57%) | 224/309, (72%) | 1e-100 |

A multiple sequence alignment is given in Table 1E, with the GPCR1 protein of the invention being shown on line 1, in a ClustalW analysis comparing GPCR1 with related protein sequences disclosed in Table 1D.

## Table 1E. Information for the ClustalW proteins:

1. SEQ ID NO:2, GPCR1
2. SEQ ID NO:41, AB038167 gustatory receptor 43
3. SEQ ID NO:42, M64377 olfactory receptor-like protein f5
4. SEQ ID NO:43, U50947 taste bud receptor protein tb 334
5. SEQ ID NO:44, AF101730 olfactory receptor
6. SEQ ID NO:45, AF101761 olfactory receptor
7. SEQ ID NO:46, AF101741 olfactory receptor

```
GPCR1          MGKPGRVNQTTVSDFLLLGLSEWPEEQPLLFGIFLGMYLVTMVGNLLIIL  50
41:AB038167    -----MENQSSVSEFFLQGISGFPEQQQLLYGLFLCMYLVTLTGNVLIIM  45
42:M64377      ---MSSTNQSSVTEFLLLGLSRQPQQQQLLFLLFLIMYLATVLGNLLIIL  47
43:U50947      -----MENQSSVSEFFLRGISGFPEQQQLLYGLFLCMYLVTLTGNVLIIL  45
44:AF101730    -----MENQSSVSEFFLQGISGFPEQQQLLYGLFLCMYLVTLTGNVLIIM  45
45:AF101761    ---MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIV  47
46:AF101741    ---MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIV  47

GPCR1          AISSDPHLHTPMYFFLANLSLTDACFTSASIPKMLANIHTQSQIISYSGC 100
41:AB038167    AIGSDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQTQRHTISYTGC  95
42:M64377      AIGTDSRLHTPMYFFLSNLSFVDVCFSSTTVPKVLANHILGSQAISFSGC  97
43:U50947      AIGSDPHLHTPMYFFLANLSFADMGLISSTVTKMLFNVQTQCHTISYTGC  95
44:AF101730    AIGSDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQTQRHTISYTGC  95
45:AF101761    LIRLDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADC  97
46:AF101741    LIRLDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADC  97

GPCR1          LAQLYFLLMFGGLDNCLLAVMAYDRYVAICQPLHYSTSMSPQLCALMLGV 150
41:AB038167    LTQMYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMRAKVCVLMLAL 145
42:M64377      LTQLYFLAVFGNMDNFLLAVMSYDRFVAICHPLHYTTKMTRQLCVLLVVG 147
43:U50947      LTQMYLFMMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMNARICVLMLIL 145
44:AF101730    LTQMYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMRAKVCVLMLAL 145
45:AF101761    LTQMYFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSLVAL 147
46:AF101741    LTQMYFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSLVAL 147

GPCR1          CWVLTNCPALMHTLLLTRVAFCAQKAIPHFYCDPSALLKLACSDTHVNEL 200
41:AB038167    CWVLTNIVALTHTLLMARLSFCVVGEIAHFFCDITPVLKLSCSDTYVNEL 195
42:M64377      SWVVANMNCLLHILLMARLSFCADNMIPHFFCDGTPLLKLSCSDTHLNEL 197
43:U50947      CWILTNVVALTHTLLMARLSFCVVGEIAHFFCDVTSVMKLSCSDTHVNEL 195
44:AF101730    CWVLTNIVALTHTLLMARLSFCVVGEIAHFFCDITPVLKLSCSDTYVNEL 195
45:AF101761    SWVLTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEW 197
46:AF101741    SWVLTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEW 197

GPCR1          MIITMGLLFLIVPLLLLIVFSYVRIFWAVFVISSPGGRWKAFSTCGSHLIV 250
41:AB038167    MVFALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTKAFSTCSSHLCV 245
42:M64377      MILTEGAVVMVTPFVCILISYIHITCAVLRVSSPRGGWKSFSTCGSHLAV 247
43:U50947      VLSGFGGTVLMVPFVSIVISYVHIVFAVLRIQSSGGGSSKAFSTCSSHLCV 245
44:AF101730    MVFALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTKAFSTCSSHLCV 245
45:AF101761    VIFIMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSHLSV 247
46:AF101741    VIFIMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSHLSV 247

GPCR1          VLLFYGSLMGVYLLPPSTYSTERESRAAVLYMVIIPTLNPFIYSLRNRDM 300
41:AB038167    VCVFYGTLFSAYLSPPSVVSTEKDIAAAAMYTVVTPMLNPFIYSLRNKDM 295
42:M64377      VCLFYGTVIAVYFNPSSSHLAGRDMAAAVMYAVVTPMLNPFIYSLRNSDM 297
43:U50947      VCVFYGTLFSVYLFPSSVEITEKDVAAAAMYTVVTPMLNPFIYSLRNKDI 295
44:AF101730    VCVFYGTLFSAYLSPPSVVSTEKDIAAAAMYTVVTPMLNPFIYSLRNKDM 295
45:AF101761    VSLFYGTIIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRNRDM 297
46:AF101741    VSLFYGTIIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRNRDM 297

GPCR1          KEALGKLFVSGKTFFL- 316
41:AB038167    KGALKRLLFHRRILSS- 311
```

```
42:M64377     KAALRKVLAMRFPSKQ- 313
43:U50947     KGALKRLLSHRRILSS- 311
44:AF101730   KGALKRLLFHRRILSS- 311
45:AF101761   KGALERVICKRKNPFLL 314
46:AF101741   KGALERVIXKRKNPFLL 314
```

The presence of identifiable domains in the protein disclosed herein was determined by searches using algorithms such as PROSITE, Blocks, Pfam, ProDomain, Prints and then determining the Interpro number by crossing the domain match (or numbers) using the Interpro website (http:www.ebi.ac.uk/interpro/).

**DOMAIN**

The results indicate that the GPCR1 protein contains the following protein domain (as defined by Interpro): domain name 7tm_1 7 transmembrane receptor (rhodopsin family). DOMAIN results for GPCR1 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections.

As discussed below, all GPCRX proteins of the invention contain significant homology to the 7tm_1 domain. This indicates that the GPCRX sequence has properties similar to those of other proteins known to contain this 7tm_1 domain and similar to the properties of these domains. The 254 amino acid domain termed 7tm_1 (SEQ ID NO:39), a seven transmembrane receptor (rhodopsin family), is shown in Table 1F.

**Table 1F. 7tm_1, 7 transmembrane receptor domain (SEQ ID NO:39)**

| gnl\|Pfam\|pfam00001, 7tm_1, 7 transmembrane receptor (rhodopsin family). |
|---|
| GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLVGALFVVNGYASILLLTAISIDRYL<br>AIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPLLFSWLRTVEEGNTTVCLIDFPEESVKRSYVLLSTLVGFVLPLLVILVC<br>YTRILRTLRKRARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRVLPTALLITLWLAYVNSCLNPI<br>IY |

The encoded GPCR1 polypeptide was identified as a member of the G protein receptor family due to the presence of a signature consensus sequence (SEQ ID NO:40) shown in Table 1G below.

**Table 1G. G-protein coupled receptors signature domain (SEQ ID NO:40)**

| | |
|---|---|
| Entry Name | G_PROTEIN_RECEPTOR |
| Entry Type | PATTERN |
| Primary Accession Number | PS00237 |
| Created / Last Updated | 01-APR-1990 / 01-JUL-1998 |
| Description | G-protein coupled receptors signature. |
| Pattern | [GSTALIVMFYWC]-[GSTANCPDE]-{EDPKRH}-x(2)-[LIVMNQGA]-x(2)-[LIVMFT]-[GSTANC]-[LIVMFYWSTAC]-[DENH]-R-[FYWCSH]-x(2)-[LIVM]. |

Table 1H lists the domain description from DOMAIN analysis results against GPCR1. This indicates that the GPCR1 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:39). For Table 1H and all successive DOMAIN sequence alignments, fully conserved single residues are indicated by black shading and "strong" semi-conserved residues are indicated by grey shading. The "strong" group of conserved amino acid residues may be any one of the following groups of amino acids: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW.

The DOMAIN results are listed in Table 1H with the statistics and domain description. An alignment of GPCR1 residues 41-290 (SEQ ID NO:2) with the full 7tm_1 domain, residues 1-254 (SEQ ID NO:39), are shown in Table 1H.

**Table 1H. DOMAIN results for GPCR1**

| PSSMs producing significant alignments: | | Score (bits) | E value |
|---|---|---|---|
| gnl\|Pfam\|pfam00001 | 7tm_1, 7 transmembrane receptor (rhodopsin family) | 118 | 4e-28 |

```
GPCR1:   44 GNLLIILAISSDPHLHTPMYFFLANLSLTDACFTSASIPKMLANIHTQSQIISYSGCLAQ  103
7tm_1:    1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV   60

GPCR1:  104 LYFLLMFGGLDNCLLAVMAYDRYVAICQPLHYSTSMSPQLCALMLGVCWVLTNCPALMHT  163
7tm_1:   61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPL  120

GPCR1:  164 LLLTRVAFCAQKAIPHFYCDPSALLKLACSDTHVNELMIITMGLLFLTVPLLLIVFSYVR  223
7tm_1:  121 LFSWLRTV---EEGNTTVCLIDF------PEESVKRSYVLLSTLVGFVLPLLVILVCYTR  171

GPCR1:  224 IFWAVFVISSPGGRWK---------AFSTCGSHLTVVLLFYG----SLMGVYLLPPSTYS  270
7tm_1:  172 ILRTLRKRARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRV  231

GPCR1:  271 TERESRAAVLYMVIIPTLNPFIY  293
```

7tm_1: 232 **LPTALLITLWLAYVNSCLNPIIY** 254

The nucleic acids and proteins of GPCR1 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, as described further herein.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR1 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR1 epitope is from about amino acids 15 to 25. In another embodiment, a GPCR1 epitope is from about amino acids 125 to 135. In further specific embodiments, GPCR1 epitopes are from about amino acids 235 to 245, from about amino acids 258 to 275 and from about amino acids 285 to 316.

**GPCR2**

A second GPCR-like protein of the invention, referred to herein as GPCR2, is an Olfactory Receptor ("OR")-like protein. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR2 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Two alternative novel GPCR2 nucleic acids and encoded polypeptides are provided, namely GPCR2a and GPCR2b.

**GPCR2a**

In one embodiment, a GPCR2 variant is the novel GPCR2a (alternatively referred to herein as GMba64p14_B), which includes the 1034 nucleotide sequence (SEQ ID NO:3) shown in Table 2A. A GPCR2a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 15-17 and ends with a TAA codon at nucleotides 945-947. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 2A, and the start and stop codons are in bold letters.

**Table 2A. GPCR2a Nucleotide Sequence (SEQ ID NO:3)**

GTGAACCCACAACT**ATG**GGAAGAAATAACCTAACAAGACCCTCTGAATTCATCCTCCTTGGACTCTC
CTCTCGACCTGAGGATCAGAAGCCGCTCTTTGCTGTGTTCCTCCCCATCTACCTTATCACAGTGATA
GGAAACCTGCTTATCATCCTGGCCATCCGCTCAGACACTCGTCTCCAGACGCCCATGTACTTCTTTC
TAAGCATCCTGTCTTTTGTTGACATTTGCTATGTGACAGTCATTATCCCTAAGATGCTGGTGAACTT
CTTATCAGAGACAAAGACCATCTCTTACAGTGAGTGTCTGACCCAGATGTACTTTTTCTTAGCCTTT
GGAAACACAGACAGTTACCTGCTAGCAGCCATGGCCATTGACCGCTATGTGGCCATATGTAATCCCT
TCCACTACATCACCATTATGAGTCACAGATGCTGTGTCCTGCTTCTGGTTCTCTCCTTCTGCATTCC
ACATTTTCACTCCCTCCTGCACATTCTTCTGACTAATCAGCTCATCTTCTGTGCCTCAAATGTCATC
CATCACTTTTTCTGCGATGATCAACCAGTGCTAAAATTGTCCTGTTCCTCCCATTTTGTCAAAGAAA
TCACAGTAATGACAGAAGGCTTGGCTGTCATAATGACCCCGTTTTCATGCATCATCATCTCTTATTT
AAGAATCCTCATCACTGTTCTGAAGATTCCTTCAGCTGCTGGAAAGCGTAAAGCATTTTCTACCTGT
GGCTCTCATCTCACAGTGGTGACCCTGTTTTATGGAAGCATTAGCTATCTCTATTTTCAGCCCCTGT
CCAACTATACTGTCAAGGATCAAATAGCAACAATTATCTACACCGTACTGACTCCTATGCTAAATCC
ATTTATCTATAGTCTGAGGAACAAAGACATGAAGCAGGGTTTGGCAAAGTTGATGCACAGGATGAAA
TGTCAG**TAA**AAGACCTAAGGTCTTAAGAGAATACCACAGATCTCTTGCCCTGGACTATAGGTTATTA
ATGGGTATGTGATTCTGAAATGATTATTA

The sequence of GPCR2a was derived by laboratory cloning of cDNA fragments, by *in silico* prediction of the sequence. The cDNA fragments covering either the full length of the DNA sequence, or part of the sequence, or both, were cloned. *In silico* prediction was based on sequences available in CuraGen's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA sequence, or some portion thereof.

The cDNA coding for the GPCR2a sequence was cloned by the polymerase chain reaction (PCR). Primers were designed based on *in silico* predictions of the full length or some portion (one or more exons) of the cDNA/protein sequence of the invention. The DNA sequence and protein sequence for a novel Olfactory Receptor-like gene were obtained by exon linking and are reported here as GPCR2a.These primers and methods used to amplify GPCR2 a cDNA are described in the Examples.

The GPCR2a polypeptide (SEQ ID NO:4) encoded by SEQ ID NO:3 is 310 aa in length, has a molecular weight of 35329.7 Daltons, and is presented using the one-letter amino acid code in Table 2B. The Psort profile for both GPCR2a and GPCR2b predicts that these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR2 polypeptide is located to the Golgi body with a certainty of 0.400, the mitochomdrial inner membrane with a certainty of 0.3828, or a mitochomdrial intermembrane space with a certainty of 0.3565. The Signal P predicts a likely cleavage site for a GPCR2 peptide is between positions 48 and 49, *i.e.*, at the dash in the sequence ILA-IR.

**Table 2B. GPCR2a protein sequence (SEQ ID NO:4)**

MGRNNLTRPSEFILLGLSSRPEDQKPLFAVFLPIYLITVIGNLLIILAIRSDTRLQTPMYFFLSILS
FVDICYVTVIIPKMLVNFLSETKTISYSECLTQMYFFLAFGNTDSYLLAAMAIDRYVAICNPFHYIT
IMSHRCCVLLLVLSFCIPHFHSLLHILLTNQLIFCASNVIHHFFCDDQPVLKLSCSSHFVKEITVMT
EGLAVIMTPFSCIIISYLRILITVLKIPSAAGKRKAFSTCGSHLTVVTLFYGSISYLYFQPLSNYTV
KDQIATIIYTVLTPMLNPFIYSLRNKDMKQGLAKLMHRMKCQ

## GPCR2b

In an alternative embodiment, a GPCR2 variant is the novel GPCR2b (alternatively referred to herein as CG56582-01), which includes the 1011 nucleotide sequence (SEQ ID NO:5) shown in Table 2C. The DNA sequence and protein sequence of GPCR2b was obtained solely by exon linking process. The GPCR2b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 25-27 and ends with a TAA codon at nucleotides 955-957, which are in bold letters in Table 2C.

**Table 2C. GPCR2b Nucleotide Sequence (SEQ ID NO:5)**

TTATCTTTACGTGAACCCACAACT**ATG**GGAAGAAATAACCTAACAAGACCCTCTGAATTCATCCTCC
TTGGACTCTCCTCTCGACCTGAGGATCAGAAGCCGCTCTTTGCTGTGTTCCTCCCCATCTACCTTAT
CACAGTGATAGGAAACCTGCTTATCATCCTGGCCATCCGCTCAGACACTCGTCTCCAGACGCCCATG
TACTTCTTTCTAAGCATCCTGTCTTTTGTTGACATTTGCTATGTGACAGTCATTATCCCTAAGATGC
TGGTGAACTTCTTATCAGAGACAAAGACCATCTCTTACAGTGAGTGTCTGACCCAGATGTACTTTTT
CTTAGCCTTTGGAAACACAGACAGTTACCTGCTAGCAGCCATGGCCATTGACCGCTATGTGGCCATA
TGTAATCCCTTCCACTACATCACCATTATGAGTCACAGATGCTGTGTCCTGCTTCTGGTTCTCTCCT
TCTGCATTCCACATTTTCACTCCCTCCTGCACATTCTTCTGACTAATCAGCTCATCTTCTGTGCCTC
AAATGTCATCCATCACTTTTTCTGCGATGATCAACCAGTGCTAAAATTGTCCTGTTCCTCCCATTTT
GTCAAAGAAATCACAGTAATGACAGAAGGCTTGGCTGTCATAATGACCCCGTTTTCATGCGTCATCA
TCTCTTATTTAAGAATCCTCATCACTGTTCTGAAGATTCCTTCAGCTGCTGGAAAGCGTAAAGCATT
TTCTACCTGTGGCTCTCATCTCACAGTGGTGACCCTGTTTTATGGAAGCATTAGCTATCTCTATTTT
CAGCCCCTGTCCAACTATACTGTCAAGGATCAAATAGCAACAATTATCTACACCGTACTGACTCCTA
TGCTAAATCCATTTATCTATAGTCTGAGGAACAAAGACATGAAGCAGGGTTTGGCAAAGTTGATGCA
CAGGATGAAATGTCAG**TAA**AAGACCTAAGGTCTTAAGAGAATACCACAGATCTCTTGNCCTGGACTA
TAGGTT

The GPCR2b protein (SEQ ID NO:6) encoded by SEQ ID NO:5 is 310 amino acids in length, has a molecular weight of 35314.35 Daltons, and is presented using the one-letter code in Table 2D. As with GPCR2a, the most likely cleavage site for a GPCR2b peptide is between amino acids positions 48 and 49, *i.e.*, at the dash in the sequence ILA-IR, based on the SignalP result.

| Table 2D. GPCR2b protein sequence (SEQ ID NO:6) |
|---|
| MGRNNLTRPSEFILLGLSSRPEDQKPLFAVFLPIYLITVIGNLLIILAIRSDTRLQTPMYFFLSILS<br>FVDICYVTVIIPKMLVNFLSETKTISYSECLTQMYFFLAFGNTDSYLLAAMAIDRYVAICNPFHYIT<br>IMSHRCCVLLLVLSFCIPHFHSLLHILLTNQLIFCASNVIHHFFCDDQPVLKLSCSSHFVKEITVMT<br>EGLAVIMTPFSCVIISYLRILITVLKIPSAAGKRKAFSTCGSHLTVVTLFYGSISYLYFQPLSNYTV<br>KDQIATIIYTVLTPMLNPFIYSLRNKDMKQGLAKLMHRMKCQ |

## GPCR2 Clones

Unless specifically addressed as GPCR2a or GPCR2b, any reference to GPCR2 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant, the residue position with respect to the "a" variant, and the residue in the "b" variant.

The amino acid sequence of GPCR2 has high homology to other proteins as shown in Table 2E.

| Table 2E. BLASTX results for GPCR2 | | | |
|---|---|---|---|
| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
| patp:AAR27868 Odorant receptor clone F5 - Rattus rattus, 313 aa. | +3 | 882 | 1.8e-87 |
| patp:AAR27876 Odorant receptor clone I15 - Rattus rattus, 314 aa. | +3 | 835 | 1.7e-82 |

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of GPCR2 has 620 of 942 bases (65%) identical to a gb:GENBANK-ID:RATOLFPROC|acc:M64377.1 mRNA from Rattus norvegicus (Rat olfactory protein mRNA, complete cds). The full amino acid sequence of theGPCR2 protein was found to have 165 of 305 amino acid residues (54%) identical to, and 224 of 305 amino acid residues (73%) similar to, the 313 amino acid residue ptnr:SWISSPROT-ACC:P23266 protein from Rattus norvegicus (Rat) (OLFACTORY RECEPTOR-LIKE PROTEIN F5).

Additional BLASTP results are shown in Table 2F.

| GPCR2 BLASTP results | | | | | |
|---|---|---|---|---|---|
| Gene Index / Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |
| M64377; P23266 | OLF5_RAT rattus norvegicus (rat). olfactory receptor-like protein f5. 7/1993 | 313 | 166/305 (54%) | 224/305, (73%) | 1e-94 |

| | | | | | |
|---|---|---|---|---|---|
| Y14442; AJ003147; U86234 | O1F1_HUMAN homo sapiens (human). olfactory receptor 1f1 (olfactory receptor 16-35) (or16-35). 10/2000 | 312 | 167/308 (54%) | 220/308, (71%) | 4e-93 |
| AF101764 | Q9TU86 gorilla gorilla (gorilla). olfactory receptor. 3/2001 | 313 | 168/305 (55%) | 214/305, (70%) | 3e-90 |
| AF087918 | Q9P1Q5 homo sapiens (human). olfactory receptor 17-7. 3/2001 | 309 | 166/307 (54%) | 219/307, (71%) | 4e-90 |
| AF101730 | Q9TUA9 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 165/309 (53%) | 223/309, (72%) | 5e-90 |

A multiple sequence alignment is given in Table 2G, with the GPCR2 protein of the invention being shown on line 1 and 2, in a ClustalW analysis comparing GPCR2 with related protein sequences of Table 2F. The residue that differs between GPCR2a and GPCR2b is marked with the (o) symbol.

**Table 2G. Information for the ClustalW proteins:**

1. SEQ ID NO:4, GPCR2a
2. SEQ ID NO:6, GPCR2a
3. SEQ ID NO:42, M64377 olfactory receptor-like protein f5
4. SEQ ID NO:44, AF101730 chimpanzee olfactory receptor
5. SEQ ID NO:47, Y14442 olfactory receptor 1f1 (or16-35)
6. SEQ ID NO:48, AF101764 gorilla olfactory receptor
7. SEQ ID NO:49, AF087918 olfactory receptor 17-7

```
GPCR2a       MGRNNLTRPSEFILLGLSSRPEDQKPLFAVFLPIYLITVIGNLLIILAIR 50
GPCR2b       MGRNNLTRPSEFILLGLSSRPEDQKPLFAVFLPIYLITVIGNLLIILAIR 50
42:M64377    MSSTNQSSVTEFLLLGLSRQPQQQQLLFLLFLIMYLATVLGNLLIILAIG 50
44:AF101730  ME--NQSSVSEFFLQGISGFPEQQQLLYGLFLCMYLVTLTGNVLIIMAIG 48
47:Y14442    MSGTNQSSVSEFLLLGLSRQPQQQHLLFVFFLSMYLATVLGNLLIILSVS 50
48:AF101764  MEGKNLTSISEFFLLGFSEQLEEQKALFGSFLFMYLVTVAGNLLIILVII 50
49:AF087918  MRENNQSSTLEFILLGVIGQQEQEDFFYILFLFIYPITLIGNLLIVLAIC 50

GPCR2a       SDTRLQTPMYFFLSILSFVDICYVTVIIPKMLVNFLSETKTISYSECLTQ 100
GPCR2b       SDTRLQTPMYFFLSILSFVDICYVTVIIPKMLVNFLSETKTISYSECLTQ 100
42:M64377    TDSRLHTPMYFFLSNLSFVDVCFSSTTVPKVLANHILGSQAISFSGCLTQ 100
44:AF101730  SDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQTQRHTISYTGCLTQ 98
47:Y14442    IDSCLHTPMYFFLSNLSFVDICFSFTTVPKMLANHILETQTISFCGCLTQ 100
48:AF101764  TDTQLHTPMYFFLANLSLADACFVSTTVPKMLANIQIQSQAISYSGCLLQ 100
49:AF087918  SDVRLHNPMYFLLANLSLVDIFFSSVTIPKMLANHLLGSKSISFGGCLTQ 100

GPCR2a       MYFFLAFGNTDSYLLAAMAIDRYVAICNPFHYITIMSHRCCVLLLVLSFC 150
GPCR2b       MYFFLAFGNTDSYLLAAMAIDRYVAICNPFHYITIMSHRCCVLLLVLSFC 150
42:M64377    LYFLAVFGNMDNFLLAVMSYDRFVAICHPLHYTTKMIRQLCVLLVVGSWV 150
44:AF101730  MYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMRAKVCVLMLALCWV 148
47:Y14442    MYFVFMFVDMDNFLLAVMAYDHFVAVCHPLHYTAKMTHQLCALLVAGLWV 150
48:AF101764  LYFFMLFVMLEAFLLAVMAYDRYVAICHPLHYILIMSPGLCVFLVSASWI 150
49:AF087918  MYFMIALGNTDSYILAAMAYDRAVAISHPLHYTTIMSPRSCIWLIAGSWV 150

GPCR2a       IPHFHSLLHILLTNQLIFCASNVIHHFFCDDQPVLKLSCSSHFVKEITVM 200
GPCR2b       IPHFHSLLHILLTNQLIFCASNVIHHFFCDDQPVLKLSCSSHFVKEITVM 200
42:M64377    VANMNCLLHILLMARLSFCADNMIPHFFCDGTPLLKLSCSDTHLNELMIL 200
```

```
44:AF101730   LTNIVALTHTLLMARLSFCVVGEIAHFFCDITPVLKLSCSDTYVNELMVF 198
47:Y14442     VANLNVLLHTLLMAPLSFCADNAITHFFCDVTPLLKLSCSNTHLNEVIIL 200
48:AF101764   MNALHSLLHTLLMNSLSFCANHEIPHFFCDIDPLLSLSCTDPFTNELVIF 200
49:AF087918   IGNANALPHTLLTASLSFCGNQEVANFYCDITPLLKLSCSDIHF-HVKMM 199

                                    o
GPCR2a        TEGLAVIMTPFSCIIISYLRILITVLKIPSAAGKRKAFSTCGSHLTVVTL 250
GPCR2b        TEGLAVIMTPFSCVIISYLRILITVLKIPSAAGKRKAFSTCGSHLTVVTL 250
42:M64377     TEGAVVMVTPFVCILISYIHITCAVLRVSSPRGGWKSFSTCGSHLAVVCL 250
44:AF101730   ALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTKAFSTCSSHLCVVCV 248
47:Y14442     SEGALVMITPFLCILASYMHITCTVLKVPSTKGRWKAFSTCGSHLAVVLL 250
48:AF101764   ITGGLTGLVCVLCLIISYTNIFSTILKIPSAQGKRKAFSTCGSHLSVVSL 250
49:AF087918   YLGVGIFSVPLLCIIVSYIRVFSTVFQVPSTKGVLKAFSTCGSHLTVVSL 249

GPCR2a        FYGSISYLYFQPLS-NYTVKDQIATIIYTVLTPMLNPFIYSLRNKDMKQG 299
GPCR2b        FYGSISYLYFQPLS-NYTVKDQIATIIYTVLTPMLNPFIYSLRNKDMKQG 299
42:M64377     FYGTVIAVYFNPSSSHLAGRDMAAAVMYAVVTPMLNPFIYSLRNSDMKAA 300
44:AF101730   FYGTLFSAYLSPPSVVSTEKDIAAAAMYTVVTPMLNPFIYSLRNKDMKGA 298
47:Y14442     FYSTIIAVYFNPLSSHSAEKDTMATVLYTVVTPMLNPFIYSLRNRYLKGA 300
48:AF101764   FFGTSFCVYFIPPSTRSAQKDTVASVMYTVVTPMLNPFIYSLRNQEIKSS 300
49:AF087918   YYGTVMGTYFRPLT-NYSLKDAVITVMYTAVTPMLNPFIYSLRNRDMKAA 298

GPCR2a        LAKLMH--RMKCQ- 310
GPCR2b        LAKLMH--RMKCQ- 310
42:M64377     LRKVLA-MRFPSKQ 313
44:AF101730   LKRLLFHRRILSS- 311
47:Y14442     LKKVIG-RVVFSV- 312
48:AF101764   LRKLIWVREIHSP- 313
49:AF087918   LRKLFN-KRISS-- 309
```

DOMAIN results for GPCR2 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. The results are listed in Table 2H with the statistics and domain description. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to have significant homology to GPCR2. An alignment of GPCR2 residues 41-290 (SEQ ID NO:4) with 7tm_1 residues 1-254 (SEQ ID NO:39) are shown in Table 2H.

**Table 2H. DOMAIN results for GPCR2**

| PSSMs producing significant alignments: | | Score (bits) | E value |
|---|---|---|---|
| gnl\|Pfam\|pfam00001 | 7tm_1, 7 transmembrane receptor (rhodopsin family) | 105 | 3e-24 |

```
GPCR2:   41 GNLLIILAIRSDTRLQTPMYFFLSILSFVDICYVTVIIPKMLVNFLSETKTISYSECLTQ  100
           .  ||||+||  |    +|+||   ||  |+  |+ ++  + |  |   +        + |
7tm_1:    1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV   60

GPCR2:  101 MYFFLAFGNTDSYLLAAMAIDRYVAICNPFHYITIMSHRCCVLLLVLSFCIPHFHSLLHI  160
              |+  |     || |++||||+|| +|  |  | + |   +|++| + +    ||  +
7tm_1:   61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPL  120

                                                                  o
GPCR2:  161 LLTNQLIFCASNVIHHFFCDDQPVLKLSCSSHFVKEITVMTEGLAVIMTPFSCIIISYLR  220
            | +           +   |      +    |+ +    |      +++      |+  ||
```

```
7tm_1: 121 LFSW---LRTVEEGNTTVCLIDFPEESVKRSYVLLSTLVGFVLPLLVILVCYTRILRTLR  177

GPCR2: 221 ---ILITVLKIPSAAGKRKAFSTCGSHLTVVTLFYGS-----ISYLYFQPLSNYTVKDQI  272
                    ||   |++ ++ |        +   |   +          +   |     +         +
7tm_1: 178 KRARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRVLPTALL  237

GPCR2: 273 ATIIYTVLTPMLNPFIY  289
             |+    +    |||  ||
7tm_1: 238 ITLWLAYVNSCLNPIIY  254
```

The GPCR2 disclosed in this invention is expressed in at least the following tissues: Apical microvilli of the retinal pigment epithelium, arterial (aortic), basal forebrain, brain, Burkitt lymphoma cell lines, corpus callosum, cardiac (atria and ventricle), caudate nucleus, CNS and peripheral tissue, cerebellum, cerebral cortex, colon, cortical neurogenic cells, endothelial (coronary artery and umbilical vein) cells, palate epithelia, eye, neonatal eye, frontal cortex, fetal hematopoietic cells, heart, hippocampus, hypothalamus, leukocytes, liver, fetal liver, lung, lung lymphoma cell lines, fetal lymphoid tissue, adult lymphoid tissue, Those that express MHC II and III nervous, medulla, subthalamic nucleus, ovary, pancreas, pituitary, placenta, pons, prostate, putamen, serum, skeletal muscle, small intestine, smooth muscle (coronary artery in aortic) spinal cord, spleen, stomach, taste receptor cells of the tongue, testis, thalamus, and thymus tissue. This information was derived by determining the tissue sources of the sequences that were included in the invention including but not limited to SeqCalling sources, Public EST sources, Literature sources, and/or RACE sources. This is by no way limiting in that olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types. Further tissue expression analysis is provided in the Examples.

The nucleic acids and proteins of GPCR2 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further above.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR2 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a

contemplated GPCR2 epitope is from about amino acids 5 to 25. In other specific embodiments, GPCR2 epitopes are from about amino acids 85 to 95, from about amino acids 180 to 195, from about amino acids 230 to 240, from about amino acids 255 to 270 and from about amino acids 285 to 310.

## GPCR3

A third GPCR-like protein of the invention, referred to herein as GPCR3, is an Olfactory Receptor ("OR")-like protein. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR3 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

The disclosed novel GPCR3 (alternatively referred to herein as GMba64p14_C) includes the 981 nucleotide sequence (SEQ ID NO:7) shown in Table 3A. A GPCR3 ORF begins with a Kozak consensus ATG initiation codon at nucleotides 15-17 and ends with a TGA codon at nucleotides 969-971. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 3A, and the start and stop codons are in bold letters.

**Table 3A. GPCR3 Nucleotide Sequence (SEQ ID NO:7)**

GCTGACTGTCACTC**ATG**ATGAGCTTTGCCCCTAATGCTTCACACTCTCCGGTTTTTTTGCTCCTTGG
GTTCTCGAGAGCTAACATCTCCTACACTCTCCTCTTCTTCCTGTTCCTGGCTATTTACCTGACCACC
ATACTGGGGAATGTGACACTGGTGCTGCTCATCTCCTGGGACTCCAGACTGCACTCACCCATGTATT
ATCTGCTTCGTGGCCTCTCTGTGATAGACATGGGGCTATCCACAGTTACACTGCCCCAGTTGCTGGC
CCATTTGGTCTCTCATTACCCAACCATTCCTGCTGCCCGCTGCTTGGCTCAGTTCTTTTTCTTCTAT
GCATTTGGGGTTACAGATACACTTGTCATTGCTGTCATGGCTCTGGATCGCTATGTGGCCATCTGTG
ACCCCCTGCACTATGCTTTGGTAATGAATCACCAACGGTGTGCCTGCTTACTAGCCTTGAGCTGGGT
GGTGTCCATACTGCACACCATGTTGCGTGTGGGACTCGTCCTGCCTCTTTGCTGGACTGGGGATGCT
GGGGGCAACGTTAACCTTCCTCACTTCTTTTGTGACCACCGGCCACTTCTGCGAGCCTCTTGTTCTG
ACATACATTCTAATGAGCTGGCCATATTCTTTGAGGGTGGCTTCCTTATGCTGGGCCCCTGTGCCCT
CATTGTACTCTCTTATGTCCGAATTGGGGCCGCTATTCTACGTTTGCCTTCAGCTGCTGGTCGCCGC
CGAGCAGTCTCCACCTGTGGATCCCACCTCACCATGGTTGGTTTCCTCTACGGCACCATCATTTGTG
TCTACTTCCAGCCTCCCTTCCAGAACTCTCAGTATCAGGACATGGTGGCTTCAGTAATGTATACTGC
CATTACACCTTTGGCCAACCCATTTGTGTATAGCCTCCACAATAAGGATGTCAAGGGTGCACTCTGC
AGGCTGCTTGAATGGGTGAAGGTAGACCCC**TGA**TTAGCCTGCT

The GPCR3 protein (SEQ ID NO:8) encoded by SEQ ID NO:7 is 318 aa in length, has a molecular weight of 35292.3 Daltons, and is presented using the one-letter amino acid code in Table 3B. The Psort profile for GPCR3 predicts that these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.640. In alternative embodiments, a GPCR3 polypeptide is located to the Golgi body with a certainty of 0.460, the

endoplasmic reticulum (membrane) with a certainty of 0.370, or the endoplasmic reticulum (lumen) with a certainty of 0.100. The Signal P predicts a likely cleavage site for a GPCR3 peptide is between positions 43 and 44, *i.e.*, at the dash in the sequence ILG-NV.

**Table 3B. Encoded GPCR3 protein sequence (SEQ ID NO:8)**

MMSFAPNASHSPVFLLLGFSRANISYTLLFFLFLAIYLTTILGNVTLVLLISWDSRLHSPMYYLL
RGLSVIDMGLSTVTLPQLLAHLVSHYPTIPAARCLAQFFFFYAFGVTDTLVIAVMALDRYVAICD
PLHYALVMNHQRCACLLALSWVVSILHTMLRVGLVLPLCWTGDAGGNVNLPHFFCDHRPLLRASC
SDIHSNELAIFFEGGFLMLGPCALIVLSYVRIGAAILRLPSAAGRRRAVSTCGSHLTMVGFLYGT
IICVYFQPPFQNSQYQDMVASVMYTAITPLANPFVYSLHNKDVKGALCRLLEWVKVDP

The amino acid sequence of GPCR3 had high homology to other proteins as shown in Table 3C.

**Table 3C. BLASTX results for GPCR3**

| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
|---|---|---|---|
| patp:AAR27876 Odorant receptor clone I15 - Rattus rattus, 314 aa. | +3 | 712 | 1.8e-69 |

GPCR3 also has homology to the proteins shown in the BLASTP data in Table 3D.

**Table 3D. GPCR3 BLASTP results**

| Gene Index / Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |
|---|---|---|---|---|---|
| X64996 | OLFD_CANFA canis familiaris (dog). olfactory receptor-like protein dtmt. 2/1994 | 313 | 139/298 (47%) | 197/298, (66%) | 5e-76 |
| AF101730 | Q9TUA9 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 141/302 (47%) | 200/302, (66%) | 3e-75 |
| AF101760; AF101761 | Q9TQX4 gorilla gorilla (gorilla). olfactory receptor. 3/2001 | 314 | 142/298 (48%) | 198/298, (66%) | 6e-75 |
| AF101741 | Q9TU A1 pan troglodytes(chimpanzee) olfactory receptor. 3/2001 | 314 | 142/298 (48%) | 198/298, (66%) | 6e-75 |
| M64392 | OLF1_RAT rattus norvegicus (rat). olfactory receptor-like protein i15. 7/1993 | 314 | 138/298 (46%) | 195/298, (65%) | 1e-74 |
| AF101739 | Q9TUA3 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 141/305 (46%) | 201/305, (66%) | 1e-74 |

A multiple sequence alignment is given in Table 3E, with the GPCR3 protein being shown on line 1 in Table 3E in a ClustalW analysis, and comparing the GPCR3 protein with

the related protein sequences shown in Table 3D. This BLASTP data is displayed graphically in the ClustalW in Table 3E.

**Table 3E. ClustalW Analysis of GPCR3**

1. SEQ ID NO:8, GPCR3
2. SEQ ID NO:44, AF101730 chimpanzee olfactory receptor
3. SEQ ID NO:45, AF101761 gorilla olfactory receptor
4. SEQ ID NO:46, AF101741 chimpanzee olfactory receptor
5. SEQ ID NO:50, X64996 olfactory receptor-like protein dtmt
6. SEQ ID NO:51, M64392 olfactory receptor-like protein i15

```
GPCR3          MMSFAPNASHSPVFLLLGFSRANISYTLLFFLFLAIYLTTILGNVTLVLL  50
44:AF101730    ----MENQSSVSEFFLQGISGFPEQQQLLYGLFLCMYLVTLIGNVLIIMA  46
45:AF101761    --MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVL  48
46:AF101741    --MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVL  48
50:X64996      --MTEKNQTVVSEFVLLGLPIDPDQRDLFYALFLAMYVTTILGNLLIIVL  48
51:M64392      --MTEENQTVISQFLLLFLPIPSEHQHVFYALFLSMYLTTVLGNLIIIIL  48

GPCR3          ISWDSRLHSPMYYLLRGLSVIDMGLSTVTLPQLLAHLVSHYPTIPAARCL 100
44:AF101730    IGSDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQTQRHTISYTGCL  96
45:AF101761    IRLDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCL  98
46:AF101741    IRLDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCL  98
50:X64996      IQLDSHLHTPMYLFLSNLSFSDLCFSSVTMPKLLQNMQSQVPSIPYAGCL  98
51:M64392      IHLDSHLHTPMYLFLSNLSFSDLCFSSVTMPKLLQNMQSQVPSIPFAGCL  98

GPCR3          AQFFFFYAFGVTDTLVIAVMALDRYVAICDPLHYALVMNHQRCACLLALS 150
44:AF101730    TQMYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMRAKVCVLMLALC 146
45:AF101761    TQMYFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSLVALS 148
46:AF101741    TQMYFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSLVALS 148
50:X64996      TQMYFFLFFGDLESFLLVAMAYDRYVAICFPLHYTTIMSPKLCFSLLVLS 148
51:M64392      TQLYFYLYFADLESFLLVAMAYDRYVAICFPLHYMSIMSPKLCVSLVVLS 148

GPCR3          WVVSILHTMLRVGLVLPLCWTGDAGGNVNLPHFFCDHRPLLRASCSDIHS 200
44:AF101730    WVLTNIVALTHTLLMARLSFCVVG----EIAHFFCDITPVLKLSCSDTYV 192
45:AF101761    WVLTTFHAMLHTLLMARLCFCADN----VIPHFFCDMSALLKLACSDTRV 194
46:AF101741    WVLTTFHAMLHTLLMARLCFCADN----VIPHFFCDMSALLKLACSDTRV 194
50:X64996      WVLTMFHAVLHTLLMARLCFCAN-----TIPHFFCDMSALLKLACSDTQV 193
51:M64392      WVLTTFHAMLHTLLMARLSFCADN----MIPHFFCDISPLLKLSCSDTHV 194

GPCR3          NELAIFFEGGFLMLGPCALIVLSYVRIGAAILRLPSAAGRRRAVSTCGSH 250
44:AF101730    NELMVFALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTKAFSTCSSH 242
45:AF101761    NEWVIFIMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSH 244
46:AF101741    NEWVIFIMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSH 244
50:X64996      NELVIFIMGGLILVIPFLLIITSYARIVSSILKVPSAIGICKVFSTCGSH 243
51:M64392      NELVIFVMGGLVIVIPFVLIIVSYARVVASILKVPSVRGIHKIFSTCGSH 244

GPCR3          LTMVGFLYGTIICVYFQPPFQNSQYQDMVASVMYTAITPLANPFVYSLHN 300
44:AF101730    LCVVCVFYGTLFSAYLSPPSVVSTEKDIAAAAMYTVVTPMLNPFIYSLRN 292
45:AF101761    LSVVSLFYGTIIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRN 294
46:AF101741    LSVVSLFYGTIIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRN 294
50:X64996      LSVVSLFYGTVIGLYLCPSANNSTVKETIMAMMYTVVTPMLNPFIYSLRN 293
51:M64392      LSVVSLFYGTIIGLYLCPSANNSTVKETVMAMMYTVVTPMLNPFIYSLRN 294

GPCR3          KDVKGALCRLLEWVKVDP-- 318
44:AF101730    KDMKGALKRLLFHRRILSS- 311
45:AF101761    RDMKGALERVICKRKNPFLL 314
46:AF101741    RDMKGALERVIXKRKNPFLL 314
```

```
50:X64996      KDMKGALRRVICRKKITFSV 313
51:M64392      RDMKEALIRVLCKKKITFCL 314
```

Table 3F lists the domain description from DOMAIN analysis results against GPCR3. This indicates that the GPCR3 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:39) itself.

**Table 3F Domain Analysis of GPCR3**

| PSSMs producing significant alignments: | Score (bits) | E value |
|---|---|---|
| gnl\|Pfam\|pfam00001 7tm_1, 7 transmembrane receptor (rhodopsin family) | 88.6 | 5e-19 |

```
GPCR3:  43 GNVTLVLLISWDSRLHSPMYYLLRGLSVIDMGLSTVTLPQLLAHLVSHYPTIPAARCLAQ 102
7tm_1:   1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV  60

GPCR3: 103 FFFFYAFGVTDTLVIAVMALDRYVAICDPLHYALVMNHQRCACLLALSWVVSILHTMLRV 162
7tm_1:  61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALL-----L 115

GPCR3: 163 GLVLPLCWTGDAGGNVNLPHFFCDHRPLLRASCSDIHSNELAIFFEGGFLMLGPC---AL 219
7tm_1: 116 SLPPLLFSWLRTVEEGN--TTVCLIDFPEESVKRSYVLLSTLVGFVLPLLVILVCYTRIL 173

GPCR3: 220 IVLSYVRIGAAILRLPSAAGRRRAVSTCGSHLTMV----GFLYGTIICVYFQPPFQNSQY 275
7tm_1: 174 RTLRKRARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRVLP 233

GPCR3: 276 QDMVASVMYTAITPLANPFVY 296
7tm_1: 234 TALLITLWLAYVNSCLNPIIY 254
```

The nucleic acids and proteins of GPCR3 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further above.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR3 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR3 epitope is from about amino acids 5 to 25. In another embodiment, a GPCR3 epitope is from about amino acids 55 to 65. In further specific embodiments, GPCR3

epitopes are from about amino acids 170 to 200, from about amino acids 235 to 250, from about amino acids 260 to 280 and from about amino acids 290 to 318.

**GPCR4**

A further GPCR-like protein of the invention, referred to herein as GPCR4, is an Olfactory Receptor ("OR")-like protein. The novel GPCR4 nucleic acid sequences were identified on chromosome 11 as described in Example 1. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR4 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Three alternative novel GPCR4 nucleic acids and encoded polypeptides are provided, namely GPCR4a, GPCR4b and GPCR4c.

**GPCR4a**

In one embodiment, a GPCR4 variant is the novel GPCR4a (alternatively referred to herein as CG55940-01), which includes the 1021 nucleotide sequence (SEQ ID NO:9) shown in Table 4A. A GPCR4a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 44-46 and ends with a TAG codon at nucleotides 977-979. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 6A, and the start and stop codons are in bold letters.

**Table 4A. GPCR4a Nucleotide Sequence (SEQ ID NO:9)**

AGCATTCTAACTGTCTTCCCTCAGCTCCAGATGCTGCAGAGTC**ATG**GAAAACCAATCCAGCATTTCT
GAATTTTTCCTCCGAGGAATATCAGCGTCTCCAGAGCAACAGCAGTCCCTCTTCGGAATTTTCCTGT
GTATGTATCTTGTCACCTTGACTGGGAACCTGCTCATCATCCTGGCCATTGGCTCTGACCTGCACCT
CCACACCCCCATGTACTTTTTCTTGGCCAACCTGTCTTTTGTTGACATGGGTTTAACGTCCTCCACA
GTTACCAAGATGCTGGTGAATATACAGACTCGGCATCACACCATCTCCTATACGGGTTGCCTCACGC
AAATGTATTTCTTTCTGATGTTTGGTGATCTAGACAGCTTCTTCCTGGCTGCCATGGCGTATGACCG
CTATGTGGCCATTTGCCACCCCCTCTGCTACTCCACAGTCATGAGGCCCCAAGTCTGTGCCCTAATG
CTTGCATTGTGCTGGGTCCTCACCAATATCGTTGCCCTGACTCACACGTTCCTCATGGCTCGGTTGT
CCTTCTGTGTGACTGGGGAAATTGCTCACTTTTTCTGTGACATCACTCCTGTCCTGAAGCTGTCATG
TTCTGACACCCACATCAACGAGATGATGGTTTTTGTCTTGGGAGGCACCGTACTCATCGTCCCCTTT
TTATGCATTGTCACCTCCTACATCCACATTGTGCCAGCTATCCTGAGGGTCCGAACCCGTGGTGGGG
TGGGCAAGGCCTTTTCCACCTGCAGTTCCCACCTCTGCGTTGTTTGTGTGTTCTATGGGACCCTCTT
CAGTGCCTACCTGTGTCCTCCCTCCATTGCCTCTGAAGAGAAGGACATTGCAGCAGCTGCAATGTAC
ACCATAGTGACTCCCATGTTGAACCCCTTTATCTATAGCCTAAGGAACAAGGACATGAAGGGGGCCC
TAAAGAGGCTCTTCAGTCACAGGAGTATTGTTTCCTCT**TAG**ATGTGGTGACAGCAACATTTAATGAA
AAGACATAGGCTTGGA

The GPCR4 protein (SEQ ID NO:10) encoded by SEQ ID NO:9 has 311 amino acid residues and is presented using the one-letter code in Table 4B. The predicted molecular weight of GPCR4 protein is approximately 34638.29 Daltons. The Psort profile for GPCR4 predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, GPCR4 is located in the Golgi body with a certainty of 0.400, the endoplasmic reticulum (membrane) with a certainty of 0.300 or microbodies (peroxisomes) with a certainty of 0.300. The Signal P predicts a likely cleavage site between positions 46 and 47, i.e., at the dash in the sequence ILA-IG.

The DNA sequence and protein sequence of GPCR4a was obtained by exon linking as described in the Example 1.

**Table 4B. Encoded GPCR4a protein sequence (SEQ ID NO:10)**

```
MENQSSISEFFLRGISASPEQQQSLFGIFLCMYLVTLTGNLLIILAIGSDLHLHTPMYFFLANLSFVDMG
LTSSTVTKMLVNIQTRHHTISYTGCLTQMYFFLMFGDLDSFFLAAMAYDRYVAICHPLCYSTVMRPQVCA
LMLALCWVLTNIVALTHTFLMARLSFCVTGEIAHFFCDITPVLKLSCSDTHINEMMVFVLGGTVLIVPFL
CIVTSYIHIVPAILRVRTRGGVGKAFSTCSSHLCVVCVFYGTLFSAYLCPPSIASEEKDIAAAAMYTIVT
PMLNPFIYSLRNKDMKGALKRLFSHRSIVSS
```

**GPCR4b**

In an alternative embodiment, a GPCR4 variant is the novel GPCR4b (alternatively referred to herein as CG55940-02), which includes the 1021 nucleotide sequence (SEQ ID NO:11) shown in Table 4C. The GPCR1b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 44-46 and ends with a TAG codon at nucleotides 977-979, which are in bold letters in Table 4C.

**Table 4C. GPCR4b Nucleotide Sequence (SEQ ID NO:11 )**

```
AGCATTCTAACTGTCTTCCCTCAGCTCCAGATGCTGCAGAGTCATGGAAAACCAATCCAGCATTTCT
GAATTTTTCCTCCGAGGAATATCAGCGTCTCCAGAGCAACAGCAGTCCCTCTTCGGAATTTTCCTGT
GTATGTATCTTGTCACCTTGACTGGGAACCTGCTCATCATCCTGGCCATTGGCTCTGACCTGCACCT
CCACACCCCCATGTACTTTTTCTTGGCCAACCTGTCTTTTGTTGACATGGGTTTAACGTCCTCCACA
GTTACCAAGATGCTGGTGAATATACAGACTCGGCATCACACCATCTCCTATACGGGTTGCCTCACGC
AAATGTATTTCTTTCTGATGTTTGGTGATCTAGACAGCTTCTTCCTGGCTGCCATGGCGTATGACCG
CTATGTGGCCATTTGCCACCCCCTCTACTACTCCACAGTCATGAGGCCCCAAGTCTGTGCCCTAATG
CTTGCATTGTGCTGGGTCCTCACCAATATCGTTGCCCTGACTCACACGTTCCTCATGGCTCGGTTGT
CCTTCTGTGTGACTGGGGAAATTGCTCACTTTTTCTGTGACATCACTCCTGTCCTGAAGCTGTCATG
TTCTGACACCCACATCAACGAGATGATGGTTTTTGTCTTGGGAGGCACCGTACTCATCGTCCCCTTT
TTATGCATTGTCACCTCCTACATCCACATTGTGCCAGCTATCCTGAGGGTCCGAACCCGTGGTGGGG
TGGGCAAGGCCTTTTCCACCTGCAGTTCCCACCTCTGCGTTGTTTGTGTGTTCTATGGGACCCTCTT
CAGTGCCTACCTGTGTCCTCCCTCCATTGCCTCTGAAGAGAAGGACATTGCAGCAGCTGCAATGTAC
ACCATAGTGACTCCCATGTTGAACCCCTTTATCTATAGCCTAAGGAACAAGGACATGAAGGGGGCCC
TAAAGAGGCTCTTCAGTCACAGGAGTATTGTTTCCTCTTAGATGTGGTGACAGCAACATTTAATGAA
AAGACATAGGCTTGGA
```

The GPCR4b protein (SEQ ID NO:12) encoded by SEQ ID NO:11 is 311 amino acid in length, has a molecular weight of 34698.32 Daltons, and is presented using the one-letter code in Table 4D. As with GPCR4a, the most likely cleavage site for a GPCR4b peptide is between amino acids 46 and 47, i.e., at the dash in the sequence ILA-IG, based on the SignalP result. The DNA sequence and protein sequence of GPCR4a was obtained by exon linking as described in the Example 1.

**Table 4D. GPCR4b protein sequence (SEQ ID NO:12 )**

```
MENQSSISEFFLRGISASPEQQQSLFGIFLCMYLVTLTGNLLIILAIGSDLHLHTPMYFFLANLSFV
DMGLTSSTVTKMLVNIQTRHHTISYTGCLTQMYFFLMFGDLDSFFLAAMAYDRYVAICHPLYYSTVM
RPQVCALMLALCWVLTNIVALTHTFLMARLSFCVTGEIAHFFCDITPVLKLSCSDTHINEMMVFVLG
GTVLIVPFLCIVTSYIHIVPAILRVRTRGGVGKAFSTCSSHLCVVCVFYGTLFSAYLCPPSIASEEK
DIAAAAMYTIVTPMLNPFIYSLRNKDMKGALKRLFSHRSIVSS
```

**GPCR4c**

In an alternative embodiment, a GPCR4 variant is the novel GPCR4c (alternatively referred to herein as GMba64p14_D), which includes the 940 nucleotide sequence (SEQ ID NO:13) shown in Table 4E. The GPCR4c ORF begins with a Kozak consensus ATG initiation codon at nucleotides 3-5 and ends with a TAG codon at nucleotides 936-938, which are in bold letters in Table 4E.

**Table 4E. GPCR4c Nucleotide Sequence (SEQ ID NO:13 )**

```
TCATGGAAAACCAATCCAGCATTTCTGAATTTTTCCTCCGAGGAATATCAGCGCCTCCAGAGCAACA
GCAGTCCCTCTTCGGAATTTTCCTGTGTATGTATCTTGTCACCTTGACTGGGAACCTGCTCATCATC
CTGGCCATTGGCTCTGACCTGCACCTCCACACCCCCATGTACTTTTTCTTGGCCAACCTGTCTTTTG
TTGACATGGGTTTAACGTCCTCCACAGTTACCAAGATGCTGGTGAATATACAGACTCGGCATCACAC
CATCTCCTATACGGGTTGCCTCACGCAAATGTATTTCTTTCTGATGTTTGGTGATCTAGACAGCTTC
TTCCTGGCTGCCATGGCGTATGACCGCTATGTGGCCATTTGCCACCCCCTCTGCTACTCCACAGTCA
TGAGGCCCCAAGTCTGTGCCCTAATGCTTGCATTGTGCTGGGTCCTCACCAATATCGTTGCCCTGAC
TCACACGTTCCTCATGGCTCGGTTGTCCTTCTGTGTGACTGGGGAAATTGCTCACTTTTTCTGTGAC
ATCACTCCTGTCCTGAAGCTGTCATGTTCTGACACCCACATCAACGAGATGATGGTTTTTGTCTTGG
GAGGCACCGTACTCATCGTCCCCTTTTTATGCATTGTCACCTCCTACATCCACATTGTGCCAGCTAT
CCTGAGGGTCCGAACCCGTGGTGGGGTGGGCAAGGCCTTTTCCACCTGCAGTTCCCACCTCTGCGTT
GTTTGTGTGTTCTATGGGACCCTCTTCAGTGCCTACCTGTGTCCTCCCTCCATTGCCTCTGAAGAGA
AGGACATTGCAGCAGCTGCAATGTACACCATAGTGACTCCCATGTTGAACCCCTTTATCTATAGCCT
AAGGAACAAGGACATGAAGGGGGCCCTAAAGAGGCTCTTCAGTCACAGGAGTATTGTTTCCTCTTAG
AT
```

The GPCR4c protein (SEQ ID NO:14) encoded by SEQ ID NO:13 is 311 amino acid in length, has a molecular weight of 34649.7 Daltons, and is presented using the one-letter code in Table 4F. As with the other GPCR4 proteins, the most likely cleavage site for a

GPCR4C peptide is between amino acids 46 and 47, i.e., at the dash in the sequence ILA-IG, based on the SignalP result.

**Table 4F. GPCR4c protein sequence (SEQ ID NO:14 )**

MENQSSISEFFLRGISAPPEQQQSLFGIFLCMYLVTLTGNLLIILAIGSDLHLHTPMYFFLANLSFV
DMGLTSSTVTKMLVNIQTRHHTISYTGCLTQMYFFLMFGDLDSFFLAAMAYDRYVAICHPLCYSTVM
RPQVCALMLALCWVLTNIVALTHTFLMARLSFCVTGEIAHFFCDITPVLKLSCSDTHINEMMVFVLG
GTVLIVPFLCIVTSYIHIVPAILRVRTRGGVGKAFSTCSSHLCVVCVFYGTLFSAYLCPPSIASEEK
DIAAAAMYTIVTPMLNPFIYSLRNKDMKGALKRLFSHRSIVSS

## GPCR4 Clones

The amino acid sequence of GPCR4 had high homology to other proteins as shown in Table 4G.

**Table 4G. BLASTX results for GPCR4**

| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
|---|---|---|---|
| patp:AAR27868 Odorant receptor clone F5 - Rattus rattus, 313 aa. | +3 | 949 | 1.4e-94 |
| patp:AAR27876 Odorant receptor clone I15 - Rattus rattus, 314 aa. | +3 | 918 | 2.7e-91 |

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of GPCR4 has 646 of 649 bases (99%) identical to a gb:GENBANK-ID:U86216|acc:U86216.1 mRNA from Homo sapiens (Homo sapiens olfactory receptor (OR1-26) gene, partial cds). The full amino acid sequence of the GPCR4 protein of the invention was found to have 265 of 311 amino acid residues (85%) identical to, and 285 of 311 amino acid residues (91%) similar to, the 311 amino acid residue ptnr:SPTREMBL-ACC:Q9JHE2 protein from Rattus norvegicus (Rat) (GUSTATORY RECEPTOR 43).

GPCR4 also has homology to the proteins shown in the BLASTP data in Table 4H.

**Table 4H. GPCR4 BLASTP results**

| Gene Index / Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |
|---|---|---|---|---|---|
| AB038167 | Q9JHE2 rattus norvegicus (rat). gustatory receptor 43. 3/2001 | 311 | 265/311 (85%) | 285/311, (92%) | 1e-155 |
| U50947 | Q62942 rattus norvegicus (rat). taste bud receptor protein tb 334. 3/2001 | 311 | 246/311 (79%) | 277/311, (89%) | 1e-147 |

| | | | | | |
|---|---|---|---|---|---|
| U86216 | O43870 homo sapiens (human). olfactory receptor (fragment). 6/2000 | 216 | 215/216 (100%) | 216/216, (100%) | 1e-124 |
| AF101763 | Q9TU87 gorilla gorilla (gorilla). olfactory receptor (fragment). 3/2001 | 338 | 184/306 (60%) | 232/306, (76%) | 1e-104 |
| M64377; P23266 | OLF5_RAT rattus norvegicus (rat). olfactory receptor-like protein f5. 7/1993 | 313 | 178/304 (59%) | 226/304, (74%) | 1e-103 |

A multiple sequence alignment is given in Table 4I, with the GPCR4 protein being shown on line 1 in Table 4I in a ClustalW analysis, and comparing the GPCR4 protein with the related protein sequences shown in Table 4H. This BLASTP data is displayed graphically in the ClustalW in Table 4I. The residues that differs between GPCR4a, GPCR4b and GPCR4c are marked with the (o) symbol.

**Table 4I. ClustalW Analysis of GPCR4**

1. SEQ ID NO:10, GPCR4a
2. SEQ ID NO:12, GPCR4b
3. SEQ ID NO:14, GPCR4c
4. SEQ ID NO:41, AB038167 gustatory receptor 43
5. SEQ ID NO:42, M64377 olfactory receptor-like protein f5
6. SEQ ID NO:43, U50947 taste bud receptor protein tb 334
7. SEQ ID NO:52, U86216 human olfactory receptor (fragment)
8. SEQ ID NO:53, AF101763 gorilla olfactory receptor

```
                                                                  o
GPCR4a          ---------------ME--NQSSISEFFLRGISASPEQQQSLFGIFLCMYL  34
GPCR4b          ---------------ME--NQSSISEFFLRGISASPEQQQSLFGIFLCMYL  34
GPCR4c          ---------------ME--NQSSISEFFLRGISAPPEQQQSLFGIFLCMYL  34
41:AB038167     ---------------ME--NQSSVSEFFLQGISGFPEQQQLLYGLFLCMYL  34
42:M64377       ---------------MSSTNQSSVTEFLLLGLSRQPQQQQLLFLLFLIMYL  36
43:U50947       ---------------ME--NQSSVSEFFLRGISGFPEQQQLLYGLFLCMYL  34
52:U86216       ------------------------------------------------------  1
53:AF101763     MGLTQDFFPPTSDLLEGGNQTSTFEFLLSGLSDQPQQQHILFLLFLWMYV  50

GPCR4a          VTLTGNLLIILAIGSDLHLHTPMYFFLANLSFVDMGLTSSTVTKMLVNIQ  84
GPCR4b          VTLTGNLLIILAIGSDLHLHTPMYFFLANLSFVDMGLTSSTVTKMLVNIQ  84
GPCR4c          VTLTGNLLIILAIGSDLHLHTPMYFFLANLSFVDMGLTSSTVTKMLVNIQ  84
41:AB038167     VTLTGNVLIIMAIGSDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQ  84
42:M64377       ATVLGNLLIILAIGTDSRLHTPMYFFLSNLSFVDVCFSSTTVPKVLANHI  86
43:U50947       VTLTGNVLIILAIGSDPHLHTPMYFFLANLSFADMGLISSTVTKMLFNVQ  84
52:U86216       ------------------------------FVDMGLTSSTVTKMLVNIQ  19
53:AF101763     VTVAGNLLIVLAIGTDTHLHTPMYFFLASLSCADIFFTSTTVPKALVNIQ  100

                                                                  o
GPCR4a          TRHHTISYTGCLTQMYFFLMFGDLDSFFLAAMAYDRYVAICHPLCYSTVM  134
GPCR4b          TRHHTISYTGCLTQMYFFLMFGDLDSFFLAAMAYDRYVAICHPLYYSTVM  134
GPCR4c          TRHHTISYTGCLTQMYFFLMFGDLDSFFLAAMAYDRYVAICHPLCYSTVM  134
41:AB038167     TQRHTISYTGCLTQMYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIM  134
42:M64377       LGSQAISFSGCLTQLYFLAVFGNMDNFLLAVMSYDRFVAICHPLHYTTKM  136
43:U50947       TQCHTISYTGCLTQMYLFMMFGDLDSFFLAVMAYDRYVAICHPLHYSTIM  134
```

```
52:U86216      TRHHTITYTGCLTQMYFFLMFGDLDSFFLAAMAYDRYVAICHPLCYSTVM  69
53:AF101763    TQSRSISYAGCLAQLYFFLTFGDMDIFLLATMAYDRYMAICHPLHYMMIM 150

GPCR4a         RPQVCALMLALCWVLTNIVALTHTFLMARLSFCVTGEIAHFFCDITPVLK 184
GPCR4b         RPQVCALMLALCWVLTNIVALTHTFLMARLSFCVTGEIAHFFCDITPVLK 184
GPCR4c         RPQVCALMLALCWVLTNIVALTHTFLMARLSFCVTGEIAHFFCDITPVLK 184
41:AB038167    RAKVCVLMLALCWVLTNIVALTHTLLMARLSFCVVGEIAHFFCDITPVLK 184
42:M64377      TRQLCVLLVVGSWVVANMNCLLHILLMARLSFCADNMIPHFFCDGTPLLK 186
43:U50947      NARICVLMLILCWILTNVVALTHTLLMARLSFCVVGEIAHFFCDVTSVMK 184
52:U86216      RPQVCALMLALCWVLTNIVALTHTFLMARLSFCVTGEIAHFFCDITPVLK 119
53:AF101763    SLHRCALLVTACWTLTSLVAMTHTFLIFRLSFCSK-IIPDFFCDLGPLMK 199

GPCR4a         LSCSDTHINEMMVFVLGGTVLIVPFLCIVTSYIHIVPAILRVRTRGGVGK 234
GPCR4b         LSCSDTHINEMMVFVLGGTVLIVPFLCIVTSYIHIVPAILRVRTRGGVGK 234
GPCR4c         LSCSDTHINEMMVFVLGGTVLIVPFLCIVTSYIHIVPAILRVRTRGGVGK 234
41:AB038167    LSCSDTYVNELMVFALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTK 234
42:M64377      LSCSDTHLNELMLLTEGAVVMVTPFVCILISYIHITCAVLRVSSPRGGWK 236
43:U50947      LSCSDTHVNELVLSGFGGTVLMVPFVSIVISYVHIVFAVLRIQSSGGSSK 234
52:U86216      LSCSDTHINEMMVFVLGGTVLIVPFLCIVTSYIHIVPAILRVRTRGGVGK 169
53:AF101763    VSCSDTQVSELVLLFLGGAVILIPFMLILVSYIRIVSAILRAPSAQGRRK 249

GPCR4a         AFSTCSSHLCVVCVFYGTLFSAYLCPPSIASE--EKDIAAAAMYTIVTPM 282
GPCR4b         AFSTCSSHLCVVCVFYGTLFSAYLCPPSIASE--EKDIAAAAMYTIVTPM 282
GPCR4c         AFSTCSSHLCVVCVFYGTLFSAYLCPPSIASE--EKDIAAAAMYTIVTPM 282
41:AB038167    AFSTCSSHLCVVCVFYGTLFSAYLSPPSVVST--EKDIAAAAMYTVVTPM 282
42:M64377      SFSTCGSHLAVVCLFYGTVIAVYFNPSSSHLAG--RDMAAAVMYAVVTPM 284
43:U50947      AFSTCSSHLCVVCVFYGTLFSVYLFPSSVETT--EKDVAAAAMYTVVTPM 282
52:U86216      AFSTCSSHLCVVCVFYGTLFSAYLCPPSIASE--EKDIAAAAMYTIVTP- 216
53:AF101763    AFSTCGSHLVVVALFFGTVIRAYLCPSSSSSNSVDEDTAAAVMYTVVTPL 299

GPCR4a         LNPFIYSLRNKDMKGALKRLFSHRSIVSS---------- 311
GPCR4b         LNPFIYSLRNKDMKGALKRLFSHRSIVSS---------- 311
GPCR4c         LNPFIYSLRNKDMKGALKRLFSHRSIVSS---------- 311
41:AB038167    LNPFIYSLRNKDMKGALKRLLFHRRILSS---------- 311
42:M64377      LNPFIYSLRNSDMKAALRKVLAMRFPSKQ---------- 313
43:U50947      LNPFIYSLRNKDIKGALKRLLSHRRILSS---------- 311
52:U86216      ------------------------------------ 216
53:AF101763    LNPFIYSLRNKDMKAAVVRLLKGRVSFSQGQGLLLRNPR 338
```

Table 4J lists the domain description from DOMAIN analysis results against GPCR4. This indicates that the GPCR4 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:39) itself.

**Table 4J Domain Analysis of GPCR4**

| PSSMs producing significant alignments: | | Score (bits) | E value |
|---|---|---|---|
| gnl\|Pfam\|pfam00001 | 7tm_1, 7 transmembrane receptor (rhodopsin family) | 116 | 2e-27 |

```
GPCR4:   39 GNLLIILAIGSDLHLHTPMYFFLANLSFVDMGLTSSTVTKMLVNIQTRHHTISYTGCLTQ  98
            ||||+||  |      |  ||    || ||+  |+     +      |  +            |
7tm_1:    1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV  60

GPCR4:   99 MYFFLMFGDLDSFFLAAMAYDRYVAICHPLCYSTVMRPQVCALMLALCWVLTNIVALTHT  158
               |++ |       | |++ |||+||  |||  |  +  |+   +++  | |||  +++|
```

```
7tm_1:   61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPL  120

GPCR4:  159 FLMARLSFCVTGEIAHFFCDITPVLKLSCSDTHINEMMVFVLGGTVLIVPFLCIVTSYIH  218
                +                          +|  |            |  +       ++|  |  |+   |
7tm_1:  121 LFSWLRTVEEGNTTVCLIDFPEESVKRS---------YVLLSTLVGFVLPLLVILVCYTR  171

GPCR4:  219 IVPAILRVRTRGGVGK---------AFSTCSSHLCVVCVFYG----TLFSAYLCPPSIAS  265
            |+   + +          |           |          +   |   +          |   +
7tm_1:  172 ILRTLRKRARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRV  231

GPCR4:  266 EEKDIAAAAMYTIVTPMLNPFIY  288
                +          |    |||  ||
7tm_1:  232 LPTALLITLWLAYVNSCLNPIIY  254
```

The GPCR4 protein predicted here is similar to the "Olfactory Receptor-Like Protein Family", some members of which end up localized at the cell surface where they exhibit activity. Therefore, it is likely that this novel GPCR4 protein is available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

The Olfactory Receptor-like GPCR4 proteins disclosed is expressed in at least the following tissues: Apical microvilli of the retinal pigment epithelium, arterial (aortic), basal forebrain, brain, Burkitt lymphoma cell lines, corpus callosum, cardiac (atria and ventricle), caudate nucleus, CNS and peripheral tissue, cerebellum, cerebral cortex, colon, cortical neurogenic cells, endothelial (coronary artery and umbilical vein) cells, palate epithelia, eye, neonatal eye, frontal cortex, fetal hematopoietic cells, heart, hippocampus, hypothalamus, leukocytes, liver, fetal liver, lung, lung lymphoma cell lines, fetal lymphoid tissue, adult lymphoid tissue, Those that express MHC II and III nervous, medulla, subthalamic nucleus, ovary, pancreas, pituitary, placenta, pons, prostate, putamen, serum, skeletal muscle, small intestine, smooth muscle (coronary artery in aortic) spinal cord, spleen, stomach, taste receptor cells of the tongue, testis, thalamus, and thymus tissue. This information was derived by determining the tissue sources of the sequences that were included in the invention including but not limited to SeqCalling sources, Public EST sources, Literature sources, and/or RACE sources.

This is by no way limiting in that olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types. Further tissue expression analysis is provided in the Examples.

The nucleic acids and proteins of GPCR4 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR4 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR4 epitope is from about amino acids 1 to 25. In additional embodiments, GPCR4 epitopes are from about amino acids 75 to 100, from about amino acids 230 to 240 and from about amino acids 285 to 311.

**GPCR5**

A fifth GPCR-like protein of the invention, referred to herein as GPCR5, is an Olfactory Receptor ("OR")-like protein. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR5 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Two alternative novel GPCR5 nucleic acids and encoded polypeptides are provided, namely GPCR5a and GPCR5b.

**GPCR5a**

In one embodiment, a GPCR5 variant is the novel GPCR5a (alternatively referred to herein as CG50385-03), which includes the 1012 nucleotide sequence (SEQ ID NO:15) shown in Table 5A. The DNA sequence and protein sequence for GPCR5a or one of its splice forms was obtained solely by exon linking. A GPCR5a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 59-61 and ends with a TGA codon at nucleotides 1001-1003, shown in bold in Table 5A.

**Table 5A. GPCR5a Nucleotide Sequence (SEQ ID NO:15)**

GGTTTGTTTCTGCCTTTTTCAATGTCCCTCTATTTCCAGCAGAGAGAAGACTGTCAGC**ATG**AAGAGG
GAGAATCAGAGCAGTGTGTCTGAGTTCCTCCTCCTGGACCTCCCCATCTGGCCAGAGCAGCAGGCTG
TGTTCTTCGCCCTGTTCTTGGGCATGTGCCTGATCACGGTGCTGGGGAACCTGCTCATCATCCTGCT
CATCCGGCTGGACTCTCACCTTCACACCCCCATGTTCTTCTTCCTCAGCCACTTGGCTCTCACTGAC
ATCTCCCTTTCATCTGTCACTGTCCCAAAGATGTTATTAAGCATGCAAACTCAGGATCAATCCATTC
TTTATGCAGGGTGTGTAACTCAGATGTATTTTTTCATATTTTTCACTGATCTAGACAATTTCCTTCT
CACTACAATGGCATACGATCGGTATGTGGCCATCTGTCACCCCCTCCGCTACACCACTATCATGAAA
GAGGGACTGTGTAACTTACTAGTCACTGTGTCCTGGATCCTCTCCTGTACCAATGCCCTGTCTCACA
CTCTCCTCCTGGCCCAGCTGTCCTTTTGTGCTGACAACACCATCCCCCATTTCTTCTGTGATCTTGT
TGCCCTACTCAAGCTCTCATGCTCAGACATCTCCCTCAATGAGCTGGTCATTTTCACAGTGGGACAG
GCAGTCATTACTCTACCACTAATATGCATCTTGATCTCTTATGGCCACATTGGGGTCACCATCCTCA

AGGCTCCATCTACTAAGGGCATCTTCAAAGCTTTGTCCACCTGTGGCTCTCACCTCTCTGTGGTGTC
TCTGTATTATGGCACAATTATTGGACTGTATTTTCTCCCCTCATCCAGTGCCTCCAGTGACAAGGAC
GTAATTGCCTCTGTGATGTACACGGTGATCACCCCATTGCTGAATCCCTTCATTTATAGCCTAAGGA
ACAGGGACATAAAGGGAGCCCCTGGAGAGACTCTTCAACAGGGCAACAGTCTTATCTCAATG**TGA**TT
TACTCTT

The GPCR5a polypeptide (SEQ ID NO:16) encoded by SEQ ID NO:15 is 314 aa in length, has a molecular weight of 34809.23 Daltons, and is presented using the one-letter amino acid code in Table 5B. The Psort profile for both GPCR5a and GPCR5b predicts that these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR5 polypeptide is located to the Golgi body with a certainty of 0.400, the endoplasmic reticulum (membrane) with a certainty of 0.300, or a microbody (peroxisome) with a certainty of 0.300. The Signal P predicts a likely cleavage site for a GPCR5 peptide is between positions 41 and 42, *i.e.*, at the slash in the sequence VLG/NL.

**Table 5B. GPCR5a protein sequence (SEQ ID NO:16)**

MKRENQSSVSEFLLLDLPIWPEQQAVFFALFLGMCLITVLG/NLLIILLIRLDSHLHTPMFFFLSHL
ALTDISLSSVTVPKMLLSMQTQDQSILYAGCVTQMYFFIFFTDLDNFLLTTMAYDRYVAICHPLRYT
TIMKEGLCNLLVTVSWILSCTNALSHTLLLAQLSFCADNTIPHFFCDLVALLKLSCSDISLNELVIF
TVGQAVITLPLICILISYGHIGVTILKAPSTKGIFKALSTCGSHLSVVSLYYGTIIGLYFLPSSSAS
SDKDVIASVMYTVITPLLNPFIYSLRNRDIKGAPGETLQQGNSLISM

**GPCR5b**

In an alternative embodiment, a GPCR5 variant is the novel GPCR5b (alternatively referred to herein as GMba64p14_E), which includes the 971 nucleotide sequence (SEQ ID NO:17) shown in Table 5C. The GPCR5b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 2-4 and ends with a TGA codon at nucleotides 941-9343, which are in bold letters in Table 5C.

**Table 5C. GPCR5b Nucleotide Sequence (SEQ ID NO:17)**

C**ATG**AAGAGGGAGAATCAGAGCAGTGTGTCTGAGTTCCTCCTCCTGGACCTCCCCATCTGGCCAGAG
CAGCAGGCTGTGTTCTTCACCCTGTTCTTGGGCATGTACCTGATCACGGTGCTGGGGAACCTGCTCA
TCATCCTGCTCATCCGGCTGGACTCTCACCTTCACACCCCCATGTTCTTCTTCCTCAGCCACTTGGC
TCTCACTGACATCTCCCTTTCATCTGTCACTGTCCCAAAGATGTTATTAAGCATGCAAACTCAGGAT
CAATCCATTCTTTATGCAGGGTGTGTAACTCAGATGTATTTTTTCATATTTTTCACTGATCTAGACA
ATTTCCTTCTCACTTCAATGGCATACGATCGGTATGTGGCCATCTGTCACCCCCTCCGCTACACCAC
TATCATGAAAGAGGGACTGTGTAACTTACTAGTCACTGTGTCCTGGATCCTCTCCTGTACCAATGCC
CTGTCTCACACTCTCCTCCTGGCCCAGCTGTCCTTTTGTGCTGACAACACCATCCCCCATTTCTTCT
GTGATCTTGTTGCCCTACTCAAGCTCTCATGCTCAGACATCTCCCTCAATGAGCTGGTCATTTTCAC
AGTGGGACAGGCAGTCATTACTCTACCACTAATATGCATCTTGATCTCTTATGGCCACATTGGGGTC
ACCATCCTCAAGGCTCCATCTACTAAGGGCATCTTCAAAGCTTTGTCCACCTGTGGCTCTCACCTCT

CTGTGGTGTCTCTGTATTATGGCACAATTATTGGACTGTATTTTCTCCCCTCATCCAGTGCCTCCAG
TGACAAGGACGTAATTGCCTCTGTGATGTACACGGTGATCACCCCATTGCTGAATCCCTTCATTTAT
AGCCTAAGGAACAGGGACATAAAGGGAGCCCTGGAGAGACTCTTCAACAGGGCAACAGTCTTATCTC
AATGACATTTACTCTTCTTTATAACAGACATAT

The GPCR5b protein (SEQ ID NO:18) encoded by SEQ ID NO:17 is 313 amino acids in length, has a molecular weight of 34958.8 Daltons, and is presented using the one-letter code in Table 5D. As with GPCR5a, the most likely cleavage site for a GPCR5b peptide is between amino acids 41 and 42, *i.e.*, at the slash in the sequence VLG/NL, based on the SignalP result.

**Table 5D. GPCR5b protein sequence (SEQ ID NO:18)**

MKRENQSSVSEFLLLDLPIWPEQQAVFFTLFLGMYLITVLG/NLLIILLIRLDSHLHTPMFFFLSHL
ALTDISLSSVTVPKMLLSMQTQDQSILYAGCVTQMYFFIFFTDLDNFLLTSMAYDRYVAICHPLRYT
TIMKEGLCNLLVTVSWILSCTNALSHTLLLAQLSFCADNTIPHFFCDLVALLKLSCSDISLNELVIF
TVGQAVITLPLICILISYGHIGVTILKAPSTKGIFKALSTCGSHLSVVSLYYGTIIGLYFLPSSSAS
SDKDVIASVMYTVITPLLNPFIYSLRNRDIKGALERLFNRATVLSQ

## GPCR5c

In an alternative embodiment, a GPCR5 variant is the novel GPCR5c (alternatively referred to herein as CG50385-01), which includes the 1051 nucleotide sequence (SEQ ID NO:108) shown in Table 5E. The GPCR5c ORF begins with a Kozak consensus ATG initiation codon at nucleotides 31-34 and ends with a TGA codon at nucleotides 970-972, which are in bold letters in Table 5E.

**Table 5E. GPCR5c Nucleotide Sequence (SEQ ID NO:108)**

TCTATTTCCAGCAGAGAGAAGACTGTCAGC**ATG**AAGAGGGAGAATCAGAGCAGTGTGTCTGAGTTCC
TCCTCCTGGACCTCCCCATCTGGCCAGAGCAGCAGGCTGTGTTCTTCACCCTGTTCTTGGGCATGTA
CCTGATCACGGTGCTGGGGAACCTGCTCATCATCCTGCTCATCCGGCTGGACTCTCACCTTCACACC
CCCATGTTCTTCTTCCTCAGCCACTTGGCTCTCACTGACATCTCCCTTTCATCTGTCACTGTCCCAA
AGATGTTATTAAGCATGCAAACTCAGGATCAATCCATTCTTTATGCAGGGTGTGTAACTCAGATGTA
TTTTTTCATATTTTTCACTGATCTAGACAATTTCCTTCTCACTTCAATGGCATACGATCGGTATGTG
GCCATCTGTCACCCCCTCCGCTACACCACTATCATGAAAGAGGGACTGTGTAACTTACTAGTCACTG
TGTCCTGGATCCTCTCCTGTACCAATGCCCTGTCTCACACTCTCCTCCTGGCCCAGCTGTCCTTTTG
TGCTGACAACACCATCCCCCATTTCTTCTGTGATCTTGTTGCCCTACTCAAGCTCTCATGCTCAGAC
ATCTCCCTCAATGAGCTGGTCATTTTCACAGTGGGACAGGCAGTCATTACTCTACCACTAATATGCA
TCTTGATCTCTTATGGCCACATTGGGGTCACCATCCTCAAGGCTCCATCTACTAAGGGCATCTTCAA
AGCTTTGTCCACCTGTGGCTCTCACCTCTCTGTGGTGTCTCTGTATTATGGCACAATTATTGGACTG
TATTTTCTCCCCTCATCCAGTGCCTCCAGTGACAAGGACGTAATTGCCTCTGTGATGTACACGGTGA
TCACCCCATTGCTGAATCCCTTCATTTATAGCCTAAGGAACAGGGACATAAAGGGAGCCCTGGAGAG
ACTCTTCAACAGGGCAACAGTCTTATCTCAA**TGA**CATTTACTCTTCTTTATAACAGACATATGTACT
GACCTATTTCCAGATCATAGATCCTTACTTCTGATCCCAGCAAGGG

The GPCR5c protein (SEQ ID NO:109) encoded by SEQ ID NO:108 is 313 amino acids in length, has a molecular weight of 34958.8 Daltons, and is presented using the one-letter code in Table 5F. As with GPCR5a, the most likely cleavage site for a GPCR5c peptide is between amino acids 41 and 42, *i.e.*, at the slash in the sequence VLG/NL, based on the SignalP result.

**Table 5F. GPCR5c protein sequence (SEQ ID NO:109)**

MKRENQSSVSEFLLLDLPIWPEQQAVFFTLFLGMYLITVLG/NLLIILLIRLDSHLHTPMFFFLSHL
ALTDISLSSVTVPKMLLSMQTQDQSILYAGCVTQMYFFIFFTDLDNFLLTSMAYDRYVAICHPLRYT
TIMKEGLCNLLVTVSWILSCTNALSHTLLLAQLSFCADNTIPHFFCDLVALLKLSCSDISLNELVIF
TVGQAVITLPLICILISYGHIGVTILKAPSTKGIFKALSTCGSHLSVVSLYYGTIIGLYFLPSSSAS
SDKDVIASVMYTVITPLLNPFIYSLRNRDIKGALERLFNRATVLSQ

**GPCR5 Clones**

Unless specifically addressed as GPCR5a or GPCR5b, any reference to GPCR5 is assumed to encompass all variants.

The amino acid sequence of GPCR2 had high homology to other proteins as shown in Table 5G.

**Table 5G. BLASTX results for GPCR5**

| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
|---|---|---|---|
| patp:AAR27876 Odorant receptor clone I15 - Rattus rattus, 314 aa. | +2 | 999 | 7.1e-100 |
| p:AAR27874 Odorant receptor clone I9 - Rattus rattus, 314 aa. | +2 | 977 | 1.5e-97 |

The novel GPCR5 nucleic acid sequences were mapped to chromosome 17. This assignment was made using mapping information associated with genomic clones, public genes and ESTs sharing sequence identity with the disclosed sequence and CuraGen Corporation's Electronic Northern bioinformatic tool.

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of GPCR5a has 692 of 988 bases (70%) identical to a gb:GENBANK-ID:HSHGM071|acc:X64994.1 mRNA from Homo sapiens (H.sapiens HGMP07I gene for olfactory receptor). The full amino acid sequence of the GPCR5a protein was found to have 192 of 314 amino acid residues (61%) identical to, and 244 of 314 amino acid residues (77%) similar to, the 314 amino acid residue ptnr:SWISSNEW-ACC:P30953 protein from Homo

sapiens (Human) (OLFACTORY RECEPTOR 1E1 (OLFACTORY RECEPTOR-LIKE PROTEIN HGMP07I) (OLFACTORY RECEPTOR 17-2/17-32) (OR17-2) (OR17-32).

Additional BLASTP results are shown in Table 5H.

**Table 5H. GPCR5 BLASTP results**

| Gene Index / Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |
|---|---|---|---|---|---|
| AF101741 | Q9TUA1 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 195/307 (64%) | 247/307, (80%) | 1e-113 |
| AF101760; AF101761 | Q9TQX4 gorilla gorilla (gorilla). olfactory receptor. 3/2001 | 314 | 195/304 (64%) | 246/304, (81%) | 1e-112 |
| AF101739 | Q9TUA3 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 193/304 (63%) | 249/304, (82%) | 1e-112 |
| AF101740 | Q9TUA2 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 194/304 (64%) | 246/304, (81%) | 1e-112 |
| AF101730 | Q9TUA9 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 193/307 (63%) | 248/307, (81%) | 1e-112 |

A multiple sequence alignment is given in Table 5J, with the GPCR5 protein of the invention being shown on line 1, in a ClustalW analysis comparing GPCR5 with related protein sequences, shown in Table 5H. The residues that differs between GPCR5a and GPCR5b are marked with the (o) symbol.

**Table 5J. Information for the ClustalW proteins:**

1. SEQ ID NO:16, GPCR5a
2. SEQ ID NO:18, GPCR5b
3. SEQ ID NO:44, AF101730 chimpanzee olfactory receptor
4. SEQ ID NO:45, AF101761 gorilla olfactory receptor
5. SEQ ID NO:46, AF101741 chimpanzee olfactory receptor
6. SEQ ID NO:54, AF101739 chimpanzee olfactory receptor
7. SEQ ID NO:55, AF101740 chimpanzee olfactory receptor

```
                                                    o       o
GPCR5a          MKRENQSSVSEFLLLDLPIWPEQQAVFFALFLGMCLITVLGNLLIILLIR  50
GPCR5b          MKRENQSSVSEFLLLDLPIWPEQQAVFFTLFLGMYLITVLGNLLIILLIR  50
44:AF101730   --MENQSSVSEFFLQGISGFPEQQQLLYGLFLCMYLVTLTGNVLIIMAIG  48
45:AF101761   MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR  50
46:AF101741   MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR  50
54:AF101739   MMEQNQTSTSDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR  50
55:AF101740   MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR  50

GPCR5a          LDSHLHTPMFFFLSHLALTDISLSSVTVPKMLLSMQTQDQSILYAGCVTQ 100
GPCR5b          LDSHLHTPMFFFLSHLALTDISLSSVTVPKMLLSMQTQDQSILYAGCVTQ 100
44:AF101730   SDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQTQRHTISYTGCLTQ  98
45:AF101761   LDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCLTQ 100
```

```
46:AF101741   LDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCLTQ 100
54:AF101739   LDSHLHTPVYLFLSNLSFSDLCFSSVTMPKLLQNMQSQNPSIPYADCLTQ 100
55:AF101740   LDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCLTQ 100

                          o
GPCR5a        MYFFIFFTDLDNFLLTTMAYDRYVAICHPLRYTTIMKEGLCNLLVTVSWI 150
GPCR5b        MYFFIFFTDLDNFLLTSMAYDRYVAICHPLRYTTIMKEGLCNLLVTVSWI 150
44:AF101730   MYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMRAKVCVLMLALCWV 148
45:AF101761   MYFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSLVALSWV 150
46:AF101741   MYFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSLVALSWV 150
54:AF101739   MYFFLYFSDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSVVALSWV 150
55:AF101740   MHFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSVVALSWV 150

GPCR5a        LSCTNALSHTLLLAQLSFCADNTIPHFFCDLVALLKLSCSDISLNELVIF 200
GPCR5b        LSCTNALSHTLLLAQLSFCADNTIPHFFCDLVALLKLSCSDISLNELVIF 200
44:AF101730   LTNIVALTHTLLMARLSFCVVGEIAHFFCDITPVLKLSCSDTYVNELMVF 198
45:AF101761   LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200
46:AF101741   LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200
54:AF101739   LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200
55:AF101740   LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200

GPCR5a        TVGQAVITLPLICILISYGHIGVTILKAPSTKGIFKALSTCGSHLSVVSL 250
GPCR5b        TVGQAVITLPLICILISYGHIGVTILKAPSTKGIFKALSTCGSHLSVVSL 250
44:AF101730   ALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTKAFSTCSSHLCVVCV 248
45:AF101761   IMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSHLSVVSL 250
46:AF101741   IMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSHLSVVSL 250
54:AF101739   IMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKALSTCGSHLSVVSL 250
55:AF101740   IMGGLIVVIPFLLILGSYARIVSSILKVPSFKGICKALSTCGSHLSVVSL 250

GPCR5a        YYGTIIGLYFLPSSSASSDKDVIASVMYTVITPLLNPFIYSLRNRDIKGA 300
GPCR5b        YYGTIIGLYFLPSSSASSDKDVIASVMYTVITPLLNPFIYSLRNRDIKGA 300
44:AF101730   FYGTLFSAYLSPPSVVSTEKDIAAAAMYTVVTPMLNPFIYSLRNKDMKGA 298
45:AF101761   FYGTIIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRNRDMKGA 300
46:AF101741   FYGTIIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRNRDMKGA 300
54:AF101739   FYGTVIGLYLCPSANSSTLKDTVMAMMYTVVTPMLNPFIYSLRNRDMKGA 300
55:AF101740   FYGTVIGLYLCPSANSSTLKDTVMAMMYTVVTPMLNPFIYSLRNRDMKGA 300

              oooooooooooooo
GPCR5a        PGETLQQGNSLISM 314
GPCR5b        L-ERLFNRATVLSQ 313
44:AF101730   LKRLLFHRRILSS- 311
45:AF101761   LERVICKRKNPFLL 314
46:AF101741   LERVIXKRKNPFLL 314
54:AF101739   LERVICKRKNPFLI 314
55:AF101740   LERVICKRKNPFLL 314
```

DOMAIN results for GPCR5 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. The results are listed in Table 5H with the statistics and domain description. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to have significant homology to GPCR5. An alignment of GPCR5 residues 41-289 with 7tm_1 residues 1-254 (SEQ ID NO:39) are shown in Table 5K.

## Table 5K. DOMAIN results for GPCR5

| PSSMs producing significant alignments: | Score (bits) | E value |
|---|---|---|
| gnl\|Pfam\|pfam00001 7tm_1, 7 transmembrane receptor (rhodopsin family) | 108 | 6e-25 |

```
GPCR5:   41 GNLLIILLIRLDSHLHTPMFFFLSHLALTDISLSSVTVPKMLLSMQTQDQSILYAGCVTQ  100
7tm_1:    1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV   60

GPCR5:  101 MYFFIFFTDLDNFLLTTMAYDRYVAICHPLRYTTIMKEGLCNLLVTVSWILSCTNALSHT  160
7tm_1:   61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPL  120

GPCR5:  161 LLLAQLSFCADNTIPHFFCD-----LVALLKLSCSDISLNELVIFTVGQAVITLPLICIL  215
7tm_1:  121 LFSWLRTVEEGNTTVCLIDFPEESVKRSYVLLSTLVGFVLPLLVILVCYTRILRTLRKRA  180

GPCR5:  216 ISYGHIGVTILKAPSTKGIFKALSTCGSHLSVVSLYYGTIIGLYFLP----SSSASSDKD  271
7tm_1:  181 RSQ-----RSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRVLPTA  235

GPCR5:  272 VIASVMYTVITPLLNPFIY  290
7tm_1:  236 LLITLWLAYVNSCLNPIIY  254
```

The olfactory receptor-like gene GPCR5a disclosed in this invention is expressed in at least the following tissues: adrenal gland, bone marrow, brain - amygdala, brain - cerebellum, brain - hippocampus, brain - substantia nigra, brain - thalamus, brain -whole, fetal brain, fetal kidney, fetal liver, fetal lung, heart, kidney, lymphoma - Raji, mammary gland, pancreas, pituitary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen, stomach, testis, thyroid, trachea and uterus. Expression information was derived from the tissue sources of the sequences that were included in the derivation of the sequence of GPCR5a (CuraGen Acc. No. CG50385-03). This is by no way limiting in that olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types. Further tissue expression analysis is provided in the Examples.

The nucleic acids and proteins of GPCR5 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further above.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to

methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR5 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR5 epitope is from about amino acids 1 to 20. In other specific embodiments, GPCR5 epitopes are from about amino acids 75 to 105, from about amino acids 115 to 130 and from about amino acids 275 to 313.

**GPCR6**

A further GPCR-like protein of the invention, referred to herein as GPCR6, is an Olfactory Receptor ("OR")-like protein. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR6 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Three alternative novel GPCR6 nucleic acids and encoded polypeptides are provided, namely GPCR6a, GPCR6b and GPCR6c.

**GPCR6a**

In one embodiment, a GPCR6 variant is the novel GPCR6a (alternatively referred to herein as ba460n11_da1), which includes the 996 nucleotide sequence (SEQ ID NO:19) shown in Table 6A. A GPCR6a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 2-4 and ends with a TGA codon at nucleotides 941-943. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 6A, and the start and stop codons are in bold letters.

**Table 6A. GPCR6a Nucleotide Sequence (SEQ ID NO:19)**

T**ATG**AGCCCTGAGAACCAGAGCAGCGTGTCCGAGTTCCTCCTTCTGGGCCTCCCCATCCGGCCAGAG
CAGCAGGCTGTGTTCTTCACCCTGTTCCTGGGCATGTACCTGACCACGGTGCTGGGGAACCTGCTCA
TCATGCTGCTCATCCAGCTGGACTCTCACCTTCACACCCCCATGTACTTCTTCCTCAGCCACTTGGC
TCTCACTGACATCTCCTTTTCATCTGTCACTGTCCCTAAGATGCTGATGGACATGCGGACTAAGTAC
AAATCGATCCTCTATGAGGAATGCATTTCTCAGATGTATTTTTTTATATTTTTTACTGACCTGGACA
GCTTCCTTATTACATCAATGGCATATGACCGATATGTTGCCATATGTCACCCTCTCCACTACACTGT
CATCATGAGGGAAGAGCTCTGTGTCTTCTTAGTGGCTGTATCTTGGATTCTGTCTTGTGCCAGCTCC
CTCTCTCACACCCTTCTCCTGACCCGGCTGTCTTTCTGTGCTGCGAACACCATCCCCCATGTCTTCT
GTGACCTTGCTGCCCTGCTCAAGCTGTCCTGCTCAGATATCTTCCTCAATGAGCTGGTCATGTTCAC
AGTAGGGGTGGTGGTCATTACCCTGCCATTCATGTGTATCCTGGTATCATATGGCTACATTGGGGCC
ACCATCCTGAGGGTCCCTTCAACCAAAGGGATCCACAAAGCATTGTCCACATGTGGCTCCCATCTCT
CTGTGGTGTCTCTCTATTATGGGTCAATATTTGGCCAGTACCTTTTCCCGACTGTAAGCAGTTCTAT
TGACAAGGATGTCATTGTGGCTCTCATGTACACGGTGGTCACACCCATGTTGAACCCCTTTATCTAC
AGCCTTAGGAACAGGGGCATGAAAGAGGCCCTTGGGAAACTCTTCAGTAGAGCAACATTTTTCTCTT
GGT**GA**CATCTGACTTTTTAAAAAATTAGAATCTCATTTTGGTTTATCTCATGTATTTG

The sequence of GPCR6a was derived by laboratory exon linking as described in Example 1.

The GPCR6a polypeptide (SEQ ID NO:20) encoded by SEQ ID NO:19 is 313 aa in length, has a molecular weight of 35325.04 Daltons, and is presented using the one-letter amino acid code in Table 6B. The Psort profile for GPCR6a, GPCR6b and GPCR6c predicts that these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.640. In alternative embodiments, a GPCR6 polypeptide is located at the Golgi body with a certainty of 0.4600, the endoplasmic reticulum (membrane) with a certainty of 0.3700, or to endoplamic reticulum (lumen) with a certainty of 0.1000. The Signal P predicts a likely cleavage site for a GPCR6 peptide is between positions 41 and 42, *i.e.*, at the dash in the sequence VLG-NL.

**Table 6B. GPCR6a protein sequence (SEQ ID NO:20)**

```
MSPENQSSVSEFLLLGLPIRPEQQAVFFTLFLGMYLTTVLGNLLIMLLIQLDSHLHTPMYF
FLSHLALTDISFSSVTVPKMLMDMRTKYKSILYEECISQMYFFIFFTDLDSFLITSMAYDR
YVAICHPLHYTVIMREELCVFLVAVSWILSCASSLSHTLLLTRLSFCAANTIPHVFCDLAA
LLKLSCSDIFLNELVMFTVGVVVITLPFMCILVSYGYIGATILRVPSTKGIHKALSTCGSH
LSVVSLYYGSIFGQYLFPTVSSSIDKDVIVALMYTVVTPMLNPFIYSLRNRGMKEALGKLF
SRATFFSW
```

**GPCR6b**

In an alternative embodiment, a GPCR6 variant is the novel GPCR6b (alternatively referred to herein as GMba64p14_F), which includes the 978 nucleotide sequence (SEQ ID NO:21) shown in Table 6C. The GPCR6b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 2-4 and ends with a TGA codon at nucleotides 941-943, which are in bold letters in Table 6C.

**Table 6C. GPCR6b Nucleotide Sequence (SEQ ID NO:21 )**

```
TATGAGCCCTGAGAACCAGAGCAGCGTGTCCGAGTTCCTCCTTCTGGGCCTCCCCATCCGGCCAGAG
CAGCAGGCTGTGTTCTTCACCCTGTTCCTGGGCATGTACCTGACCACGGTGCTGGGGAACCTGCTCA
TCATGCTGCTCATCCAGCTGGACTCTCACCTTCACACCCCCATGTACTTCTTCCTCAGCCACTTGGC
TCTCACTGACATCTCCTTTTCATCTGTCACTGTCCCTAAGATGCTGATGGACATGCGGACTAAGTAC
AAATCGATCCTCTATGAGGAATGCATTTCTCAGATGTATTTTTTTATATTTTTTACTGACCTGGACA
GCTTCCTTATTACATCAATGGCATATGACCGATATGTTGCCATATGTCACCCTCTCCACTACACTGT
CATCATGAGGGAAGAGCTCTGTGTCTTCTTAGTGGCTGTATCTTGGATTCTGTCTTGTGCCAGCTCC
CTCTCTCACACCCTTCTCCTGACCCGGCTGTCTTTCTGTGCTGCGAACACCATCCCCCATGTCTTCT
GTGACCTTGCTGCCCTGCTCAAGCTGTCCTGCTCAGATATCTTCCTCAATGAGCTGGTCATGTTCAC
AGTAGGGGTGGTGGTCATTACCCTGCCATTCATGTGTATCCTGGTATCATATGGCTACATTGGGGCC
ACCATCCTGAGGGTCCCTTCAACCAAAGGGATCCACAAAGCATTGTCCACATGTGGCTCCCATCTCT
CTGTGGTGTCTCTCTATTATGGGTCAATATTTGGCCAGTACCTTTTCCCGACTGTAAGCAGTTCTAT
TGACAAGGATGTCATTGTGGCTCTCATGTACACGGTGGTCACACCCATGTTGAACCCCTTTATCTAC
AGCCTTAGGAACAGGGACATGAAAGAGGCCCTTGGGAAACTCTTCAGTAGAGCAACATTTTTCTCTT
```

GGTGACATCTGACTTTTTAAAAAATTAGAATCTCATTTTG

The GPCR6b protein (SEQ ID NO:22) encoded by SEQ ID NO:21 is 313 amino acid in length, has a molecular weight of 35384.5 Daltons, and is presented using the one-letter code in Table 6D. As with GPCR6a, the most likely cleavage site for a GPCR6b peptide is between amino acids 48 and 49, *i.e.*, at the dash in the sequence ILA-IH, based on the SignalP result.

**Table 6D. GPCR6b protein sequence (SEQ ID NO:22 )**

MSPENQSSVSEFLLLGLPIRPEQQAVFFTLFLGMYLTTVLGNLLIMLLIQLDSHLHTPMYFFLSHLA
LTDISFSSVTVPKMLMDMRTKYKSILYEECISQMYFFIFFTDLDSFLITSMAYDRYVAICHPLHYTV
IMREELCVFLVAVSWILSCASSLSHTLLLTRLSFCAANTIPHVFCDLAALLKLSCSDIFLNELVMFT
VGVVVITLPFMCILVSYGYIGATILRVPSTKGIHKALSTCGSHLSVVSLYYGSIFGQYLFPTVSSSI
DKDVIVALMYTVVTPMLNPFIYSLRNRDMKEALGKLFSRATFFSW

**GPCR6c**

In an alternative embodiment, a GPCR6 variant is the novel GPCR6c (alternatively referred to herein as ba460n11_da2 or 147307499), which includes the 925 nucleotide sequence (SEQ ID NO:23) shown in Table 6E. GPCR6c is a partial ORF starting at nucleotide position 27 relative to SEQ ID NOs:19 and 21, and ends with a TGA codon at nucleotides 915-917, which are in bold letters in Table 6E.

**Table 6E. GPCR6c Nucleotide Sequence (SEQ ID NO:23 )**

TGTCCGAGTTCCTCCTTCTGGGCCTCCCCATCCGGCCAGAGCAGCAGGCTGTGTTCTTCACCCTGTT
CCTGGGCATGTACCTGACCACGGTGCTGGGGAACCTGCTCATCATGCTGCTCATCCAGCTGGACTCT
CACCTTCACACCCCCATGTACTTCTTCCTCAGCCACTTGGCTCTCACTGACATCTCCTTTTCATCTG
TCACTGTCCCTAAGATGCTGATGGACATGCGGACTAAGTACAAATCGATCCTCTATGAGGAATGCAT
TTCTCAGATGTATTTTTTTATATTTTTTACTGACCTGGACAGCTTCCTTATTACATCAATGACATAT
GACCGATATGTTGCCATATGTCACCCTCTCCACTACACTGTCATCATGAGGGAAGAGCTCTGTGTCT
TCTTAGTGGCTGTATCTTGGATTCTGTCTTGTGCCAGCTCCCTCTCTCACACCCTTCTCCTGACCCG
GCTGTCTTTCTGTGCTGCGAACACCATCCCCCATGTCTTCTGTGACCTTGCTGCCCTGCTCAAGCTG
TCCTGCTCAGATATCTTCCTCAATGAGCTGGTCATGTTCACAGTAGGGGTGGTGGTCATTACCCAGC
CATTCATGTGTATCCTGGTATCATATGGCTACATTGGGGCCACCATCCTGAGGGTCCCTTCAACCAA
AGGGATCCACAAAGCATTGTCCACATGTGGCTCCCATCTCTCTGTGGTGTCTCTCTATTATGGGTCA
ATATTTGGCCAGTACCTTTTCCCGACTGTAAGCAGTTCTATTGACAAGGATGTCATTGTGGCTCTCA
TGTACACGGTGGTCACACCCATGTTGAACCCCTTTATCTACAGCCTTAGGAACAGGGACATGAAAGA
GGCCCTTGGGAAACTCTTCAGTAGAGCAACATTTTTCTCTTGG**TGA**CATCTGAC

The GPCR6c protein (SEQ ID NO:24) encoded by SEQ ID NO:23 is 304 amino acid in length, has a molecular weight of 34468.10 Daltons, and is presented using the one-letter code in Table 1F. As with the other GPCR6 proteins, the most likely cleavage site for a GPCR6c peptide is at the dash in the sequence ILA-IH, based on the SignalP result.

| Table 6F. GPCR6c protein sequence (SEQ ID NO:24 ) |
|---|
| SEFLLLGLPIRPEQQAVFFTLFLGMYLTTVLGNLLIMLLIQLDSHLHTPMYFFLSHLALTDISFSSV<br>TVPKMLMDMRTKYKSILYEECISQMYFFIFFTDLDSFLITSMTYDRYVAICHPLHYTVIMREELCVF<br>LVAVSWILSCASSLSHTLLLTRLSFCAANTIPHVFCDLAALLKLSCSDIFLNELVMFTVGVVVITQP<br>FMCILVSYGYIGATILRVPSTKGIHKALSTCGSHLSVVSLYYGSIFGQYLFPTVSSSIDKDVIVALM<br>YTVVTPMLNPFIYSLRNRDMKEALGKLFSRATFFSW |

The DNA sequence and protein sequence of GPCR6c was obtained solely by exon linking process.

**GPCR6 Clones**

Unless specifically addressed as GPCR6a, GPCR6b or GPCR6c, any reference to GPCR6 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant, the residue position with respect to the "a" variant, and the residue in the "b" variant.

The amino acid sequence of GPCR6 had high homology to other proteins as shown in Table 6G.

| Table 6G. BLASTX results for GPCR6 | | | |
|---|---|---|---|
| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
| patp:AAR27876 Odorant receptor clone I15 - Rattus rattus, 314 aa. | +2 | 1012 | 3.0e-101 |

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of GPCR6c has 486/487 (99%) identical to a GENBANK-ID:AF179737| acc:AF179737.1 Pan troglodytes olfactory receptor(PTR212) gene, partial cds - Pan troglodytes, 487 bp. Furthermore, the full amino acid sequence of the GPCR6a protein has 191 of 313 amino acid residues (61%) identical to, and 251 of 313 amino acid residues (80%) similar to, the 314 amino acid residue ptnr:SWISSNEW-ACC:P30953 Olfactory Receptor 1E1 protein from human (OLFACTORY RECEPTOR-LIKE PROTEIN HGMP07I) (OLFACTORY RECEPTOR 17-2) (OR17-2).

Additional BLASTP results are shown in Table 6H.

| Table 6H - GPCR6 BLASTP results | | | | | |
|---|---|---|---|---|---|
| Gene Index/ Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |

| | | | | | |
|---|---|---|---|---|---|
| AF101730 | Q9TUA9 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 193/313 (62%) | 255/313, (81%) | 1e-112 |
| AF101739 | Q9TUA3 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 189/307 (62%) | 253/307, (82%) | 1e-111 |
| AF101741 | Q9TUA1 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 190/307 (62%) | 250/307, (81%) | 1e-111 |
| AF101760; AF101761 | Q9TQX4 gorilla gorilla (gorilla). olfactory receptor. 3/2001 | 314 | 190/307 (62%) | 250/307, (81%) | 1e-111 |
| M64392 | OLF1_RAT rattus norvegicus (rat). olfactory receptor-like protein i15. 7/1993 | 314 | 188/301 (62%) | 245/301, (81%) | 1e-110 |
| AF101740 | Q9TUA2 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 189/307 (62%) | 250/307, (81%) | 1e-110 |

A multiple sequence alignment is given in Table 6I, with the GPCR6 protein of the invention being shown on lines 1-3, in a ClustalW analysis comparing GPCR6 with the related protein sequences listed in Table 6H.

**Table 6I. Information for the ClustalW proteins:**

1. SEQ ID NO:20, GPCR6a
2. SEQ ID NO:22, GPCR6b
3. SEQ ID NO:24, GPCR6c
4. SEQ ID NO:44, AF101730 chimpanzee olfactory receptor
5. SEQ ID NO:45, AF101761 gorilla olfactory receptor
6. SEQ ID NO:46, AF101741 chimpanzee olfactory receptor
7. SEQ ID NO:51, M64392 olfactory receptor-like protein i15
8. SEQ ID NO:54, AF101739 chimpanzee olfactory receptor
9. SEQ ID NO:55, AF101740 chimpanzee olfactory receptor

```
GPCR6a          MSPENQSSVSEFLLLGLPIRPEQQAVFFTLFLGMYLTTVLGNLLIMLLIQ 50
GPCR6b          MSPENQSSVSEFLLLGLPIRPEQQAVFFTLFLGMYLTTVLGNLLIMLLIQ 50
GPCR6c          ----------SEFLLLGLPIRPEQQAVFFTLFLGMYLTTVLGNLLIMLLIQ 41
44:AF101730     --MENQSSVSEFFLQGISGFPEQQQLLYGLFLCMYLVTLTGNVLIIMAIG 48
45:AF101761     MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR 50
46:AF101741     MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR 50
51:M64392       MTEENQTVISQFLLLFLPIPSEHQHVFYALFLSMYLTTVLGNLIIILLIH 50
54:AF101739     MMEQNQTSTSDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR 50
55:AF101740     MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR 50

GPCR6a          LDSHLHTPMYFFLSHLALTDISFSSVTVPKMLMDMRTKYKSILYEECISQ 100
GPCR6b          LDSHLHTPMYFFLSHLALTDISFSSVTVPKMLMDMRTKYKSILYEECISQ 100
GPCR6c          LDSHLHTPMYFFLSHLALTDISFSSVTVPKMLMDMRTKYKSILYEECISQ 91
44:AF101730     SDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQTQRHIISYTGCLTQ 98
45:AF101761     LDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCLTQ 100
46:AF101741     LDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCLTQ 100
51:M64392       LDSHLHTPMYLFLSNLSFSDLCFSSVTMPKLLQNMQSQVPSIPFAGCLTQ 100
54:AF101739     LDSHLHTPVYLFLSNLSFSDLCFSSVTMPKLLQNMQSQNPSIPYADCLTQ 100
55:AF101740     LDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCLTQ 100

GPCR6a          MYFFIFFTDLDSFLITSMAYDRYVAICHPLHYTVIMREELCVFLVAVSWI 150
```

```
GPCR6b      MYFFIFFTDLDSFLITSMAYDRYVAICHPLHYTVIMREELCVFLVAVSWI 150
GPCR6c      MYFFIFFTDLDSFLITSMTYDRYVAICHPLHYTVIMREELCVFLVAVSWI 141
44:AF101730 MYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMRAKVCVLMLALCWV 148
45:AF101761 MYFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSLVALSWV 150
46:AF101741 MYFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSLVALSWV 150
51:M64392   LYFYLYFADLESFLLVAMAYDRYVAICFPLHYMSIMSPKLCVSLVVLSWV 150
54:AF101739 MYFFLYFSDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSVVALSWV 150
55:AF101740 MHFFLLFGDLESFLLVAMAYDRYVAICFPLHYTAIMSPMLCLSVVALSWV 150

GPCR6a      LSCASSLSHTLLLTRLSFCAANTIPHVFCDLAALLKLSCSDIFLNELVMF 200
GPCR6b      LSCASSLSHTLLLTRLSFCAANTIPHVFCDLAALLKLSCSDIFLNELVMF 200
GPCR6c      LSCASSLSHTLLLTRLSFCAANTIPHVFCDLAALLKLSCSDIFLNELVMF 191
44:AF101730 LTNIVALTHTLLMARLSFCVVGEIAHFFCDITPVLKLSCSDTYVNELMVF 198
45:AF101761 LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200
46:AF101741 LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200
51:M64392   LTTFHAMLHTLLMARLSFCADNMIPHFFCDISPLLKLSCSDTHVNELVIF 200
54:AF101739 LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200
55:AF101740 LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200

GPCR6a      TVGVVVITLPFMCILVSYGYIGATILRVPSTKGIHKALSTCGSHLSVVSL 250
GPCR6b      TVGVVVITLPFMCILVSYGYIGATILRVPSIKGIHKALSTCGSHLSVVSL 250
GPCR6c      TVGVVVITQPFMCILVSYGYIGATILRVPSTKGIHKALSTCGSHLSVVSL 241
44:AF101730 ALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTKAFSTCSSHLCVVCV 248
45:AF101761 IMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSHLSVVSL 250
46:AF101741 IMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSHLSVVSL 250
51:M64392   VMGGLVIVIPFVLIIVSYARVVASILKVPSVRGIHKIFSTCGSHLSVVSL 250
54:AF101739 IMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKALSTCGSHLSVVSL 250
55:AF101740 IMGGLIVVIPFLLILGSYARIVSSILKVPSFKGICKALSTCGSHLSVVSL 250

GPCR6a      YYGSIFGQYLFPTVSSSIDKDVIVALMYTVVTPMLNPFIYSLRNRGMKEA 300
GPCR6b      YYGSIFGQYLFPTVSSSIDKDVIVALMYTVVTPMLNPFIYSLRNRDMKEA 300
GPCR6c      YYGSIFGQYLFPTVSSSIDKDVIVALMYTVVTPMLNPFIYSLRNRDMKEA 291
44:AF101730 FYGTLFSAYLSPPSVVSTEKDIAAAAMYTVVTPMLNPFIYSLRNKDMKGA 298
45:AF101761 FYGTIIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRNRDMKGA 300
46:AF101741 FYGTIIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRNRDMKGA 300
51:M64392   FYGTIIGLYLCPSANNSTVKETVMAMMYTVVTPMLNPFIYSLRNRDMKEA 300
54:AF101739 FYGTVIGLYLCPSANSSTLKDTVMAMMYTVVTPMLNPFIYSLRNRDMKGA 300
55:AF101740 FYGTVIGLYLCPSANSSTLKDTVMAMMYTVVTPMLNPFIYSLRNRDMKGA 300

GPCR6a      LGKLFS-RATFFSW 313
GPCR6b      LGKLFS-RATFFSW 313
GPCR6c      LGKLFS-RATFFSW 304
44:AF101730 LKRLLFHRRILSS- 311
45:AF101761 LERVICKRKNPFLL 314
46:AF101741 LERVIXKRKNPFLL 314
51:M64392   LIRVLCKKKITFCL 314
54:AF101739 LERVICKRKNPFLI 314
55:AF101740 LERVICKRKNPFLL 314
```

DOMAIN results for GPCR6 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. The results are listed in Table 6G with the statistics and domain description. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to have significant homology to GPCR6. An alignment of GPCR6 residues 41-290 (SEQ ID NO:20) with 7tm_1 residues 1-254 (SEQ ID NO:39) are shown in Table 6J.

**Table 6J. DOMAIN results for GPCR6**

| PSSMs producing significant alignments: | | Score (bits) | E value |
|---|---|---|---|
| gnl\|Pfam\|pfam00001 | 7tm_1, 7 transmembrane receptor (rhodopsin family) | 97.4 | 1e-21 |

```
GPCR6:   41 GNLLIMLLIQLDSHLHTPMYFFLSHLALTDISFSSVTVPKMLMDMRTKYKSILYEECISQ  100
7tm_1:    1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV   60

GPCR6:  101 MYFFIFFTDLDSFLITSMAYDRYVAICHPLHYTVIMREELCVFLVAVSWILSCASSLSHT  160
7tm_1:   61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPL  120

GPCR6:  161 LLLTRLSFCAANTIPHVFCDLAALLKLSCSDIFLNELVMFTVGVVVITLPFMCILVSYGY  220
7tm_1:  121 LFSWLRTVEEGNTTVCLIDFPEESVKRSYVLLSTLVGFVLPLLVILVCYTRILRTLRKRA  180

GPCR6:  221 IGATILRVPSTKGIHKALSTCGSHLSVVSLYYGSIFGQYLFPTVSSS----IDKDVIVAL  276
7tm_1:  181 RSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRVLPTALLITL  240

GPCR6:  277 MYTVVTPMLNPFIY  290
7tm_1:  241 WLAYVNSCLNPIIY  254
```

The GPCR6c disclosed in this invention is expressed in at least the following tissues:testis. This information was derived by determining the tissue sources of the sequences that were included in the invention including but not limited to SeqCalling sources, Public EST sources, Literature sources, and/or RACE sources. This is by no way limiting in that olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types. Further tissue expression analysis is provided in the Examples.

The GPCR6c disclosed in this invention maps to chromosome 9. This information was assigned using OMIM, the electronic northern bioinformatic tool implemented by CuraGen Corporation, public ESTs, public literature references and/or genomic clone homologies. This was executed to derive the chromosomal mapping of the SeqCalling assemblies, Genomic clones, literature references and/or EST sequences that were included in the invention.

The nucleic acids and proteins of GPCR6 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX

Antibodies" section below. The disclosed GPCR6 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In various embodiments, a contemplated GPCR6 epitope is from about amino acids 1 to 20, from about amino acids 75 to 95, from about amino acids 225 to 235 and from about amino acids 280 to 313.

## GPCR7

A further GPCR-like protein of the invention, referred to herein as GPCR7, is an Olfactory Receptor ("OR")-like protein. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR7 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Two alternative novel GPCR7 nucleic acids and encoded polypeptides are provided, namely GPCR7a and GPCR7b.

### GPCR7a

In one embodiment, a GPCR7 variant is the novel GPCR7a (alternatively referred to herein as CG57809-01), which includes the 936 nucleotide sequence (SEQ ID NO:25) shown in Table 7A. A GPCR7a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 4-6 and ends with a TGA codon at nucleotides 934-936. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 7A, and the start and stop codons are in bold letters.

**Table 7A. GPCR7a Nucleotide Sequence (SEQ ID NO:25)**

TCC**ATG**GGAATGTCCAACCTGACAAGACTCTCTGAATTTATTCTCTTGGGACTCTCCTCTCGGTCTG
AAGACCAGAGGCCACTCTTTGCCCTCTTTCTTATCATATACCTGGTCACTTTGATGGGAAATCTGCT
CATCATCTTGGCTATCCACTCTGATCCTCGACTTCAAAACCCTATGTATTTTTTCCTAAGCATCTTG
TCCTTTGCTGATATTTGCTACACAACAGTCATAGTCCCAAAGATGCTCGTGAACTTCTTATCAGAGA
AAAAGACCATTTCCTATGCTGAATGTCTGGCACAGATGTATTTCTTCCTGGTTTTTGGAAACATAGA
TAGTTATCTCCTGGCGGCTATGGCCATCAACCGCTGTGTAGCCATTTGTAACCCATTCCATTATGTC
ACTGTTATGAACCGCAGATGCTGTGTGTTGCTACTAGCATTCCCCATCACTTTCTCCTATTTCCACT
CCCTCCTACATGTCCTCCTGGTGAATCGGCTCACCTTTTGTACATCAAATGTTATCCATCATTTTTT
TTGTGATGTCAACCCTGTGCTGAAACTGTCCTGCTCCTCCACCTTTGTCAATGAAATTGTGGCCATG
ACAGAAGGGCTGGCCTCTGTGATGGCTCCATTTGTCTGTATCATCATCTCTTATCTAAGAATTCTCA
TCGCTGTTCTCAAGATTCCCTCAGCAGCTGGAAAACACAAAGCCTTCTCCACCTGCAGCTCCCATCT
CACTGTGGTGATTCTGTTTTATGGGAGTATTAGCTATGTCTATTTGCAGCCTTTGTCCAGCTATACT
GTCAAGGACCGAATAGCAACAATCAACTACACTGTGTTGACATCAGTGTTGAACCCATTTATCTACA
GTTTAAGAAACAAAGACATGAAACGGGGCTTACAGAAATTGATAAACAAGATTAAGTCTCAA**TGA**

The GPCR7a protein (SEQ ID NO:26) encoded by SEQ ID NO:25 has 310 amino acid residues and is presented using the one-letter code in Table 7B. The predicted molecular

weight of GPCR7a protein is 35079.05 Daltons. The Psort profile for GPCR7a predicts that this sequence has a signal peptide and is likely to be localized at the mitochondrial inner membrane with a certainty of 0.6046. In alternative embodiments, GPCR7 is located in the plasma membrane with a certainty of 0.600, in the mitochondrial intermembrane space with a certainty of 0.4615 or a Golgi body with a certainty of 0.400. The Signal P predicts a likely cleavage site between positions 48 and 49, *i.e.*, at the dash in the sequence ILA-IH.

**Table 7B. Encoded GPCR7a protein sequence (SEQ ID NO:26)**

```
MGMSNLTRLSEFILLGLSSRSEDQRPLFALFLIIYLVTLMGNLLIILAIHSDPRLQNPMYFFLSILSFAD
ICYTTVIVPKMLVNFLSEKKTISYAECLAQMYFFLVFGNIDSYLLAAMAINRCVAICNPFHYVTVMNRRC
CVLLLAFPITFSYFHSLLHVLLVNRLTFCTSNVIHHFFCDVNPVLKLSCSSTFVNEIVAMTEGLASVMAP
FVCIIISYLRILIAVLKIPSAAGKHKAFSTCSSHLTVVILFYGSISYVYLQPLSSYTVKDRIATINYTVL
TSVLNPFIYSLRNKDMKRGLQKLINKIKSQ
```

The DNA sequence and protein sequnece of GPCR7a was obtained solely by exon linking process.

**GPCR7b**

In an alternative embodiment, a GPCR7 variant is the novel GPCR7b (alternatively referred to herein as GMba64p14_G), which includes the 936 nucleotide sequence (SEQ ID NO:27) shown in Table 7C. The GPCR7b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 4-6 and ends with a TGA codon at nucleotides 934-936, which are in bold letters in Table 1C.

**Table 7C. GPCR7b Nucleotide Sequence (SEQ ID NO:27)**

```
TCCATGGGAATGTCCAACCTGACAAGACTCTCTGAATTTATTCTCTTGGGACTCTCCTCTCGGTCTG
AAGACCAGAGGCCACTCTTTGCCCTCTTTCTTATCATATACCTGGTCACTTTGATGGGAAATCTGCT
CATCATCTTGGCTATCCACTCTGATCCTCGACTTCAAAACCCTATGTATTTTTTCCTAAGCATCTTG
TCCTTTGCTGATATTTGCTACACAACAGTCATAGTCCCAAAGATGCTCGTGAACTTCTTATCAGAGA
AAAAGACCATTTCCTATGCTGAATGTCTGGCACAGATGTATTTCTTCCTGGTTTTTGGAAACATAGA
TAGTTATCTCCTGGCGGCTATGGCCATCAACCGCTGTGTAGCCATTTGTAACCCATTCCATTATGTC
ACTGTTATGAACCGCAGATGCTGTGTGTTGCTACTAGCATTCCCCATCACTTTCTCCTATTTCCACT
CTCTCCTACATGTCCTCCTGGTGAATCGGCTCACCTTTTGTACATCAAATGTTATCCATCATTTTTT
TTGTGATGTCAACCCTGTGCTGAAACTGTCCTGCTCCTCCACCTTTGTCAATGAAATTGTGGCCATG
ACAGAAGGGCTGGCCTCTGTGATGGCTCCATTTGTCTGTATCATCATCTCTTATCTAAGAATTCTC
ATCGCTGTTCTCAAGATTCCCTCAGCAGCTGGAAAACACAAAGCCTTCTCCACCTGCAGCTC
CCATCTCACTGTGGTGATTCTGTTTTATGGGAGTATTAGCTATGTCTATTTGCAGCCTTTGTCCAGC
TATACTGTCAAGGACCGAATAGCAACAATCAACTACACTGTGTTGACATCAGTGTTGAACCCATTTA
TCTACAGTTTAAGAAACAAAGACATGAAACGGGGCTTACAGAAATTGATAAACAAGATTAAGTCTCA
ATGA
```

The GPCR7b protein (SEQ ID NO:28) encoded by SEQ ID NO:27 is 310 amino acid in length, and is presented using the one-letter code in Table 7D. The GPCR7a and GPCR7b

polypeptides are identical, due to a silent codon change between the GPCR7 nucleotide sequences.

**Table 7D. GPCR7b protein sequence (SEQ ID NO:28)**

MGMSNLTRLSEFILLGLSSRSEDQRPLFALFLIIYLVTLMGNLLIILAIHSDPRLQNPMYFFLSILS
FADICYTTVIVPKMLVNFLSEKKTISYAECLAQMYFFLVFGNIDSYLLAAMAINRCVAICNPFHYVT
VMNRRCCVLLLAFPITFSYFHSLLHVLLVNRLTFCTSNVIHHFFCDVNPVLKLSCSSTFVNEIVAMT
EGLASVMAPFVCIIISYLRILIAVLKIPSAAGKHKAFSTCSSHLTVVILFYGSISYVYLQPLSSYTV
KDRIATINYTVLTSVLNPFIYSLRNKDMKRGLQKLINKIKSQ

**GPCR7 Clones**

Unless specifically addressed as GPCR7a or GPCR7b, any reference to GPCR7 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant, the residue position with respect to the "a" variant, and the residue in the "b" variant. For example, the GPCR7 nucleic acid sequences differ at the following position: T471C. The encoded GPCR7a and GPCR7b polypeptides are identical.

The amino acid sequence of GPCR7 had high homology to other proteins as shown in Table 7E.

**Table 7E. BLASTX results for GPCR7**

| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
|---|---|---|---|
| patp:AAR27868 Odorant receptor clone F5 - Rattus rattus, 313 aa. | +1 | 850 | 4.3e-84 |

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of GPCR7a has 602 of 922 bases (65%) identical to a gb:GENBANK-ID:RATOLFPROC|acc:M64377.1 mRNA from Rattus norvegicus (Rat olfactory protein mRNA, complete cds). The full amino acid sequence of the GPCR7a protein was found to have 161 of 305 amino acid residues (52%) identical to, and 219 of 305 amino acid residues (71%) similar to, the 313 amino acid residue ptnr:SWISSPROT-ACC:P23266 protein from Rattus norvegicus (Rat) (OLFACTORY RECEPTOR-LIKE PROTEIN F5).

GPCR7 also has homology to the proteins shown in the BLASTP data in Table 7F.

**Table 7F. GPCR7 BLASTP results**

| Gene Index / Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |
|---|---|---|---|---|---|
| M64377; P23266 | OLF5_RAT rattus norvegicus (rat). olfactory receptor-like protein f5. 7/1993 | 313 | 161/305 (53%) | 219/305, (72%) | 8e-91 |
| Y14442; AJ003147; U86234 | O1F1_HUMAN homo sapiens (human). olfactory receptor 1f1 (olfactory receptor 16-35) (or16-35). 10/2000 | 312 | 157/308 (51%) | 215/308, (70%) | 7e-88 |
| M64392 | OLF1_RAT rattus norvegicus (rat). olfactory receptor-like protein i15. 7/1993 | 314 | 156/309 (50%) | 220/309, (71%) | 3e-87 |
| AB038167 | Q9JHE2 rattus norvegicus (rat). gustatory receptor 43. 3/2001 | 311 | 162/303 (53%) | 211/303, (70%) | 5e-87 |
| AF101767 | Q9TU84 pongo pygmaeus (orangutan). olfactory receptor. 3/2001 | 313 | 160/304 (53%) | 212/304, (70%) | 5e-87 |

A multiple sequence alignment is given in Table 7G, with the GPCR7 protein being shown on line 1, in a ClustalW analysis comparing the protein of the invention with the related protein sequences shown in Table 7F. This BLASTP data is displayed graphically in the ClustalW in Table 7G.

**Table 7G. ClustalW Analysis of GPCR7**

1. SEQ ID NO:26 and 28, GPCR7
2. SEQ ID NO:41, AB038167 gustatory receptor 43
3. SEQ ID NO:42, M64377 olfactory receptor-like protein f5
4. SEQ ID NO:47, Y14442 olfactory receptor 1f1 (or16-35)
5. SEQ ID NO:51, M64392 olfactory receptor-like protein i15
6. SEQ ID NO:56, AF101767 orangutan olfactory receptor

```
GPCR7          MGMSNLTRLSEFILLGLSSRSEDQRPLFALFLIIYLVTLMGNLLIILAIH 50
41:AB038167    --MENQSSVSEFFLQGISGFPEQQQLLYGLFLCMYLVTLIGNVLIIMAIG 48
42:M64377      MSSTNQSSVTEFLLLGLSRQPQQQQLLFLLFLIMYLATVLGNLLIILAIG 50
47:Y14442      MSGTNQSSVSEFLLLGLSRQPQQQHLLFVFFLSMYLATVLGNLLIILSVS 50
51:M64392      MTEENQTVISQFLLLFLPIPSEHQHVFYALFLSMYLITVLGNLIIILLIH 50
56:AF101767    MDGGNQSEGSEFLLLGMSESPEQQRILFWMFLSMYLVTVLGNVLIILAIS 50

GPCR7          SDPRLQNPMYFFLSILSFADICYTTVIVPKMLVNFLSEKKTISYAECLAQ 100
41:AB038167    SDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQTQRHTISYTGCLTQ 98
42:M64377      TDSRLHTPMYFFLSNLSFVDVCFSSTTVPKVLANHILGSQAISFSGCLTQ 100
47:Y14442      IDSCLHTPMYFFLSNLSFVDICFSFTTVPKMLANHILETQTISFCGCLTQ 100
51:M64392      LDSHLHTPMYLFLSNLSFSDLCFSSVTMPKLLQNMQSQVPSIPFAGCLTQ 100
56:AF101767    SDSRLHTPMYFFLANLSFTDLFFVTNTIPKMLVNLQSQDKAISYAGCLTQ 100

GPCR7          MYFFLVFGNIDSYLLAAMAINRCVAICNPFHYVTVMNRRCCVLLLAFPIT 150
41:AB038167    MYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMRAKVCVLMLALCWV 148
```

```
42:M64377     LYFLAVFGNMDNFLLAVMSYDRFVAICHPLHYTTKMTRQLCVLLVVGSWV 150
47:Y14442     MYFVFMFVDMDNFLLAVMAYDHFVAVCHPLHYTAKMTHQLCALLVAGLWV 150
51:M64392     LYFYLYFADLESFLLVAMAYDRYVAICFPLHYMSIMSPKLCVSLVVLSWV 150
56:AF101767   LYFLLSLVTLDNLILAVMAYDRYVAICCPLHYVTAMSPRLCILLLSLCWV 150

GPCR7         FSYFHSLLHVLLVNRLTFCTSNVIHHFFCDVNPVLKLSCSSTFVNEIVAM 200
41:AB038167   LTNIVALTHTLLMARLSFCVVGEIAHFFCDITPVLKLSCSDTYVNELMVF 198
42:M64377     VANMNCLLHILLMARLSFCADNMIPHFFCDGTPLLKLSCSDTHLNELMIL 200
47:Y14442     VANLNVLLHTLLMAPLSFCADNAITHFFCDVTPLLKLSCSNTHLNEVIIL 200
51:M64392     LTTFHAMLHTLLMARLSFCADNMIPHFFCDISPLLKLSCSDTHVNELVIF 200
56:AF101767   FSVLYGLIHTLLMTRVTFCGSRKIHYLFCEMYFLLRLACSNIQINHTVLX 200

GPCR7         TEGLASVMAPFVCIIISYLRILIAVLKIPSAAGKHKAFSTCSSHLTVVIL 250
41:AB038167   ALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTKAFSTCSSHLCVVCV 248
42:M64377     TEGAVVMVTPFVCILISYIHITCAVLRVSSPRGGWKSFSTCGSHLAVVCL 250
47:Y14442     SEGALVMITPFLCILASYMHITCTVLKVPSTKGRWKAFSTCGSHLAVVLL 250
51:M64392     VMGGLVIVIPFVLIIVSYARVVASILKVPSVRGIHKIFSTCGSHLSVVSL 250
56:AF101767   ATGCFLFLIPLGFMIXSYARIVRAILRIPSATGKYKAFSTCASHLAVVSL 250

GPCR7         FYGSISYVYLQPLSSYTV-KDRIATINYTVLTSVLNPFIYSLRNKDMKRG 299
41:AB038167   FYGTLFSAYLSPPSVVSTEKDIAAAAMYTVVTPMLNPFIYSLRNKDMKGA 298
42:M64377     FYGTVIAVYFNPSSSHLAGRDMAAAVMYAVVTPMLNPFIYSLRNSDMKAA 300
47:Y14442     FYSTIIAVYFNPLSSHSAEKDTMATVLYTVVTPMLNPFIYSLRNRYLKGA 300
51:M64392     FYGTIIGLYLCPSANNSTVKETVMAMMYTVVTPMLNPFIYSLRNRDMKEA 300
56:AF101767   FYGTLGMVYLQPLQIYST-KDSVATVMYAVVTPMMNPFIYSLRNKDIHGA 299

GPCR7         LQKLINKIKSQ--- 310
41:AB038167   LKRLLFHRRILSS- 311
42:M64377     LRKVLAMRFPSKQ- 313
47:Y14442     LKKVIGRVVFSV-- 312
51:M64392     LIRVLCKKKITFCL 314
56:AF101767   LGRLLQGKAFQKLT 313
```

Table 7H lists the domain description from DOMAIN analysis results against GPCR7. This indicates that the GPCR7 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:39) itself.

**Table 7H. Domain Analysis of GPCR7**

| PSSMs producing significant alignments: | | Score (bits) | E value |
|---|---|---|---|
| gnl\|Pfam\|pfam00001 | 7tm_1, 7 transmembrane receptor (rhodopsin family) | 98.2 | 6e-22 |

```
GPCR7:   41 GNLLIILAIHSDPRLQNPMYFFLSILSFADICYTTVIVPKMLVNFLSEKKTISYAECLAQ  100
7tm_1:    1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV   60

GPCR7:  101 MYFFLVFGNIDSYLLAAMAINRCVAICNPFHYVTVMNRRCCVLLLAFPITFSYFHSLLHV  160
7tm_1:   61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPL  120

GPCR7:  161 LLVNRLTFCTSNVIHHFFCDVNPVLKLSCSSTFVNEIVAMTEGLASVMAPFVCIIISYLR  220
7tm_1:  121 LFSWLRTVEEGNTTVCLIDFPEESVKRS---------YVLLSTLVGFVLPLLVILVCYTR  171
```

```
GPCR7: 221 ILIAV---------LKIPSAAGKHKAFSTCSSHLTVVILFYGSISYVYLQPLSSYTVKDR  271
           ||   +          ||   |++ +   |        +   |+ +       + |   |    ++
7tm_1: 172 ILRTLRKRARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRV  231

GPCR7: 272 I-----ATINYTVLTSVLNPFIY  289
           +       |+    + | |||  ||
7tm_1: 232 LPTALLITLWLAYVNSCLNPIIY  254
```

The nucleic acids and proteins of GPCR7 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein.

The novel GPCR7 nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR7 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR7 epitope is from about amino acids 5 to 25. In additional embodiments, GPCR7 epitopes are from about amino acids 50 to 60, from about amino acids 80 to 100, from about amino acids 130 to 145, from about amino acids 230 to 240, from about amino acids 260 to 270 and from about amino acids 290 to 310.

**GPCR8**

An eighth GPCR-like protein of the invention, referred to herein as GPCR8, is an Olfactory Receptor ("OR")-like protein. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR8 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

The GPCR disclosed in this invention maps to chromosome 9. This information was assigned using OMIM, the electronic northern bioinformatic tool implemented by CuraGen Corporation, public ESTs, public literature references and/or genomic clone homologies. This was executed to derive the chromosomal mapping of the SeqCalling assemblies, Genomic clones, literature references and/or EST sequences that were included in the invention.

Four alternative novel GPCR8 nucleic acids, namely GPCR8a, GPCR8b, GPCR8c and GPCR8d, and their encoded GPCR8 polypeptides are provided.

**GPCR8a**

The first disclosed novel GPCR8 clone is GPCR8a (also referred to as ba542k23), which has a nucleic acid (SEQ ID NO:29) of 1920 nucleotides as shown in Table 8A. An ORF begins with an ATG initiation codon at nucleotides 38-40 and ends with a TAG codon at nucleotides 965-967.

The following genomic clone was identified as having high homology to olfactory receptor-like protein (HS6M1-6). The start and stop codons in Table 8A are in bold letters and the putative untranslated regions upstream from the initiation site and downstream from the termination codon are underlined.

**Table 8A. GPCR8a Nucleotide Sequence (SEQ ID NO:29)**

ATGACAAAACTTCTTGTTTATAACTGAGCCCAAGTCA**ATG**GAAAGAATCAACCACACCAGCAGTGTCTCCGAG
TTTATCCTCCTGGGACTCTCCTCCCGGCCTGAGGACCAAAAGACACTCTTTGTTCTCTTCCTCATCGTGTACC
TGGTCACCATAACAGGGAACCTGCTCATCATCCTGGCCATTCGCTTCAACCCCCATCTTCAGACCCCTATGTA
TTTCTTCTTGAGTTTTTTGTCTCTCACTGATATTTGCTTTACAACAAGCGTTGTCCCCAAGATGCTGATGAAC
TTCCTGTCAGAAAAGAAGACCATCTCCTATGCTGGGTGTCTGACACAGATGTATTTTCTCTATGCCTTGGGCA
ACAGTGACAGCTGCCTTCTGGCAATCATGGCCTTTGACCGCTATGTGGCCGTCTGTGACCCTTTCCACTATGT
CACCACCATGAGCCACCACCACTGTGTCCTGCTGGTGGCCTTCTCCTGCTCATTTCCTCACCTCCACTCACTC
CTGCACACACTTCTGCTGAATCGTCTCACCTTCTGTGACTCCAATGTTATCCACCACTTTCTCTGTGACCTCA
GCCCTGTGCTGAAATTGTCCTGCTCGTCCATATTTGTCAATGAAATTGTGCAGATGACAGAAGCACCTATTGT
TTTGGTGACTCCTTTTCTCTGCATTGCTTTCTCTTATATACGAATCCTCACTACAGTTCTCAAGATTCCCTCT
ACTTCTGGGAAACGCAAAGCCTTCTCCACCTGTGGTTTTTACCTCACCGTGGTGACGCTCTTTTATGGAAGCA
TCTTCTGTGTCTATTTACAGCCCCCATCCACCTACGCTGTCAAGGACCACGTGGCAACAATTGTTTACACAGT
TTTGTCATCCATGCTCAATCCTTTTATCTACAGCCTGAGAAACAAAGACCTGAAACAGGGCCTGAGGAAGCTT
ATGAGCAAGAGATCC**TAG**GAAGCACCCTCTTGAAAAACTCGTAAGTGGAATCTGCTCAACTTGGACGTGTTTT
CTACTGGTTTCTGGTGAACA

The GPCR8a protein (SEQ ID NO:30) encoded by SEQ ID NO:29 has 309 amino acid residues and is presented using the one-letter code in Table 8B. The predicted molecular weight of GPCR8 protein is 35035.77 Daltons. The Psort profile for all GPCR8 polypeptides predicts that this sequence has a signal peptide and is likely to be localized at the mitochondrial inner membrane with a certainty of 0.7099. In alternative embodiments, GPCR8 is located in the plasma membrane with a certainty of 0.600, in a Golgi body with a certainty of 0.4000 or the mitochondrial intermembrane space with a certainty of 0.3306. The Signal P predicts a likely cleavage site between positions 49 and 50, *i.e.*, at the dash in the sequence ILA-IR.

| Table 8B. Encoded GPCR8a protein sequence (SEQ ID NO:30) |
|---|
| MERINHTSSVSEFILLGLSSRPEDQKTLFVLFLIVYLVTITGNLLIILAIRFNPHLQTPMYFFLSFLSLTDICF<br>TTSVVPKMLMNFLSEKKTISYAGCLTQMYFLYALGNSDSCLLAIMAFDRYVAVCDPFHYVTTMSHHHCVLLVAF<br>SCSFPHLHSLLHTLLLNRLTFCDSNVIHHFLCDLSPVLKLSCSSIFVNEIVQMTEAPIVLVTPFLCIAFSYIRI<br>LTTVLKIPSTSGKRKAFSTCGFYLTVVTLFYGSIFCVYLQPPSTYAVKDHVATIVYTVLSSMLNPFIYSLRNKD<br>LKQGLRKLMSKRS |

Genomic sequence GPCR8a on chromosome 9 was identified by TBLASTN using CuraGen Corporation's sequence file for members of GPCR family, run against the genomic daily files made available by GenBank or obtained from Human Genome Project Sequencing centers. It was then extended experimentally by the Exon linking process (nucleotides 233-1042 of the sequence of the invention GPCR8a) and in silico by using genomic clone AL162254 (contributed nucleotides 1-232 of the sequence of the invention GPCR8a) to generate the full length sequence as described above. Therefore, apart from Curagen's Exon Linking process, sequence from genomic clone AL162254 was included in the invention.

**GPCR8b**

In an alternative embodiment, a GPCR8 variant is the novel GPCR8b (alternatively referred to herein as 148540666 or CG50259-01), which includes the 1033 nucleotide sequence (SEQ ID NO:31) shown in Table 8C. The GPCR8b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 29-31 and ends with a TAG codon at nucleotides 956-958, which are in bold letters in Table 8C.

| Table 8C. GPCR8b Nucleotide Sequence (SEQ ID NO:31) |
|---|
| CTTCTTGTTTATAACTGAGCCCAAGTCA**ATG**GAAAGAATCAACCATACCAGCAGTGTCTCCGAGTTT<br>ATCCTCCTGGGACTCTCCTCCCGGCCTGAGGACCAAAAGCCACTCTTTGTTCTCTTCCTCATCGTGT<br>ACCTGGTCACCATAACAGGGAACCAGCTCATCATCCTGGCCATTCGCTTCAACCCCCATCTTCAGAC<br>CCCTATGTATTTCTTCTTGAGTTTTTTGTCTCTCACTGATATTTGCTTTACAACAAGCGTTGTCCCC<br>AAGATGCTGATGAACTTCCTGTCAGAAAAGAAGACCATCTCCTATGCTGGGTGTCTGACACAGATGT<br>ATTTTCTCTATGCCTTGGGCAACAGTGACAGCTGCCTTCTGGCAATCATGGCCTTTGACCGCTATGT<br>GGCCGTCTGTGACCCTTTCCACTATGTCACCACCATGAGCCACCACCACTGTGTCCTGCTGGTGGCC<br>TTCTCCTGCTCATTTCCTCACCTCCACTCACTCCTGCACACACTTCTGCTGAATCGTCTCACCTTCT<br>GTGACTCCAATGTTATCCACCACTTTCTCTGTGACCTCAGCCCTGTGCTGAAATTGTCCTGCTCGTC<br>CATATTTGTCAATGAAATTGTGCAGATGACAGAAGCACCTATTGTTTTGGTGACTCCTTTTCTCTGC<br>ATTGCTTTCTCTTATATACGAATCCTCACTACAGTTCTCAAGATTCCCTCTACTTCTGGGAAACGCA<br>AAGCCTTCTCCACCTGTGGTTTTTACCTCACCGTGGTGACGCTCTTTTATGGAAGCATCTTCTGTGT<br>CTATTTACAGCCCCCATCCACCTACGCTGTCAAGGACCACGTGGCAACAATTGTTTACACAGTTTTG<br>TCATCCATGCTCAATCCTTTTATCTACAGCCTGAGAAACAAAGACCTGAAACAGGGCCTGAGGAAGC<br>TTATGAGCAAGAGATCC**TAG**GAAGCACCCTCTTGAAAAACTCGTAAGTGGAATCTGCTCAACTTGGA<br>CGTGTTTTCTACTGGTTTCTGGTGAACA |

The GPCR8b protein (SEQ ID NO:32) encoded by SEQ ID NO:31 is 309 amino acids in length, has a molecular weight of 35046.76 Daltons, and is presented using the one-letter

code in Table 8D. As with all GPCR8 proteins, the most likely cleavage site for a GPCR8b peptide is between amino acids 49 and 50, *i.e.*, at the dash in the sequence ILA-IR, based on the SignalP result. The DNA sequence and protein sequnece of GPCR8b was obtained by exon linking process.

**Table 8D. GPCR8b protein sequence (SEQ ID NO:32)**

```
MERINHTSSVSEFILLGLSSRPEDQKPLFVLFLIVYLVTITGNQLIILAIRFNPHLQTPMY
FFLSFLSLTDICFTTSVVPKMLMNFLSEKKTISYAGCLTQMYFLYALGNSDSCLLAIMAFD
RYVAVCDPFHYVTTMSHHHCVLLVAFSCSFPHLHSLLHTLLLNRLTFCDSNVIHHFLCDLS
PVLKLSCSSIFVNEIVQMTEAPIVLVTPFLCIAFSYIRILTTVLKIPSTSGKRKAFSTCGF
YLTVVTLFYGSIFCVYLQPPSTYAVKDHVATIVYTVLSSMLNPFIYSLRNKDLKQGLRKLM
SKRS
```

**GPCR8c**

In an alternative embodiment, a GPCR8 variant is the novel GPCR8c (alternatively referred to herein as AL162254-da1), which includes the 969 nucleotide sequence (SEQ ID NO:33) shown in Table 8E. The partial GPCR8c ORF begins 6bp downstream from the Kozak consensus ATG initiation codon of the other GPCR8 nucleic acids of the invention, and ends with a TAG codon at nucleotides 922-924, which are in bold letters in Table 8C.

**Table 8E. GPCR8c Nucleotide Sequence (SEQ ID NO:33)**

**AGA**ATCAACCACACCAGCAGTGTCTCCGAGTTTATCCTCCTGGGACTCTCCTCCCGGCCTGAGGACC
AAAAGACACTCTTTGTTCTCTTCCTCATCGTGTACCTGGTCACCATAACAGGGAACCTGCTCATCAT
CCTGGCCATTCGCTTCAACCCCCATCTTCAGACCCCTATGTATTTCTTCTTGAGTTTTCTGTCTCTC
ACTGATATTTGCTTTACAACAAGCGTTGTCCCCAAGATGCTGATGAACTTCCTGTCAGAAAAGAAGA
CCATCTCCTATGCTGGGTGTCTGACACAGATGTATTTTCTCTATGCCTTGGGCAACAGTGACAGCTG
CCTTCTGGCAGTCATGGCCTTTGACCGCTATGTGGCCGTCTGTGACCCTTTCCACTATGTCACCACC
ATGAGCCACCACCACTGTGTCCTGCTGGTGGCCTTCTCCTGCTCATTTCCTCACCTCCACTCACTCC
TGCACACACTTCTGCTGAATCGTCTCACCTTCTGTGACTCCAATGTTATCCACCACTTTCTCTGTGA
CCTCAGCCCTGTGCTGAAATTGTCCTGCTCTTCCATATTTGTCAATGAAATTGTGCAGATGACAGAA
GCACCTATTGTTTTGGTGACTCGTTTTCTCTGCATTGCTTTCTCTTATATACGAATCCTCACTACAG
TTCTCAAGATTCCCTCTACTTCTGGGAAACGCAAAGCCTTCTCCACCTGTGGTTTTTACCTCACCGT
GGTGACGCTCTTTTATGGAAGCATCTTCTGTGTCTATTTACAGCCCCCATCCACCTACGCTGTCAAG
GACCACGTGGCAACAATTGTTTACACAGTTTTGTCATCCATGCTCAATCCTTTTATCTACAGCCTGA
GAAACAAAGACCTGAAACAGGGCCTGAGGAAGCTTATGAGCAAGAGATCC**TAG**GAAGCACCCTCTTG
AAAAACTCGTAAGTGGAATCTGCTCAACTTG

The GPCR8c protein fragment (SEQ ID NO:34) encoded by SEQ ID NO:33 is 307 amino acid in length, has a molecular weight of 34820.52 Daltons, and is presented using the one-letter code in Table 8D. The GPCR8c protein fragment lacks the first two amino acids encoded by the other GPCR8 nucleic acid sequences of the invention. The most likely GPCR8c cleavage site remains between the amino acids in the sequence ILA-IR, based on the

SignalP result. The DNA sequence and protein sequnece of GPCR8c was obtained solely by exon linking process.

| Table 8F. GPCR8c protein sequence (SEQ ID NO:34) |
|---|
| RINHTSSVSEFILLGLSSRPEDQKTLFVLFLIVYLVTITGNLLIILAIRFNPHLQTPMYFFLSFLSL<br>TDICFTTSVVPKMLMNFLSEKKTISYAGCLTQMYFLYALGNSDSCLLAVMAFDRYVAVCDPFHYVTT<br>MSHHHCVLLVAFSCSFPHLHSLLHTLLLNRLTFCDSNVIHHFLCDLSPVLKLSCSSIFVNEIVQMTE<br>APIVLVTRFLCIAFSYIRILTTVLKIPSTSGKRKAFSTCGFYLTVVTLFYGSIFCVYLQPPSTYAVK<br>DHVATIVYTVLSSMLNPFIYSLRNKDLKQGLRKLMSKRS |

**GPCR8d**

In an alternative embodiment, a GPCR8 variant is the novel GPCR8d (alternatively referred to herein as GMba64p14_H ), which includes the 955 nucleotide sequence (SEQ ID NO:35) shown in Table 8G. The GPCR8d ORF begins with a Kozak consensus ATG initiation codon at nucleotides 4-6 and ends with a TAG codon at nucleotides 931-933, which are in bold letters in Table 8G.

| Table 8G. GPCR8d Nucleotide Sequence (SEQ ID NO:35) |
|---|
| TCA**ATG**GAAAGAATCAACCACACCAGCAGTGTCTCCGAGTTTATCCTCCTGGGACTCTCCTCCCGGC<br>CTGAGGACCAAAAGACACTCTTTGTTCTCTTCCTCATCGTGTACCTGGTCACCATAACAGGGAACCT<br>GCTCATCATCCTGGCCATTCGCTTCAACCCCCATCTTCAGACCCCTATGTATTTCTTCTTGAGTTTT<br>CTGTCTCTCACTGATATTTGCTTTACAACAAGCGTTGTCCCCAAGATGCTGATGAACTTCCTGTCAG<br>AAAAGAAGACCATCTCCTATGCTGGGTGTCTGACACAGATGTATTTTCTCTATGCCTTGGGCAACAG<br>TGACAGCTGCCTTCTGGCAGTCATGGCCTTTGACCGCTATGTGGCCGTCTGTGACCCTTTCCACTAT<br>GTCACCACCATGAGCCACCACCACTGTGTCCTGCTGGTGGCCTTCTCCTGCTCATTTCCTCACCTCC<br>ACTCACTCCTGCACACACTTCTGCTGAATCGTCTCACCTTCTGTGACTCCAATGTTATCCACCACTT<br>TCTCTGTGACCTCAGCCCTGTGCTGAAATTGTCCTGCTCTTCCATATTTGTCAATGAAATTGTGCAG<br>ATGACAGAAGCACCTATTGTTTTGGTGACTCGTTTTCTCTGCATTGCTTTCTCTTATATACGAATCC<br>TCACTACAGTTCTCAAGATTCCCTCTACTTCTGGGAAACGCAAAGCCTTCTCCACCTGTGGTTTTTA<br>CCTCACCGTGGTGACGCTCTTTTATGGAAGCATCTTCTGTGTCTATTTACAGCCCCCATCCACCTAC<br>GCTGTCAAGGACCACGTGGCAACAATTGTTTACACAGTTTTGTCATCCATGCTCAATCCTTTTATCT<br>ACAGCCTGAGAAACAAAGACCTGAAACAGGGCCTGAGGAAGCTTATGAGCAAGAGATCC**TAG**GAAGC<br>ACCCTCTTGAAAAACTC |

The GPCR8d protein (SEQ ID NO:36) encoded by SEQ ID NO:35 is 309 amino acid in length, has a molecular weight of 35080.82 Daltons, and is presented using the one-letter code in Table 8H. As with all other GPCR8 polypeptides, the most likely cleavage site for a GPCR8d peptide is between amino acids 49 and 50, *i.e.*, at the dash in the sequence ILA-IR, based on the SignalP result.

| Table 8H. GPCR8d protein sequence (SEQ ID NO:36) |
|---|
| MERINHTSSVSEFILLGLSSRPEDQKTLFVLFLIVYLVTITGNLLIILAIRFNPHLQTPMYFFLSFL<br>SLTDICFTTSVVPKMLMNFLSEKKTISYAGCLTQMYFLYALGNSDSCLLAVMAFDRYVAVCDPFHYV<br>TTMSHHHCVLLVAFSCSFPHLHSLLHTLLLNRLTFCDSNVIHHFLCDLSPVLKLSCSSIFVNEIVQM<br>TEAPIVLVTRFLCIAFSYIRILTTVLKIPSTSGKRKAFSTCGFYLTVVTLFYGSIFCVYLQPPSTYA<br>VKDHVATIVYTVLSSMLNPFIYSLRNKDLKQGLRKLMSKRS |

## GPCR8 Clones

Unless specifically addressed as GPCR8a, GPCR8b, GPCR8c or GPCR8d, any reference to GPCR8 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant, the residue position with respect to the "a" variant, and the residue in the "b", "c" or "d" variant.

The amino acid sequence of GPCR8 had high homology to other proteins as shown in Table 8I.

**Table 8I. BLASTX results for GPCR8**

| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
|---|---|---|---|
| patp:AAR27868 Odorant receptor clone F5 - Rattus rattus, 313 aa. | +1 | 849 | 5.5e-84 |

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of GPCR8 has 619 of 922 bases (67%) identical to a Rat olfactory protein mRNA from Rattus norvegicus (GENBANK-ID:RATOLFPROC|acc:M64377). The full amino acid sequence of the GPCR8 protein of the invention was found to have 162 of 302 amino acid residues (53%) identical to, and 220 of 302 amino acid residues (72%) similar to the 313 amino acid residue OLFACTORY RECEPTOR-LIKE PROTEIN F5 from Rattus norvegicus (SWISSPROT-ACC:P23266). GPCR8 also has homology to the proteins shown in the BLASTP data in Table 8J.

**Table 8J. GPCR8 BLASTP results**

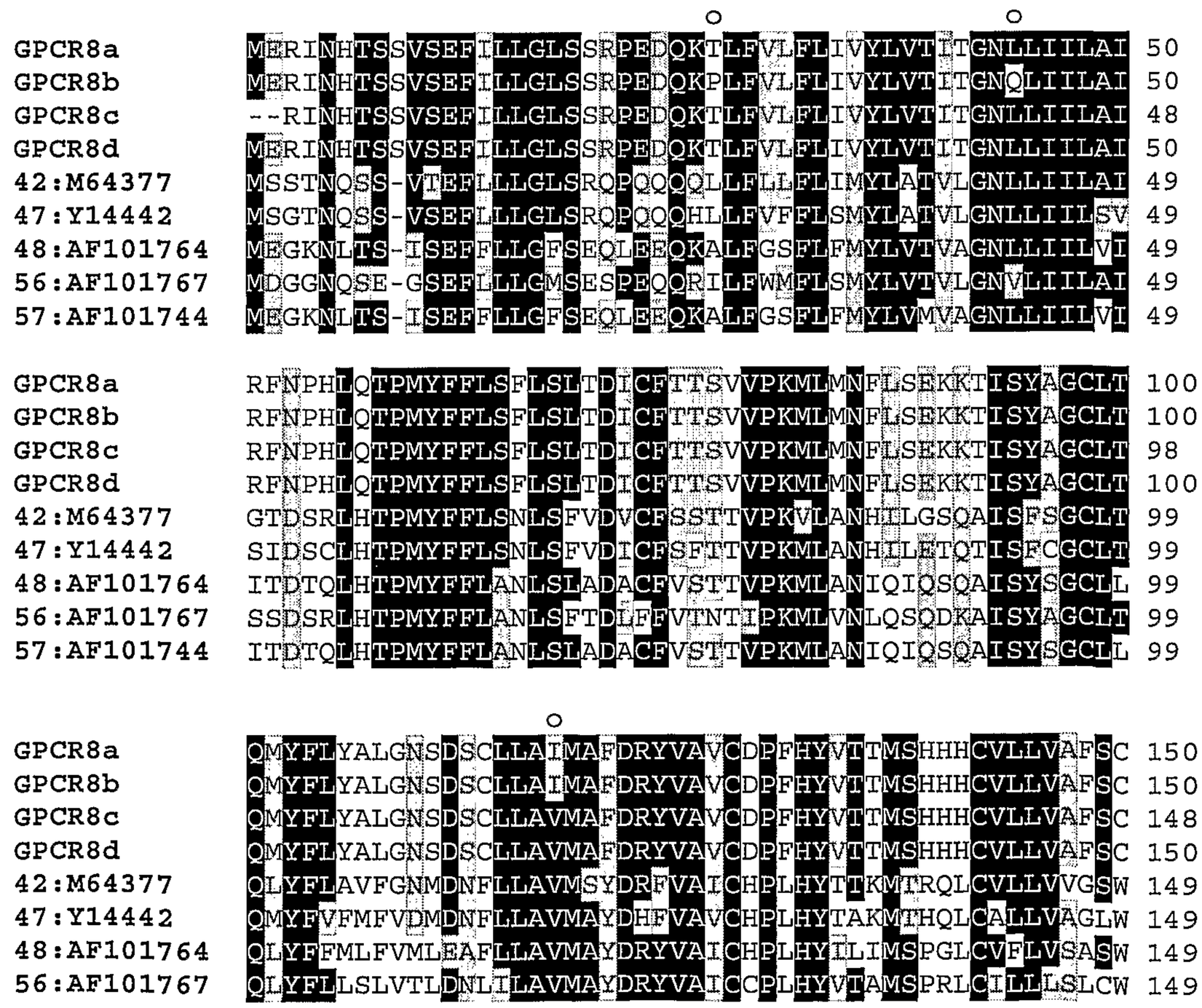
| Gene Index / Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |
|---|---|---|---|---|---|
| Y14442; AJ003147; U86234 | O1F1_HUMAN homo sapiens (human). olfactory receptor 1f1 (olfactory receptor 16-35) (or16-35). 10/2000 | 312 | 167/301 (55%) | 219/301, (73%) | 4e-92 |
| M64377; P23266 | OLF5_RAT rattus norvegicus (rat). olfactory receptor-like protein f5. 7/1993 | 313 | 162/302 (54%) | 219/302, (73%) | 1e-90 |

| | | | | | |
|---|---|---|---|---|---|
| AF101764 | Q9TU86 gorilla gorilla (gorilla). olfactory receptor. 3/2001 | 313 | 164/299 (55%) | 211/299, (71%) | 5e-89 |
| AF101767 | Q9TU84 pongo pygmaeus (orangutan). olfactory receptor. 3/2001 | 313 | 156/302 (52%) | 213/302, (71%) | 1e-86 |
| AF101744 | Q9TU99 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 313 | 159/299 (53%) | 207/299, (69%) | 4e-86 |

A multiple sequence alignment is given in Table 8K, with the GPCR8 proteins of the invention being shown on lines 1-4, in a ClustalW analysis comparing the GPCR8 polypeptides with related protein sequences shown in Table 8J. This BLASTP data is displayed graphically in the Clustal W in Table 8K.

**Table 8K. ClustalW Analysis of GPCR8**

1. SEQ ID NO:30, GPCR8a
2. SEQ ID NO:32, GPCR8b
3. SEQ ID NO:34,GPCR8c
4. SEQ ID NO:36, GPCR8d
5. SEQ ID NO:42, M64377 olfactory receptor-like protein f5
6. SEQ ID NO:47, Y14442 olfactory receptor 1f1 (or16-35)
7. SEQ ID NO:48, AF101764 gorilla olfactory receptor
8. SEQ ID NO:56, AF101767 orangutan olfactory receptor
9. SEQ ID NO:57, AF101744 chimpanzee olfactory receptor



```
57:AF101744   QLYFFMLFVMLEAFLLAVMAYDHYVAICHPLHYILIMSPGLCVFLVSASW 149

GPCR8a        SFPHLHSLLHTLLLNRLTFCDSNVIHHFLCDLSPVLKLSCSSIFVNEIVQ 200
GPCR8b        SFPHLHSLLHTLLLNRLTFCDSNVIHHFLCDLSPVLKLSCSSIFVNEIVQ 200
GPCR8c        SFPHLHSLLHTLLLNRLTFCDSNVIHHFLCDLSPVLKLSCSSIFVNEIVQ 198
GPCR8d        SFPHLHSLLHTLLLNRLTFCDSNVIHHFLCDLSPVLKLSCSSIFVNEIVQ 200
42:M64377     VVANMNCLLHILLMARLSFCADNMIPHFFCDGTPLLKLSCSDTHLNELMI 199
47:Y14442     VVANLNVLLHTLLMAPLSFCADNAITHFFCDVTPLLKLSCSNTHLNEVII 199
48:AF101764   IMNALHSLLHTLLMNSLSFCANHEIPHFFCDIDPLLSLSCTDPFTNELVI 199
56:AF101767   VFSVLYGLIHTLLMTRVTFCGSRKIHYLFCEMYFLLRLACSNIQINHTVL 199
57:AF101744   IMNALYSLLHTLLMNSLSFCANHEIPHFFCDIDPLLSLSCADPFTNELVI 199

GPCR8a        MTEAPIVLVTPFLCIAFSYIRILTTVLKIPSTSGKRKAFSTCGFYLTVVT 250
GPCR8b        MTEAPIVLVTPFLCIAFSYIRILTTVLKIPSTSGKRKAFSTCGFYLTVVT 250
GPCR8c        MTEAPIVLVTRFLCIAFSYIRILTTVLKIPSTSGKRKAFSTCGFYLTVVT 248
GPCR8d        MTEAPIVLVTRFLCIAFSYIRILTTVLKIPSTSGKRKAFSTCGFYLTVVT 250
42:M64377     LTEGAVVMVTPFVCILISYIHITCAVLRVSSPRGGWKSFSTCGSHLAVVC 249
47:Y14442     LSEGALVMITPFLCILASYMHITCTVLKVPSTKGRWKAFSTCGSHLAVVL 249
48:AF101764   FITGGLTGLVCVLCLIISYTNIFSTILKIPSAQGKRKAFSTCGSHLSVVS 249
56:AF101767   XATGCFIFLIPLGFMIXSYARIVRAILRIPSATGKYKAFSTCASHLAVVS 249
57:AF101744   FITGGLTGLICVLCLIISYTNVFSTILKIPSAQGKRKAFSTCSSHLSVVS 249

GPCR8a        LFYGSIFCVYLQPPSTY-AVKDHVATIVYTVLSSMLNPFIYSLRNKDLKQ 299
GPCR8b        LFYGSIFCVYLQPPSTY-AVKDHVATIVYTVLSSMLNPFIYSLRNKDLKQ 299
GPCR8c        LFYGSIFCVYLQPPSTY-AVKDHVATIVYTVLSSMLNPFIYSLRNKDLKQ 297
GPCR8d        LFYGSIFCVYLQPPSTY-AVKDHVATIVYTVLSSMLNPFIYSLRNKDLKQ 299
42:M64377     LFYGTVIAVYFNPSSSHLAGRDMAAAVMYAVVTPMLNPFIYSLRNSDMKA 299
47:Y14442     LFYSTIIAVYFNPLSSHSAEKDTMATVLYTVVTPMLNPFIYSLRNRYLKG 299
48:AF101764   LFFGISFCVYFIPPSTRSAQKDTVASVMYTVVTPMLNPFIYSLRNQEIKS 299
56:AF101767   LFYGTLGMVYLQPLQTY-STKDSVATVMYAVVTPMMNPFIYSLRNKDIHG 298
57:AF101744   LFXGISFCVYFSPPSTRXAQKDTVASVMYTVVTPMLNPFIYSLRNQEIKS 299

GPCR8a        GLRKLMSKRS----- 309
GPCR8b        GLRKLMSKRS----- 309
GPCR8c        GLRKLMSKRS----- 307
GPCR8d        GLRKLMSKRS----- 309
42:M64377     ALRKVLAMRFPSKQ- 313
47:Y14442     ALKKVIGRVVFSV-- 312
48:AF101764   SLRKLIWVREIHSP- 313
56:AF101767   ALGRLLQGKAFQKLT 313
57:AF101744   SLRKLIWVRKIHSP- 313
```

Table 8L lists the domain description from DOMAIN analysis results against GPCR8. This indicates that the GPCR8 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:39) itself.

**Table 8L Domain Analysis of GPCR8**

| PSSMs producing significant alignments: | | Score (bits) | E value |
|---|---|---|---|
| gnl\|Pfam\|pfam00001 | 7tm_1, 7 transmembrane receptor (rhodopsin family) | 103 | 2e-23 |

```
GPCR8:   42 GNLLIILAIRFNPHLQTPMYFFLSFLSLTDICFTTSVVPKMLMNFLSEKKTISYAGCLTQ 101
            ||||+||  |      |+||   ||   |++ |+  |   ++  |   |    +         | |
7tm_1:    1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV  60
```

```
GPCR8: 102 MYFLYALGNSDSCLLAIMAFDRYVAVCDPFHYVTTMSHHHCVLLVAFSCSFPHLHSLLHT 161
7tm_1:  61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPL 120

GPCR8: 162 LLLNRLTFCDSNVIHHFLCDLSPVLKLSCSSIFVNEIVQMTEAPIVLVTPFLCIAFSYIR 221
7tm_1: 121 LFSWLRTVEEGN---TTVCLIDFPEESVKRSYVLLSTL------VGFVLPLLVILVCYTR 171

GPCR8: 222 IL---------TTVLKIPSTSGKRKAFSTCGFYLTVVTLFYGS-----IFCVYLQPPSTY 267
7tm_1: 172 ILRTLRKRARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRV 231

GPCR8: 268 AVKDHVATIVYTVLSSMLNPFIY 290
7tm_1: 232 LPTALLITLWLAYVNSCLNPIIY 254
```

The GPCR8a is expressed in at least some of the following tissues: adrenal gland, bone marrow, brain - amygdala, brain - cerebellum, brain - hippocampus, brain - substantia nigra, brain - thalamus, brain -whole, fetal brain, fetal kidney, fetal liver, fetal lung, heart, kidney, lymphoma - Raji, mammary gland, pancreas, pituitary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen, stomach, testis, thyroid, trachea, uterus.

The GPCR8c is expressed in at least the following tissues: Apical microvilli of the retinal pigment epithelium, arterial (aortic), basal forebrain, brain, Burkitt lymphoma cell lines, corpus callosum, cardiac (atria and ventricle), caudate nucleus, CNS and peripheral tissue, cerebellum, cerebral cortex, colon, cortical neurogenic cells, endothelial (coronary artery and umbilical vein) cells, palate epithelia, eye, neonatal eye, frontal cortex, fetal hematopoietic cells, heart, hippocampus, hypothalamus, leukocytes, liver, fetal liver, lung, lung lymphoma cell lines, fetal lymphoid tissue, adult lymphoid tissue, Those that express MHC II and III nervous, medulla, subthalamic nucleus, ovary, pancreas, pituitary, placenta, pons, prostate, putamen, serum, skeletal muscle, small intestine, smooth muscle (coronary artery in aortic) spinal cord, spleen, stomach, taste receptor cells of the tongue, testis, thalamus, and thymus tissue. This information was derived by determining the tissue sources of the sequences that were included in the invention including but not limited to SeqCalling sources, Public EST sources, Genomic Clone sources, Literature sources, and/or RACE sources.

In addition, the GPCR8 gene is predicted to be expressed in brain because of the expression pattern of many odorant receptor in that organ. The GPCR8 gene is also predicted to be expressed in the following tissues because of the expression pattern of (GENBANK-ID:

gb:GENBANK-ID:RATOLFPROC|acc:M64377.1) a closely related {Rat olfactory protein mRNA, complete cds homolog in species Rattus norvegicus: ventromedial hypothalamus, brain cortex, frontal cortex, cerebellum, pons, striatum, and thalamus testis, brain cortical structures, including the anterior cingulate, posterior cingulate, and frontoparietal, somatosensory, and piriform cortices, Olfactory tubercle, the islands of Calleja, ventromedial and posterior nuclei of the hypothalamus, the medial septal nucleus, the nucleus of the diagonal band, and the ventral tegmental area.

This is by no way limiting in that olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types. Further tissue expression analysis is provided in the Examples.

The GPCR8 nucleic acids and proteins are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein. A monoclonal antibody targeting a GPCR8 protein, specifically its extracellular region, will have a therapeutic role in treating cancer. It will also have a role in treating angiogenesis related diseases. Being a GPCR, it could be used to screen for small molecule drug to treat cancer.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR8 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR8 epitope is from about amino acids 1 to 25. In additional embodiments, GPCR8 epitopes are from about amino acids 80 to 100, from about amino acids 130 to 140, from about amino acids 225 to 245 and from about amino acids 285 to 309.

**GPCR9**

A novel GPCR9 (also referred to as GMba64p14_I) clone was identified. The GPCR9 nucleic acid (SEQ ID NO:37) of 958 nucleotides is shown in Table 9A. The GPCR9 ORF begins with an ATG initiation codon at nucleotides 1-3 and ends with a TAG codon at nucleotides 939-941. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 9A, and the start and stop codons are in bold letters.

**Table 9A. GPCR9 Nucleotide Sequence (SEQ ID NO:37)**

**ATG**GACAACAGCAACTGGACCAGTGTGTCCCATTTTGTTCTCTTGGGCATTTCCACCCACCCAGAA
GAGCAAATCCCACTCTTCCTTGTTTTCTCACTCATGTACGCAATCAATATTTCTGGCAACTTGGCC
ATCATCACACTGATTCTCTCTGCTCCACGCCTCCACATCCCCATGTACATCTTCCTCAGTAACTTG
GCCTTGACAGACATCTGCTTCACCTCCACCACGGTCCCCAAGATGCTGCAGATTATTTTCTCCCCT
ACAAAGGTAATTTCCTACACAGGCTGTTTAGCCCAAACTTATTTCTTCATTTGCTTCGCCGTCATG
GAAAACTTCATCCTGGCTGTGATGGCCTATGACAGGTACATTGCCATCTGCCACCCTTTCCACTAC
ACTATGATCCTGACTAGAATGCTGTGTGTGAAGATGGTGGTCATGTGCCATGCTCTCTCCCACCTT
CATGCCATGCTGCATACCTTTCTCATAGGCCAACTAATCTTCTGTGCAGATAACAGAATCCCCCAC
TTCTTCTGTGACCTCTACGCTCTGATGAAGATCTCCTGCACCAGCACCTACCTCAACACCCTTATG
ATTCACACAGAAGGTGCTGTTGTAATCAGTGGAGCTCTGGCCTTCATTACTGCCTCCTATGCCTGC
ATCATCCTGGTGGTCCTCCGGATCCCCTCAGCCAAGGGCAGGTGGAAAACCTTTTCTACCTGCGGC
TCCCACCTCACTGTGGTGGCCATATTCTATGGCACCCTCAGTTGGGTCTACTTCCGGCCCCTTTCC
AGCTATTCAGTGACCAAGGGTCGCATTATAACAGTCGTGTACACAGTGGTGACTCCCATGCTGAAC
CCCTTCATCTACAGCCTGAGGAATGGGGATGTCAAGGGAGGCTTCATGAAATGGATGAGCAGAATG
CAGACTTTTTTCTT**TAG**ATAAAACCCCAAACACA

The GPCR9a polypeptide (SEQ ID NO:38) encoded by SEQ ID NO:37 is 314 aa in length, has a molecular weight of 35597.1 Daltons, and is presented using the one-letter amino acid code in Table 9B. The Psort profile for GPCR9 predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR9 polypeptide is located to the Golgi body with a certainty of 0.400, the endoplasmic reticulum (membrane) with a certainty of 0.300, or a microbody (peroxisome) with a certainty of 0.300. The Signal P software program predicts no likely signal cleavage site for a GPCR9 peptide.

**Table 9B. GPCR9a protein sequence (SEQ ID NO:38)**

MDNSNWTSVSHFVLLGISTHPEEQIPLFLVFSLMYAINISGNLAIITLILSAPRLHIPMYIFLSNLA
LTDICFTSTTVPKMLQIIFSPTKVISYTGCLAQTYFFICFAVMENFILAVMAYDRYIAICHPFHYTM
ILTRMLCVKMVVMCHALSHLHAMLHTFLIGQLIFCADNRIPHFFCDLYALMKISCTSTYLNTLMIHT
EGAVVISGALAFITASYACIILVVLRIPSAKGRWKTFSTCGSHLTVVAIFYGTLSWVYFRPLSSYSV
TKGRIITVVYTVVTPMLNPFIYSLRNGDVKGGFMKWMSRMQTFFFR

The amino acid sequence of GPCR9 had high homology to other proteins as shown in Table 9C.

**Table 9C. BLASTX results for GPCR9**

| Sequences producing High-scoring Segment Pairs: | Reading Frame | High Score | Smallest Sum Prob P(N) |
|---|---|---|---|
| patp:AAR27868 Odorant receptor clone F5 - Rattus rattus, 313 aa. | +1 | 841 | 3.9e-83 |

Additional BLASTP results for GPCR9 are shown in Table 9D.

**Table 9D. GPCR9 BLASTP results**

| Gene Index / Identifier | Protein / Organism | Length (aa) | Identity (%) | Positives (%) | Expect |
|---|---|---|---|---|---|
| M64377; P23266 | OLF5_RAT rattus norvegicus (rat). olfactory receptor-like protein f5. 7/1993 | 313 | 161/303 (53%) | 210/303, (69%) | 4e-90 |
| Y14442; AJ003147; U86234 | O1F1_HUMAN homo sapiens (human). olfactory receptor 1f1 (olfactory receptor 16-35) (or16-35). 10/2000 | 312 | 162/308 (53%) | 208/308, (68%) | 1e-89 |
| X89667 | Q15612 homo sapiens (human). putative olfactory receptor (fragment). 6/2000 | 157 | 155/157 ([redacted] | 156/157, ([redacted]) | [redacted] |
| AF101749 | Q9TU94 gorilla gorilla (gorilla). olfactory receptor. 3/2001 | 314 | 152/312 (49%) | 212/312, (68%) | 5e-87 |
| AF101730 | Q9TUA9 pan troglodytes (chimpanzee). olfactory receptor. 3/2001 | 314 | 151/312 (48%) | 211/312, (68%) | 2e-86 |
| AF101764 | Q9TU86 gorilla gorilla (gorilla). olfactory receptor. 3/2001 | 313 | 155/303 (51%) | 207/303, (68%) | 3e-86 |

A multiple sequence alignment is given in Table 9E, with the GPCR9 protein of the invention being shown on line 1, in a ClustalW analysis comparing GPCR9 with the related protein sequences listed in Table 9D.

**Table 9E. Information for the ClustalW proteins:**

1. SEQ ID NO:38, GPCR9
2. SEQ ID NO:42, M64377 olfactory receptor-like protein f5
3. SEQ ID NO:44, AF101730 chimpanzee olfactory receptor
4. SEQ ID NO:47, Y14442 olfactory receptor 1f1 (or16-35)
5. SEQ ID NO:58, X89667 putative olfactory receptor (fragment)
6. SEQ ID NO:59, AF101749 gorilla olfactory receptor

```
GPCR9          MDNSNWISVSHFVLLGISTHPEEQIPLFLVFSLMYAINISGNLAIITLIL 50
42:M64377      MSSTNQSSVTEFLLLGLSRQPQQQQLLFLLFLIMYLATVLGNLLIILAIG 50
44:AF101730    --MENQSSVSEFFLQGISGFPEQQQLLYGLFLCMYLVTLTGNVLIIMAIG 48
47:Y14442      MSGTNQSSVSEFLLLGLSRQPQQQHLLFVFFLSMYLATVLGNLLIILSVS 50
58:X89667      -------------------------------------------------- 1
59:AF101749    MMGQNQTSISDFLLLGLPIQPEQQNLCYALFLAMYLTTLLGNLLIIVLIR 50

GPCR9          SAPRLHIPMYIFLSNLALTDICFTSTTVPKMLQIIFSPTKVISYTGCLAQ 100
42:M64377      TDSRLHTPMYFFLSNLSFVDVCFSSTTVPKVLANHILGSQAISFSGCLTQ 100
44:AF101730    SDPHLHTPMYFFLANLSFADMGLISSTVTQMLFNVQTQRHTISYTGCLTQ 98
47:Y14442      IDSCLHTPMYFFLSNLSFVDICFSFTTVPKMLANHILETQTISFCGCLTQ 100
58:X89667      -------------------------------------------------- 1
59:AF101749    LDSHLHTPMYLFLSNLSFSDLCFSSVTIPKLLQNMQNQDPSIPYADCLTQ 100

GPCR9          TYFFICFAVMENFILAVMAYDRYIAICHPFHYTMILTRMLCVKMVVMCHA 150
```

```
          42:M64377     LYFLAVFGNMDNFLLAVMSYDRFVAICHPLHYTTKMTRQLCVLLVVGSWV 150
          44:AF101730   MYFFLMFGDLDSFFLAVMAYDRYVAICHPLHYSTIMRAKVCVLMLALCWV 148
          47:Y14442     MYFVFMFVDMDNFLLAVMAYDHFVAVCHPLHYTAKMTHQLCALLVAGLWV 150
          58:X89667     -------------------------FCHPFHYTMILTRMLCVKMVVMCHA 25
5         59:AF101749   MYFFLLFGDLESFLLVAMAYDRYVAICFALHYTAIMSPMLCLSLVALSWV 150

          GPCR9         LSHLHAMLHTFLIGQLIFCADNRIPHFFCDLYALMKISCTSTYLNTLMIH 200
          42:M64377     VANMNCLLHILLMARLSFCADNMIPHFFCDGTPLLKLSCSDTHLNELMIL 200
          44:AF101730   LTNIVALTHTLLMARLSFCVVGEIAHFFCDITPVLKLSCSDTYVNELMVF 198
10        47:Y14442     VANLNVLLHTLLMAPLSFCADNAITHFFCDVTPLLKLSCSNTHLNEVIIL 200
          58:X89667     LSHLHAMLHTFLMGQLIFCADNRIPHFFCDLYALMKISCTSTYLNTLMIH 75
          59:AF101749   LTTFHAMLHTLLMARLCFCADNVIPHFFCDMSALLKLACSDTRVNEWVIF 200

          GPCR9         TEGAVVISGALAFITASYACIILVVLRIPSAKGRWKTFSTCGSHLTVVAI 250
15        42:M64377     TEGAVVMVTPFVCILISYIHITCAVLRVSSPRGGWKSFSTCGSHLAVVCL 250
          44:AF101730   ALGGTVLMLPFICIVISYIHIVFAILRVRTPGGGTKAFSTCSSHLCVVCV 248
          47:Y14442     SEGALVMITPFLCILASYMHITCTVLKVPSTKGRWKAFSTCGSHLAVVLL 250
          58:X89667     TEGAVVISGALAFITASYACIILVVLRIPSAKGRWKTFSTCGSHLTVVAI 125
          59:AF101749   IMGGLIVVIPFLLILGSYARIVSSILKVPSSKGICKAFSTCGSHLSVVSI 250
20
          GPCR9         FYGTLSWVYFRPLSSYSVTKGRIITVVYTVVTPMLNPFIYSLRNGDVKGG 300
          42:M64377     FYGTVIAVYFNPSSSHLAGRDMAAAVMYAVVTPMLNPFIYSLRNSDMKAA 300
          44:AF101730   FYGTLFSAYLSPPSVVSTEKDIAAAAMYTVVTPMLNPFIYSLRNKDMKGA 298
          47:Y14442     FYSTIIAVYFNPLSSHSAEKDTMATVLYTVVTPMLNPFIYSLRNRYLKGA 300
25        58:X89667     FYGTLSWVYFRPLSSYSVTKGRIITVVYTVVT------------------- 157
          59:AF101749   FYGTVIGLYLCPSANSSTLKETVMAMMYTVVTPMLNPFIYSLRNGDMKGA 300

          GPCR9         FMKWMSRMQTFFFR 314
          42:M64377     LRKVLAMRFPSKQ- 313
30        44:AF101730   LKRLLFHRRILSS- 311
          47:Y14442     LKKVIGRVVFSV-- 312
          58:X89667     -------------- 157
          59:AF101749   LSRVIHQKKTFFSL 314
```

35 DOMAIN results for GPCR9 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. The results are listed in Table 9F with the statistics and domain description. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to have significant homology to GPCR9. An alignment of GPCR9 residues 41-287 (SEQ ID
40 NO:26) with 7tm_1 residues 1-254 (SEQ ID NO:39) are shown in Table 9F.

**Table 9F. DOMAIN results for GPCR9**

| PSSMs producing significant alignments: | | Score (bits) | E value |
|---|---|---|---|
| gnl\|Pfam\|pfam00001 | 7tm_1, 7 transmembrane receptor (rhodopsin family) | 95.9 | 3e-21 |

```
45  GPCR9:   41 GNLAIITLILSAPRLHIPMYIFLSNLALTDICFTSTTVPKMLQIIFSPTKVISYTGCLAQ  100
                |||  +|  +||    +|   |   |||  |||+  |+  |   |   |   |   +       |       |
    7tm_1:    1 GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLV   60

    GPCR9:  101 TYFFICFAVMENFILAVMAYDRYIAICHPFHYTMILTRMLCVKMVVMCHALSHLHAMLHT  160
                   |+          +|   ++ |||+||  ||   |   | |        ++++    |+  |  ++
    7tm_1:   61 GALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPL  120
50
```

```
GPCR9: 161 FLIGQLIFCADNRIPHFFCDLYALMKISCTSTYLNTLMIHTEGAVVISGALAFI--TASY 218
                      |              +|    +    |+||+       +||       |   |
7tm_1: 121 LFSWLRTVEEGNTTVCLIDFPEESVKR--SYVLLSTLVGFVLPLLVILVCYTRILRTLRK 178

GPCR9: 219 ACIILVVLRIPSAKGRWKTFSTCGSHLTVVAIFYG----TLSWVYFRPLSSYSVTKGRII 274
                 |+   |+   |            +   |   +        |               +    +|
7tm_1: 179 RARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRVLPTALLI 238

GPCR9: 275 TVVYTVVTPMLNPFIY 290
           |+    |   |||  ||
7tm_1: 239 TLWLAYVNSCLNPIIY 254
```

The nucleic acids and proteins of GPCR9 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further above.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR9 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR9 epitope is from about amino acids 5 to 20. In other specific embodiments, GPCR9 epitopes are from about amino acids 240 to 250, from about amino acids 255 to 265 and from about amino acids 285 to 314.

**GPCRX Nucleic Acids and Polypeptides**

A summary of the GPCRX nucleic acids and proteins of the invention is provided in Table 10.

**TABLE 10: Summary Of Nucleic Acids And Proteins Of The Invention**

| GPCRX Name | Tables | Clone; Description of Homolog | Nucleic Acid SEQ ID NO | Amino Acid SEQ ID NO |
|---|---|---|---|---|
| GPCR1 | 1A, 1B | **GMba64p14_A** | 1 | 2 |
| GPCR 2a | 2A, 2B | **GMba64p14_B** | 3 | 4 |
| GPCR 2b | 2C, 2D | **CG56582-01** | 5 | 6 |
| GPCR 3 | 3A, 3B | **GMba64p14_C** | 7 | 8 |

| | | | | |
|---|---|---|---|---|
| GPCR 4a | 4A, 4B | **CG55940-01** | 9 | 10 |
| GPCR 4b | 4C, 4D | **CG55940-02** | 11 | 12 |
| GPCR 4c | 4E, 4F | **GMba64p14_D** | 13 | 14 |
| GPCR 5a | 5A, 5B | **CG50385-03** | 15 | 16 |
| GPCR 5b | 5C, 5D | **GMba64p14_E** | 17 | 18 |
| GPCR5c | 5E, 5F | **CG50385-01** | 108 | 109 |
| GPCR 6a | 6A, 6B | **ba460n11_da1** | 19 | 20 |
| GPCR 6b | 6C, 6D | **GMba64p14_F** | 21 | 22 |
| GPCR 6c | 6E, 6F | **ba460n11_da2; 147307499** | 23 | 24 |
| GPCR 7a | 7A, 7B | **CG57809-01** | 25 | 26 |
| GPCR 7b | 7C, 7D | **GMba64p14_G** | 27 | 28 |
| GPCR 8a | 8A, 8B | **ba542k23** | 29 | 30 |
| GPCR 8b | 8C, 8D | **CG50259-01; 148540666** | 31 | 32 |
| GPCR 8c | 8E, 8F | **AL162254-da1** | 33 | 34 |
| GPCR 8d | 8G, 8I | **GMba64p14_H** | 35 | 36 |
| GPCR 9 | 9A, 9B | **GMba64p14_I** | 37 | 38 |

One aspect of the invention pertains to isolated nucleic acid molecules that encode GPCRX polypeptides or biologically active portions thereof. Also included in the invention are nucleic acid fragments sufficient for use as hybridization probes to identify GPCRX-encoding nucleic acids (*e.g.*, GPCRX mRNAs) and fragments for use as PCR primers for the amplification and/or mutation of GPCRX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule may be single-stranded or double-stranded, but preferably is comprised double-stranded DNA.

An GPCRX nucleic acid can encode a mature GPCRX polypeptide. As used herein, a "mature" form of a polypeptide or protein disclosed in the present invention is the product of a naturally occurring polypeptide or precursor form or proprotein. The naturally occurring polypeptide, precursor or proprotein includes, by way of nonlimiting example, the full-length gene product, encoded by the corresponding gene. Alternatively, it may be defined as the polypeptide, precursor or proprotein encoded by an ORF described herein. The product "mature" form arises, again by way of nonlimiting example, as a result of one or more naturally occurring processing steps as they may take place within the cell, or host cell, in which the gene product arises. Examples of such processing steps leading to a "mature" form of a polypeptide or protein include the cleavage of the N-terminal methionine residue encoded by the initiation codon of an ORF, or the proteolytic cleavage of a signal peptide or leader

sequence. Thus a mature form arising from a precursor polypeptide or protein that has residues 1 to N, where residue 1 is the N-terminal methionine, would have residues 2 through N remaining after removal of the N-terminal methionine. Alternatively, a mature form arising from a precursor polypeptide or protein having residues 1 to N, in which an N-terminal signal sequence from residue 1 to residue M is cleaved, would have the residues from residue M+1 to residue N remaining. Further as used herein, a "mature" form of a polypeptide or protein may arise from a step of post-translational modification other than a proteolytic cleavage event. Such additional processes include, by way of non-limiting example, glycosylation, myristoylation or phosphorylation. In general, a mature polypeptide or protein may result from the operation of only one of these processes, or a combination of any of them.

The term "probes", as utilized herein, refers to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as approximately, *e.g.*, 6,000 nt, depending upon the specific use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are generally obtained from a natural or recombinant source, are highly specific, and much slower to hybridize than shorter-length oligomer probes. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

The term "isolated" nucleic acid molecule, as utilized herein, is one, which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5'- and 3'-termini of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated GPCRX nucleic acid molecules can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell/tissue from which the nucleic acid is derived (*e.g.*, brain, heart, liver, spleen, etc.). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, or a complement of this aforementioned nucleotide sequence, can be isolated using

standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 as a hybridization probe, GPCRX molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, *et al.*, (eds.), 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 & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to GPCRX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment of the invention, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, or a complement thereof. Oligonucleotides may be chemically synthesized and may also be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, or a portion of this nucleotide sequence (*e.g.*, a fragment that can be used as a probe or primer or a fragment encoding a biologically-active portion of an GPCRX polypeptide). A nucleic acid molecule that is complementary to the nucleotide sequence shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 is one that is sufficiently complementary to the nucleotide sequence shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 that it can hydrogen bond with little or no mismatches to the nucleotide

sequence shown SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, thereby forming a stable duplex.

As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, van der Waals, hydrophobic interactions, and the like. A physical interaction can be either direct or indirect. Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively, and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type. Homologs are nucleic acid sequences or amino acid sequences of a particular gene that are derived from different species.

Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the invention, in various embodiments, by at least about 70%, 80%, or 95% identity (with a preferred identity of 80-95%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under

stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of GPCRX polypeptides. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the invention, homologous nucleotide sequences include nucleotide sequences encoding for an GPCRX polypeptide of species other than humans, including, but not limited to: vertebrates, and thus can include, *e.g.*, frog, mouse, rat, rabbit, dog, cat cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the exact nucleotide sequence encoding human GPCRX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, as well as a polypeptide possessing GPCRX biological activity. Various biological activities of the GPCRX proteins are described below.

As used herein, "identical" residues correspond to those residues in a comparison between two sequences where the equivalent nucleotide base or amino acid residue in an alignment of two sequences is the same residue. Residues are alternatively described as "similar" or "positive" when the comparisons between two sequences in an alignment show that residues in an equivalent position in a comparison are either the same amino acid or a conserved amino acid as defined below.

An GPCRX polypeptide is encoded by the open reading frame ("ORF") of an GPCRX nucleic acid. An ORF corresponds to a nucleotide sequence that could potentially be translated into a polypeptide. A stretch of nucleic acids comprising an ORF is uninterrupted by a stop codon. An ORF that represents the coding sequence for a full protein begins with an ATG "start" codon and terminates with one of the three "stop" codons, namely, TAA, TAG, or TGA. For the purposes of this invention, an ORF may be any part of a coding sequence, with or without a start codon, a stop codon, or both. For an ORF to be considered as a good candidate for coding for a *bona fide* cellular protein, a minimum size requirement is often set, *e.g.,* a stretch of DNA that would encode a protein of 50 amino acids or more.

The nucleotide sequences determined from the cloning of the human GPCRX genes allows for the generation of probes and primers designed for use in identifying and/or cloning GPCRX homologues in other cell types, *e.g.* from other tissues, as well as GPCRX homologues from other vertebrates. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 consecutive sense strand nucleotide sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37; or an anti-sense strand nucleotide sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37; or of a naturally occurring mutant of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37.

Probes based on the human GPCRX nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.* the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissues which mis-express an GPCRX protein, such as by measuring a level of an GPCRX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting GPCRX mRNA levels or determining whether a genomic GPCRX gene has been mutated or deleted.

"A polypeptide having a biologically-active portion of an GPCRX polypeptide" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically-active portion of GPCRX" can be prepared by isolating a portion SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 that encodes a polypeptide having an GPCRX biological activity (the biological activities of the GPCRX proteins are described below), expressing the encoded portion of GPCRX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of GPCRX.

**GPCRX Nucleic Acid and Polypeptide Variants**

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequences shown SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 due to degeneracy of the genetic code and thus encode the same GPCRX

proteins as that encoded by the nucleotide sequences shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38.

In addition to the human GPCRX nucleotide sequences shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the GPCRX polypeptides may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the GPCRX genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame (ORF) encoding an GPCRX protein, preferably a vertebrate GPCRX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the GPCRX genes. Any and all such nucleotide variations and resulting amino acid polymorphisms in the GPCRX polypeptides, which are the result of natural allelic variation and that do not alter the functional activity of the GPCRX polypeptides, are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding GPCRX proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the GPCRX cDNAs of the invention can be isolated based on their homology to the human GPCRX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37. In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500, 750, 1000, 1500, or 2000 or more nucleotides in length. In yet another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended

to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.,* nucleic acids encoding GPCRX proteins derived from species other than human) or other related sequences (*e.g.,* paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5 °C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at Tm, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (*e.g.*, 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in Ausubel, *et al*., (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions are hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C, followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequences of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a

"naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well-known within the art. See, *e.g.*, Ausubel, et *al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990; GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequences of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as employed for cross-species hybridizations). See, *e.g.,* Ausubel, *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981. *Proc Natl Acad Sci USA* **78**: 6789-6792.

**Conservative Mutations**

In addition to naturally-occurring allelic variants of GPCRX sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 thereby leading to changes in the amino acid sequences of the encoded GPCRX proteins, without altering the functional ability of said GPCRX proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20,

22, 24, 26, 28, 30, 32, 34, 36 and 38. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequences of the GPCRX proteins without altering their biological activity, whereas an "essential" amino acid residue is required for such biological activity. For example, amino acid residues that are conserved among the GPCRX proteins of the invention are predicted to be particularly non-amenable to alteration. Amino acids for which conservative substitutions can be made are well-known within the art.

Another aspect of the invention pertains to nucleic acid molecules encoding GPCRX proteins that contain changes in amino acid residues that are not essential for activity. Such GPCRX proteins differ in amino acid sequence from SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38 yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 45% homologous to the amino acid sequences of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38. Preferably, the protein encoded by the nucleic acid molecule is at least about 60% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38; more preferably at least about 70% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38; still more preferably at least about 80% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38; even more preferably at least about 90% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38; and most preferably at least about 95% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38.

An isolated nucleic acid molecule encoding an GPCRX protein homologous to the protein of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted, non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue

having a similar side chain. Families of amino acid residues having similar side chains have been defined within the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted non-essential amino acid residue in the GPCRX protein is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an GPCRX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for GPCRX biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

The relatedness of amino acid families may also be determined based on side chain interactions. Substituted amino acids may be fully conserved "strong" residues or fully conserved "weak" residues. The "strong" group of conserved amino acid residues may be any one of the following groups: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW, wherein the single letter amino acid codes are grouped by those amino acids that may be substituted for each other. Likewise, the "weak" group of conserved residues may be any one of the following: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, VLIM, HFY, wherein the letters within each group represent the single letter amino acid code.

In one embodiment, a mutant GPCRX protein can be assayed for (*i*) the ability to form protein:protein interactions with other GPCRX proteins, other cell-surface proteins, or biologically-active portions thereof, (*ii*) complex formation between a mutant GPCRX protein and an GPCRX ligand; or (*iii*) the ability of a mutant GPCRX protein to bind to an intracellular target protein or biologically-active portion thereof; (*e.g.* avidin proteins).

In yet another embodiment, a mutant GPCRX protein can be assayed for the ability to regulate a specific biological function (*e.g.*, regulation of insulin release).

**Antisense Nucleic Acids**

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein (*e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence). In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire GPCRX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of an GPCRX protein of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, or antisense nucleic acids complementary to an GPCRX nucleic acid sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an GPCRX protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding the GPCRX protein. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding the GPCRX protein disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of GPCRX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of GPCRX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of GPCRX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally-occurring nucleotides or

variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids (*e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used).

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an GPCRX protein to thereby inhibit expression of the protein (*e.g.*, by inhibiting transcription and/or translation). The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface (*e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens). The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient nucleic acid molecules, vector

constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other. See, *e.g.*, Gaultier, *et al.*, 1987. *Nucl. Acids Res.* **15**: 6625-6641. The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (see, *e.g.*, Inoue, *et al.* 1987. *Nucl. Acids Res.* **15**: 6131-6148) or a chimeric RNA-DNA analogue (see, *e.g.*, Inoue, *et al.*, 1987. *FEBS Lett.* **215**: 327-330.

**Ribozymes and PNA Moieties**

Nucleic acid modifications include, by way of non-limiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In one embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes as described in Haselhoff and Gerlach 1988. *Nature* 334: 585-591) can be used to catalytically cleave GPCRX mRNA transcripts to thereby inhibit translation of GPCRX mRNA. A ribozyme having specificity for an GPCRX-encoding nucleic acid can be designed based upon the nucleotide sequence of an GPCRX cDNA disclosed herein (*i.e.*, SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an GPCRX-encoding mRNA. See, *e.g.*, U.S. Patent 4,987,071 to Cech, *et al.* and U.S. Patent 5,116,742 to Cech, *et al.* GPCRX mRNA can also be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, GPCRX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the GPCRX nucleic acid (*e.g.*, the GPCRX promoter and/or enhancers) to form triple helical structures that prevent transcription

of the GPCRX gene in target cells. *See, e.g.,* Helene, 1991. *Anticancer Drug Des.* 6: 569-84; Helene, *et al.* 1992. *Ann. N.Y. Acad. Sci.* 660: 27-36; Maher, 1992. *Bioassays* 14: 807-15.

In various embodiments, the GPCRX nucleic acids can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids. *See, e.g.,* Hyrup, *et al.*, 1996. *Bioorg Med Chem* 4: 5-23. As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics (*e.g.*, DNA mimics) in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup, *et al.*, 1996. *supra*; Perry-O'Keefe, *et al.*, 1996. *Proc. Natl. Acad. Sci. USA* 93: 14670-14675.

PNAs of GPCRX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of GPCRX can also be used, for example, in the analysis of single base pair mutations in a gene (*e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, $S_1$ nucleases (*see,* Hyrup, *et al.,* 1996.*supra*); or as probes or primers for DNA sequence and hybridization (*see,* Hyrup, *et al.*, 1996, *supra*; Perry-O'Keefe, *et al.,* 1996. *supra*).

In another embodiment, PNAs of GPCRX can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of GPCRX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes (*e.g.*, RNase H and DNA polymerases) to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (*see,* Hyrup, et al., 1996. *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup, *et al.,* 1996. *supra* and Finn, *et al.*, 1996. *Nucl Acids Res* 24: 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and

modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA. *See, e.g.,* Mag, *et al.*, 1989. *Nucl Acid Res* 17: 5973-5988. PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment. *See, e.g.,* Finn, *et al.*, 1996. *supra.* Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. *See, e.g.,* Petersen, *et al.*, 1975. *Bioorg. Med. Chem. Lett.* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (*see, e.g.*, Letsinger, *et al.*, 1989. *Proc. Natl. Acad. Sci. U.S.A.* 86: 6553-6556; Lemaitre, *et al.*, 1987. *Proc. Natl. Acad. Sci.* 84: 648-652; PCT Publication No. WO88/09810) or the blood-brain barrier (*see, e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (*see, e.g.,* Krol, *et al.*, 1988. *BioTechniques* 6:958-976) or intercalating agents (*see, e.g.*, Zon, 1988. *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, and the like.

**GPCRX Polypeptides**

A polypeptide according to the invention includes a polypeptide including the amino acid sequence of GPCRX polypeptides whose sequences are provided in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38. The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residues shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38 while still encoding a protein that maintains its GPCRX activities and physiological functions, or a functional fragment thereof.

In general, an GPCRX variant that preserves GPCRX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above.

One aspect of the invention pertains to isolated GPCRX proteins, and biologically-active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-GPCRX antibodies. In one embodiment, native GPCRX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, GPCRX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, an GPCRX protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or protein or biologically-active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the GPCRX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of GPCRX proteins in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly-produced. In one embodiment, the language "substantially free of cellular material" includes preparations of GPCRX proteins having less than about 30% (by dry weight) of non-GPCRX proteins (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-GPCRX proteins, still more preferably less than about 10% of non-GPCRX proteins, and most preferably less than about 5% of non-GPCRX proteins. When the GPCRX protein or biologically-active portion thereof is recombinantly-produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the GPCRX protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of GPCRX proteins in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of GPCRX proteins having less than about 30% (by dry weight) of chemical precursors or non-GPCRX chemicals, more preferably less than about 20% chemical precursors or non-GPCRX chemicals, still more preferably less than about 10% chemical precursors or non-GPCRX chemicals, and most preferably less than about 5% chemical precursors or non-GPCRX chemicals.

Biologically-active portions of GPCRX proteins include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequences of the GPCRX proteins (*e.g.*, the amino acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38) that include fewer amino acids than the full-length GPCRX proteins, and exhibit at least one activity of an GPCRX protein. Typically, biologically-active portions comprise a domain or motif with at least one activity of the GPCRX protein. A biologically-active portion of an GPCRX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acid residues in length.

Moreover, other biologically-active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native GPCRX protein.

In an embodiment, the GPCRX protein has an amino acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38. In other embodiments, the GPCRX protein is substantially homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, and retains the functional activity of the protein of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail, below. Accordingly, in another embodiment, the GPCRX protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, and retains the functional activity of the GPCRX proteins of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38.

**Determining Homology Between Two or More Sequences**

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. *See,* Needleman and Wunsch, 1970. *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37.

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.,* A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.,* the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

**Chimeric and Fusion Proteins**

The invention also provides GPCRX chimeric or fusion proteins. As used herein, an GPCRX "chimeric protein" or "fusion protein" comprises an GPCRX polypeptide operatively-linked to a non-GPCRX polypeptide. An "GPCRX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to an GPCRX protein (SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38), whereas a "non-GPCRX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the GPCRX protein, *e.g.*, a protein that is different from the GPCRX protein and that is derived from the same or a different organism. Within an GPCRX fusion protein the GPCRX polypeptide can correspond to all or a portion of an GPCRX protein. In one embodiment, an GPCRX fusion protein comprises at least one

biologically-active portion of an GPCRX protein. In another embodiment, an GPCRX fusion protein comprises at least two biologically-active portions of an GPCRX protein. In yet another embodiment, an GPCRX fusion protein comprises at least three biologically-active portions of an GPCRX protein. Within the fusion protein, the term "operatively-linked" is intended to indicate that the GPCRX polypeptide and the non-GPCRX polypeptide are fused in-frame with one another. The non-GPCRX polypeptide can be fused to the N-terminus or C-terminus of the GPCRX polypeptide.

In one embodiment, the fusion protein is a GST-GPCRX fusion protein in which the GPCRX sequences are fused to the C-terminus of the GST (glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant GPCRX polypeptides.

In another embodiment, the fusion protein is an GPCRX protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of GPCRX can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is an GPCRX-immunoglobulin fusion protein in which the GPCRX sequences are fused to sequences derived from a member of the immunoglobulin protein family. The GPCRX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between an GPCRX ligand and an GPCRX protein on the surface of a cell, to thereby suppress GPCRX-mediated signal transduction *in vivo*. The GPCRX-immunoglobulin fusion proteins can be used to affect the bioavailability of an GPCRX cognate ligand. Inhibition of the GPCRX ligand/GPCRX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the GPCRX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-GPCRX antibodies in a subject, to purify GPCRX ligands, and in screening assays to identify molecules that inhibit the interaction of GPCRX with an GPCRX ligand.

An GPCRX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate,

alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (*see, e.g.,* Ausubel, *et al.* (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). An GPCRX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the GPCRX protein.

**GPCRX Agonists and Antagonists**

The invention also pertains to variants of the GPCRX proteins that function as either GPCRX agonists (*i.e.,* mimetics) or as GPCRX antagonists. Variants of the GPCRX protein can be generated by mutagenesis (*e.g.*, discrete point mutation or truncation of the GPCRX protein). An agonist of the GPCRX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the GPCRX protein. An antagonist of the GPCRX protein can inhibit one or more of the activities of the naturally occurring form of the GPCRX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the GPCRX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the GPCRX proteins.

Variants of the GPCRX proteins that function as either GPCRX agonists (*i.e.,* mimetics) or as GPCRX antagonists can be identified by screening combinatorial libraries of mutants (*e.g.*, truncation mutants) of the GPCRX proteins for GPCRX protein agonist or antagonist activity. In one embodiment, a variegated library of GPCRX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of GPCRX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential GPCRX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of GPCRX sequences therein. There are a variety of methods which can be used to produce

libraries of potential GPCRX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential GPCRX sequences. Methods for synthesizing degenerate oligonucleotides are well-known within the art. *See, e.g.,* Narang, 1983. *Tetrahedron* 39: 3; Itakura, *et al.*, 1984. *Annu. Rev. Biochem.* 53: 323; Itakura, *et al.*, 1984. *Science* 198: 1056; Ike, *et al.*, 1983. *Nucl. Acids Res.* 11: 477.

**Polypeptide Libraries**

In addition, libraries of fragments of the GPCRX protein coding sequences can be used to generate a variegated population of GPCRX fragments for screening and subsequent selection of variants of an GPCRX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an GPCRX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double-stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with $S_1$ nuclease, and ligating the resulting fragment library into an expression vector. By this method, expression libraries can be derived which encodes N-terminal and internal fragments of various sizes of the GPCRX proteins.

Various techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of GPCRX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify GPCRX variants. *See, e.g.,* Arkin and Yourvan, 1992. *Proc. Natl. Acad. Sci. USA* 89: 7811-7815; Delgrave, *et al.*, 1993. *Protein Engineering* 6:327-331.

**Anti-GPCRX Antibodies**

Also included in the invention are antibodies to GPCRX proteins, or fragments of GPCRX proteins. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, $F_{ab}$, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an $F_{ab}$ expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as $IgG_1$, $IgG_2$, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated GPCRX-related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of GPCRX-related protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human GPCRX-related protein sequence will indicate which regions of a GPCRX-related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without

Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow and Lane, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

**Polyclonal Antibodies**

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide

primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

**Monoclonal Antibodies**

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, *J. Immunol.*, 133:3001 (1984); Brodeur et al., MONOCLONAL ANTIBODY PRODUCTION TECHNIQUES AND APPLICATIONS, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, *Anal. Biochem.*, 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary

(CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, *Nature* 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

**Humanized Antibodies**

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeyen et al., *Science*, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)).

**Human Antibodies**

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (*Bio/Technology* 10, 779-783 (1992)); Lonberg et al. (*Nature* 368 856-859 (1994)); Morrison ( *Nature* 368, 812-13 (1994)); Fishwild et al,( *Nature Biotechnology* 14, 845-51 (1996)); Neuberger (*Nature Biotechnology* 14, 826 (1996)); and Lonberg and Huszar (*Intern. Rev. Immunol.* 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's

complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

**$F_{ab}$ Fragments and Single Chain Antibodies**

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of $F_{ab}$ expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective

identification of monoclonal $F_{ab}$ fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an $F_{ab}$ fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an $F_{ab}$ fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) $F_v$ fragments.

**Bispecific Antibodies**

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, *Nature*, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., *Methods in Enzymology*, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part

of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., *Science* 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., *J. Exp. Med.* 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., *J. Immunol.* 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody"

technology described by Hollinger et al., *Proc. Natl. Acad. Sci. USA* 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain ($V_H$) connected to a light-chain variable domain ($V_L$) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the $V_H$ and $V_L$ domains of one fragment are forced to pair with the complementary $V_L$ and $V_H$ domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., *J. Immunol.* 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., *J. Immunol.* 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

**Heteroconjugate Antibodies**

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

**Effector Function Engineering**

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

**Immunoconjugates**

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include $^{212}Bi$, $^{131}I$, $^{131}In$, $^{90}Y$, and $^{186}Re$.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-

2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of an GPCRX protein is facilitated by generation of hybridomas that bind to the fragment of an GPCRX protein possessing such a domain. Thus, antibodies that are specific for a desired domain within an GPCRX protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-GPCRX antibodies may be used in methods known within the art relating to the localization and/or quantitation of an GPCRX protein (*e.g.,* for use in measuring levels of the GPCRX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for GPCRX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds (hereinafter "Therapeutics").

An anti-GPCRX antibody (*e.g.*, monoclonal antibody) can be used to isolate an GPCRX polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-GPCRX antibody can facilitate the purification of natural GPCRX polypeptide from cells and of recombinantly-produced GPCRX polypeptide expressed in host cells. Moreover, an anti-GPCRX antibody can be used to detect GPCRX protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the GPCRX protein. Anti-GPCRX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to,

for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include $^{125}I$, $^{131}I$, $^{35}S$ or $^{3}H$.

**GPCRX Recombinant Expression Vectors and Host Cells**

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an GPCRX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively-linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the

basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably-linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell).

The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, GPCRX proteins, mutant forms of GPCRX proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of GPCRX proteins in prokaryotic or eukaryotic cells. For example, GPCRX proteins can be expressed in bacterial cells such as *Escherichia coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *Escherichia coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (*i*) to increase expression of recombinant protein; (*ii*) to increase the solubility of the recombinant protein; and (*iii*) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the

recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. *See, e.g.,* Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (*see, e.g.,* Wada, *et al.*, 1992. *Nucl. Acids Res.* 20: 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the GPCRX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *Saccharomyces cerivisae* include pYepSec1 (Baldari, *et al.*, 1987. *EMBO J.* 6: 229-234), pMFa (Kurjan and Herskowitz, 1982. *Cell* 30: 933-943), pJRY88 (Schultz *et al.*, 1987. *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

Alternatively, GPCRX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith, *et al.*, 1983. *Mol. Cell. Biol.* 3: 2156-2165) and the pVL series (Lucklow and Summers, 1989. *Virology* 170: 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987. *Nature* 329: 840) and pMT2PC (Kaufman, *et al.*, 1987. *EMBO J.* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are

derived from polyoma, adenovirus 2, cytomegalovirus, and simian virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see, *e.g.*, Chapters 16 and 17 of Sambrook, *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert, *et al.*, 1987. *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989. *EMBO J.* 8: 729-733) and immunoglobulins (Banerji, *et al.*, 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33: 741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund, *et al.*, 1985. *Science* 230: 912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss, 1990. *Science* 249: 374-379) and the α-fetoprotein promoter (Campes and Tilghman, 1989. *Genes Dev.* 3: 537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively-linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to GPCRX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes *see, e.g.,* Weintraub, *et al.*, "Antisense RNA as a molecular tool for genetic analysis," *Reviews-Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, GPCRX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding GPCRX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) GPCRX protein. Accordingly, the invention further provides methods for producing GPCRX protein using the host cells of the invention. In one

embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding GPCRX protein has been introduced) in a suitable medium such that GPCRX protein is produced. In another embodiment, the method further comprises isolating GPCRX protein from the medium or the host cell.

**Transgenic GPCRX Animals**

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which GPCRX protein-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous GPCRX sequences have been introduced into their genome or homologous recombinant animals in which endogenous GPCRX sequences have been altered. Such animals are useful for studying the function and/or activity of GPCRX protein and for identifying and/or evaluating modulators of GPCRX protein activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous GPCRX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing GPCRX-encoding nucleic acid into the male pronuclei of a fertilized oocyte (*e.g.*, by microinjection, retroviral infection) and allowing the oocyte to develop in a pseudopregnant female foster animal. The human GPCRX cDNA sequences of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a non-human homologue of the human GPCRX gene, such as a mouse GPCRX gene, can be isolated based on hybridization to the human GPCRX cDNA (described further *supra*) and used as a transgene. Intronic sequences and polyadenylation signals can

also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably-linked to the GPCRX transgene to direct expression of GPCRX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan, 1986. In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the GPCRX transgene in its genome and/or expression of GPCRX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene-encoding GPCRX protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of an GPCRX gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the GPCRX gene. The GPCRX gene can be a human gene (*e.g.*, the cDNA of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37), but more preferably, is a non-human homologue of a human GPCRX gene. For example, a mouse homologue of human GPCRX gene of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 can be used to construct a homologous recombination vector suitable for altering an endogenous GPCRX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous GPCRX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous GPCRX gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous GPCRX protein). In the homologous recombination vector, the altered portion of the GPCRX gene is flanked at its 5'- and 3'-termini by additional nucleic acid of the GPCRX gene to allow for homologous recombination to occur between the exogenous GPCRX gene carried by the vector and an endogenous GPCRX gene in an embryonic stem cell. The additional flanking GPCRX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of

flanking DNA (both at the 5'- and 3'-termini) are included in the vector. *See, e.g.,* Thomas, *et al.*, 1987. *Cell* 51: 503 for a description of homologous recombination vectors. The vector is ten introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced GPCRX gene has homologously-recombined with the endogenous GPCRX gene are selected. *See, e.g.*, Li, *et al.*, 1992. *Cell* 69: 915.

The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. *See, e.g.*, Bradley, 1987. In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously-recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously-recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley, 1991. *Curr. Opin. Biotechnol.* 2: 823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, *See, e.g.*, Lakso, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae. See,* O'Gorman, *et al.*, 1991. *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, *et al.*, 1997. *Nature* 385: 810-813. In brief, a cell (*e.g.*, a somatic cell) from the transgenic animal can be isolated and induced to exit the growth cycle and enter $G_0$ phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The

offspring borne of this female foster animal will be a clone of the animal from which the cell (*e.g.*, the somatic cell) is isolated.

**Pharmaceutical Compositions**

The GPCRX nucleic acid molecules, GPCRX proteins, and anti-GPCRX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (*i.e.,* topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid (EDTA); buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, an GPCRX protein or anti-GPCRX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and

swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.,* U.S. Patent No. 5,328,470) or by stereotactic injection (*see, e.g.*, Chen, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

**Screening and Detection Methods**

The isolated nucleic acid molecules of the invention can be used to express GPCRX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect GPCRX mRNA (*e.g.*, in a biological sample) or a genetic lesion in an GPCRX gene, and to modulate GPCRX activity, as described further, below. In addition, the GPCRX proteins can be used to screen drugs or compounds that modulate the GPCRX protein activity or expression as well as to treat disorders characterized by insufficient or excessive production of GPCRX protein or production of GPCRX protein forms that have decreased or aberrant activity compared to GPCRX wild-type protein (*e.g.*; diabetes (regulates insulin release); obesity (binds and transport lipids); metabolic disturbances associated with obesity, the metabolic syndrome X as well as anorexia and wasting disorders associated with chronic diseases and various cancers, and infectious disease(possesses anti-microbial activity) and the

various dyslipidemias. In addition, the anti-GPCRX antibodies of the invention can be used to detect and isolate GPCRX proteins and modulate GPCRX activity. In yet a further aspect, the invention can be used in methods to influence appetite, absorption of nutrients and the disposition of metabolic substrates in both a positive and negative fashion.

The invention further pertains to novel agents identified by the screening assays described herein and uses thereof for treatments as described, *supra*.

**Screening Assays**

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) that bind to GPCRX proteins or have a stimulatory or inhibitory effect on, *e.g.*, GPCRX protein expression or GPCRX protein activity. The invention also includes compounds identified in the screening assays described herein.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of the membrane-bound form of an GPCRX protein or polypeptide or biologically-active portion thereof. The test compounds of the invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds. *See, e.g.,* Lam, 1997. *Anticancer Drug Design* 12: 145.

A "small molecule" as used herein, is meant to refer to a composition that has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be, *e.g.*, nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic or inorganic molecules. Libraries of chemical and/or biological mixtures, such as fungal, bacterial, or algal extracts, are known in the art and can be screened with any of the assays of the invention.

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt, *et al.*, 1993. *Proc. Natl. Acad. Sci. U.S.A.* 90: 6909; Erb, *et al.*, 1994. *Proc. Natl. Acad. Sci. U.S.A.* 91: 11422; Zuckermann, *et al.*, 1994. *J. Med. Chem.* 37: 2678; Cho, *et al.*, 1993. *Science* 261: 1303; Carrell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33:

2059; Carell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2061; and Gallop, *et al.*, 1994. *J. Med. Chem.* 37: 1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten, 1992. *Biotechniques* 13: 412-421), or on beads (Lam, 1991. *Nature* 354: 82-84), on chips (Fodor, 1993. *Nature* 364: 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, U.S. Patent 5,233,409), plasmids (Cull, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 1865-1869) or on phage (Scott and Smith, 1990. *Science* 249: 386-390; Devlin, 1990. *Science* 249: 404-406; Cwirla, *et al.*, 1990. *Proc. Natl. Acad. Sci. U.S.A.* 87: 6378-6382; Felici, 1991. *J. Mol. Biol.* 222: 301-310; Ladner, U.S. Patent No. 5,233,409.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of GPCRX protein, or a biologically-active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to an GPCRX protein determined. The cell, for example, can of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the GPCRX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the GPCRX protein or biologically-active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with $^{125}I$, $^{35}S$, $^{14}C$, or $^{3}H$, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically-labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of GPCRX protein, or a biologically-active portion thereof, on the cell surface with a known compound which binds GPCRX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an GPCRX protein, wherein determining the ability of the test compound to interact with an GPCRX protein comprises determining the ability of the test compound to preferentially bind to GPCRX protein or a biologically-active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of GPCRX protein, or a biologically-active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the GPCRX protein or

biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of GPCRX or a biologically-active portion thereof can be accomplished, for example, by determining the ability of the GPCRX protein to bind to or interact with an GPCRX target molecule. As used herein, a "target molecule" is a molecule with which an GPCRX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses an GPCRX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. An GPCRX target molecule can be a non-GPCRX molecule or an GPCRX protein or polypeptide of the invention. In one embodiment, an GPCRX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.* a signal generated by binding of a compound to a membrane-bound GPCRX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with GPCRX.

Determining the ability of the GPCRX protein to bind to or interact with an GPCRX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the GPCRX protein to bind to or interact with an GPCRX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular $Ca^{2+}$, diacylglycerol, $IP_3$, etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising an GPCRX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the invention is a cell-free assay comprising contacting an GPCRX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to bind to the GPCRX protein or biologically-active portion thereof. Binding of the test compound to the GPCRX protein can be determined either directly or indirectly as described above. In one such embodiment, the assay comprises contacting the GPCRX protein or biologically-active portion thereof with a known compound which binds GPCRX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an GPCRX

protein, wherein determining the ability of the test compound to interact with an GPCRX protein comprises determining the ability of the test compound to preferentially bind to GPCRX or biologically-active portion thereof as compared to the known compound.

In still another embodiment, an assay is a cell-free assay comprising contacting GPCRX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.* stimulate or inhibit) the activity of the GPCRX protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of GPCRX can be accomplished, for example, by determining the ability of the GPCRX protein to bind to an GPCRX target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of GPCRX protein can be accomplished by determining the ability of the GPCRX protein further modulate an GPCRX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as described, *supra*.

In yet another embodiment, the cell-free assay comprises contacting the GPCRX protein or biologically-active portion thereof with a known compound which binds GPCRX protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an GPCRX protein, wherein determining the ability of the test compound to interact with an GPCRX protein comprises determining the ability of the GPCRX protein to preferentially bind to or modulate the activity of an GPCRX target molecule.

The cell-free assays of the invention are amenable to use of both the soluble form or the membrane-bound form of GPCRX protein. In the case of cell-free assays comprising the membrane-bound form of GPCRX protein, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of GPCRX protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl) dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the invention, it may be desirable to immobilize either GPCRX protein or its target molecule to facilitate separation of

complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to GPCRX protein, or interaction of GPCRX protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, GST-GPCRX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or GPCRX protein, and the mixture is incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described, *supra*. Alternatively, the complexes can be dissociated from the matrix, and the level of GPCRX protein binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the GPCRX protein or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated GPCRX protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well-known within the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with GPCRX protein or target molecules, but which do not interfere with binding of the GPCRX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or GPCRX protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the GPCRX protein or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the GPCRX protein or target molecule.

In another embodiment, modulators of GPCRX protein expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of GPCRX mRNA or protein in the cell is determined. The level of expression of GPCRX mRNA or

protein in the presence of the candidate compound is compared to the level of expression of GPCRX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of GPCRX mRNA or protein expression based upon this comparison. For example, when expression of GPCRX mRNA or protein is greater (*i.e.*, statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of GPCRX mRNA or protein expression. Alternatively, when expression of GPCRX mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of GPCRX mRNA or protein expression. The level of GPCRX mRNA or protein expression in the cells can be determined by methods described herein for detecting GPCRX mRNA or protein.

In yet another aspect of the invention, the GPCRX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos, *et al.*, 1993. *Cell* 72: 223-232; Madura, *et al.*, 1993. *J. Biol. Chem.* 268: 12046-12054; Bartel, *et al.*, 1993. *Biotechniques* 14: 920-924; Iwabuchi, *et al.*, 1993. *Oncogene* 8: 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact with GPCRX ("GPCRX-binding proteins" or "GPCRX-bp") and modulate GPCRX activity. Such GPCRX-binding proteins are also likely to be involved in the propagation of signals by the GPCRX proteins as, for example, upstream or downstream elements of the GPCRX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for GPCRX is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming an GPCRX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with GPCRX.

The invention further pertains to novel agents identified by the aforementioned screening assays and uses thereof for treatments as described herein.

**Detection Assays**

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. By way of example, and not of limitation, these sequences can be used to: (*i*) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (*ii*) identify an individual from a minute biological sample (tissue typing); and (*iii*) aid in forensic identification of a biological sample. Some of these applications are described in the subsections, below.

**Chromosome Mapping**

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the GPCRX sequences, SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, or fragments or derivatives thereof, can be used to map the location of the GPCRX genes, respectively, on a chromosome. The mapping of the GPCRX sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, GPCRX genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the GPCRX sequences. Computer analysis of the GPCRX, sequences can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the GPCRX sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but in which human cells can, the one human chromosome that contains the gene encoding the needed enzyme will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small

number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. *See, e.g.,* D'Eustachio, *et al.*, 1983. *Science* 220: 919-924. Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the GPCRX sequences to design oligonucleotide primers, sub-localization can be achieved with panels of fragments from specific chromosomes.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases, will suffice to get good results at a reasonable amount of time. For a review of this technique, *see,* Verma, *et al.*, HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, *e.g.,* in McKusick, MENDELIAN INHERITANCE IN MAN, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage

analysis (co-inheritance of physically adjacent genes), described in, *e.g.*, Egeland, *et al.*, 1987. *Nature*, 325: 783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the GPCRX gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

**Tissue Typing**

The GPCRX sequences of the invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Patent No. 5,272,057).

Furthermore, the sequences of the invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the GPCRX sequences described herein can be used to prepare two PCR primers from the 5'- and 3'-termini of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the invention can be used to obtain such identification sequences from individuals and from tissue. The GPCRX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

**Predictive Medicine**

The invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the invention relates to diagnostic assays for determining GPCRX protein and/or nucleic acid expression as well as GPCRX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant GPCRX expression or activity. The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with GPCRX protein, nucleic acid expression or activity. For example, mutations in an GPCRX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with GPCRX protein, nucleic acid expression, or biological activity.

Another aspect of the invention provides methods for determining GPCRX protein, nucleic acid expression or activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the

genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of GPCRX in clinical trials.

These and other agents are described in further detail in the following sections.

**Diagnostic Assays**

An exemplary method for detecting the presence or absence of GPCRX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting GPCRX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes GPCRX protein such that the presence of GPCRX is detected in the biological sample. An agent for detecting GPCRX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to GPCRX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length GPCRX nucleic acid, such as the nucleic acid of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to GPCRX mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting GPCRX protein is an antibody capable of binding to GPCRX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently-labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently-labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect GPCRX mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of GPCRX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of GPCRX protein include enzyme linked

immunosorbent assays (ELISAs), Western blots, immunoprecipitations, and immunofluorescence. *In vitro* techniques for detection of GPCRX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of GPCRX protein include introducing into a subject a labeled anti-GPCRX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting GPCRX protein, mRNA, or genomic DNA, such that the presence of GPCRX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of GPCRX protein, mRNA or genomic DNA in the control sample with the presence of GPCRX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of GPCRX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting GPCRX protein or mRNA in a biological sample; means for determining the amount of GPCRX in the sample; and means for comparing the amount of GPCRX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect GPCRX protein or nucleic acid.

**Prognostic Assays**

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant GPCRX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with GPCRX protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the invention provides a method for identifying a disease or disorder associated with aberrant GPCRX expression or activity in which a test sample is obtained from a subject and GPCRX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of GPCRX protein or nucleic acid is diagnostic for a

subject having or at risk of developing a disease or disorder associated with aberrant GPCRX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant GPCRX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder. Thus, the invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant GPCRX expression or activity in which a test sample is obtained and GPCRX protein or nucleic acid is detected (*e.g.*, wherein the presence of GPCRX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant GPCRX expression or activity).

The methods of the invention can also be used to detect genetic lesions in an GPCRX gene, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding an GPCRX-protein, or the misexpression of the GPCRX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of: (*i*) a deletion of one or more nucleotides from an GPCRX gene; (*ii*) an addition of one or more nucleotides to an GPCRX gene; (*iii*) a substitution of one or more nucleotides of an GPCRX gene, (*iv*) a chromosomal rearrangement of an GPCRX gene; (*v*) an alteration in the level of a messenger RNA transcript of an GPCRX gene, (*vi*) aberrant modification of an GPCRX gene, such as of the methylation pattern of the genomic DNA, (*vii*) the presence of a non-wild-type splicing pattern of a messenger RNA transcript of an GPCRX gene, (*viii*) a non-wild-type level of an GPCRX protein, (*ix*) allelic loss of an GPCRX gene, and (*x*) inappropriate post-translational modification of an GPCRX protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in an GPCRX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional

means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (*see, e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (*see, e.g.*, Landegran, *et al.*, 1988. *Science* 241: 1077-1080; and Nakazawa, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 360-364), the latter of which can be particularly useful for detecting point mutations in the GPCRX-gene (*see,* Abravaya, *et al.*, 1995. *Nucl. Acids Res.* 23: 675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to an GPCRX gene under conditions such that hybridization and amplification of the GPCRX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (*see,* Guatelli, *et al.*, 1990. *Proc. Natl. Acad. Sci. USA* 87: 1874-1878), transcriptional amplification system (*see,* Kwoh, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 1173-1177); Qβ Replicase (*see,* Lizardi, *et al*, 1988. *BioTechnology* 6: 1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in an GPCRX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (*see, e.g.,* U.S. Patent No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in GPCRX can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high-density arrays containing

hundreds or thousands of oligonucleotides probes. *See, e.g.,* Cronin, *et al.*, 1996. *Human Mutation* 7: 244-255; Kozal, *et al.*, 1996. *Nat. Med.* 2: 753-759. For example, genetic mutations in GPCRX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, *et al.*, *supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the GPCRX gene and detect mutations by comparing the sequence of the sample GPCRX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert, 1977. *Proc. Natl. Acad. Sci. USA* 74: 560 or Sanger, 1977. *Proc. Natl. Acad. Sci. USA* 74: 5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (*see, e.g.,* Naeve, *et al.*, 1995. *Biotechniques* 19: 448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen, *et al.*, 1996. *Adv. Chromatography* 36: 127-162; and Griffin, *et al.*, 1993. *Appl. Biochem. Biotechnol.* 38: 147-159).

Other methods for detecting mutations in the GPCRX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes. *See, e.g.,* Myers, *et al.*, 1985. *Science* 230: 1242. In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type GPCRX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with $S_1$ nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched

regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. *See, e.g.,* Cotton, *et al.,* 1988. *Proc. Natl. Acad. Sci. USA* 85: 4397; Saleeba, *et al.,* 1992. *Methods Enzymol.* 217: 286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in GPCRX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches. *See, e.g.,* Hsu, *et al.*, 1994. *Carcinogenesis* 15: 1657-1662. According to an exemplary embodiment, a probe based on an GPCRX sequence, *e.g.*, a wild-type GPCRX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. *See, e.g.,* U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in GPCRX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids. *See, e.g.,* Orita, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA*: 86: 2766; Cotton, 1993. *Mutat. Res.* 285: 125-144; Hayashi, 1992. *Genet. Anal. Tech. Appl.* 9: 73-79.

Single-stranded DNA fragments of sample and control GPCRX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. *See, e.g.,* Keen, *et al.*, 1991. *Trends Genet.* 7: 5.

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). *See, e.g.,* Myers, *et al.,* 1985. *Nature* 313: 495. When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich

DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. *See, e.g.,* Rosenbaum and Reissner, 1987. *Biophys. Chem.* 265: 12753.

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. *See, e.g.,* Saiki, *et al.*, 1986. *Nature* 324: 163; Saiki, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization; *see, e.g.,* Gibbs, *et al.*, 1989. *Nucl. Acids Res.* 17: 2437-2448) or at the extreme 3'-terminus of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (*see, e.g.,* Prossner, 1993. *Tibtech.* 11: 238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. *See, e.g.,* Gasparini, *et al.,* 1992. *Mol. Cell Probes* 6: 1. It is anticipated that in certain embodiments amplification may also be performed using *Taq* ligase for amplification. *See, e.g.,* Barany, 1991. *Proc. Natl. Acad. Sci. USA* 88: 189. In such cases, ligation will occur only if there is a perfect match at the 3'-terminus of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving an GPCRX gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which GPCRX is expressed may be utilized in the prognostic assays described herein. However, any

biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

**Pharmacogenomics**

Agents, or modulators that have a stimulatory or inhibitory effect on GPCRX activity (*e.g.*, GPCRX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers.) In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of GPCRX protein, expression of GPCRX nucleic acid, or mutation content of GPCRX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum, 1996. *Clin. Exp. Pharmacol. Physiol.*, 23: 983-985; Linder, 1997. *Clin. Chem.*, 43: 254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of GPCRX protein, expression of GPCRX nucleic acid, or mutation content of GPCRX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with an GPCRX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

**Monitoring of Effects During Clinical Trials**

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of GPCRX (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase GPCRX gene expression, protein levels, or upregulate GPCRX activity, can be monitored in clinical trails of subjects exhibiting decreased GPCRX gene expression, protein levels, or downregulated GPCRX activity. Alternatively, the effectiveness of an agent

determined by a screening assay to decrease GPCRX gene expression, protein levels, or downregulate GPCRX activity, can be monitored in clinical trails of subjects exhibiting increased GPCRX gene expression, protein levels, or upregulated GPCRX activity. In such clinical trials, the expression or activity of GPCRX and, preferably, other genes that have been implicated in, for example, a cellular proliferation or immune disorder can be used as a "read out" or markers of the immune responsiveness of a particular cell.

By way of example, and not of limitation, genes, including GPCRX, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) that modulates GPCRX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of GPCRX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of GPCRX or other genes. In this manner, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (*i*) obtaining a pre-administration sample from a subject prior to administration of the agent; (*ii*) detecting the level of expression of an GPCRX protein, mRNA, or genomic DNA in the preadministration sample; (*iii*) obtaining one or more post-administration samples from the subject; (*iv*) detecting the level of expression or activity of the GPCRX protein, mRNA, or genomic DNA in the post-administration samples; (*v*) comparing the level of expression or activity of the GPCRX protein, mRNA, or genomic DNA in the pre-administration sample with the GPCRX protein, mRNA, or genomic DNA in the post administration sample or samples; and (*vi*) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of GPCRX to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased

administration of the agent may be desirable to decrease expression or activity of GPCRX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

**Methods of Treatment**

The invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant GPCRX expression or activity. The disorders include cardiomyopathy, atherosclerosis, hypertension, congenital heart defects, aortic stenosis, atrial septal defect (ASD), atrioventricular (A-V) canal defect, ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma, obesity, transplantation, adrenoleukodystrophy, congenital adrenal hyperplasia, prostate cancer, neoplasm; adenocarcinoma, lymphoma, uterus cancer, fertility, hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies, graft versus host disease, AIDS, bronchial asthma, Crohn's disease; multiple sclerosis, treatment of Albright Hereditary Ostoeodystrophy, and other diseases, disorders and conditions of the like.

These methods of treatment will be discussed more fully, below.

**Disease and Disorders**

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to: (*i*) an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; (*ii*) antibodies to an aforementioned peptide; (*iii*) nucleic acids encoding an aforementioned peptide; (*iv*) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.,* due to a heterologous insertion within the coding sequences of coding sequences to an aforementioned peptide) that are utilized to "knockout" endoggenous function of an aforementioned peptide by homologous recombination (*see, e.g.,* Capecchi, 1989. *Science* 244: 1288-1292); or (*v*) modulators ( *i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between an aforementioned peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with

Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.,* from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of an aforementioned peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.,* by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.,* Northern assays, dot blots, *in situ* hybridization, and the like).

**Prophylactic Methods**

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant GPCRX expression or activity, by administering to the subject an agent that modulates GPCRX expression or at least one GPCRX activity. Subjects at risk for a disease that is caused or contributed to by aberrant GPCRX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the GPCRX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending upon the type of GPCRX aberrancy, for example, an GPCRX agonist or GPCRX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein. The prophylactic methods of the invention are further discussed in the following subsections.

**Therapeutic Methods**

Another aspect of the invention pertains to methods of modulating GPCRX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of GPCRX protein activity associated with the cell. An agent that modulates GPCRX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of an GPCRX protein, a peptide, an GPCRX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more GPCRX protein activity. Examples of

such stimulatory agents include active GPCRX protein and a nucleic acid molecule encoding GPCRX that has been introduced into the cell. In another embodiment, the agent inhibits one or more GPCRX protein activity. Examples of such inhibitory agents include antisense GPCRX nucleic acid molecules and anti-GPCRX antibodies. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of an GPCRX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, up-regulates or down-regulates) GPCRX expression or activity. In another embodiment, the method involves administering an GPCRX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant GPCRX expression or activity.

Stimulation of GPCRX activity is desirable in situations in which GPCRX is abnormally downregulated and/or in which increased GPCRX activity is likely to have a beneficial effect. One example of such a situation is where a subject has a disorder characterized by aberrant cell proliferation and/or differentiation (*e.g.*, cancer or immune associated disorders). Another example of such a situation is where the subject has a gestational disease (*e.g.*, preclampsia).

**Determination of the Biological Effect of the Therapeutic**

In various embodiments of the invention, suitable *in vitro* or *in vivo* assays are performed to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

**Prophylactic and Therapeutic Uses of the Compositions of the Invention**

The GPCRX nucleic acids and proteins of the invention are useful in potential prophylactic and therapeutic applications implicated in a variety of disorders including, but not limited to: metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers.

As an example, a cDNA encoding the GPCRX protein of the invention may be useful in gene therapy, and the protein may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the invention will have efficacy for treatment of patients suffering from: metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, hematopoietic disorders, and the various dyslipidemias.

Both the novel nucleic acid encoding the GPCRX protein, and the GPCRX protein of the invention, or fragments thereof, may also be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. A further use could be as an anti-bacterial molecule (*i.e.,* some peptides have been found to possess anti-bacterial properties). These materials are further useful in the generation of antibodies which immunospecifically-bind to the novel substances of the invention for use in therapeutic or diagnostic methods.

**EXAMPLES**

**Example 1. Identification of GPCRX clones**

All novel GPCRX target sequences identified in the present invention were subjected to the exon linking process to confirm the sequence. PCR primers were designed by starting at the most upstream sequence available, for the forward primer, and at the most downstream sequence available for the reverse primer. Table 11A shows the sequences of the PCR primers used for obtaining different clones. In each case, the sequence was examined, walking inward from the respective termini toward the coding sequence, until a suitable sequence that is either unique or highly selective was encountered, or, in the case of the reverse primer, until the stop

codon was reached. Such primers were designed based on in silico predictions for the full length cDNA, part (one or more exons) of the DNA or protein sequence of the target sequence, or by translated homology of the predicted exons to closely related human sequences from other species. These primers were then employed in PCR amplification based on the following pool of human cDNAs: adrenal gland, bone marrow, brain - amygdala, brain - cerebellum, brain - hippocampus, brain - substantia nigra, brain - thalamus, brain -whole, fetal brain, fetal kidney, fetal liver, fetal lung, heart, kidney, lymphoma - Raji, mammary gland, pancreas, pituitary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen, stomach, testis, thyroid, trachea, uterus. Usually the resulting amplicons were gel purified, cloned and sequenced to high redundancy. The PCR product derived from exon linking was cloned into the pCR2.1 vector from Invitrogen. The resulting bacterial clone has an insert covering the entire open reading frame cloned into the pCR2.1 vector. Table 11B shows a list of these bacterial clones. The resulting sequences from all clones were assembled with themselves, with other fragments in CuraGen Corporation's database and with public ESTs. Fragments and ESTs were included as components for an assembly when the extent of their identity with another component of the assembly was at least 95% over 50 bp. In addition, sequence traces were evaluated manually and edited for corrections if appropriate. These procedures provide the sequence reported herein.

**Table 11A. PCR Primers for Exon Linking**

| GPCR Clone | Primer 1 | SEQ ID NO | Primer 2 |
|---|---|---|---|
| GPCR2b | 5′-GGAGA AATTT TGGAC AGATT GTATC AGA-3′ | 60 | 5′-AATCA TTTCA GAATC ACATA CCCAT TA |
| GPCR4a | 5′-CTGAA TTTTT CCTCC GAGGA ATATC -3′ | 62 | 5′-TCTAA GAGGA AACAA TACTC CTGTG AC- |
| GPCR5a | 5'-GTTTG TTTCT GCCTT TTTCA ATGTC C-3' | 64 | 5'-AAGAG TAAAT GTCAT TGAGA TAAGA CT |
| GPCR6c | 5′-TATGA GCCCT GAGAA CCAGA GCAG - 3′ | 66 | 5′-TCAAA TACAT GAGAT AAACC AAAAT GA |
| GPCR8a | 5'-TTTTT GTCTC TCACT GATAT TTGC-3' | 68 | 5'-TGTTC ACCAG AAACC AGTAG AAAAC AC- |
| GPCR8b | 5′-CTTCT TGTTT ATAAC TGAGC CCAAG TC-3′ | 70 | 5′-AACCA GTAGA AAACA CGTCC AAGTT G- |

**Table 11B. Physical Clones for PCR products**

| GPCR Clone | Bacterial Clone |
|---|---|
| GPCR2b | Gmba64p14_B.698299.P3 |
| GPCR4b | 112184::30262209_EXT.698299.G6 |
| GPCR4c | 112861::Gmba64p14_D.698299.F4 |
| GPCR6c | 119741::ba460n11_da1.698324.G14 |

| | |
|---|---|
| GPCR7a | 112867::GMba64p14_G.698299.J2 |
| GPCR8b | 112869::Gmba64p14_H.698372.F1 |
| GPCR8c | 114181::AL162254.698252.P10 |

**Example 2. Quantitative expression analysis of clones in various cells and tissues**

The quantitative expression of various clones was assessed using microtiter plates containing RNA samples from a variety of normal and pathology-derived cells, cell lines and tissues using real time quantitative PCR (RTQ PCR; TAQMAN®). RTQ PCR was performed on a Perkin-Elmer Biosystems ABI PRISM® 7700 Sequence Detection System. Various collections of samples are assembled on the plates, and referred to as Panel 1 (containing cells and cell lines from normal and cancer sources), Panel 2 (containing samples derived from tissues, in particular from surgical samples, from normal and cancer sources), Panel 3 (containing samples derived from a wide variety of cancer sources), Panel 4 (containing cells and cell lines from normal cells and cells related to inflammatory conditions) and Panel CNSD.01 (containing samples from normal and diseased brains).

First, the RNA samples were normalized to constitutively expressed genes such as β-actin and GAPDH. RNA (~50 ng total or ~1 ng polyA+) was converted to cDNA using the TAQMAN® Reverse Transcription Reagents Kit (PE Biosystems, Foster City, CA; Catalog No. N808-0234) and random hexamers according to the manufacturer's protocol. Reactions were performed in 20 ul and incubated for 30 min. at 48$^0$C. cDNA (5 ul) was then transferred to a separate plate for the TAQMAN® reaction using β-actin and GAPDH TAQMAN® Assay Reagents (PE Biosystems; Catalog Nos. 4310881E and 4310884E, respectively) and TAQMAN® universal PCR Master Mix (PE Biosystems; Catalog No. 4304447) according to the manufacturer's protocol. Reactions were performed in 25 ul using the following parameters: 2 min. at 50$^0$C; 10 min. at 95$^0$C; 15 sec. at 95$^0$C/1 min. at 60$^0$C (40 cycles). Results were recorded as CT values (cycle at which a given sample crosses a threshold level of fluorescence) using a log scale, with the difference in RNA concentration between a given sample and the sample with the lowest CT value being represented as 2 to the power of delta CT. The percent relative expression is then obtained by taking the reciprocal of this RNA difference and multiplying by 100. The average CT values obtained for ß-actin and GAPDH were used to normalize RNA samples. The RNA sample generating the highest CT value required no further diluting, while all other samples were diluted relative to this sample according to their β-actin /GAPDH average CT values.

Normalized RNA (5 ul) was converted to cDNA and analyzed via TAQMAN® using One Step RT-PCR Master Mix Reagents (PE Biosystems; Catalog No. 4309169) and gene-specific primers according to the manufacturer's instructions. Probes and primers were designed for each assay according to Perkin Elmer Biosystem's *Primer Express* Software package (version I for Apple Computer's Macintosh Power PC) or a similar algorithm using the target sequence as input. Default settings were used for reaction conditions and the following parameters were set before selecting primers: primer concentration = 250 nM, primer melting temperature ($T_m$) range = 58°-60° C, primer optimal Tm = 59° C, maximum primer difference = 2° C, probe does not have 5' G, probe $T_m$ must be 10° C greater than primer $T_m$, amplicon size 75 bp to 100 bp. The probes and primers selected (see below) were synthesized by Synthegen (Houston, TX, USA). Probes were double purified by HPLC to remove uncoupled dye and evaluated by mass spectroscopy to verify coupling of reporter and quencher dyes to the 5' and 3' ends of the probe, respectively. Their final concentrations were: forward and reverse primers, 900 nM each, and probe, 200nM.

PCR conditions: Normalized RNA from each tissue and each cell line was spotted in each well of a 96 well PCR plate (Perkin Elmer Biosystems). PCR cocktails including two probes (a probe specific for the target clone and another gene-specific probe multiplexed with the target probe) were set up using 1X TaqMan™ PCR Master Mix for the PE Biosystems 7700, with 5 mM MgCl2, dNTPs (dA, G, C, U at 1:1:1:2 ratios), 0.25 U/ml AmpliTaq Gold™ (PE Biosystems), and 0.4 U/μl RNase inhibitor, and 0.25 U/μl reverse transcriptase. Reverse transcription was performed at 48° C for 30 minutes followed by amplification/PCR cycles as follows: 95° C 10 min, then 40 cycles of 95° C for 15 seconds, 60° C for 1 minute.

In the results for Panel 1, the following abbreviations are used:

ca. = carcinoma,
* = established from metastasis,
met = metastasis,
s cell var= small cell variant,
non-s = non-sm =non-small,
squam = squamous,
pl. eff = pl effusion = pleural effusion,
glio = glioma,
astro = astrocytoma, and
neuro = neuroblastoma.

**Panel 2**

The plates for Panel 2 generally include 2 control wells and 94 test samples composed of RNA or cDNA isolated from human tissue procured by surgeons working in close cooperation with the National Cancer Institute's Cooperative Human Tissue Network (CHTN) or the National Disease Research Initiative (NDRI). The tissues are derived from human malignancies and in cases where indicated many malignant tissues have "matched margins" obtained from noncancerous tissue just adjacent to the tumor. These are termed normal adjacent tissues and are denoted "NAT" in the results below. The tumor tissue and the "matched margins" are evaluated by two independent pathologists (the surgical pathologists and again by a pathologists at NDRI or CHTN). This analysis provides a gross histopathological assessment of tumor differentiation grade. Moreover, most samples include the original surgical pathology report that provides information regarding the clinical stage of the patient. These matched margins are taken from the tissue surrounding (i.e. immediately proximal) to the zone of surgery (designated "NAT", for normal adjacent tissue, in Table RR). In addition, RNA and cDNA samples were obtained from various human tissues derived from autopsies performed on elderly people or sudden death victims (accidents, etc.). These tissues were ascertained to be free of disease and were purchased from various commercial sources such as Clontech (Palo Alto, CA), Research Genetics, and Invitrogen.

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

**Panel 3D**

The plates of Panel 3D are comprised of 94 cDNA samples and two control samples. Specifically, 92 of these samples are derived from cultured human cancer cell lines, 2 samples of human primary cerebellar tissue and 2 controls. The human cell lines are generally obtained from ATCC (American Type Culture Collection), NCI or the German tumor cell bank and fall into the following tissue groups: Squamous cell carcinoma of the tongue, breast cancer, prostate cancer, melanoma, epidermoid carcinoma, sarcomas, bladder carcinomas, pancreatic cancers, kidney cancers, leukemias/lymphomas, ovarian/uterine/cervical, gastric,

colon, lung and CNS cancer cell lines. In addition, there are two independent samples of cerebellum. These cells are all cultured under standard recommended conditions and RNA extracted using the standard procedures. The cell lines in panel 3D and 1.3D are of the most common cell lines used in the scientific literature.

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

**Panel 4**

Panel 4 includes samples on a 96 well plate (2 control wells, 94 test samples) composed of RNA (Panel 4r) or cDNA (Panel 4d) isolated from various human cell lines or tissues related to inflammatory conditions. Total RNA from control normal tissues such as colon and lung (Stratagene, La Jolla, CA) and thymus and kidney (Clontech) were employed. Total RNA from liver tissue from cirrhosis patients and kidney from lupus patients was obtained from BioChain (Biochain Institute, Inc., Hayward, CA). Intestinal tissue for RNA preparation from patients diagnosed as having Crohn's disease and ulcerative colitis was obtained from the National Disease Research Interchange (NDRI) (Philadelphia, PA).

Astrocytes, lung fibroblasts, dermal fibroblasts, coronary artery smooth muscle cells, small airway epithelium, bronchial epithelium, microvascular dermal endothelial cells, microvascular lung endothelial cells, human pulmonary aortic endothelial cells, human umbilical vein endothelial cells were all purchased from Clonetics (Walkersville, MD) and grown in the media supplied for these cell types by Clonetics. These primary cell types were activated with various cytokines or combinations of cytokines for 6 and/or 12-14 hours, as indicated. The following cytokines were used; IL-1 beta at approximately 1-5 ng/ml, TNF alpha at approximately 5-10 ng/ml, IFN gamma at approximately 20-50 ng/ml, IL-4 at approximately 5-10 ng/ml, IL-9 at approximately 5-10 ng/ml, IL-13 at approximately 5-10 ng/ml. Endothelial cells were sometimes starved for various times by culture in the basal media from Clonetics with 0.1% serum.

Mononuclear cells were prepared from blood of employees at CuraGen Corporation, using Ficoll. LAK cells were prepared from these cells by culture in DMEM 5% FCS

(Hyclone), 100 μM non essential amino acids (Gibco/Life Technologies, Rockville, MD), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco) and Interleukin 2 for 4-6 days. Cells were then either activated with 10-20 ng/ml PMA and 1-2 μg/ml ionomycin, IL-12 at 5-10 ng/ml, IFN gamma at 20-50 ng/ml and IL-18 at 5-10 ng/ml for 6 hours. In some cases, mononuclear cells were cultured for 4-5 days in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco) with PHA (phytohemagglutinin) or PWM (pokeweed mitogen) at approximately 5 μg/ml. Samples were taken at 24, 48 and 72 hours for RNA preparation. MLR (mixed lymphocyte reaction) samples were obtained by taking blood from two donors, isolating the mononuclear cells using Ficoll and mixing the isolated mononuclear cells 1:1 at a final concentration of approximately $2x10^6$ cells/ml in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol (5.5 x $10^{-5}$ M) (Gibco), and 10 mM Hepes (Gibco). The MLR was cultured and samples taken at various time points ranging from 1- 7 days for RNA preparation.

Monocytes were isolated from mononuclear cells using CD14 Miltenyi Beads, +**ve** VS selection columns and a Vario Magnet according to the manufacturer's instructions. Monocytes were differentiated into dendritic cells by culture in DMEM 5% fetal calf serum (FCS) (Hyclone, Logan, UT), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco), 50 ng/ml GMCSF and 5 ng/ml IL-4 for 5-7 days. Macrophages were prepared by culture of monocytes for 5-7 days in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), 10 mM Hepes (Gibco) and 10% AB Human Serum or MCSF at approximately 50 ng/ml. Monocytes, macrophages and dendritic cells were stimulated for 6 and 12-14 hours with lipopolysaccharide (LPS) at 100 ng/ml. Dendritic cells were also stimulated with anti-CD40 monoclonal antibody (Pharmingen) at 10 μg/ml for 6 and 12-14 hours.

CD4 lymphocytes, CD8 lymphocytes and NK cells were also isolated from mononuclear cells using CD4, CD8 and CD56 Miltenyi beads, positive VS selection columns and a Vario Magnet according to the manufacturer's instructions. CD45RA and CD45RO CD4 lymphocytes were isolated by depleting mononuclear cells of CD8, CD56, CD14 and CD19 cells using CD8, CD56, CD14 and CD19 Miltenyi beads and positive selection. Then

CD45RO beads were used to isolate the CD45RO CD4 lymphocytes with the remaining cells being CD45RA CD4 lymphocytes. CD45RA CD4, CD45RO CD4 and CD8 lymphocytes were placed in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco) and plated at $10^6$ cells/ml onto Falcon 6 well tissue culture plates that had been coated overnight with 0.5 μg/ml anti-CD28 (Pharmingen) and 3 ug/ml anti-CD3 (OKT3, ATCC) in PBS. After 6 and 24 hours, the cells were harvested for RNA preparation. To prepare chronically activated CD8 lymphocytes, we activated the isolated CD8 lymphocytes for 4 days on anti-CD28 and anti-CD3 coated plates and then harvested the cells and expanded them in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco) and IL-2. The expanded CD8 cells were then activated again with plate bound anti-CD3 and anti-CD28 for 4 days and expanded as before. RNA was isolated 6 and 24 hours after the second activation and after 4 days of the second expansion culture. The isolated NK cells were cultured in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco) and IL-2 for 4-6 days before RNA was prepared.

To obtain B cells, tonsils were procured from NDRI. The tonsil was cut up with sterile dissecting scissors and then passed through a sieve. Tonsil cells were then spun down and resupended at $10^6$ cells/ml in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco). To activate the cells, we used PWM at 5 μg/ml or anti-CD40 (Pharmingen) at approximately 10 μg/ml and IL-4 at 5-10 ng/ml. Cells were harvested for RNA preparation at 24,48 and 72 hours.

To prepare the primary and secondary Th1/Th2 and Tr1 cells, six-well Falcon plates were coated overnight with 10 μg/ml anti-CD28 (Pharmingen) and 2 μg/ml OKT3 (ATCC), and then washed twice with PBS. Umbilical cord blood CD4 lymphocytes (Poietic Systems, German Town, MD) were cultured at $10^5$-$10^6$ cells/ml in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), 10 mM Hepes (Gibco) and IL-2 (4 ng/ml). IL-12 (5 ng/ml) and anti-IL4 (1 μg/ml) were used to direct to Th1, while IL-4 (5 ng/ml) and anti-IFN gamma (1 μg/ml) were used to direct to Th2 and IL-10 at 5 ng/ml was used to direct to Tr1. After 4-5 days, the

activated Th1, Th2 and Tr1 lymphocytes were washed once in DMEM and expanded for 4-7 days in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), 10 mM Hepes (Gibco) and IL-2 (1 ng/ml). Following this, the activated Th1, Th2 and Tr1 lymphocytes were re-stimulated for 5 days with anti-CD28/OKT3 and cytokines as described above, but with the addition of anti-CD95L (1 μg/ml) to prevent apoptosis. After 4-5 days, the Th1, Th2 and Tr1 lymphocytes were washed and then expanded again with IL-2 for 4-7 days. Activated Th1 and Th2 lymphocytes were maintained in this way for a maximum of three cycles. RNA was prepared from primary and secondary Th1, Th2 and Tr1 after 6 and 24 hours following the second and third activations with plate bound anti-CD3 and anti-CD28 mAbs and 4 days into the second and third expansion cultures in Interleukin 2.

The following leukocyte cells lines were obtained from the ATCC: Ramos, EOL-1, KU-812. EOL cells were further differentiated by culture in 0.1 mM dbcAMP at 5 x$10^5$ cells/ml for 8 days, changing the media every 3 days and adjusting the cell concentration to 5 x$10^5$ cells/ml. For the culture of these cells, we used DMEM or RPMI (as recommended by the ATCC), with the addition of 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), 10 mM Hepes (Gibco). RNA was either prepared from resting cells or cells activated with PMA at 10 ng/ml and ionomycin at 1 μg/ml for 6 and 14 hours. Keratinocyte line CCD106 and an airway epithelial tumor line NCI-H292 were also obtained from the ATCC. Both were cultured in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x $10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco). CCD1106 cells were activated for 6 and 14 hours with approximately 5 ng/ml TNF alpha and 1 ng/ml IL-1 beta, while NCI-H292 cells were activated for 6 and 14 hours with the following cytokines: 5 ng/ml IL-4, 5 ng/ml IL-9, 5 ng/ml IL-13 and 25 ng/ml IFN gamma.

For these cell lines and blood cells, RNA was prepared by lysing approximately $10^7$ cells/ml using Trizol (Gibco BRL). Briefly, 1/10 volume of bromochloropropane (Molecular Research Corporation) was added to the RNA sample, vortexed and after 10 minutes at room temperature, the tubes were spun at 14,000 rpm in a Sorvall SS34 rotor. The aqueous phase was removed and placed in a 15 ml Falcon Tube. An equal volume of isopropanol was added and left at –20 degrees C overnight. The precipitated RNA was spun down at 9,000 rpm for 15 min in a Sorvall SS34 rotor and washed in 70% ethanol. The pellet was redissolved in 300 μl of RNAse-free water and 35 μl buffer (Promega) 5 μl DTT, 7 μl RNAsin and 8 μl DNAse

were added. The tube was incubated at 37 degrees C for 30 minutes to remove contaminating genomic DNA, extracted once with phenol chloroform and re-precipitated with 1/10 volume of 3 M sodium acetate and 2 volumes of 100% ethanol. The RNA was spun down and placed in RNAse free water. RNA was stored at –80 degrees C.

**Panel CNSD.01**

The plates for Panel CNSD.01 include two control wells and 94 test samples comprised of cDNA isolated from postmortem human brain tissue obtained from the Harvard Brain Tissue Resource Center. Brains are removed from calvaria of donors between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in liquid nitrogen vapor. All brains are sectioned and examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains two brains from each of the following diagnoses: Alzheimer's disease, Parkinson's disease, Huntington's disease, Progressive Supernuclear Palsy, Depression, and "Normal controls". Within each of these brains, the following regions are represented: cingulate gyrus, temporal pole, globus palladus, substantia nigra, Brodman Area 4 (primary motor strip), Brodman Area 7 (parietal cortex), Brodman Area 9 (prefrontal cortex), and Brodman area 17 (occipital cortex). Not all brain regions are represented in all cases; e.g., Huntington's disease is characterized in part by neurodegeneration in the globus palladus, thus this region is impossible to obtain from confirmed Huntington's cases. Likewise Parkinson's disease is characterized by degeneration of the substantia nigra making this region more difficult to obtain. Normal control brains were examined for neuropathology and found to be free of any pathology consistent with neurodegeneration.

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

In the labels employed to identify tissues in the CNS panel, the following abbreviations are used:

PSP = Progressive supranuclear palsy

Sub Nigra = Substantia nigra

Glob Palladus= Globus palladus

Temp Pole = Temporal pole

Cing Gyr = Cingulate gyrus

BA 4 = Brodman Area 4

**Example 2A. GPCR1 (also known as ba64p14-A or CG56853-01):**

Expression of gene GPCR1 was assessed using the primer-probe set Ag1257 described in Table 12A. Results of the RTQ-PCR runs are shown in Table 12B,12C, and 12D.

**Table 12A. Probe Name Ag1257**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-AGCCATCCCTCATTTCTATTGT-3' | 59 | 22 | 545 | 72 |
| Probe | FAM-5'-TCCTAGTGCTCTCCTGAAGCTTGCCT-3'-TAMRA | 68.7 | 26 | 569 | 73 |
| Reverse | 5'-TCAGCTCGTTTACATGGGTATC-3' | 59.1 | 22 | 600 | 74 |

**Table 12B. Panel 1.2**

| Tissue Name | Relative Expression(%) | |
|---|---|---|
| | 1.2tm1423f_ag1257 | 1.2tm2143f_ag1257 |
| Endothelial cells | 2.8 | 0.9 |
| Fetal heart | 0.0 | 0.0 |
| Pancreas | 0.6 | 0.0 |
| Pancreatic ca. CAPAN 2 | 3.6 | 0.5 |
| Adrenal Gland (new lot*) | 2.1 | 0.0 |
| Thyroid | 1.9 | 0.0 |
| Salivary gland | 5.8 | 1.1 |
| Pituitary gland | 3.0 | 0.0 |
| Brain (fetal) | 4.6 | 0.0 |
| Brain (whole) | 9.2 | 0.5 |
| Brain (amygdala) | 4.7 | 0.4 |
| Brain (cerebellum) | 5.3 | 0.0 |
| Brain (hippocampus) | 19.3 | 1.7 |
| Brain (thalamus) | 0.9 | 0.8 |
| Cerebral Cortex | 3.4 | 1.4 |
| Spinal cord | 1.5 | 0.0 |
| CNS ca. (glio/astro) U87-MG | 25.3 | 6.5 |
| CNS ca. (glio/astro) U-118-MG | 0.5 | 0.0 |
| CNS ca. (astro) SW1783 | 0.9 | 0.1 |

| | | |
|---|---|---|
| CNS ca.* (neuro; met ) SK-N-AS | 2.3 | 0.0 |
| CNS ca. (astro) SF-539 | 0.0 | 0.2 |
| CNS ca. (astro) SNB-75 | 0.6 | 0.0 |
| CNS ca. (glio) SNB-19 | 7.1 | 1.4 |
| CNS ca. (glio) U251 | 2.6 | 0.3 |
| CNS ca. (glio) SF-295 | 0.0 | 0.0 |
| Heart | 0.9 | 0.9 |
| Skeletal Muscle (new lot*) | 1.3 | 0.0 |
| Bone marrow | 0.8 | 0.7 |
| Thymus | 0.7 | 0.0 |
| Spleen | 3.7 | 0.0 |
| Lymph node | 7.9 | 0.2 |
| Colorectal | 1.3 | 0.9 |
| Stomach | 17.9 | 0.0 |
| Small intestine | 2.2 | 0.0 |
| Colon ca. SW480 | 1.8 | 0.0 |
| Colon ca.* (SW480 met)SW620 | 3.5 | 1.3 |
| Colon ca. HT29 | 0.6 | 0.0 |
| Colon ca. HCT-116 | 0.0 | 0.0 |
| Colon ca. CaCo-2 | 0.0 | 0.0 |
| 83219 CC Well to Mod Diff (ODO3866) | 6.0 | 1.8 |
| Colon ca. HCC-2998 | 1.7 | 0.2 |
| Gastric ca.* (liver met) NCI-N87 | 1.3 | 0.0 |
| Bladder | 18.9 | 5.2 |
| Trachea | 1.1 | 0.2 |
| Kidney | 5.6 | 3.2 |
| Kidney (fetal) | 1.6 | 0.0 |
| Renal ca. 786-0 | 0.0 | 0.4 |
| Renal ca. A498 | 7.5 | 1.6 |
| Renal ca. RXF 393 | 2.1 | 0.0 |
| Renal ca. ACHN | 0.0 | 0.0 |
| Renal ca. UO-31 | 3.3 | 0.0 |
| Renal ca. TK-10 | 38.2 | 14.6 |
| Liver | 1.3 | 0.4 |
| Liver (fetal) | 0.7 | 0.3 |
| Liver ca. (hepatoblast) HepG2 | 2.0 | 0.0 |
| Lung | 2.8 | 0.2 |
| Lung (fetal) | 0.9 | 0.0 |
| Lung ca. (small cell) LX-1 | 81.2 | 37.1 |
| Lung ca. (small cell) NCI-H69 | 30.6 | 5.6 |
| Lung ca. (s.cell var.) SHP-77 | 2.1 | 0.4 |
| Lung ca. (large cell)NCI-H460 | 5.8 | 1.8 |
| Lung ca. (non-sm. cell) A549 | 50.3 | 9.2 |

| | | |
|---|---|---|
| Lung ca. (non-s.cell) NCI-H23 | 59.9 | 17.4 |
| Lung ca (non-s.cell) HOP-62 | 1.2 | 0.0 |
| Lung ca. (non-s.cl) NCI-H522 | 0.0 | 0.0 |
| Lung ca. (squam.) SW 900 | 0.2 | 0.0 |
| Lung ca. (squam.) NCI-H596 | 10.0 | 0.2 |
| Mammary gland | 1.0 | 1.0 |
| Breast ca.* (pl. effusion) MCF-7 | 0.2 | 0.3 |
| Breast ca.* (pl.ef) MDA-MB-231 | 0.0 | 0.0 |
| Breast ca.* (pl. effusion) T47D | 7.0 | 0.8 |
| Breast ca. BT-549 | 2.5 | 0.3 |
| Breast ca. MDA-N | 0.2 | 0.7 |
| Ovary | 0.8 | 0.0 |
| Ovarian ca. OVCAR-3 | 1.1 | 0.0 |
| Ovarian ca. OVCAR-4 | 1.5 | 1.4 |
| Ovarian ca. OVCAR-5 | 37.1 | 20.2 |
| Ovarian ca. OVCAR-8 | 22.2 | 2.0 |
| Ovarian ca. IGROV-1 | 12.9 | 2.4 |
| Ovarian ca.* (ascites) SK-OV-3 | 17.8 | 4.2 |
| Uterus | 1.5 | 0.2 |
| Placenta | 3.0 | 0.0 |
| Prostate | 3.3 | 0.6 |
| Prostate ca.* (bone met)PC-3 | 0.7 | 0.0 |
| Testis | 29.3 | 0.2 |
| Melanoma Hs688(A).T | 10.7 | 4.9 |
| Melanoma* (met) Hs688(B).T | 12.2 | 2.9 |
| Melanoma UACC-62 | 1.1 | 0.0 |
| Melanoma M14 | 11.5 | 2.6 |
| Melanoma LOX IMVI | 0.0 | 0.0 |
| Melanoma* (met) SK-MEL-5 | 0.3 | 0.2 |
| Adipose | 100.0 | 100.0 |

**Table 12C. Panel 4D**

| Tissue Name | Relative Expression(%) | |
|---|---|---|
| | 4Dtm2117f_ ag1257 | 4Dtm2163f_ ag1257 |
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 4.9 | 0.0 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 0.0 | 0.0 |

| | | |
|---|---|---|
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93565_primary Th1_resting dy 4-6 in IL-2 | 3.3 | 0.0 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 0.0 | 0.0 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 0.0 | 0.0 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 16.7 | 0.0 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 0.0 | 6.6 |
| 93354_CD4_none | 19.3 | 0.0 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 4.6 | 7.9 |
| 93103_LAK cells_resting | 0.0 | 0.0 |
| 93788_LAK cells_IL-2 | 0.0 | 0.0 |
| 93787_LAK cells_IL-2+IL-12 | 0.0 | 0.0 |
| 93789_LAK cells_IL-2+IFN gamma | 0.0 | 0.0 |
| 93790_LAK cells_IL-2+ IL-18 | 0.0 | 0.0 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 0.0 |
| 93578_NK Cells IL-2_resting | 0.0 | 0.0 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 0.0 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 0.0 | 0.0 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 0.0 | 0.0 |
| 93249_Ramos (B cell)_none | 0.0 | 0.0 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 0.0 |
| 93349_B lymphocytes_PWM | 0.0 | 0.0 |
| 93350_B lymphoytes_CD40L and IL-4 | 12.0 | 8.1 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 0.0 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 4.7 | 0.0 |
| 93356_Dendritic Cells_none | 0.0 | 0.0 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 0.0 |
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 6.9 |
| 93774_Monocytes_resting | 0.0 | 0.0 |
| 93776_Monocytes_LPS 50 ng/ml | 0.0 | 0.0 |
| 93581_Macrophages_resting | 0.0 | 0.0 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 0.0 |
| 93098_HUVEC (Endothelial)_none | 0.0 | 0.0 |
| 93099_HUVEC (Endothelial)_starved | 0.0 | 0.0 |
| 93100_HUVEC (Endothelial)_IL-1b | 0.0 | 0.0 |
| 93779_HUVEC (Endothelial)_IFN gamma | 0.0 | 10.1 |
| 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 0.0 | 0.0 |
| 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 0.0 | 0.0 |

| | | |
|---|---|---|
| 93781_HUVEC (Endothelial)_IL-11 | 0.0 | 0.0 |
| 93583_Lung Microvascular Endothelial Cells_none | 0.0 | 0.0 |
| 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92662_Microvascular Dermal endothelium_none | 0.0 | 7.1 |
| 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 0.0 | 4.7 |
| 93347_Small Airway Epithelium_none | 0.0 | 0.0 |
| 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92668_Coronery Artery SMC_resting | 0.0 | 0.0 |
| 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93107_astrocytes_resting | 0.0 | 0.0 |
| 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92666_KU-812 (Basophil)_resting | 0.0 | 0.0 |
| 92667_KU-812 (Basophil)_PMA/ionoycin | 0.0 | 3.0 |
| 93579_CCD1106 (Keratinocytes)_none | 0.0 | 0.0 |
| 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 | 0.0 |
| 93791_Liver Cirrhosis | 52.5 | 46.7 |
| 93792_Lupus Kidney | 0.0 | 0.0 |
| 93577_NCI-H292 | 5.8 | 0.0 |
| 93358_NCI-H292_IL-4 | 0.0 | 0.0 |
| 93360_NCI-H292_IL-9 | 0.0 | 0.0 |
| 93359_NCI-H292_IL-13 | 0.0 | 0.0 |
| 93357_NCI-H292_IFN gamma | 0.0 | 0.0 |
| 93777_HPAEC_- | 0.0 | 0.0 |
| 93778_HPAEC_IL-1 beta/TNA alpha | 0.0 | 0.0 |
| 93254_Normal Human Lung Fibroblast_none | 0.0 | 8.0 |
| 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 0.0 | 0.0 |
| 93257_Normal Human Lung Fibroblast_IL-4 | 0.0 | 4.1 |
| 93256_Normal Human Lung Fibroblast_IL-9 | 0.0 | 8.6 |
| 93255_Normal Human Lung Fibroblast_IL-13 | 10.9 | 29.7 |
| 93258_Normal Human Lung Fibroblast_IFN gamma | 22.8 | 34.9 |
| 93106_Dermal Fibroblasts CCD1070_resting | 0.0 | 7.0 |
| 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 11.5 | 4.6 |
| 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 12.2 | 8.3 |
| 93772_dermal fibroblast_IFN gamma | 100.0 | 48.6 |
| 93771_dermal fibroblast_IL-4 | 60.7 | 37.4 |
| 93259_IBD Colitis 1** | 91.4 | 100.0 |
| 93260_IBD Colitis 2 | 0.0 | 0.0 |
| 93261_IBD Crohns | 8.4 | 0.0 |

| | | |
|---|---|---|
| 735010_Colon_normal | 0.0 | 0.0 |
| 735019_Lung_none | 0.0 | 0.0 |
| 64028-1_Thymus_none | 14.0 | 9.0 |
| 64030-1_Kidney_none | 0.0 | 0.0 |

**Table 12D. Panel 4R**

| Tissue Name | Relative Expression(%) 4Rtm2154f_ag1257 | Tissue Name | Relative Expression(%) 4Rtm2154f_ag1257 |
|---|---|---|---|
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 93100_HUVEC (Endothelial)_IL-1b | 6.2 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 0.0 | 93779_HUVEC (Endothelial)_IFN gamma | 0.0 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 0.0 | 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 0.0 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 1.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 93781_HUVEC (Endothelial)_IL-11 | 0.0 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 6.2 | 93583_Lung Microvascular Endothelial Cells_none | 0.0 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 0.0 | 92662_Microvascular Dermal endothelium_none | 5.2 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93565_primary Th1_resting dy 4-6 in IL-2 | 12.4 | 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 0.0 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 0.0 | 93347_Small Airway Epithelium_none | 0.0 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 0.0 | 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 9.0 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 3.4 | 92668_Coronery Artery SMC_resting | 0.0 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 19.8 | 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 0.0 | 93107_astrocytes_resting | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4- | 17.7 | 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |

| | | | |
|---|---|---|---|
| 6 in IL-2 | | | |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 0.0 | 92666_KU-812 (Basophil)_resting | 0.0 |
| 93354_CD4_none | 0.0 | 92667_KU-812 (Basophil)_PMA/ionoycin | 0.0 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 10.3 | 93579_CCD1106 (Keratinocytes)_none | 0.0 |
| 93103_LAK cells_resting | 0.0 | 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 |
| 93788_LAK cells_IL-2 | 0.0 | 93791_Liver Cirrhosis | 30.1 |
| 93787_LAK cells_IL-2+IL-12 | 0.0 | 93792_Lupus Kidney | 0.0 |
| 93789_LAK cells_IL-2+IFN gamma | 0.0 | 93577_NCI-H292 | 0.0 |
| 93790_LAK cells_IL-2+ IL-18 | 11.9 | 93358_NCI-H292_IL-4 | 0.0 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 93360_NCI-H292_IL-9 | 12.5 |
| 93578_NK Cells IL-2_resting | 0.0 | 93359_NCI-H292_IL-13 | 0.0 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 93357_NCI-H292_IFN gamma | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 93777_HPAEC_- | 0.0 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 93778_HPAEC_IL-1 beta/TNA alpha | 0.0 |
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 93254_Normal Human Lung Fibroblast_none | 14.6 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 0.0 | 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 17.2 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 15.4 | 93257_Normal Human Lung Fibroblast_IL-4 | 0.0 |
| 93249_Ramos (B cell)_none | 0.0 | 93256_Normal Human Lung Fibroblast_IL-9 | 16.5 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 93255_Normal Human Lung Fibroblast_IL-13 | 0.0 |
| 93349_B lymphocytes_PWM | 0.0 | 93258_Normal Human Lung Fibroblast_IFN gamma | 24.1 |
| 93350_B lymphoytes_CD40L and IL-4 | 7.4 | 93106_Dermal Fibroblasts CCD1070_resting | 0.0 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 0.0 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 0.0 | 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 0.0 |
| 93356_Dendritic Cells_none | 6.4 | 93772_dermal fibroblast_IFN gamma | 100.0 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 93771_dermal fibroblast_IL-4 | 13.3 |

| | | | |
|---|---|---|---|
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 93259_IBD Colitis 1** | 66.0 |
| 93774_Monocytes_resting | 0.0 | 93260_IBD Colitis 2 | 0.0 |
| 93776_Monocytes_LPS 50 ng/ml | 4.6 | 93261_IBD Crohns | 0.0 |
| 93581_Macrophages_resting | 11.3 | 735010_Colon_normal | 11.0 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 735019_Lung_none | 0.0 |
| 93098_HUVEC (Endothelial)_none | 0.0 | 64028-1_Thymus_none | 0.0 |
| 93099_HUVEC (Endothelial)_starved | 52.1 | 64030-1_Kidney_none | 3.4 |

**Panel 1.2 Summary**: Ag1257 The results from replicate experiments using the same primer/probe set are in reasonable agreement. Expression of the GPCR1 gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, highest expression in found in a lung cancer cell line. Expression of the GPCR1 gene is high in cultured cell lines derived from cancers as compared to normal controls. Specifically, there is higher GPCR1 gene expression in 3/6 ovarian cancer cell lines, 5/10 lung cancer cell lines and a single kidney cancer cell line, indicating that this gene may play a role in the pathogenesis of these diseases. Thus, therapeutic modulation of the GPCR1 gene may be of utility in the treatment of kidney, ovarian and lung cancer. Among normal tissues, the GPCR1 gene is expressed in salivary gland, bladder, and kidney. In addition, in one of the replicate experiments the GPCR1 gene is expressed at low levels in CNS, particularly in the substantia nigra and hippocampus. Several neurotransmitter receptors are GPCRs, including the dopamine receptor family, the serotonin receptor family, the GABAB receptor, muscarinic acetylcholine receptors, and others. The hippocampus is an area of the brain which is critical for long-term memory formation, shows marked neurodegeneration in Alzheimer's disease, and has been implicated in the pathophysiology of schizophrenia, bipolar disorder and depression. Therefore, therapeutic modulation and/or activation/antagonism of the GPCR1 protein may have beneficial effects in one or more of these diseases. Similarly, as the substantia nigra degenerates in Parkinson's disease, modulation of this protein may be useful in the treatment of this disease as well.

**Panel 4D/4R Summary**: Ag1257 The expression profile of the GPCR1 transcript was examined three different times using the same probe/primer set and the results are in good agreement. This transcript encodes a GPCR that is highly expressed in gamma interferon treated dermal fibroblasts and induced in a dermal fibroblast cell line treated with this

cytokines. It is also induced in gamma interferon treated lung fibroblasts. This profile indicates that the expression of this GPCR may be up regulated as a result of asthma, emphysema, allergy, psoriasis, and viral infections when gamma interferon is present. Therefore, antibody or small molecule therapeutics that block the function of the GPCR encoded by the GPCR1 gene could reduce or inhibit the inflammation and tissue remodeling due to inflammation associated with these diseases. Please note that expression detected in the colitis 1 sample is skewed by genomic DNA contamination.

**Example 2B. GPCR2a (also known as ba64p14-B):**

Expression of gene GPCR2a was assessed using the primer-probe set Ag1258, described in Table 13A. Results of the RTQ-PCR run are shown in Table 13B.

**Table 13A. Probe Name Ag1258**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-ATCCCTAAGATGCTGGTGAACT-3' | 59.1 | 22 | 313 | 75 |
| Probe | TET-5'-CCATCTCTTACAGTGAGTGTCTGACCCA-3'-TAMRA | 67.7 | 28 | 353 | 76 |
| Reverse | 5'-TGTGTTTCCAAAGGCTAAGAAA-3' | 58.9 | 22 | 390 | 77 |

**Table 13B. Panel 1.2**

| Tissue Name | Relative Expression(%) 1.2tm1413t_ ag1258 | Tissue Name | Relative Expression(%) 1.2tm1413t_ ag1258 |
|---|---|---|---|
| Endothelial cells | 5.8 | Renal ca. 786-0 | 0.0 |
| Fetal Heart | 0.0 | Renal ca. A498 | 2.0 |
| Pancreas | 7.7 | Renal ca. RXF 393 | 0.5 |
| Pancreatic ca. CAPAN 2 | 0.0 | Renal ca. ACHN | 0.0 |
| Adrenal Gland (new lot*) | 3.4 | Renal ca. UO-31 | 0.9 |
| Thyroid | 2.2 | Renal ca. TK-10 | 0.0 |
| Salivary gland | 4.9 | Liver | 0.4 |
| Pituitary gland | 14.8 | Liver (fetal) | 2.9 |
| Brain (fetal) | 2.8 | Liver ca. (hepatoblast) HepG2 | 0.0 |
| Brain (whole) | 2.3 | Lung | 0.7 |
| Brain (amygdala) | 3.6 | Lung (fetal) | 1.1 |
| Brain (cerebellum) | 0.2 | Lung ca. (small cell) LX-1 | 0.0 |
| Brain (hippocampus) | 8.9 | Lung ca. (small cell) NCI-H69 | 10.7 |
| Brain (thalamus) | 5.3 | Lung ca. (s.cell var.) SHP-77 | 0.5 |
| Cerebral Cortex | 0.0 | Lung ca. (large cell)NCI-H460 | 10.3 |
| Spinal cord | 5.0 | Lung ca. (non-sm. cell) A549 | 3.5 |

| | | | |
|---|---|---|---|
| CNS ca. (glio/astro) U87-MG | 0.0 | Lung ca. (non-s.cell) NCI-H23 | 0.0 |
| CNS ca. (glio/astro) U-118-MG | 0.0 | Lung ca (non-s.cell) HOP-62 | 0.0 |
| CNS ca. (astro) SW1783 | 1.0 | Lung ca. (non-s.cl) NCI-H522 | 36.1 |
| CNS ca.* (neuro; met ) SK-N-AS | 0.0 | Lung ca. (squam.) SW 900 | 12.2 |
| CNS ca. (astro) SF-539 | 1.1 | Lung ca. (squam.) NCI-H596 | 4.0 |
| CNS ca. (astro) SNB-75 | 0.0 | Mammary gland | 1.0 |
| CNS ca. (glio) SNB-19 | 1.5 | Breast ca.* (pl. effusion) MCF-7 | 0.0 |
| CNS ca. (glio) U251 | 1.7 | Breast ca.* (pl.ef) MDA-MB-231 | 0.0 |
| CNS ca. (glio) SF-295 | 0.6 | Breast ca.* (pl. effusion) T47D | 17.2 |
| Heart | 1.2 | Breast ca. BT-549 | 1.8 |
| Skeletal Muscle (new lot*) | 17.1 | Breast ca. MDA-N | 0.0 |
| Bone marrow | 0.2 | Ovary | 2.3 |
| Thymus | 1.6 | Ovarian ca. OVCAR-3 | 0.1 |
| Spleen | 0.9 | Ovarian ca. OVCAR-4 | 0.0 |
| Lymph node | 1.6 | Ovarian ca. OVCAR-5 | 10.7 |
| Colorectal | 1.8 | Ovarian ca. OVCAR-8 | 1.1 |
| Stomach | 4.4 | Ovarian ca. IGROV-1 | 0.8 |
| Small intestine | 1.6 | Ovarian ca.* (ascites) SK-OV-3 | 1.4 |
| Colon ca. SW480 | 0.0 | Uterus | 1.4 |
| Colon ca.* (SW480 met)SW620 | 0.0 | Placenta | 4.4 |
| Colon ca. HT29 | 0.0 | Prostate | 14.0 |
| Colon ca. HCT-116 | 0.0 | Prostate ca.* (bone met)PC-3 | 0.0 |
| Colon ca. CaCo-2 | 0.0 | Testis | 45.7 |
| 83219 CC Well to Mod Diff (ODO3866) | 4.0 | Melanoma Hs688(A).T | 1.1 |
| Colon ca. HCC-2998 | 0.0 | Melanoma* (met) Hs688(B).T | 5.4 |
| Gastric ca.* (liver met) NCI-N87 | 0.0 | Melanoma UACC-62 | 0.0 |
| Bladder | 16.8 | Melanoma M14 | 3.1 |
| Trachea | 0.6 | Melanoma LOX IMVI | 0.0 |
| Kidney | 8.2 | Melanoma* (met) SK-MEL-5 | 0.0 |
| Kidney (fetal) | 10.6 | Adipose | 100.0 |

**Panel 1.2 Summary**: Ag1258 Expression of the GPCR2a gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, low to moderate expression (CT values = 32-35) of the GPCR2a gene is detected in a number of normal tissues, including endothelial cells, pancreas, skeletal muscle, adrenal gland, salivary gland, pituitary gland, brain (amygdala, hippocampus and thalamus), spinal cord, stomach, bladder, kidney (adult and fetal), placenta, prostate, and testis. In addition, expression of the GPCR2a gene is high in 6/10 lung cancer cell lines compared to normal lung tissue. Thus,

therapeutic modulation of the GPCR2a gene might be useful in the treatment of lung cancer or the gene may alternatively be useful in the diagnosis of lung cancer. Expression in the amygdala, hippocampus, thalamus and spinal cord suggests that the GPCR2a gene may play a role in normal nervous system function and may be disregulated in neurological diseases. As mentioned previously, the GPCR2a gene also shows low to moderate expression in skeletal muscle (CT value = 32.5) and pancreas (CT value = 33.6). Skeletal muscle and pancreatic beta cells are insulin-responsive tissues, indicating that this gene product may be regulated by insulin and important for metabolic control of the body. In addition, the GPCR2a gene shows moderate expression in the pituitary, which controls much endocrine secretion through response to hypophysiotrophic hormones (such as thyrotropin-releasing hormone, somatostatin, somatocrinin, gonadotropin-releasing hormone, corticotropin-releasing hormone) in the posterior pituitary, and response to peripheral hormones (e.g., estrogen, testosterone, etc) in the anterior pituitary. There are a number of diseases associated with pituitary pathophysiology (hyper- and hypothyroidism, gigantism, dwarfism, acromegaly, Addison's disease, Cushing's disease, diabetes insipidus) and therapeutic modulation, antagoinsm, or stimulation of the GPCR encoded by the GPCR2a gene may be useful in the treatment of one or more of these diseases. In addition, therapeutic modulation of the GPCR2a gene product might be useful in the treatment Type 1 and 2 diabetes and all other endocrinopathies involving the pancreas and pituitary.

**Panel 4D Summary**: Ag1258 Expression of the GPCR2a gene is low to undetectable (Ct values >35) in all of the samples on this panel.

**Example 2C. GPCR3 (also known as ba64p14-C):**

Expression of gene GPCR3 was assessed using the primer-probe set Ag1259 described in Table 14A. Results of the RTQ-PCR runs is shown in Table 14B.

**Table 14A. Probe Name Ag1259**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-GGTCTCTCATTACCCAACCATT-3' | 59 | 22 | 330 | 78 |
| Probe | FAM-5'-CCCGCTGCTTGGCTCAGTTCTTTT-3'-TAMRA | 70.4 | 24 | 359 | 79 |
| Reverse | 5'-TGTAACCCCAAATGCATAGAAG-3' | 59 | 22 | 384 | 80 |

**Table 14B. Panel 1.2**

| Tissue Name | Relative Expression(%) 1.2tm1425f_ ag1259 | Tissue Name | Relative Expression(%) 1.2tm1425f_ ag1259 |
|---|---|---|---|
| Endothelial cells | 0.3 | Renal ca. 786-0 | 0.0 |
| Fetal heart | 0.2 | Renal ca. A498 | 0.2 |
| Pancreas | 0.0 | Renal ca. RXF 393 | 0.0 |
| Pancreatic ca. CAPAN 2 | 0.0 | Renal ca. ACHN | 0.0 |
| Adrenal Gland (new lot*) | 1.4 | Renal ca. UO-31 | 0.1 |
| Thyroid | 0.0 | Renal ca. TK-10 | 0.0 |
| Salivary gland | 1.8 | Liver | 0.2 |
| Pituitary gland | 0.1 | Liver (fetal) | 0.3 |
| Brain (fetal) | 0.6 | Liver ca. (hepatoblast) HepG2 | 0.0 |
| Brain (whole) | 0.7 | Lung | 0.7 |
| Brain (amygdala) | 1.0 | Lung (fetal) | 0.1 |
| Brain (cerebellum) | 0.1 | Lung ca. (small cell) LX-1 | 0.0 |
| Brain (hippocampus) | 3.9 | Lung ca. (small cell) NCI-H69 | 2.4 |
| Brain (thalamus) | 0.3 | Lung ca. (s.cell var.) SHP-77 | 0.0 |
| Cerebral Cortex | 1.1 | Lung ca. (large cell)NCI-H460 | 0.3 |
| Spinal cord | 0.4 | Lung ca. (non-sm. cell) A549 | 2.4 |
| CNS ca. (glio/astro) U87-MG | 0.0 | Lung ca. (non-s.cell) NCI-H23 | 0.4 |
| CNS ca. (glio/astro) U-118-MG | 0.0 | Lung ca (non-s.cell) HOP-62 | 0.0 |
| CNS ca. (astro) SW1783 | 0.1 | Lung ca. (non-s.cl) NCI-H522 | 0.0 |
| CNS ca.* (neuro; met ) SK-N-AS | 0.0 | Lung ca. (squam.) SW 900 | 0.2 |
| CNS ca. (astro) SF-539 | 0.0 | Lung ca. (squam.) NCI-H596 | 0.0 |
| CNS ca. (astro) SNB-75 | 0.0 | Mammary gland | 0.5 |
| CNS ca. (glio) SNB-19 | 0.3 | Breast ca.* (pl. effusion) MCF-7 | 0.0 |
| CNS ca. (glio) U251 | 0.8 | Breast ca.* (pl.ef) MDA-MB-231 | 0.0 |
| CNS ca. (glio) SF-295 | 0.0 | Breast ca.* (pl. effusion) T47D | 1.4 |
| Heart | 0.0 | Breast ca. BT-549 | 0.0 |
| Skeletal Muscle (new lot*) | 0.0 | Breast ca. MDA-N | 0.0 |
| Bone marrow | 0.5 | Ovary | 0.0 |
| Thymus | 0.0 | Ovarian ca. OVCAR-3 | 0.2 |
| Spleen | 0.6 | Ovarian ca. OVCAR-4 | 0.0 |
| Lymph node | 0.5 | Ovarian ca. OVCAR-5 | 0.8 |
| Colorectal | 0.4 | Ovarian ca. OVCAR-8 | 0.6 |
| Stomach | 0.2 | Ovarian ca. IGROV-1 | 0.0 |
| Small intestine | 0.0 | Ovarian ca.* (ascites) SK-OV-3 | 0.0 |
| Colon ca. SW480 | 0.1 | Uterus | 0.4 |

| | | | |
|---|---|---|---|
| Colon ca.* (SW480 met)SW620 | 0.0 | Placenta | 0.1 |
| Colon ca. HT29 | 0.0 | Prostate | 2.8 |
| Colon ca. HCT-116 | 0.0 | Prostate ca.* (bone met)PC-3 | 0.0 |
| Colon ca. CaCo-2 | 0.0 | Testis | 0.5 |
| 83219 CC Well to Mod Diff (ODO3866) | 1.8 | Melanoma Hs688(A).T | 0.2 |
| Colon ca. HCC-2998 | 0.0 | Melanoma* (met) Hs688(B).T | 0.7 |
| Gastric ca.* (liver met) NCI-N87 | 0.3 | Melanoma UACC-62 | 0.0 |
| Bladder | 3.2 | Melanoma M14 | 0.8 |
| Trachea | 0.0 | Melanoma LOX IMVI | 0.0 |
| Kidney | 0.5 | Melanoma* (met) SK-MEL-5 | 0.0 |
| Kidney (fetal) | 4.3 | Adipose | **100.0** |

**Panel 1.2 Summary**: Ag1259 Expression of the GPCR3 gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, this gene is expressed at low levels in only a few normal tissues including hippocampus, bladder, and fetal kidney. No overexpression of the GPCR3 gene is detected in any of the cancer cell lines on this panel. There appears to be low but significant expression in fetal kidney when compared to adult kidney. Thus, the GPCR3 gene could play a role in kidney development and the therapeutic modulation of this gene might have utility in the treatment of disorders of the kidney. This gene product might also be used to distinguish fetal kidney from other tissues.

**Panel 4D Summary**: Ag1259 Expression of the GPCR3 gene is low to undetectable (CT values >35) in all of the samples on this panel except in IBD colitis 1; however, this sample is believed to be contaminated with genomic DNA and must therefore be disregarded and thus the data not shown.

**Example 2D. GPCR4c (also known as ba64p14-D):**

Expression of gene GPCR4c was assessed using the primer-probe set Ag1260 described in Table 15A. Results of the RTQ-PCR runs are shown in Table 15B and 15C.

**Table 15A. Probe Name Ag1260**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-CCAACCTGTCTTTTGTTGACAT-3' | 59 | 22 | 228 | 81 |
| Probe | TET-5'-CGTCCTCCACAGTTACCAAGATGCTG-3'-TAMRA | 68.9 | 26 | 258 | 82 |
| Reverse | 5'-CCGTATAGGAGATGGTGTGATG-3' | 59.3 | 22 | 302 | 83 |

**Table 15B. Panel 1.2**

| Tissue Name | Relative Expression(%) 1.2tm1423t_ ag1260 | Tissue Name | Relative Expression(%) 1.2tm1423t_ ag1260 |
|---|---|---|---|
| Endothelial cells | 1.0 | Renal ca. 786-0 | 0.0 |
| Fetal heart | 0.0 | Renal ca. A498 | 0.3 |
| Pancreas | 0.0 | Renal ca. RXF 393 | 0.0 |
| Pancreatic ca. CAPAN 2 | 0.0 | Renal ca. ACHN | 0.0 |
| Adrenal Gland (new lot*) | 0.0 | Renal ca. UO-31 | 0.0 |
| Thyroid | 0.0 | Renal ca. TK-10 | 42.6 |
| Salivary gland | 0.0 | Liver | 0.0 |
| Pituitary gland | 0.0 | Liver (fetal) | 0.0 |
| Brain (fetal) | 0.0 | Liver ca. (hepatoblast) HepG2 | 0.0 |
| Brain (whole) | 0.0 | Lung | 0.0 |
| Brain (amygdala) | 0.0 | Lung (fetal) | 0.0 |
| Brain (cerebellum) | 0.0 | Lung ca. (small cell) LX-1 | 100.0 |
| Brain (hippocampus) | 0.0 | Lung ca. (small cell) NCI-H69 | 0.0 |
| Brain (thalamus) | 0.0 | Lung ca. (s.cell var.) SHP-77 | 0.0 |
| Cerebral Cortex | 0.0 | Lung ca. (large cell)NCI-H460 | 1.0 |
| Spinal cord | 0.0 | Lung ca. (non-sm. cell) A549 | 5.8 |
| CNS ca. (glio/astro) U87-MG | 3.4 | Lung ca. (non-s.cell) NCI-H23 | 2.8 |
| CNS ca. (glio/astro) U-118-MG | 0.0 | Lung ca (non-s.cell) HOP-62 | 0.0 |
| CNS ca. (astro) SW1783 | 0.0 | Lung ca. (non-s.cl) NCI-H522 | 0.0 |
| CNS ca.* (neuro; met ) SK-N-AS | 0.0 | Lung ca. (squam.) SW 900 | 0.0 |
| CNS ca. (astro) SF-539 | 0.0 | Lung ca. (squam.) NCI-H596 | 0.0 |
| CNS ca. (astro) SNB-75 | 0.0 | Mammary gland | 0.0 |
| CNS ca. (glio) SNB-19 | 0.0 | Breast ca.* (pl. effusion) MCF-7 | 0.0 |
| CNS ca. (glio) U251 | 0.0 | Breast ca.* (pl.ef) MDA-MB-231 | 0.0 |
| CNS ca. (glio) SF-295 | 0.0 | Breast ca.* (pl. effusion) T47D | 0.2 |
| Heart | 0.0 | Breast ca. BT-549 | 0.0 |
| Skeletal Muscle (new lot*) | 0.0 | Breast ca. MDA-N | 0.0 |
| Bone marrow | 0.0 | Ovary | 0.1 |
| Thymus | 0.0 | Ovarian ca. OVCAR-3 | 0.0 |
| Spleen | 0.0 | Ovarian ca. OVCAR-4 | 0.0 |
| Lymph node | 0.5 | Ovarian ca. OVCAR-5 | 7.5 |
| Colorectal | 0.0 | Ovarian ca. OVCAR-8 | 11.1 |
| Stomach | 0.8 | Ovarian ca. IGROV-1 | 0.9 |
| Small intestine | 0.0 | Ovarian ca.* (ascites) SK-OV-3 | 0.3 |
| Colon ca. SW480 | 0.0 | Uterus | 0.0 |
| Colon ca.* (SW480 met)SW620 | 0.0 | Placenta | 0.0 |

| | | | |
|---|---|---|---|
| Colon ca. HT29 | 0.0 | Prostate | 0.0 |
| Colon ca. HCT-116 | 0.0 | Prostate ca.* (bone met)PC-3 | 0.0 |
| Colon ca. CaCo-2 | 0.0 | Testis | 7.5 |
| 83219 CC Well to Mod Diff (ODO3866) | 6.5 | Melanoma Hs688(A).T | 1.3 |
| Colon ca. HCC-2998 | 0.0 | Melanoma* (met) Hs688(B).T | 1.8 |
| Gastric ca.* (liver met) NCI-N87 | 0.0 | Melanoma UACC-62 | 0.0 |
| Bladder | 0.1 | Melanoma M14 | 0.0 |
| Trachea | 0.0 | Melanoma LOX IMVI | 0.0 |
| Kidney | 0.0 | Melanoma* (met) SK-MEL-5 | 0.0 |
| Kidney (fetal) | 0.0 | Adipose | 73.2 |

**Table 15C. Panel 4D**

| Tissue Name | Relative Expression(%) | |
|---|---|---|
| | 4Dtm2120t_ ag1260 | 4dtm4371t_ ag1260 |
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 7.0 | 0.0 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93565_primary Th1_resting dy 4-6 in IL-2 | 6.2 | 19.6 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 0.0 | 0.0 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 0.0 | 0.0 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 6.4 | 0.0 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 0.0 | 30.1 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 6.2 | 0.0 |
| 93354_CD4_none | 0.0 | 0.0 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 4.5 | 0.0 |
| 93103_LAK cells_resting | 0.0 | 0.0 |
| 93788_LAK cells_IL-2 | 0.0 | 0.0 |
| 93787_LAK cells_IL-2+IL-12 | 0.0 | 6.0 |
| 93789_LAK cells_IL-2+IFN gamma | 10.3 | 0.0 |
| 93790_LAK cells_IL-2+ IL-18 | 8.2 | 31.2 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 0.0 |
| 93578_NK Cells IL-2_resting | 0.0 | 0.0 |

| | | |
|---|---|---|
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 0.0 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 0.0 | 0.0 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 45.1 | 17.1 |
| 93249_Ramos (B cell)_none | 0.0 | 0.0 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 0.0 |
| 93349_B lymphocytes_PWM | 0.0 | 0.0 |
| 93350_B lymphoytes_CD40L and IL-4 | 0.0 | 0.0 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 0.0 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 11.2 | 0.0 |
| 93356_Dendritic Cells_none | 9.5 | 0.0 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 6.1 | 0.0 |
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 8.8 |
| 93774_Monocytes_resting | 0.0 | 0.0 |
| 93776_Monocytes_LPS 50 ng/ml | 2.8 | 0.0 |
| 93581_Macrophages_resting | 4.7 | 22.2 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 0.0 |
| 93098_HUVEC (Endothelial)_none | 5.8 | 9.0 |
| 93099_HUVEC (Endothelial)_starved | 5.7 | 31.9 |
| 93100_HUVEC (Endothelial)_IL-1b | 6.8 | 25.3 |
| 93779_HUVEC (Endothelial)_IFN gamma | 5.8 | 44.8 |
| 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 0.0 | 7.5 |
| 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 0.0 | 11.9 |
| 93781_HUVEC (Endothelial)_IL-11 | 12.2 | 0.0 |
| 93583_Lung Microvascular Endothelial Cells_none | 0.0 | 0.0 |
| 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92662_Microvascular Dermal endothelium_none | 0.0 | 0.0 |
| 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 0.0 | 0.0 |
| 93347_Small Airway Epithelium_none | 0.0 | 0.0 |
| 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92668_Coronery Artery SMC_resting | 0.0 | 0.0 |
| 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93107_astrocytes_resting | 0.0 | 0.0 |
| 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92666_KU-812 (Basophil)_resting | 0.0 | 0.0 |
| 92667_KU-812 (Basophil)_PMA/ionoycin | 0.0 | 0.0 |
| 93579_CCD1106 (Keratinocytes)_none | 0.0 | 0.0 |

| | | |
|---|---|---|
| 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 | 0.0 |
| 93791_Liver Cirrhosis | 32.8 | 73.2 |
| 93792_Lupus Kidney | 0.0 | 0.0 |
| 93577_NCI-H292 | 0.0 | 0.0 |
| 93358_NCI-H292_IL-4 | 0.0 | 0.0 |
| 93360_NCI-H292_IL-9 | 0.0 | 11.0 |
| 93359_NCI-H292_IL-13 | 6.3 | 0.0 |
| 93357_NCI-H292_IFN gamma | 0.0 | 0.0 |
| 93777_HPAEC_- | 0.0 | 0.0 |
| 93778_HPAEC_IL-1 beta/TNA alpha | 0.0 | 0.0 |
| 93254_Normal Human Lung Fibroblast_none | 0.0 | 11.3 |
| 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 0.0 | 0.0 |
| 93257_Normal Human Lung Fibroblast_IL-4 | 5.3 | 19.3 |
| 93256_Normal Human Lung Fibroblast_IL-9 | 8.1 | 23.7 |
| 93255_Normal Human Lung Fibroblast_IL-13 | 6.2 | 19.5 |
| 93258_Normal Human Lung Fibroblast_IFN gamma | 37.9 | 75.3 |
| 93106_Dermal Fibroblasts CCD1070_resting | 8.5 | 11.0 |
| 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 9.3 | 18.7 |
| 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 10.9 | 16.6 |
| 93772_dermal fibroblast_IFN gamma | 58.6 | 54.7 |
| 93771_dermal fibroblast_IL-4 | 100.0 | 100.0 |
| 93259_IBD Colitis 1** | 69.7 | 0.0 |
| 93260_IBD Colitis 2 | 0.0 | 0.0 |
| 93261_IBD Crohns | 6.2 | 0.0 |
| 735010_Colon_normal | 0.0 | 5.1 |
| 735019_Lung_none | 10.4 | 0.0 |
| 64028-1_Thymus_none | 9.7 | 18.7 |
| 64030-1_Kidney_none | 0.0 | 0.0 |

**Panel 1.2 Summary**: Ag1260 Expression of the GPCR4c gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, there appears to be specific expression in cultured cell lines derived from several types of cancers. No expression of the ba64p14-E gene is detected in any normal tissues with the exception of testis. However, there is moderate expression in ovarian and lung cancer cell lines indicating that this gene may play a role in the pathogenesis of these diseases. Thus, therapeutic modulation of this gene may be of utility in the treatment of ovarian and lung cancer.

**Panel 4D Summary**: Ag1260 The expression profile of the GPCR4c transcript was examined two different times using the same probe/primer set and the results are reasonably concordant. This gene encodes a GPCR that is highly expressed in gamma interferon and IL-4

treated dermal fibroblasts and induced in a dermal fibroblast cell line treated with these cytokines. It is also induced in gamma interferon treated lung fibroblasts. This profile indicates that the expression of this GPCR may be up regulated during asthma, emphysema and allergy (in which high levels of IL-4 are present) or during psoriasis and viral infections (when high levels of gamma interferon are present). Antibody or small molecule therapeutics that block the function of the GPCR encoded by the GPCR4c gene could therefore reduce or inhibit the inflammation and tissue remodeling due to inflammation associated with these diseases.

**Example E. GPCR5b (also known as ba64p14-E or CG50385-01):**

Expression of gene GPCR5b was assessed using the primer-probe sets Ag1261, Ag1261b and Ag1261c, described in Tables 16A, 16B and 16C. Results of the RTQ-PCR runs are shown in Table 16D, 16E, 16F, 16G, 16H and 16I.

**Table 16A. Probe Name Ag1261**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-GACGTAATTGCCTCTGTGATGT-3' | 59.1 | 22 | 841 | 84 |
| Probe | FAM-5'-ACCCCATTGCTGAATCCCTTCATTTA-3'-TAMRA | 68.2 | 26 | 874 | 85 |
| Reverse | 5'-CCCTTTATGTCCCTGTTCCTTA-3' | 59.4 | 22 | 905 | 86 |

**Table 16B. Probe Name Ag1261b**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-TCACTGACATCTCCCTTTCATC-3' | 59.1 | 22 | 232 | 87 |
| Probe | TET-5'-TGTCACTGTCCCAAAGATGTTATTAAGCA-3'-TAMRA | 68.2 | 29 | 254 | 88 |
| Reverse | 5'-GAATGGATTGATCCTGAGTTTG-3' | 59.4 | 22 | 306 | 89 |

**Table 16C. Probe Name Ag1261c**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-CTAAGGGCATCTTCAAAGCTTT-3' | 59.1 | 22 | 721 | 90 |
| Probe | TET-5'-CTGTGGCTCTCACCTCTCTGTGGTGT-3'-TAMRA | 69.2 | 26 | 749 | 91 |
| Reverse | 5'-GAGGGGAGAAAATACAGTCCAA-3' | 59.5 | 22 | 817 | 92 |

**Table 16D. Panel 1.2**

| Tissue Name | Relative Expression(%) |
|---|---|

| | 1.2tm1426f_ ag1261 | 1.2tm2185f_ ag1261 |
|---|---|---|
| Endothelial cells | 2.7 | 1.4 |
| Fetal Heart | 0.0 | 0.0 |
| Pancreas | 0.0 | 0.0 |
| Pancreatic ca. CAPAN 2 | 1.7 | 0.7 |
| Adrenal Gland (new lot*) | 0.0 | 0.0 |
| Thyroid | 0.0 | 0.0 |
| Salivary gland | 0.0 | 0.3 |
| Pituitary gland | 0.0 | 0.0 |
| Brain (fetal) | 0.0 | 0.0 |
| Brain (whole) | 0.0 | 0.0 |
| Brain (amygdala) | 0.0 | 0.0 |
| Brain (cerebellum) | 0.0 | 0.0 |
| Brain (hippocampus) | 2.5 | 0.8 |
| Brain (thalamus) | 0.0 | 0.0 |
| Cerebral Cortex | 0.0 | 0.0 |
| Spinal cord | 0.0 | 0.0 |
| CNS ca. (glio/astro) U87-MG | 3.0 | 4.2 |
| CNS ca. (glio/astro) U-118-MG | 0.0 | 0.0 |
| CNS ca. (astro) SW1783 | 0.0 | 0.0 |
| CNS ca.* (neuro; met ) SK-N-AS | 0.3 | 0.0 |
| CNS ca. (astro) SF-539 | 0.0 | 0.0 |
| CNS ca. (astro) SNB-75 | 0.0 | 0.0 |
| CNS ca. (glio) SNB-19 | 0.0 | 0.0 |
| CNS ca. (glio) U251 | 0.0 | 0.0 |
| CNS ca. (glio) SF-295 | 0.0 | 0.0 |
| Heart | 0.0 | 0.0 |
| Skeletal Muscle (new lot*) | 0.0 | 0.0 |
| Bone marrow | 0.0 | 0.6 |
| Thymus | 0.0 | 0.0 |
| Spleen | 0.0 | 0.0 |
| Lymph node | 0.0 | 0.0 |
| Colorectal | 0.3 | 0.1 |
| Stomach | 0.6 | 0.2 |
| Small intestine | 0.0 | 0.2 |
| Colon ca. SW480 | 0.0 | 0.0 |
| Colon ca.* (SW480 met)SW620 | 0.2 | 0.5 |
| Colon ca. HT29 | 0.0 | 0.0 |
| Colon ca. HCT-116 | 0.0 | 0.0 |
| Colon ca. CaCo-2 | 0.0 | 0.0 |
| 83219 CC Well to Mod Diff (ODO3866) | 3.0 | 1.7 |
| Colon ca. HCC-2998 | 0.0 | 0.0 |
| Gastric ca.* (liver met) NCI-N87 | 0.0 | 0.0 |

| | | |
|---|---|---|
| Bladder | 0.0 | 1.9 |
| Trachea | 0.0 | 0.0 |
| Kidney | 0.0 | 0.0 |
| Kidney (fetal) | 0.0 | 0.0 |
| Renal ca. 786-0 | 0.0 | 0.0 |
| Renal ca. A498 | 0.0 | 0.2 |
| Renal ca. RXF 393 | 0.0 | 0.0 |
| Renal ca. ACHN | 0.0 | 0.0 |
| Renal ca. UO-31 | 0.0 | 0.0 |
| Renal ca. TK-10 | 15.4 | 9.5 |
| Liver | 0.0 | 0.3 |
| Liver (fetal) | 0.0 | 0.0 |
| Liver ca. (hepatoblast) HepG2 | 0.0 | 0.0 |
| Lung | 0.0 | 0.0 |
| Lung (fetal) | 0.0 | 0.0 |
| Lung ca. (small cell) LX-1 | 16.3 | 21.9 |
| Lung ca. (small cell) NCI-H69 | 12.2 | 2.1 |
| Lung ca. (s.cell var.) SHP-77 | 0.5 | 0.2 |
| Lung ca. (large cell)NCI-H460 | 1.3 | 0.8 |
| Lung ca. (non-sm. cell) A549 | 33.7 | 3.5 |
| Lung ca. (non-s.cell) NCI-H23 | 8.2 | 9.2 |
| Lung ca (non-s.cell) HOP-62 | 0.0 | 0.0 |
| Lung ca. (non-s.cl) NCI-H522 | 0.0 | 0.0 |
| Lung ca. (squam.) SW 900 | 0.0 | 0.0 |
| Lung ca. (squam.) NCI-H596 | 0.3 | 0.0 |
| Mammary gland | 0.0 | 0.4 |
| Breast ca.* (pl. effusion) MCF-7 | 0.0 | 0.0 |
| Breast ca.* (pl.ef) MDA-MB-231 | 0.0 | 0.0 |
| Breast ca.* (pl. effusion) T47D | 1.1 | 0.9 |
| Breast ca. BT-549 | 0.0 | 0.3 |
| Breast ca. MDA-N | 0.0 | 0.0 |
| Ovary | 0.0 | 0.0 |
| Ovarian ca. OVCAR-3 | 0.0 | 0.0 |
| Ovarian ca. OVCAR-4 | 0.0 | 0.0 |
| Ovarian ca. OVCAR-5 | 6.1 | 3.3 |
| Ovarian ca. OVCAR-8 | 10.8 | 5.0 |
| Ovarian ca. IGROV-1 | 0.0 | 1.0 |
| Ovarian ca.* (ascites) SK-OV-3 | 0.2 | 0.5 |
| Uterus | 0.0 | 0.0 |
| Placenta | 0.0 | 0.0 |
| Prostate | 0.0 | 0.4 |
| Prostate ca.* (bone met)PC-3 | 0.0 | 0.0 |
| Testis | 0.0 | 0.0 |

| | | |
|---|---|---|
| Melanoma Hs688(A).T | 5.4 | 5.8 |
| Melanoma* (met) Hs688(B).T | 5.6 | 3.9 |
| Melanoma UACC-62 | 0.0 | 0.0 |
| Melanoma M14 | 3.3 | 0.9 |
| Melanoma LOX IMVI | 0.0 | 0.0 |
| Melanoma* (met) SK-MEL-5 | 0.0 | 0.0 |
| Adipose | 100.0 | 100.0 |

**Table 16E. Panel 1.3D**

| Tissue Name | Relative Expression(%) 1.3dx4tm5586t_ag1261b_a2 | Tissue Name | Relative Expression(%) 1.3dx4tm5586t_ag1261b_a2 |
|---|---|---|---|
| Liver adenocarcinoma | 16.4 | Kidney (fetal) | 0.0 |
| Pancreas | 0.0 | Renal ca. 786-0 | 0.0 |
| Pancreatic ca. CAPAN 2 | 10.2 | Renal ca. A498 | 0.0 |
| Adrenal gland | 0.0 | Renal ca. RXF 393 | 14.4 |
| Thyroid | 0.0 | Renal ca. ACHN | 0.0 |
| Salivary gland | 0.0 | Renal ca. UO-31 | 0.0 |
| Pituitary gland | 0.0 | Renal ca. TK-10 | 4.9 |
| Brain (fetal) | 0.0 | Liver | 0.0 |
| Brain (whole) | 11.1 | Liver (fetal) | 0.0 |
| Brain (amygdala) | 3.6 | Liver ca. (hepatoblast) HepG2 | 0.0 |
| Brain (cerebellum) | 0.0 | Lung | 0.0 |
| Brain (hippocampus) | 9.0 | Lung (fetal) | 0.0 |
| Brain (substantia nigra) | 0.0 | Lung ca. (small cell) LX-1 | 36.5 |
| Brain (thalamus) | 0.0 | Lung ca. (small cell) NCI-H69 | 0.0 |
| Cerebral Cortex | 0.0 | Lung ca. (s.cell var.) SHP-77 | 0.0 |
| Spinal cord | 0.0 | Lung ca. (large cell)NCI-H460 | 0.0 |
| CNS ca. (glio/astro) U87-MG | 13.6 | Lung ca. (non-sm. cell) A549 | 6.3 |
| CNS ca. (glio/astro) U-118-MG | 0.0 | Lung ca. (non-s.cell) NCI-H23 | 8.5 |
| CNS ca. (astro) SW1783 | 0.0 | Lung ca (non-s.cell) HOP-62 | 0.0 |
| CNS ca.* (neuro; met ) SK-N-AS | 0.0 | Lung ca. (non-s.cl) NCI-H522 | 0.0 |
| CNS ca. (astro) SF-539 | 0.0 | Lung ca. (squam.) SW 900 | 0.0 |
| CNS ca. (astro) SNB-75 | 0.0 | Lung ca. (squam.) NCI-H596 | 0.0 |
| CNS ca. (glio) SNB-19 | 0.0 | Mammary gland | 0.0 |
| CNS ca. (glio) U251 | 0.0 | Breast ca.* (pl. effusion) MCF-7 | 0.0 |
| CNS ca. (glio) SF-295 | 0.0 | Breast ca.* (pl.ef) MDA-MB-231 | 0.0 |
| Heart (fetal) | 0.0 | Breast ca.* (pl. effusion) T47D | 0.0 |
| Heart | 0.0 | Breast ca. BT-549 | 5.8 |
| Fetal Skeletal | 0.0 | Breast ca. MDA-N | 0.0 |

| | | | |
|---|---|---|---|
| Skeletal muscle | 0.0 | Ovary | 0.0 |
| Bone marrow | 0.0 | Ovarian ca. OVCAR-3 | 0.0 |
| Thymus | 0.0 | Ovarian ca. OVCAR-4 | 0.0 |
| Spleen | 100.0 | Ovarian ca. OVCAR-5 | 3.7 |
| Lymph node | 0.0 | Ovarian ca. OVCAR-8 | 0.0 |
| Colorectal | 4.3 | Ovarian ca. IGROV-1 | 0.0 |
| Stomach | 8.9 | Ovarian ca.* (ascites) SK-OV-3 | 18.1 |
| Small intestine | 0.0 | Uterus | 0.0 |
| Colon ca. SW480 | 3.8 | Placenta | 0.0 |
| Colon ca.* (SW480 met)SW620 | 4.0 | Prostate | 0.0 |
| Colon ca. HT29 | 0.0 | Prostate ca.* (bone met)PC-3 | 0.0 |
| Colon ca. HCT-116 | 0.0 | Testis | 0.0 |
| Colon ca. CaCo-2 | 0.0 | Melanoma Hs688(A).T | 0.0 |
| 83219 CC Well to Mod Diff (ODO3866) | 3.8 | Melanoma* (met) Hs688(B).T | 0.0 |
| Colon ca. HCC-2998 | 0.0 | Melanoma UACC-62 | 0.0 |
| Gastric ca.* (liver met) NCI-N87 | 0.0 | Melanoma M14 | 0.0 |
| Bladder | 3.6 | Melanoma LOX IMVI | 0.0 |
| Trachea | 0.0 | Melanoma* (met) SK-MEL-5 | 0.0 |
| Kidney | 0.0 | Adipose | 0.0 |

**Table 16F. Panel 4D (Part 1)**

| Tissue Name | Relative Expression(%) | |
|---|---|---|
| | 4Dtm2164f_ag1261 | 4dtm2191f_ag1261 |
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 9.0 | 0.0 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93565_primary Th1_resting dy 4-6 in IL-2 | 0.0 | 0.0 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 0.0 | 0.0 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 0.0 | 0.0 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 8.2 | 15.2 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 0.0 | 0.0 |

| | | |
|---|---|---|
| 93354_CD4_none | 0.0 | 11.0 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 0.0 | 0.0 |
| 93103_LAK cells_resting | 0.0 | 0.0 |
| 93788_LAK cells_IL-2 | 0.0 | 0.0 |
| 93787_LAK cells_IL-2+IL-12 | 9.7 | 0.0 |
| 93789_LAK cells_IL-2+IFN gamma | 11.2 | 0.0 |
| 93790_LAK cells_IL-2+ IL-18 | 11.7 | 0.0 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 0.0 |
| 93578_NK Cells IL-2_resting | 0.0 | 0.0 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 0.0 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 0.0 | 0.0 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 37.9 | 36.1 |
| 93249_Ramos (B cell)_none | 0.0 | 0.0 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 0.0 |
| 93349_B lymphocytes_PWM | 0.0 | 0.0 |
| 93350_B lymphoytes_CD40L and IL-4 | 0.0 | 0.0 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 0.0 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 0.0 | 0.0 |
| 93356_Dendritic Cells_none | 0.0 | 0.0 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 0.0 |
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 0.0 |
| 93774_Monocytes_resting | 0.0 | 0.0 |
| 93776_Monocytes_LPS 50 ng/ml | 0.0 | 0.0 |
| 93581_Macrophages_resting | 0.0 | 0.0 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 0.0 |
| 93098_HUVEC (Endothelial)_none | 0.0 | 0.0 |
| 93099_HUVEC (Endothelial)_starved | 0.0 | 12.5 |
| 93100_HUVEC (Endothelial)_IL-1b | 0.0 | 0.0 |
| 93779_HUVEC (Endothelial)_IFN gamma | 0.0 | 0.0 |
| 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 0.0 | 0.0 |
| 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 0.0 | 0.0 |
| 93781_HUVEC (Endothelial)_IL-11 | 0.0 | 0.0 |
| 93583_Lung Microvascular Endothelial Cells_none | 0.0 | 0.0 |
| 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92662_Microvascular Dermal endothelium_none | 0.0 | 0.0 |
| 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 0.0 | 0.0 |
| 93347_Small Airway Epithelium_none | 0.0 | 0.0 |

| | | |
|---|---|---|
| 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92668_Coronery Artery SMC_resting | 0.0 | 0.0 |
| 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93107_astrocytes_resting | 0.0 | 0.0 |
| 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92666_KU-812 (Basophil)_resting | 0.0 | 0.0 |
| 92667_KU-812 (Basophil)_PMA/ionoycin | 0.0 | 0.0 |
| 93579_CCD1106 (Keratinocytes)_none | 0.0 | 0.0 |
| 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 | 0.0 |
| 93791_Liver Cirrhosis | 21.6 | 19.2 |
| 93792_Lupus Kidney | 0.0 | 0.0 |
| 93577_NCI-H292 | 0.0 | 0.0 |
| 93358_NCI-H292_IL-4 | 0.0 | 0.0 |
| 93360_NCI-H292_IL-9 | 0.0 | 0.0 |
| 93359_NCI-H292_IL-13 | 0.0 | 0.0 |
| 93357_NCI-H292_IFN gamma | 0.0 | 0.0 |
| 93777_HPAEC_- | 0.0 | 0.0 |
| 93778_HPAEC_IL-1 beta/TNA alpha | 0.0 | 0.0 |
| 93254_Normal Human Lung Fibroblast_none | 0.0 | 0.0 |
| 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 0.0 | 0.0 |
| 93257_Normal Human Lung Fibroblast_IL-4 | 0.0 | 0.0 |
| 93256_Normal Human Lung Fibroblast_IL-9 | 10.7 | 0.0 |
| 93255_Normal Human Lung Fibroblast_IL-13 | 15.2 | 15.6 |
| 93258_Normal Human Lung Fibroblast_IFN gamma | 25.2 | 31.9 |
| 93106_Dermal Fibroblasts CCD1070_resting | 0.0 | 0.0 |
| 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 9.8 | 0.0 |
| 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 23.3 | 12.3 |
| 93772_dermal fibroblast_IFN gamma | 47.6 | 43.8 |
| 93771_dermal fibroblast_IL-4 | 100.0 | 66.9 |
| 93259_IBD Colitis 1** | 76.8 | 100.0 |
| 93260_IBD Colitis 2 | 0.0 | 0.0 |
| 93261_IBD Crohns | 0.0 | 0.0 |
| 735010_Colon_normal | 0.0 | 0.0 |
| 735019_Lung_none | 0.0 | 0.0 |
| 64028-1_Thymus_none | 0.0 | 22.8 |
| 64030-1_Kidney_none | 0.0 | 0.0 |

**Table 16G. Panel 4D (Part 2)**

| Tissue Name | Relative Expression(%) | | |
|---|---|---|---|
| | 4Dtm2260t_ ag1261c | 4dtm2270t_ ag1261c | 4dx4tm5554t_ ag1261b_a1 |

| | | | |
|---|---|---|---|
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 0.5 | 0.0 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 4.9 | 2.4 | 1.5 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 4.9 | 4.5 | 1.0 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 1.5 | 0.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 0.0 | 0.0 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 0.0 | 0.4 | 0.0 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 7.7 | 0.3 | 0.0 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 1.2 | 0.0 |
| 93565_primary Th1_resting dy 4-6 in IL-2 | 0.0 | 0.8 | 0.0 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 10.3 | 1.4 | 3.1 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 0.0 | 0.0 | 0.0 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 10.7 | 1.5 | 0.6 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 6.2 | 2.6 | 5.2 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 0.0 | 6.0 | 3.2 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 4.5 | 11.2 | 1.1 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 0.0 | 1.0 | 0.4 |
| 93354_CD4_none | 0.0 | 10.3 | 5.5 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 0.0 | 0.6 | 0.0 |
| 93103_LAK cells_resting | 0.0 | 0.6 | 0.6 |
| 93788_LAK cells_IL-2 | 0.0 | 1.8 | 0.0 |
| 93787_LAK cells_IL-2+IL-12 | 6.8 | 29.1 | 2.4 |
| 93789_LAK cells_IL-2+IFN gamma | 0.0 | 14.0 | 5.5 |
| 93790_LAK cells_IL-2+ IL-18 | 0.0 | 3.6 | 1.6 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 0.0 | 0.0 |
| 93578_NK Cells IL-2_resting | 0.0 | 3.5 | 2.2 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 4.9 | 2.1 | 1.2 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 5.5 | 0.3 | 0.0 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 | 1.3 |
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 0.4 | 0.0 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 0.0 | 0.0 | 0.0 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 27.7 | 7.0 | 0.7 |
| 93249_Ramos (B cell)_none | 0.0 | 0.0 | 0.0 |
| 93250_Ramos (B cell)_ionomycin | 3.4 | 0.6 | 0.0 |
| 93349_B lymphocytes_PWM | 0.0 | 0.5 | 0.0 |
| 93350_B lymphoytes_CD40L and IL-4 | 0.0 | 3.6 | 0.0 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 1.7 | 0.0 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 5.2 | 0.5 | 0.0 |
| 93356_Dendritic Cells_none | 4.8 | 21.6 | 2.3 |

| | | | |
|---|---|---|---|
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 0.0 | 0.0 |
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 0.7 | 3.3 |
| 93774_Monocytes_resting | 0.0 | 0.4 | 0.8 |
| 93776_Monocytes_LPS 50 ng/ml | 0.0 | 0.0 | 0.0 |
| 93581_Macrophages_resting | 4.5 | 2.9 | 1.3 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 0.0 | 0.0 |
| 93098_HUVEC (Endothelial)_none | 4.8 | 3.7 | 1.8 |
| 93099_HUVEC (Endothelial)_starved | 31.0 | 4.6 | 6.9 |
| 93100_HUVEC (Endothelial)_IL-1b | 0.0 | 0.9 | 1.5 |
| 93779_HUVEC (Endothelial)_IFN gamma | 5.2 | 1.9 | 2.5 |
| 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 0.0 | 1.1 | 2.4 |
| 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 10.4 | 5.9 | 0.0 |
| 93781_HUVEC (Endothelial)_IL-11 | 0.0 | 0.0 | 0.6 |
| 93583_Lung Microvascular Endothelial Cells_none | 0.0 | 0.0 | 0.0 |
| 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 | 0.0 |
| 92662_Microvascular Dermal endothelium_none | 0.0 | 0.0 | 0.0 |
| 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 | 0.0 |
| 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 0.0 | 0.0 | 0.0 |
| 93347_Small Airway Epithelium_none | 0.0 | 0.0 | 0.0 |
| 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 | 0.0 |
| 92668_Coronery Artery SMC_resting | 0.0 | 0.3 | 0.0 |
| 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.4 | 0.0 |
| 93107_astrocytes_resting | 0.0 | 0.0 | 0.0 |
| 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 | 0.0 |
| 92666_KU-812 (Basophil)_resting | 0.0 | 0.8 | 0.0 |
| 92667_KU-812 (Basophil)_PMA/ionoycin | 4.2 | 5.8 | 0.9 |
| 93579_CCD1106 (Keratinocytes)_none | 0.0 | 0.0 | 0.6 |
| 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 | 6.5 | 0.0 |
| 93791_Liver Cirrhosis | 34.9 | 3.5 | 19.1 |
| 93792_Lupus Kidney | 0.0 | 0.0 | 1.6 |
| 93577_NCI-H292 | 0.0 | 0.0 | 0.0 |
| 93358_NCI-H292_IL-4 | 0.0 | 0.0 | 0.0 |
| 93360_NCI-H292_IL-9 | 0.0 | 0.2 | 1.2 |
| 93359_NCI-H292_IL-13 | 0.0 | 0.0 | 0.0 |
| 93357_NCI-H292_IFN gamma | 0.0 | 0.4 | 0.7 |
| 93777_HPAEC_- | 0.0 | 0.0 | 0.0 |
| 93778_HPAEC_IL-1 beta/TNA alpha | 0.0 | 0.6 | 0.0 |
| 93254_Normal Human Lung Fibroblast_none | 0.0 | 4.7 | 2.6 |

| | | | |
|---|---|---|---|
| 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 6.6 | 2.9 | 6.2 |
| 93257_Normal Human Lung Fibroblast_IL-4 | 9.3 | 27.2 | 4.1 |
| 93256_Normal Human Lung Fibroblast_IL-9 | 6.3 | 14.7 | 3.1 |
| 93255_Normal Human Lung Fibroblast_IL-13 | 12.2 | 20.7 | 1.1 |
| 93258_Normal Human Lung Fibroblast_IFN gamma | 52.1 | 13.7 | 17.0 |
| 93106_Dermal Fibroblasts CCD1070_resting | 23.2 | 14.7 | 13.6 |
| 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 36.3 | 39.8 | 4.6 |
| 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 11.5 | 3.7 | 1.6 |
| 93772_dermal fibroblast_IFN gamma | 95.3 | 54.0 | 21.5 |
| 93771_dermal fibroblast_IL-4 | 100.0 | 100.0 | 26.5 |
| 93259_IBD Colitis 1** | 87.7 | 45.7 | 100.0 |
| 93260_IBD Colitis 2 | 0.0 | 5.1 | 4.5 |
| 93261_IBD Crohns | 0.0 | 0.6 | 2.8 |
| 735010_Colon_normal | 0.0 | 3.4 | 1.7 |
| 735019_Lung_none | 5.6 | 1.2 | 0.0 |
| 64028-1_Thymus_none | 5.2 | 12.1 | 1.7 |
| 64030-1_Kidney_none | 0.0 | 1.3 | 0.0 |

**Table 16H. Panel 4R Summary**

| Tissue Name | Relative Expression(%) | |
|---|---|---|
| | 4Rtm2178f_ag1261 | 4Rtm2261t_ag1261c |
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 16.7 | 5.7 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 0.0 | 4.7 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 8.7 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 0.0 | 2.8 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93565_primary Th1_resting dy 4-6 in IL-2 | 0.0 | 4.4 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 4.3 | 0.0 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 5.1 | 0.0 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 5.2 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 25.2 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 7.0 | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 31.4 | 0.0 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 0.0 | 0.0 |
| 93354_CD4_none | 18.2 | 0.0 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 11.7 | 0.0 |

| | | |
|---|---|---|
| 93103_LAK cells_resting | 9.4 | 0.0 |
| 93788_LAK cells_IL-2 | 5.0 | 12.2 |
| 93787_LAK cells_IL-2+IL-12 | 6.0 | 3.7 |
| 93789_LAK cells_IL-2+IFN gamma | 0.0 | 14.2 |
| 93790_LAK cells_IL-2+ IL-18 | 27.0 | 8.0 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 0.0 |
| 93578_NK Cells IL-2_resting | 5.2 | 4.5 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 5.2 | 0.0 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 3.0 | 0.0 |
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 0.0 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 3.6 | 0.0 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 68.3 | 71.7 |
| 93249_Ramos (B cell)_none | 0.0 | 0.0 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 0.0 |
| 93349_B lymphocytes_PWM | 2.9 | 0.0 |
| 93350_B lymphoytes_CD40L and IL-4 | 0.0 | 0.0 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 0.0 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 0.0 | 0.0 |
| 93356_Dendritic Cells_none | 0.0 | 4.8 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 0.0 |
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 5.5 |
| 93774_Monocytes_resting | 0.0 | 4.6 |
| 93776_Monocytes_LPS 50 ng/ml | 0.0 | 0.0 |
| 93581_Macrophages_resting | 13.2 | 0.0 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 0.0 |
| 93098_HUVEC (Endothelial)_none | 8.0 | 5.0 |
| 93099_HUVEC (Endothelial)_starved | 49.7 | 19.8 |
| 93100_HUVEC (Endothelial)_IL-1b | 8.2 | 0.0 |
| 93779_HUVEC (Endothelial)_IFN gamma | 0.0 | 0.0 |
| 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 6.5 | 4.0 |
| 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 0.0 | 5.6 |
| 93781_HUVEC (Endothelial)_IL-11 | 0.0 | 0.0 |
| 93583_Lung Microvascular Endothelial Cells_none | 0.0 | 0.0 |
| 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92662_Microvascular Dermal endothelium_none | 0.0 | 0.0 |
| 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 0.0 | 0.0 |
| 93347_Small Airway Epithelium_none | 0.0 | 0.0 |
| 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 12.7 |
| 92668_Coronery Artery SMC_resting | 0.0 | 0.0 |

| | | |
|---|---|---|
| 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93107_astrocytes_resting | 0.0 | 0.0 |
| 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 3.1 |
| 92666_KU-812 (Basophil)_resting | 0.0 | 0.0 |
| 92667_KU-812 (Basophil)_PMA/ionoycin | 15.0 | 4.0 |
| 93579_CCD1106 (Keratinocytes)_none | 0.0 | 0.0 |
| 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 | 0.0 |
| 93791_Liver Cirrhosis | 52.8 | 39.5 |
| 93792_Lupus Kidney | 0.0 | 0.0 |
| 93577_NCI-H292 | 0.0 | 3.6 |
| 93358_NCI-H292_IL-4 | 0.0 | 0.0 |
| 93360_NCI-H292_IL-9 | 0.0 | 0.0 |
| 93359_NCI-H292_IL-13 | 0.0 | 0.0 |
| 93357_NCI-H292_IFN gamma | 0.0 | 0.0 |
| 93777_HPAEC_- | 36.1 | 0.0 |
| 93778_HPAEC_IL-1 beta/TNA alpha | 6.7 | 0.0 |
| 93254_Normal Human Lung Fibroblast_none | 24.1 | 5.9 |
| 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 0.0 | 7.2 |
| 93257_Normal Human Lung Fibroblast_IL-4 | 30.1 | 0.0 |
| 93256_Normal Human Lung Fibroblast_IL-9 | 47.3 | 11.8 |
| 93255_Normal Human Lung Fibroblast_IL-13 | 0.0 | 0.0 |
| 93258_Normal Human Lung Fibroblast_IFN gamma | 40.3 | 32.5 |
| 93106_Dermal Fibroblasts CCD1070_resting | 24.1 | 12.5 |
| 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 54.0 | 19.2 |
| 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 7.2 | 5.1 |
| 93772_dermal fibroblast_IFN gamma | 100.0 | 100.0 |
| 93771_dermal fibroblast_IL-4 | 42.0 | 54.0 |
| 93259_IBD Colitis 1** | 44.8 | 90.8 |
| 93260_IBD Colitis 2 | 0.0 | 0.0 |
| 93261_IBD Crohns | 0.0 | 0.0 |
| 735010_Colon_normal | 0.0 | 0.0 |
| 735019_Lung_none | 4.7 | 11.1 |
| 64028-1_Thymus_none | 27.0 | 0.0 |
| 64030-1_Kidney_none | 5.4 | 3.9 |

**Table 16I. Panel 4.1D**

| Tissue Name | Relative Expression(%) 4.1dx4tm6164f_ag1261_a1 | Tissue Name | Relative Expression(%) 4.1dx4tm6164f_ag1261_a1 |
|---|---|---|---|
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 93100_HUVEC (Endothelial)_IL-1b | 3.6 |

| | | | |
|---|---|---|---|
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 1.0 | 93779_HUVEC (Endothelial)_IFN gamma | 0.0 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 0.0 | 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 0.0 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 0.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 93781_HUVEC (Endothelial)_IL-11 | 0.0 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 0.0 | 93583_Lung Microvascular Endothelial Cells_none | 0.0 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 0.0 | 92662_Microvascular Dermal endothelium_none | 0.0 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93565_primary Th1_resting dy 4-6 in IL-2 | 0.0 | 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 0.0 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 0.0 | 93347_Small Airway Epithelium_none | 0.0 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 0.0 | 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 92668_Coronery Artery SMC_resting | 0.0 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 1.4 | 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 0.7 | 93107_astrocytes_resting | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 0.7 | 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 0.0 | 92666_KU-812 (Basophil)_resting | 0.0 |
| 93354_CD4_none | 0.9 | 92667_KU-812 (Basophil)_PMA/ionoycin | 0.0 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 0.0 | 93579_CCD1106 (Keratinocytes)_none | 0.0 |
| 93103_LAK cells_resting | 0.0 | 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 |
| 93788_LAK cells_IL-2 | 0.0 | 93791_Liver Cirrhosis | 0.0 |
| 93787_LAK cells_IL-2+IL-12 | 0.0 | 93577_NCI-H292 | 1.0 |
| 93789_LAK cells_IL-2+IFN | 0.8 | 93358_NCI-H292_IL-4 | 0.0 |

| | | | |
|---|---|---|---|
| gamma | | | |
| 93790_LAK cells_IL-2+ IL-18 | 0.0 | 93360_NCI-H292_IL-9 | 0.0 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.7 | 93359_NCI-H292_IL-13 | 0.0 |
| 93578_NK Cells IL-2_resting | 0.0 | 93357_NCI-H292_IFN gamma | 2.1 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 93777_HPAEC_- | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 0.7 | 93778_HPAEC_IL-1 beta/TNA alpha | 0.0 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 93254_Normal Human Lung Fibroblast_none | 1.6 |
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 0.0 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 0.0 | 93257_Normal Human Lung Fibroblast_IL-4 | 0.7 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 8.4 | 93256_Normal Human Lung Fibroblast_IL-9 | 5.4 |
| 93249_Ramos (B cell)_none | 0.0 | 93255_Normal Human Lung Fibroblast_IL-13 | 6.1 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 93258_Normal Human Lung Fibroblast_IFN gamma | 9.5 |
| 93349_B lymphocytes_PWM | 0.0 | 93106_Dermal Fibroblasts CCD1070_resting | 8.3 |
| 93350_B lymphoytes_CD40L and IL-4 | 0.0 | 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 6.0 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 3.1 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 4.7 | 93772_dermal fibroblast_IFN gamma | 14.2 |
| 93356_Dendritic Cells_none | 0.5 | 93771_dermal fibroblast_IL-4 | 13.3 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 93892_Dermal fibroblasts_none | 11.4 |
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 99202_Neutrophils_TNFa+LPS | 2.4 |
| 93774_Monocytes_resting | 0.0 | 99203_Neutrophils_none | 1.4 |
| 93776_Monocytes_LPS 50 ng/ml | 0.0 | 735010_Colon_normal | 2.9 |
| 93581_Macrophages_resting | 0.0 | 735019_Lung_none | 9.8 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 64028-1_Thymus_none | 24.1 |
| 93098_HUVEC (Endothelial)_none | 0.5 | 64030-1_Kidney_none | 100.0 |
| 93099_HUVEC (Endothelial)_starved | 0.7 | | |

**Panel 1.2 Summary**: Ag1261 The results from replicate experiments using the same primer/probe set are in concordance with each other. Expression of the GPCR5b gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, there appears to be specific expression in cultured cell lines derived from several types of cancers. No expression of the GPCR5b gene is detected in any normal tissues. However, there is moderate expression in ovarian, lung and kidney cancer cell lines indicating that this gene may play a role in the pathogenesis of these diseases. Thus, therapeutic modulation of this gene may be of utility in the treatment of kidney, ovarian and lung cancer.

**Panel 1.3D Summary**: Ag1261b Expression of the GPCR5b gene on this panel is limited to spleen (CT = 34.6). This may suggest that the GPCR5b gene is involved in immune function or a function of the spleen not related to its main immune function. Thus, therapeutic modulation of the expression of this gene might be of use in the treatment of immune disorders or other disorders related to the spleen. The GPCR5b gene product may be used to distinguish spleen from other tissues.

**Panel 2D Summary**: Ag1261 Significant variability in GPCR5b gene expression levels is detected in multiple experiments, most likely due to low level of expression detected. This gene appears to be expressed at a very low level in panel 2D. There is, however, detectable expression in samples of ovarian, breast, kidney and colon cancers when compared to their normal controls. Thus, therapeutic modulation of expression of this gene might be of use in the treatment of these cancer types.

**Panel 2.2 Summary**: Ag1261b Expression of the GPCR5b gene is low to undetectable (Ct values >35) in all of the samples on this panel.

**Panels 4D/4.1D/4R Summary**: Ag1261/1261b/1261c The expression profile of the GPCR5b transcript was examined eight different times using three different probe/primer sets. This transcript is expressed in fibroblasts and in colitis 1. The level of expression is low in some of the runs, but the pattern is consistent. The GPCR5b gene encodes a GPCR that is highly expressed in gamma interferon and IL-4 treated dermal fibroblasts and is also induced in a dermal fibroblast cell line treated with these cytokines. It is also induced in gamma interferon treated lung fibroblasts. This profile indicates that the expression of this GPCR may be up regulated in during asthma, emphysema and allergy in which high levels of IL-4 are present or during psoriasis, colitis or viral infections when high levels of gamma interferon are present. Antibody or small molecule therapeutics that block the function of the GPCR encoded by the GPCR5b gene could therefore reduce or inhibit the inflammation and tissue

remodeling due to inflammation. *See* Bisping G., Lugering N., Lutke-Brintrup S., Pauels H.G., Schurmann G., Domschke W., Kucharzik T. (2001) Patients with inflammatory bowel disease (IBD) reveal increased induction capacity of intracellular interferon-gamma (IFN-gamma) in peripheral CD8+ lymphocytes co-cultured with intestinal epithelial cells. *Clin. Exp. Immunol.* 123: 15-22.

Intestinal epithelial cells seem to play a key role during IBD. The network of cellular interactions between epithelial cells and lamina propria mononuclear cells is still incompletely understood. In the following co-culture model we investigated the influence of intestinal epithelial cells on cytokine expression of T cytotoxic and T helper cells from patients with IBD and healthy controls. Peripheral blood mononuclear cells (PBMC) were purified by a Ficoll-Hypaque gradient followed by co-incubation with epithelial cells in multiwell cell culture insert plates in direct contact as well as separated by transwell filters. We used Caco-2 cells as well as freshly isolated colonic epithelia obtained from surgical specimens. Three-colour immunofluorescence flow cytometry was performed after collection, stimulation and staining of PBMC with anti-CD4, anti-CD8, anti-IFN-gamma and anti-IL-4. Patients with IBD (Crohn's disease (CD), n = 12; ulcerative colitis (UC), n = 16) and healthy controls (n = 10) were included in the study. After 24 h of co-incubation with Caco-2 cells we found a significant increase of IFN-gamma-producing CD8+ lymphocytes in patients with IBD. In contrast, healthy controls did not respond to the epithelial stimulus. No significant differences could be found between CD and UC or active and inactive disease. A significant increase of IFN-gamma+/CD8+ lymphocytes in patients with UC was also seen after direct co-incubation with primary cultures of colonic crypt cells. The observed epithelial-lymphocyte interaction seems to be MHC I-restricted. No significant epithelial cell-mediated effects on cytokine expression were detected in the PBMC CD4+ subsets. Patients with IBD-even in an inactive state of disease-exert an increased capacity for IFN-gamma induction in CD8+ lymphocytes mediated by intestinal epithelial cells. This mechanism may be important during chronic intestinal inflammation, as in the case of altered mucosal barrier function epithelial cells may become targets for IFN-gamma-producing CD8+ lymphocytes. See, *e.g.*, PMID: 11167992.

**Example F. GPCR6b (also known as ba64p14-F or CG57034-01):**

Expression of gene GPCR6b was assessed using the primer-probe sets Ag1262 (identical sequence as Ag2026) and Ag2370, described in Tables 17A and 17B. Results of the RTQ-PCR runs are shown in Table 17C and 17D.

**Table 17A. Probe Name Ag1262/Ag2026**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-GGACATGCGGACTAAGTACAAA-3' | 59.2 | 22 | 253 | 93 |
| Probe | TET-5'-CGATCCTCTATGAGGAATGCATTTCTCA-3'-TAMRA | 68.2 | 28 | 276 | 94 |
| Reverse | 5'-GAAGCTGTCCAGGTCAGTAAAA-3' | 58.5 | 22 | 322 | 95 |

**Table 17B. Probe Name Ag2370**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-AGGCTGTAGATAAAGGGGTTCA-3' | 59.2 | 22 | 76 | 96 |
| Probe | TET-5'-TGAGAGCCACAATGACATCCTTGTCA-3'-TAMRA | 69.1 | 26 | 122 | 97 |
| Reverse | 5'-TTCCCGACTGTAAGCAGTTCTA-3' | 59 | 22 | 149 | 98 |

**Table 17C. Panel 1.2**

| Tissue Name | Relative Expression(%) 1.2tm1423t_ag1262 | Tissue Name | Relative Expression(%) 1.2tm1423t_ag1262 |
|---|---|---|---|
| Endothelial cells | 5.6 | Renal ca. 786-0 | 0 |
| Fetal heart | 0 | Renal ca. A498 | 5.6 |
| Pancreas | 1.6 | Renal ca. RXF 393 | 1.2 |
| Pancreatic ca. CAPAN 2 | 6.5 | Renal ca. ACHN | 0 |
| Adrenal Gland (new lot*) | 0 | Renal ca. UO-31 | 2.3 |
| Thyroid | 0 | Renal ca. TK-10 | 68.8 |
| Salivary gland | 2.5 | Liver | 1.9 |
| Pituitary gland | 3.1 | Liver (fetal) | 0.9 |
| Brain (fetal) | 0 | Liver ca. (hepatoblast) HepG2 | 0.6 |
| Brain (whole) | 5.9 | Lung | 0.2 |
| Brain (amygdala) | 2.2 | Lung (fetal) | 0 |
| Brain (cerebellum) | 0.9 | Lung ca. (small cell) LX-1 | 100 |
| Brain (hippocampus) | 7.3 | Lung ca. (small cell) NCI-H69 | 17.2 |
| Brain (thalamus) | 0.5 | Lung ca. (s.cell var.) SHP-77 | 1.8 |
| Cerebral Cortex | 2.1 | Lung ca. (large cell)NCI-H460 | 8.5 |
| Spinal cord | 2.7 | Lung ca. (non-sm. cell) A549 | 41.2 |
| CNS ca. (glio/astro) U87-MG | 18.8 | Lung ca. (non-s.cell) NCI-H23 | 21.2 |
| CNS ca. (glio/astro) U-118-MG | 0 | Lung ca (non-s.cell) HOP-62 | 1.5 |
| CNS ca. (astro) SW1783 | 0.5 | Lung ca. (non-s.cl) NCI-H522 | 0 |
| CNS ca.* (neuro; met ) SK-N-AS | 1.4 | Lung ca. (squam.) SW 900 | 0 |
| CNS ca. (astro) SF-539 | 0 | Lung ca. (squam.) NCI-H596 | 2.8 |

| | | | |
|---|---|---|---|
| CNS ca. (astro) SNB-75 | 0.3 | Mammary gland | 3.2 |
| CNS ca. (glio) SNB-19 | 4 | Breast ca.* (pl. effusion) MCF-7 | 1.5 |
| CNS ca. (glio) U251 | 1.5 | Breast ca.* (pl.ef) MDA-MB-231 | 0 |
| CNS ca. (glio) SF-295 | 0.2 | Breast ca.* (pl. effusion) T47D | 8.7 |
| Heart | 0.3 | Breast ca. BT-549 | 2.5 |
| Skeletal Muscle (new lot*) | 0.8 | Breast ca. MDA-N | 1.6 |
| Bone marrow | 0.8 | Ovary | 0.5 |
| Thymus | 0.9 | Ovarian ca. OVCAR-3 | 1.4 |
| Spleen | 0.7 | Ovarian ca. OVCAR-4 | 3 |
| Lymph node | 3.9 | Ovarian ca. OVCAR-5 | 34.6 |
| Colorectal | 1.7 | Ovarian ca. OVCAR-8 | 22.4 |
| Stomach | 15.5 | Ovarian ca. IGROV-1 | 11.5 |
| Small intestine | 1.5 | Ovarian ca.* (ascites) SK-OV-3 | 11.7 |
| Colon ca. SW480 | 1.3 | Uterus | 0 |
| Colon ca.* (SW480 met)SW620 | 7.7 | Placenta | 3.4 |
| Colon ca. HT29 | 0.7 | Prostate | 3.1 |
| Colon ca. HCT-116 | 0.2 | Prostate ca.* (bone met)PC-3 | 1.8 |
| Colon ca. CaCo-2 | 0.4 | Testis | 20.4 |
| 83219 CC Well to Mod Diff (ODO3866) | 5.6 | Melanoma Hs688(A).T | 15.6 |
| Colon ca. HCC-2998 | 0 | Melanoma* (met) Hs688(B).T | 15 |
| Gastric ca.* (liver met) NCI-N87 | 0.9 | Melanoma UACC-62 | 1.9 |
| Bladder | 13.8 | Melanoma M14 | 7.6 |
| Trachea | 0.9 | Melanoma LOX IMVI | 0 |
| Kidney | 1.8 | Melanoma* (met) SK-MEL-5 | 0.2 |
| Kidney (fetal) | 0.4 | Adipose | 90.1 |

**Table 17D. Panel 4D**

| Tissue Name | Relative Expression(%) 4dx4tm4241f_ag2026_b2 | Relative Expression(%) 4Dx4tm4622t_ag2370_a1 |
|---|---|---|
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 0.0 | 28.0 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 0.0 | 5.3 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 0.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 5.2 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 0.0 | 5.9 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 0.0 |

| | | |
|---|---|---|
| 93565_primary Th1_resting dy 4-6 in IL-2 | 0.0 | 11.9 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 0.0 | 12.9 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 0.0 | 4.6 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 9.2 | 6.3 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 0.0 | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 7.1 | 0.0 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 4.8 | 0.0 |
| 93354_CD4_none | 18.4 | 0.0 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 0.0 | 0.0 |
| 93103_LAK cells_resting | 4.2 | 5.9 |
| 93788_LAK cells_IL-2 | 0.0 | 5.6 |
| 93787_LAK cells_IL-2+IL-12 | 3.5 | 9.6 |
| 93789_LAK cells_IL-2+IFN gamma | 20.3 | 0.0 |
| 93790_LAK cells_IL-2+ IL-18 | 24.1 | 0.0 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 0.0 |
| 93578_NK Cells IL-2_resting | 7.6 | 3.8 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 5.3 | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 0.0 |
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 0.0 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 1.7 | 0.0 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 7.7 | 40.7 |
| 93249_Ramos (B cell)_none | 0.0 | 0.0 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 0.0 |
| 93349_B lymphocytes_PWM | 0.0 | 0.0 |
| 93350_B lymphoytes_CD40L and IL-4 | 0.0 | 0.0 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 0.0 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 0.0 | 2.1 |
| 93356_Dendritic Cells_none | 0.0 | 0.0 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 5.9 |
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 0.0 |
| 93774_Monocytes_resting | 0.0 | 0.0 |
| 93776_Monocytes_LPS 50 ng/ml | 0.0 | 0.0 |
| 93581_Macrophages_resting | 4.3 | 12.7 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 0.0 |
| 93098_HUVEC (Endothelial)_none | 24.5 | 0.0 |
| 93099_HUVEC (Endothelial)_starved | 31.6 | 19.7 |
| 93100_HUVEC (Endothelial)_IL-1b | 9.2 | 5.8 |
| 93779_HUVEC (Endothelial)_IFN gamma | 0.0 | 0.0 |
| 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 5.8 | 0.0 |
| 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 0.0 | 6.0 |
| 93781_HUVEC (Endothelial)_IL-11 | 4.1 | 6.0 |

| | | |
|---|---|---|
| 93583_Lung Microvascular Endothelial Cells_none | 0.0 | 0.0 |
| 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92662_Microvascular Dermal endothelium_none | 0.0 | 0.0 |
| 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 0.0 | 6.6 |
| 93347_Small Airway Epithelium_none | 0.0 | 0.0 |
| 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92668_Coronery Artery SMC_resting | 0.0 | 6.4 |
| 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 93107_astrocytes_resting | 0.0 | 0.0 |
| 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 | 0.0 |
| 92666_KU-812 (Basophil)_resting | 0.0 | 0.0 |
| 92667_KU-812 (Basophil)_PMA/ionoycin | 0.0 | 0.0 |
| 93579_CCD1106 (Keratinocytes)_none | 3.9 | 0.0 |
| 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 | 0.0 |
| 93791_Liver Cirrhosis | 21.1 | 6.1 |
| 93792_Lupus Kidney | 0.0 | 0.0 |
| 93577_NCI-H292 | 0.0 | 4.3 |
| 93358_NCI-H292_IL-4 | 0.0 | 0.0 |
| 93360_NCI-H292_IL-9 | 0.0 | 0.0 |
| 93359_NCI-H292_IL-13 | 4.7 | 0.0 |
| 93357_NCI-H292_IFN gamma | 0.0 | 0.0 |
| 93777_HPAEC_- | 0.0 | 0.0 |
| 93778_HPAEC_IL-1 beta/TNA alpha | 0.0 | 0.0 |
| 93254_Normal Human Lung Fibroblast_none | 4.2 | 6.0 |
| 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 3.4 | 0.0 |
| 93257_Normal Human Lung Fibroblast_IL-4 | 32.0 | 4.3 |
| 93256_Normal Human Lung Fibroblast_IL-9 | 12.3 | 7.1 |
| 93255_Normal Human Lung Fibroblast_IL-13 | 26.0 | 28.1 |
| 93258_Normal Human Lung Fibroblast_IFN gamma | 40.2 | 64.2 |
| 93106_Dermal Fibroblasts CCD1070_resting | 5.1 | 23.4 |
| 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 29.5 | 6.4 |
| 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 18.1 | 6.2 |
| 93772_dermal fibroblast_IFN gamma | 65.4 | 100.0 |
| 93771_dermal fibroblast_IL-4 | 100.0 | 67.4 |
| 93259_IBD Colitis 1** | 2.8 | 0.0 |
| 93260_IBD Colitis 2 | 0.0 | 0.0 |
| 93261_IBD Crohns | 4.8 | 0.0 |
| 735010_Colon_normal | 0.0 | 0.0 |

| | | |
|---|---|---|
| 735019_Lung_none | 0.0 | 6.0 |
| 64028-1_Thymus_none | 0.0 | 0.0 |
| 64030-1_Kidney_none | 0.0 | 0.0 |

**Panel 1.2 Summary**: Ag1262 Expression of the GPCR6b gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, highest expression is detected in a lung cancer cell line. Interestingly, there appears to be high expression in 5/10 lung cancer cell lines and 4/6 ovarian cancer cell lines, indicating that the expression of this gene maybe associated with the pathology of these diseases. In contrast, there is little expression of the GPCR6b gene in the samples of normal lung and ovary. Thus, the therapeutic modulation of the GPCR6b gene may show utility in the treatment of lung and ovarian cancer. Alternatively, this gene may be a useful marker in the diagnosis of lung and ovarian cancers. Low expression of the GPCR6b gene is also detected in the following normal tissues: hippocampus, stomach, bladder, and testis.

**Panel 1.3D Summary**: Ag2026/Ag2370 Expression of the GPCR6b gene is low to undetectable (Ct values >35) in all of the samples on this panel.

**Panel 2.2 Summary**: Ag2370 Expression of the GPCR6b gene is low to undetectable (Ct values >35) in all of the samples on this panel.

**Panel 4D Summary**: Ag1262 Expression of the GPCR6b gene is low to undetectable (Ct values >35) in all of the samples on this panel. Ag2370/Ag2026 The GPCR6b transcript is expressed in IL-4 and gamma interferon treated fibroblasts. Induction of the transcript by IL-4 or by gamma interferon is also seen in the CCD1070 cell line and in lung fibroblasts. The induction of the transcript in lung fibroblasts by IL-4 is not detectable using the Ag2370 probe and primer set. The GPCR encoded for by this transcript may therefore be important in fibroblast responses to gamma interferon or IL-4. Blocking the function of this GPCR with antibody or small molecule therapeutics may reduce or inhibit inflammation in these tissues and be important in the treatment of psoriasis, allergy, asthma, and emphysema.

**Example G. GPCR7b (also known as ba64p14-G or CG57809-01):**

Expression of gene GPCR7b was assessed using the primer-probe set Ag1263, described in Table 18A. Results of the RTQ-PCR runs are shown in Tables 18B and 18C.

**Table 18A. Probe Name Ag1263**

| Primers | Sequences | TM | Length | Start Positio | SEQ ID NO |
|---|---|---|---|---|---|

| | | | | n | |
|---|---|---|---|---|---|
| Forward | 5'-ACTCTCCTCTCGGTCTGAAGAC-3' | 59.1 | 22 | 96 | 99 |
| Probe | FAM-5'-AGAGGCCACTCTTTGCCCTCTTTCTT-3'-TAMRA | 68.4 | 26 | 119 | 100 |
| Reverse | 5'-TAGCCAAGATGATGAGCAGATT-3' | 59 | 22 | 172 | 101 |

**Table 18B. Panel 1.2**

| Tissue Name | Relative Expression(%) 1.2tm1457f_ag1263 | Tissue Name | Relative Expression(%) 1.2tm1457f_ag1263 |
|---|---|---|---|
| Endothelial cells | 0.5 | Renal ca. 786-0 | 0.0 |
| Fetal Heart | 0.1 | Renal ca. A498 | 0.4 |
| Pancreas | 0.5 | Renal ca. RXF 393 | 0.2 |
| Pancreatic ca. CAPAN 2 | 0.0 | Renal ca. ACHN | 0.1 |
| Adrenal Gland (new lot*) | 0.5 | Renal ca. UO-31 | 0.8 |
| Thyroid | 0.1 | Renal ca. TK-10 | 0.4 |
| Salivary gland | 3.2 | Liver | 1.4 |
| Pituitary gland | 0.5 | Liver (fetal) | 0.2 |
| Brain (fetal) | 0.7 | Liver ca. (hepatoblast) HepG2 | 0.2 |
| Brain (whole) | 11.0 | Lung | 1.1 |
| Brain (amygdala) | 1.2 | Lung (fetal) | 0.0 |
| Brain (cerebellum) | 1.2 | Lung ca. (small cell) LX-1 | 0.1 |
| Brain (hippocampus) | 10.6 | Lung ca. (small cell) NCI-H69 | 8.6 |
| Brain (thalamus) | 0.9 | Lung ca. (s.cell var.) SHP-77 | 0.4 |
| Cerebral Cortex | 1.1 | Lung ca. (large cell)NCI-H460 | 3.8 |
| Spinal cord | 0.2 | Lung ca. (non-sm. cell) A549 | 1.9 |
| CNS ca. (glio/astro) U87-MG | 0.3 | Lung ca. (non-s.cell) NCI-H23 | 0.7 |
| CNS ca. (glio/astro) U-118-MG | 0.2 | Lung ca (non-s.cell) HOP-62 | 0.9 |
| CNS ca. (astro) SW1783 | 0.5 | Lung ca. (non-s.cl) NCI-H522 | 0.1 |
| CNS ca.* (neuro; met ) SK-N-AS | 0.0 | Lung ca. (squam.) SW 900 | 0.2 |
| CNS ca. (astro) SF-539 | 0.1 | Lung ca. (squam.) NCI-H596 | 1.2 |
| CNS ca. (astro) SNB-75 | 0.0 | Mammary gland | 0.2 |
| CNS ca. (glio) SNB-19 | 1.7 | Breast ca.* (pl. effusion) MCF-7 | 0.0 |
| CNS ca. (glio) U251 | 1.0 | Breast ca.* (pl.ef) MDA-MB-231 | 0.0 |
| CNS ca. (glio) SF-295 | 0.1 | Breast ca.* (pl. effusion) T47D | 2.0 |
| Heart | 1.1 | Breast ca. BT-549 | 1.4 |
| Skeletal Muscle (new lot*) | 0.5 | Breast ca. MDA-N | 0.3 |
| Bone marrow | 0.2 | Ovary | 0.0 |
| Thymus | 0.1 | Ovarian ca. OVCAR-3 | 1.0 |
| Spleen | 0.2 | Ovarian ca. OVCAR-4 | 0.4 |

| | | | |
|---|---|---|---|
| Lymph node | 0.5 | Ovarian ca. OVCAR-5 | 2.8 |
| Colorectal | 0.2 | Ovarian ca. OVCAR-8 | 1.0 |
| Stomach | 3.0 | Ovarian ca. IGROV-1 | 0.3 |
| Small intestine | 0.7 | Ovarian ca.* (ascites) SK-OV-3 | 1.0 |
| Colon ca. SW480 | 0.0 | Uterus | 0.3 |
| Colon ca.* (SW480 met)SW620 | 0.0 | Placenta | 0.5 |
| Colon ca. HT29 | 0.0 | Prostate | 1.1 |
| Colon ca. HCT-116 | 0.0 | Prostate ca.* (bone met)PC-3 | 0.0 |
| Colon ca. CaCo-2 | 0.0 | Testis | 2.7 |
| 83219 CC Well to Mod Diff (ODO3866) | 4.2 | Melanoma Hs688(A).T | 0.2 |
| Colon ca. HCC-2998 | 0.1 | Melanoma* (met) Hs688(B).T | 1.4 |
| Gastric ca.* (liver met) NCI-N87 | 0.0 | Melanoma UACC-62 | 0.0 |
| Bladder | 2.5 | Melanoma M14 | 2.9 |
| Trachea | 1.2 | Melanoma LOX IMVI | 0.0 |
| Kidney | 1.2 | Melanoma* (met) SK-MEL-5 | 0.0 |
| Kidney (fetal) | 2.2 | Adipose | 100.0 |

**Table 18C. Panel 4D**

| Tissue Name | Relative Expression (%) 4Dtm2165f_ag1263 | Tissue Name | Relative Expression (%) 4Dtm2165f_ag1263 |
|---|---|---|---|
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 0.0 | 93100_HUVEC (Endothelial)_IL-1b | 0.0 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 2.9 | 93779_HUVEC (Endothelial)_IFN gamma | 12.4 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 5.6 | 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 0.0 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 2.5 | 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 0.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 0.0 | 93781_HUVEC (Endothelial)_IL-11 | 2.9 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 3.3 | 93583_Lung Microvascular Endothelial Cells_none | 6.9 |
| 93568_primary Th1_anti-CD28/anti-CD3 | 0.0 | 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 0.0 | 92662_Microvascular Dermal endothelium_none | 3.2 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 2.2 | 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 2.3 |

| | | | |
|---|---|---|---|
| 93565_primary Th1_resting dy 4-6 in IL-2 | 4.4 | 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 2.6 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 3.5 | 93347_Small Airway Epithelium_none | 0.0 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 6.8 | 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 3.0 | 92668_Coronery Artery SMC_resting | 0.0 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 0.0 | 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 5.9 | 93107_astrocytes_resting | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 2.4 | 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 3.1 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 9.0 | 92666_KU-812 (Basophil)_resting | 0.0 |
| 93354_CD4_none | 0.0 | 92667_KU-812 (Basophil)_PMA/ionoycin | 0.0 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 0.0 | 93579_CCD1106 (Keratinocytes)_none | 0.0 |
| 93103_LAK cells_resting | 0.0 | 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 0.0 |
| 93788_LAK cells_IL-2 | 0.0 | 93791_Liver Cirrhosis | 11.4 |
| 93787_LAK cells_IL-2+IL-12 | 3.3 | 93792_Lupus Kidney | 3.1 |
| 93789_LAK cells_IL-2+IFN gamma | 0.0 | 93577_NCI-H292 | 2.7 |
| 93790_LAK cells_IL-2+ IL-18 | 1.4 | 93358_NCI-H292_IL-4 | 0.0 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 93360_NCI-H292_IL-9 | 0.0 |
| 93578_NK Cells IL-2_resting | 1.7 | 93359_NCI-H292_IL-13 | 0.0 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 93357_NCI-H292_IFN gamma | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 93777_HPAEC_- | 8.4 |
| 93111_Mixed Lymphocyte Reaction_Two Way MLR | 39.2 | 93778_HPAEC_IL-1 beta/TNA alpha | 5.7 |
| 93112_Mononuclear Cells (PBMCs)_resting | 2.6 | 93254_Normal Human Lung Fibroblast_none | 2.2 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 0.0 | 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 0.0 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 4.5 | 93257_Normal Human Lung Fibroblast_IL-4 | 3.1 |

| | | | |
|---|---|---|---|
| 93249_Ramos (B cell)_none | 0.0 | 93256_Normal Human Lung Fibroblast_IL-9 | 0.0 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 93255_Normal Human Lung Fibroblast_IL-13 | 0.0 |
| 93349_B lymphocytes_PWM | 0.0 | 93258_Normal Human Lung Fibroblast_IFN gamma | 0.0 |
| 93350_B lymphoytes_CD40L and IL-4 | 3.1 | 93106_Dermal Fibroblasts CCD1070_resting | 1.9 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 4.7 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin | 0.0 | 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 0.0 |
| 93356_Dendritic Cells_none | 0.0 | 93772_dermal fibroblast_IFN gamma | 8.3 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 93771_dermal fibroblast_IL-4 | 0.0 |
| 93775_Dendritic Cells_anti-CD40 | 0.0 | 93259_IBD Colitis 1** | 100.0 |
| 93774_Monocytes_resting | 0.0 | 93260_IBD Colitis 2 | 0.0 |
| 93776_Monocytes_LPS 50 ng/ml | 0.0 | 93261_IBD Crohns | 0.0 |
| 93581_Macrophages_resting | 7.9 | 735010_Colon_normal | 0.0 |
| 93582_Macrophages_LPS 100 ng/ml | 0.0 | 735019_Lung_none | 0.0 |
| 93098_HUVEC (Endothelial)_none | 2.4 | 64028-1_Thymus_none | 2.3 |
| 93099_HUVEC (Endothelial)_starved | 2.8 | 64030-1_Kidney_none | 0.0 |

**Panel 1.2 Summary**: Ag1263 Expression of the GPCR7b gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, low expression (CTs 32-35) of the GPCR7b gene is detected in brain tissues, salivary gland, stomach, bladder and testis as well as in a few cultured cell lines derived from lung cancer and ovarian cancer. In addition, its expression appears to be absent in cell lines derived from brain cancer. Thus, the therapeutic down-regulation of this gene using a monoclonal antibody or small molecule therapeutic may be of use in the treatment of lung or ovarian cancer, while increased expression of this gene may be of use in the treatment of brain cancer. The protein encoded by the GPCR7b gene is a GPCR that shows high expression levels in the brain, especially in the hippocampus. Several neurotransmitter receptors are GPCRs, including the dopamine receptor family, the serotonin receptor family, the GABAB receptor, muscarinic acetylcholine receptors, and others. The hippocampus is an area of the brain that is critical for long-term memory formation, and shows marked neurodegeneration in Alzheimer's disease

(and to a lesser extent, Parkinson's). The hippocampus has also been implicated in major psychiatric disorders such as schizophrenia and bipolar depression. Thus, the expression levels of the GPCR7b gene may be a marker for one or more of these diseases, and therapeutic manipulation of this gene or its protein product may have beneficial effects in memory enhancement, Alzheimer's disease, Parkinson's disease, schizophrenia, bipolar disorder, or depression.

**Panel 4D Summary**: Ag1263 Expression of the GPCR7b gene in this panel is skewed by genomic DNA contamination in the IBD colitis 1 sample. Disregarding this sample, the gene GPCR7b is expressed at very low levels in a mixed lymphocyte reaction. This observation suggests that the GPCR may be important in the events that occur in late stages of T cell activation. These may include clonal expansion, polarization, and expression of specific classes of adhesion molecules such as ligands for P or E-selectin. Antagonistic antibody or small molecule therapeutics may therefore block the function or recirculating capabilities of activated T cells and reduce the ability of T cells to initiate inflammation. These therapeutics could be important in the treatment of asthma, allergy, psoriasis, arthritis, GVHD and other diseases that involve the activation of T cells. Small molecule therapies that mimic the ligand interaction with the GPCR (agonists) may be important in initiating an immune response and increasing the efficacy of immunizations.

**Panel CNSD.01 Summary**: Ag1263 Expression of the GPCR7b gene is low to undetectable (CT values > 35) across all the samples in this panel. However the results obtained in Panel 1.2 suggests that the gene is mostly expressed in the hippocampus, which is not represented in this panel.

**Example H. GPCR8d (also known as ba64p14-H and CG50259-01):**

Expression of gene ba64p14-H was assessed using the primer-probe set Ag1264, described in Table 19A. Results of the RTQ-PCR run are shown in Table 19B.

**Table 19A. Probe Name Ag1264**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-TCCAATGTTATCCACCACTTTC-3' | 58.8 | 22 | 571 | 102 |
| Probe | TET-5'-CCCTGTGCTGAAATTGTCCTGCTCTT-3'-TAMRA | 69.4 | 26 | 606 | 103 |
| Reverse | 5'-TCTGCACAATTTCATTGACAAA-3' | 59.2 | 22 | 637 | 104 |

**Table 19B. Panel 1.2**

| Tissue Name | Relative Expression(%) 1.2tm1425t_ ag1264 | Tissue Name | Relative Expression(%) 1.2tm1425t_ ag1264 |
|---|---|---|---|
| Endothelial cells | 2.4 | Renal ca. 786-0 | 0.5 |
| Fetal heart | 0.3 | Renal ca. A498 | 3.4 |
| Pancreas | 0.7 | Renal ca. RXF 393 | 0.0 |
| Pancreatic ca. CAPAN 2 | 0.5 | Renal ca. ACHN | 0.0 |
| Adrenal Gland (new lot*) | 3.5 | Renal ca. UO-31 | 0.3 |
| Thyroid | 0.9 | Renal ca. TK-10 | 10.4 |
| Salivary gland | 8.5 | Liver | 2.0 |
| Pituitary gland | 3.6 | Liver (fetal) | 0.9 |
| Brain (fetal) | 1.3 | Liver ca. (hepatoblast) HepG2 | 0.0 |
| Brain (whole) | 6.1 | Lung | 1.7 |
| Brain (amygdala) | 5.7 | Lung (fetal) | 1.2 |
| Brain (cerebellum) | 1.9 | Lung ca. (small cell) LX-1 | 8.2 |
| Brain (hippocampus) | 9.3 | Lung ca. (small cell) NCI-H69 | 3.8 |
| Brain (thalamus) | 1.9 | Lung ca. (s.cell var.) SHP-77 | 0.2 |
| Cerebral Cortex | 13.8 | Lung ca. (large cell)NCI-H460 | 4.2 |
| Spinal cord | 4.4 | Lung ca. (non-sm. cell) A549 | 7.4 |
| CNS ca. (glio/astro) U87-MG | 4.0 | Lung ca. (non-s.cell) NCI-H23 | 3.0 |
| CNS ca. (glio/astro) U-118-MG | 0.5 | Lung ca (non-s.cell) HOP-62 | 1.6 |
| CNS ca. (astro) SW1783 | 1.1 | Lung ca. (non-s.cl) NCI-H522 | 0.5 |
| CNS ca.* (neuro; met ) SK-N-AS | 1.0 | Lung ca. (squam.) SW 900 | 0.2 |
| CNS ca. (astro) SF-539 | 0.1 | Lung ca. (squam.) NCI-H596 | 0.9 |
| CNS ca. (astro) SNB-75 | 0.0 | Mammary gland | 3.5 |
| CNS ca. (glio) SNB-19 | 3.0 | Breast ca.* (pl. effusion) MCF-7 | 0.4 |
| CNS ca. (glio) U251 | 2.7 | Breast ca.* (pl.ef) MDA-MB-231 | 0.2 |
| CNS ca. (glio) SF-295 | 0.4 | Breast ca.* (pl. effusion) T47D | 2.4 |
| Heart | 5.0 | Breast ca. BT-549 | 0.8 |
| Skeletal Muscle (new lot*) | 2.1 | Breast ca. MDA-N | 0.2 |
| Bone marrow | 0.4 | Ovary | 0.7 |
| Thymus | 1.2 | Ovarian ca. OVCAR-3 | 1.2 |
| Spleen | 2.3 | Ovarian ca. OVCAR-4 | 2.9 |
| Lymph node | 3.2 | Ovarian ca. OVCAR-5 | 10.0 |
| Colorectal | 2.5 | Ovarian ca. OVCAR-8 | 7.7 |
| Stomach | 6.5 | Ovarian ca. IGROV-1 | 3.9 |
| Small intestine | 2.2 | Ovarian ca.* (ascites) SK-OV-3 | 1.6 |
| Colon ca. SW480 | 0.0 | Uterus | 3.1 |
| Colon ca.* (SW480 met)SW620 | 0.0 | Placenta | 4.2 |

| | | | |
|---|---|---|---|
| Colon ca. HT29 | 0.0 | Prostate | 4.5 |
| Colon ca. HCT-116 | 0.2 | Prostate ca.* (bone met)PC-3 | 0.6 |
| Colon ca. CaCo-2 | 0.1 | Testis | 12.9 |
| 83219 CC Well to Mod Diff (ODO3866) | 7.0 | Melanoma Hs688(A).T | 3.2 |
| Colon ca. HCC-2998 | 0.2 | Melanoma* (met) Hs688(B).T | 2.7 |
| Gastric ca.* (liver met) NCI-N87 | 0.5 | Melanoma UACC-62 | 0.0 |
| Bladder | 5.7 | Melanoma M14 | 1.7 |
| Trachea | 2.3 | Melanoma LOX IMVI | 0.0 |
| Kidney | 3.2 | Melanoma* (met) SK-MEL-5 | 0.4 |
| Kidney (fetal) | 3.6 | Adipose | 100.0 |

**Panel 1.2 Summary**: Ag1264 Expression of the GPCR8d gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, low to moderate expression (CT values = 31-35) of the GPCR8d gene is detected in a wide variety of normal tissues, including endothelial cells, heart, skeletal muscle, liver, adrenal gland, salivary gland, pituitary gland, brain (amygdala, cerebellum, hippocampus, cerebral cortex and thalamus), spinal cord, spleen, lymph node, colon, small intestine, stomach, bladder, kidney (adult and fetal), trachea, mammary gland, uterus, placenta, prostate, and testis. This gene is expressed in the adrenal (CT value = 33.9) and pituitary (CT value = 33.8) glands. Therefore, the GPCR8d gene product may be involved in an adrenal-pituitary axis and may be of utility as a drug target for the treatment of diseases that are regulated by this neuroendocrine axis. The GPCR encoded by the GPCR8d gene also shows expression in several regions of the brain, with the highest expression (CT value = 31.5) detected in the substantia nigra, an area of the brain that shows neurodegeneration in Parkinson's disease. Therefore, therapeutic modulation or selective stimulation/antagonism of this receptor may have beneficial effects in the treatment of this disease. In addition, it appears that the GPCR8d gene is predominantly expressed in clusters of cultured cell lines derived from ovarian and lung cancer, but is expressed to a small degree in normal ovary and lung tissue. These data imply that the therapeutic modulation of this gene might be of use in the treatment of ovarian and lung cancers. Alternatively, the GPCR8d gene may be a useful marker in the diagnosis of lung and ovarian cancers.

**Panel 4D Summary**: Ag1264 Expression of the GPCR8d gene is low to undetectable (Ct values >35) in all of the samples on this panel except in IBD colitis 1; however, this sample is believed to be contaminated with genomic DNA and must therefore be disregarded.

**Example I. GPCR9 (also known as ba64p14-I):**

Expression of gene GPCR9 was assessed using the primer-probe set Ag1265, described in Table 20A. Results of the RTQ-PCR run are shown in Table 20B and Table 20C.

**Table 20A. Probe Name Ag1265**

| Primers | Sequences | TM | Length | Start Position | SEQ ID NO |
|---|---|---|---|---|---|
| Forward | 5'-GCCATCATCACACTGATTCTCT-3' | 59.2 | 22 | 195 | 105 |
| Probe | FAM-5'-CCACATCCCCATGTACATCTTCCTCA-3'-TAMRA | 68.9 | 26 | 230 | 106 |
| Reverse | 5'-GATGTCTGTCAAGGCCAAGTTA-3' | 59.2 | 22 | 257 | 107 |

5

**Table 20B. Panel 1.2**

| Tissue Name | Relative Expression(%) 1.2tm1427f_ ag1265 | Tissue Name | Relative Expression(%) 1.2tm1427f_ ag1265 |
|---|---|---|---|
| Endothelial cells | 17.3 | Renal ca. 786-0 | 0.3 |
| Fetal heart | 0.0 | Renal ca. A498 | 6.2 |
| Pancreas | 6.3 | Renal ca. RXF 393 | 0.0 |
| Pancreatic ca. CAPAN 2 | 10.0 | Renal ca. ACHN | 0.0 |
| Adrenal Gland (new lot*) | 8.5 | Renal ca. UO-31 | 2.6 |
| Thyroid | 7.2 | Renal ca. TK-10 | 57.0 |
| Salivary gland | 8.1 | Liver | 10.4 |
| Pituitary gland | 19.6 | Liver (fetal) | 3.5 |
| Brain (fetal) | 9.4 | Liver ca. (hepatoblast) HepG2 | 2.0 |
| Brain (whole) | 16.3 | Lung | 5.8 |
| Brain (amygdala) | 13.2 | Lung (fetal) | 3.6 |
| Brain (cerebellum) | 11.0 | Lung ca. (small cell) LX-1 | 98.6 |
| Brain (hippocampus) | 19.9 | Lung ca. (small cell) NCI-H69 | 36.6 |
| Brain (thalamus) | 6.0 | Lung ca. (s.cell var.) SHP-77 | 3.4 |
| Cerebral Cortex | 17.2 | Lung ca. (large cell)NCI-H460 | 7.6 |
| Spinal cord | 8.7 | Lung ca. (non-sm. cell) A549 | 28.1 |
| CNS ca. (glio/astro) U87-MG | 11.6 | Lung ca. (non-s.cell) NCI-H23 | 35.6 |
| CNS ca. (glio/astro) U-118-MG | 1.5 | Lung ca (non-s.cell) HOP-62 | 5.5 |
| CNS ca. (astro) SW1783 | 0.5 | Lung ca. (non-s.cl) NCI-H522 | 0.3 |
| CNS ca.* (neuro; met ) SK-N-AS | 0.0 | Lung ca. (squam.) SW 900 | 1.8 |
| CNS ca. (astro) SF-539 | 2.3 | Lung ca. (squam.) NCI-H596 | 12.1 |
| CNS ca. (astro) SNB-75 | 1.1 | Mammary gland | 9.2 |
| CNS ca. (glio) SNB-19 | 15.7 | Breast ca.* (pl. effusion) MCF-7 | 1.9 |

| | | | |
|---|---|---|---|
| CNS ca. (glio) U251 | 3.6 | Breast ca.* (pl.ef) MDA-MB-231 | 0.0 |
| CNS ca. (glio) SF-295 | 7.5 | Breast ca.* (pl. effusion) T47D | 13.0 |
| Heart | 7.8 | Breast ca. BT-549 | 3.2 |
| Skeletal Muscle (new lot*) | 4.1 | Breast ca. MDA-N | 2.3 |
| Bone marrow | 6.5 | Ovary | 0.2 |
| Thymus | 3.8 | Ovarian ca. OVCAR-3 | 3.8 |
| Spleen | 8.1 | Ovarian ca. OVCAR-4 | 3.8 |
| Lymph node | 13.8 | Ovarian ca. OVCAR-5 | 27.0 |
| Colorectal | 7.4 | Ovarian ca. OVCAR-8 | 34.9 |
| Stomach | 14.4 | Ovarian ca. IGROV-1 | 14.7 |
| Small intestine | 5.8 | Ovarian ca.* (ascites) SK-OV-3 | 13.8 |
| Colon ca. SW480 | 6.4 | Uterus | 8.8 |
| Colon ca.* (SW480 met)SW620 | 5.9 | Placenta | 14.1 |
| Colon ca. HT29 | 3.6 | Prostate | 7.5 |
| Colon ca. HCT-116 | 1.9 | Prostate ca.* (bone met)PC-3 | 4.0 |
| Colon ca. CaCo-2 | 4.1 | Testis | 21.8 |
| 83219 CC Well to Mod Diff (ODO3866) | 15.8 | Melanoma Hs688(A).T | 20.7 |
| Colon ca. HCC-2998 | 3.4 | Melanoma* (met) Hs688(B).T | 16.0 |
| Gastric ca.* (liver met) NCI-N87 | 5.7 | Melanoma UACC-62 | 0.0 |
| Bladder | 26.2 | Melanoma M14 | 7.2 |
| Trachea | 3.3 | Melanoma LOX IMVI | 0.2 |
| Kidney | 11.1 | Melanoma* (met) SK-MEL-5 | 0.0 |
| Kidney (fetal) | 8.1 | Adipose | 100.0 |

**Table 20C. Panel 4D**

| Tissue Name | Relative Expression(%) 4Dtm2166f_ag1265 | Tissue Name | Relative Expression(%) 4Dtm2166f_ag1265 |
|---|---|---|---|
| 93768_Secondary Th1_anti-CD28/anti-CD3 | 4.0 | 93100_HUVEC (Endothelial)_IL-1b | 0.0 |
| 93769_Secondary Th2_anti-CD28/anti-CD3 | 4.3 | 93779_HUVEC (Endothelial)_IFN gamma | 11.7 |
| 93770_Secondary Tr1_anti-CD28/anti-CD3 | 2.4 | 93102_HUVEC (Endothelial)_TNF alpha + IFN gamma | 3.9 |
| 93573_Secondary Th1_resting day 4-6 in IL-2 | 0.0 | 93101_HUVEC (Endothelial)_TNF alpha + IL4 | 2.0 |
| 93572_Secondary Th2_resting day 4-6 in IL-2 | 13.4 | 93781_HUVEC (Endothelial)_IL-11 | 3.6 |
| 93571_Secondary Tr1_resting day 4-6 in IL-2 | 1.8 | 93583_Lung Microvascular Endothelial Cells_none | 8.5 |

| | | | |
|---|---|---|---|
| 93568_primary Th1_anti-CD28/anti-CD3 | 4.2 | 93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 2.0 |
| 93569_primary Th2_anti-CD28/anti-CD3 | 6.7 | 92662_Microvascular Dermal endothelium_none | 23.7 |
| 93570_primary Tr1_anti-CD28/anti-CD3 | 0.0 | 92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 8.1 |
| 93565_primary Th1_resting dy 4-6 in IL-2 | 7.9 | 93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) ** | 10.4 |
| 93566_primary Th2_resting dy 4-6 in IL-2 | 2.1 | 93347_Small Airway Epithelium_none | 0.0 |
| 93567_primary Tr1_resting dy 4-6 in IL-2 | 4.6 | 93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 4.9 |
| 93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3 | 4.5 | 92668_Coronery Artery SMC_resting | 2.0 |
| 93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3 | 12.7 | 92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 0.0 |
| 93251_CD8 Lymphocytes_anti-CD28/anti-CD3 | 0.0 | 93107_astrocytes_resting | 0.0 |
| 93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2 | 2.3 | 93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml) | 4.2 |
| 93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28 | 2.2 | 92666_KU-812 (Basophil)_resting | 0.0 |
| 93354_CD4_none | 4.0 | 92667_KU-812 (Basophil)_PMA/ionoycin | 2.6 |
| 93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11 | 9.1 | 93579_CCD1106 (Keratinocytes)_none | 0.0 |
| 93103_LAK cells_resting | 8.6 | 93580_CCD1106 (Keratinocytes)_TNFa and IFNg ** | 2.3 |
| 93788_LAK cells_IL-2 | 7.1 | 93791_Liver Cirrhosis | 23.8 |
| 93787_LAK cells_IL-2+IL-12 | 6.0 | 93792_Lupus Kidney | 2.2 |
| 93789_LAK cells_IL-2+IFN gamma | 5.9 | 93577_NCI-H292 | 0.0 |
| 93790_LAK cells_IL-2+ IL-18 | 4.5 | 93358_NCI-H292_IL-4 | 0.0 |
| 93104_LAK cells_PMA/ionomycin and IL-18 | 0.0 | 93360_NCI-H292_IL-9 | 0.0 |
| 93578_NK Cells IL-2_resting | 0.0 | 93359_NCI-H292_IL-13 | 0.0 |
| 93109_Mixed Lymphocyte Reaction_Two Way MLR | 6.9 | 93357_NCI-H292_IFN gamma | 0.0 |
| 93110_Mixed Lymphocyte Reaction_Two Way MLR | 0.0 | 93777_HPAEC_- | 7.2 |
| 93111_Mixed Lymphocyte | 0.0 | 93778_HPAEC_IL-1 beta/TNA | 0.0 |

| Reaction_Two Way MLR | | alpha | |
|---|---|---|---|
| 93112_Mononuclear Cells (PBMCs)_resting | 0.0 | 93254_Normal Human Lung Fibroblast_none | 17.1 |
| 93113_Mononuclear Cells (PBMCs)_PWM | 4.5 | 93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml) | 6.2 |
| 93114_Mononuclear Cells (PBMCs)_PHA-L | 11.0 | 93257_Normal Human Lung Fibroblast_IL-4 | 2.3 |
| 93249_Ramos (B cell)_none | 2.2 | 93256_Normal Human Lung Fibroblast_IL-9 | 3.3 |
| 93250_Ramos (B cell)_ionomycin | 0.0 | 93255_Normal Human Lung Fibroblast_IL-13 | 26.4 |
| 93349_B lymphocytes_PWM | 0.0 | 93258_Normal Human Lung Fibroblast_IFN gamma | 20.9 |
| 93350_B lymphoytes_CD40L and IL-4 | 0.0 | 93106_Dermal Fibroblasts CCD1070_resting | 3.4 |
| 92665_EOL-1 (Eosinophil)_dbcAMP differentiated | 0.0 | 93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml | 6.5 |
| 93248_EOL-1 (Eosinophil)_dbcAMP/ PMAionomycin | 2.7 | 93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml | 1.6 |
| 93356_Dendritic Cells_none | 2.1 | 93772_dermal fibroblast_IFN gamma | 74.2 |
| 93355_Dendritic Cells_LPS 100 ng/ml | 0.0 | 93771_dermal fibroblast_IL-4 | 100.0 |
| 93775_Dendritic Cells_anti-CD40 | 6.4 | 93259_IBD Colitis 1** | 90.8 |
| 93774_Monocytes_resting | 0.0 | 93260_IBD Colitis 2 | 3.5 |
| 93776_Monocytes_LPS 50 ng/ml | 0.0 | 93261_IBD Crohns | 0.0 |
| 93581_Macrophages_resting | 7.6 | 735010_Colon_normal | 0.0 |
| 93582_Macrophages_LPS 100 ng/ml | 1.1 | 735019_Lung_none | 0.0 |
| 93098_HUVEC (Endothelial)_none | 6.9 | 64028-1_Thymus_none | 18.0 |
| 93099_HUVEC (Endothelial)_starved | 23.7 | 64030-1_Kidney_none | 5.5 |

**Panel 1.2 Summary**: Ag1265 Expression of the GPCR9 gene in this panel is skewed by genomic DNA contamination in the adipose sample. Disregarding this sample, moderate to low expression (CT values = 29-35) of the GPCR9 gene is detected in most tissues on this panel. Interestingly, there appears to be high expression in several lung and ovarian cancer cell lines, indicating that the expression of this gene maybe associated with the pathology of these diseases. In contrast, there is little expression of the GPCR9 gene in the samples of normal lung and ovary. Thus, therapeutic modulation of the GPCR9 gene may show utility in

the treatment of lung and ovarian cancer. Alternatively, the GPCR9 gene may be a useful marker in the diagnosis of lung and ovarian cancers. In addition, this gene is moderately expressed in pituitary, pancreas, liver and adrenal gland, tissues that play a role in normal metabolic and neuroendocrine function. The pituitary gland controls much endocrine secretion through response to hypophysiotrophic hormones (such as thyrotropin-releasing hormone, somatostatin, somatocrinin, gonadotropin-releasing hormone, corticotropin-releasing hormone) in the posterior pituitary, and response to peripheral hormones (e.g., estrogen, testosterone, etc) in the anterior pituitary. There are a number of diseases associated with pituitary pathophysiology (hyper- and hypothyroidism, gigantism, dwarfism, acromegaly, Addison's disease, Cushing's disease, diabetes insipidus) and therapeutic modulation, antagoinsm, or stimulation of this gene may be useful in the treatment of one or more of these diseases.

**Panel 4D Summary**: Ag1265 The GPCR9 transcript is expressed in untreated lung fibroblasts, starved and untreated HUVECs, and untreated microvascular dermal endothelial cells. Therefore, the GPCR9 gene, encoding for a GPCR, may be important for maintaining normal homeostasis. Additionally, it may be selectively down regulated by pro-inflammatory cytokines such as IL-1 and TNF alpha. Other cytokines such as IL-13 and gamma interferon do not appear to reduce expression. Small molecule therapeutics (antagonistic) could reduce the expression or activity of the GPCR9 protein to encourage an immune response. This type of adjuvant effect would complement immunization or treatment of bacterial and viral infections. Alternatively, small molecule therapies with agonistic function (ligand-like molecules) toward the GPCR9 gene product could reduce or block inflammatory responses for diseases such as psoriasis, emphysema, asthma, allergy, and arthritis.

**Example 3. SNP analysis of GPCRX clones**

**SeqCallingTM Technology:** cDNA was derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, cell lines, primary cells or tissue cultured primary cells and cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression for example, growth factors, chemokines, steroids. The cDNA thus derived was then sequenced using CuraGen's proprietary SeqCalling technology. Sequence traces were evaluated manually and edited for corrections if appropriate. cDNA sequences from all samples were assembled with themselves

and with public ESTs using bioinformatics programs to generate CuraGen's human SeqCalling database of SeqCalling assemblies. Each assembly contains one or more overlapping cDNA sequences derived from one or more human samples. Fragments and ESTs were included as components for an assembly when the extent of identity with another component of the assembly was at least 95% over 50 bp. Each assembly can represent a gene and/or its variants such as splice forms and/or single nucleotide polymorphisms (SNPs) and their combinations.

Variant sequences are included in this application. A variant sequence can include a single nucleotide polymorphism (SNP). A SNP can, in some instances, be referred to as a "cSNP" to denote that the nucleotide sequence containing the SNP originates as a cDNA. A SNP can arise in several ways. For example, a SNP may be due to a substitution of one nucleotide for another at the polymorphic site. Such a substitution can be either a transition or a transversion. A SNP can also arise from a deletion of a nucleotide or an insertion of a nucleotide, relative to a reference allele. In this case, the polymorphic site is a site at which one allele bears a gap with respect to a particular nucleotide in another allele. SNPs occurring within genes may result in an alteration of the amino acid encoded by the gene at the position of the SNP. Intragenic SNPs may also be silent, however, in the case that a codon including a SNP encodes the same amino acid as a result of the redundancy of the genetic code. SNPs occurring outside the region of a gene, or in an intron within a gene, do not result in changes in any amino acid sequence of a protein but may result in altered regulation of the expression pattern for example, alteration in temporal expression, physiological response regulation, cell type expression regulation, intensity of expression, stability of transcribed message.

**Method of novel SNP Identification:** SNPs are identified by analyzing sequence assemblies using CuraGen's proprietary SNPTool algorithm. SNPTool identifies variation in assemblies with the following criteria: SNPs are not analyzed within 10 base pairs on both ends of an alignment; Window size (number of bases in a view) is 10; The allowed number of mismatches in a window is 2; Minimum SNP base quality (PHRED score) is 23; Minimum number of changes to score an SNP is 2/assembly position. SNPTool analyzes the assembly and displays SNP positions, associated individual variant sequences in the assembly, the depth of the assembly at that given position, the putative assembly allele frequency, and the SNP sequence variation. Sequence traces are then selected and brought into view for manual validation. The consensus assembly sequence is imported into CuraTools along with variant sequence changes to identify potential amino acid changes resulting from the SNP sequence variation. Comprehensive SNP data analysis is then exported into the SNPCalling database.

**Method of novel SNP Confirmation:** SNPs are confirmed employing a validated method know as Pyrosequencing (Pyrosequencing, Westborough, MA). Detailed protocols for Pyrosequencing can be found in: Alderborn et al. Determination of Single Nucleotide Polymorphisms by Real-time Pyrophosphate DNA Sequencing. (2000). *Genome Research*. 10, Issue 8, August. 1249-1265. In brief, Pyrosequencing is a real time primer extension process of genotyping. This protocol takes double-stranded, biotinylated PCR products from genomic DNA samples and binds them to streptavidin beads. These beads are then denatured producing single stranded bound DNA. SNPs are characterized utilizing a technique based on an indirect bioluminometric assay of pyrophosphate (PPi) that is released from each dNTP upon DNA chain elongation. Following Klenow polymerase-mediated base incorporation, PPi is released and used as a substrate, together with adenosine 5'-phosphosulfate (APS), for ATP sulfurylase, which results in the formation of ATP. Subsequently, the ATP accomplishes the conversion of luciferin to its oxi-derivative by the action of luciferase. The ensuing light output becomes proportional to the number of added bases, up to about four bases. To allow processivity of the method dNTP excess is degraded by apyrase, which is also present in the starting reaction mixture, so that only dNTPs are added to the template during the sequencing. The process has been fully automated and adapted to a 96-well format, which allows rapid screening of large SNP panels. The DNA and protein sequences for the novel single nucleotide polymorphic variants are reported. Variants are reported individually but any combination of all or a select subset of variants are also included. In addition, the positions of the variant bases and the variant amino acid residues are underlined.

**Results**

Variants are reported individually but any combination of all or a select subset of variants are also included.

**GPCR2**

There is one variant reported for GPCR2. As shown in Table 21, variant 13374791 is an A to G SNP at 654 bp of the nucleotide sequence that results in an Ile to Val change at amino acid 214 of protein sequence. Because GPCR2b has 10 extra nucleotides at its 5' terminal compared to GPCR2a, it is important to note that the base positions of cSNPs are determined using the nucleotide sequence of GPCR2a (also known as Gmba64p14B)as the reference sequence.

| Table 21. cSNP for GPCR2 | | | | |
|---|---|---|---|---|
| Name of Variant | Base Position of cSNP | Wild Type | Variant | Amino Acid Change |
| 13374791 | 654 | A | G | Ile -> Val |

**GPCR3**

There are 3 variants reported for GPCR3. As shown in Table 22, variant 13374811 is a T to C SNP at 628 bp of the nucleotide sequence that results in an Ile to Thr change at amino acid 205 of protein sequence, variant 13374812 is an A to G SNP at 152 bp of the nucleotide sequence that results in no change in the protein sequence (silent), and variant 13374813 is a T to C SNP at 109 bp of the nucleotide sequence that results in a Leu to Pro change at amino acid 32 of protein sequence.

| Table 22. cSNPs for GPCR3 | | | | |
|---|---|---|---|---|
| Name of Variant | Base Position of cSNP | Wild Type | Variant | Amino Acid Change |
| 13374811 | 628 | T | C | Ile -> Ser |
| 13374812 | 152 | A | G | None |
| 13374813 | 109 | T | C | Leu -> Pro |

**GPCR4**

There are 9 variants reported for GPCR4. As shown in Table 23, variant 13374125 is a G to A SNP at 340 bp of the nucleotide sequence that results in a Met to Ile change at amino acid 99 of protein sequence, variant 13374126 is a G to T SNP at 402 bp of the nucleotide sequence that results in an Arg to Leu change at amino acid 120 of protein sequence, variant 13374127 is a C to T SNP at 435 bp of the nucleotide sequence that results in a Ser to Phe change at amino acid 131 of protein sequence, variant 13374128 is a T to A SNP at 489 bp of the nucleotide sequence that results in a Leu to His change at amino acid 149 of protein sequence, variant 13374129 is a G to T SNP at 653 bp of the nucleotide sequence that results in a Val to Leu change at amino acid 204 of protein sequence, variant 13374130 is a T to A SNP at 927 bp of the nucleotide sequence that results in a Met to Lys change at amino acid 295 of protein sequence, variant 13374808 is a C to A SNP at 313 bp of the nucleotide sequence that results in no change in the protein sequence (silent), variant 13374809 is a G to C SNP at 795 bp of the nucleotide sequence that results in a Gly to Ala change at amino acid

251 of protein sequence, and variant 13374810 is an A to G SNP at 944 bp of the nucleotide sequence that results in an Arg to Gly change at amino acid 301 of protein sequence.

| Table 23. cSNPs for GPCR4 | | | | |
|---|---|---|---|---|
| Name of Variant | Base Position of cSNP | Wild Type | Variant | Amino Acid Change |
| 13374125 | 340 | G | A | Met -> Ile |
| 13374126 | 402 | G | T | Arg -> Leu |
| 13374127 | 435 | C | T | Ser -> Phe |
| 13374128 | 489 | T | A | Leu -> His |
| 13374129 | 653 | G | T | Val -> Leu |
| 13374130 | 927 | T | A | Met -> Lys |
| 13374808 | 313 | C | A | None |
| 13374809 | 795 | G | C | Gly -> Ala |
| 13374810 | 944 | A | G | Arg -> Gly |

**GPCR5c**

There are 6 variants reported for GPCR5c. As shown in Table 24, variant 13374799 is a C to T SNP at 100 bp of the nucleotide sequence that results in a Gln to Stop Codon change at amino acid 24 of protein sequence, variant 13374800 is an A to G SNP at 293 bp of the nucleotide sequence that results in a Gln to Arg change at amino acid 88 of protein sequence, variant 13374801 is a T to A SNP at 379 bp of the nucleotide sequence that results in a Ser to Thr change at amino acid 117 of protein sequence, variant 13374802 is an A to G SNP at 413 bp of the nucleotide sequence that results in a His to Arg change at amino acid 128 of protein sequence, variant 13374803 is a T to C SNP at 502 bp of the nucleotide sequence that results in a Ser to Pro change at amino acid 158 of protein sequence, and variant 13374804 is a T to C SNP at 532 bp of the nucleotide sequence that results in a Phe to Leu change at amino acid 168 of protein sequence.

| **Table 24. cSNPs for GPCR5c** | | | | |
|---|---|---|---|---|
| **Name of Variant** | **Base Position of cSNP** | **Wild Type** | **Variant** | **Amino Acid Change** |
| 13374799 | 100 | C | T | Gln -> Stop |
| 13374800 | 293 | A | G | Gln -> Arg |
| 13374801 | 379 | T | A | Ser -> Thr |
| 13374802 | 413 | A | G | His -> Arg |
| 13374803 | 502 | T | C | Ser -> Pro |

| | | | | |
|---|---|---|---|---|
| 13374804 | 532 | T | C | Phe -> Leu |

**GPCR6**

There are 3 variants reported for GPCR6. As shown in Table 25, variant 13374794 is a C to T SNP at 105 bp of the nucleotide sequence that results in a Leu to Phe change at amino acid 35 of protein sequence, variant 13374795 is a T to C SNP at 220 bp of the nucleotide sequence that results in a Leu to Pro change at amino acid 73 of protein sequence, and variant 13374796 is an A to G SNP at 381 bp of the nucleotide sequence that results in a Met to Val change at amino acid 127 of protein sequence. Because the nucleotide sequences of GPCR6a, 6b, and 6c vary from each other, it is important to note that the base positions of cSNPs are determined using the nucleotide sequence of GPCR6c (also known as Gmba64p14n11_da1 or 147307499) as the reference sequence.

**Table 25. cSNPs for GPCR6**

| Name of Variant | Base Position of cSNP | Wild Type | Variant | Amino Acid Change |
|---|---|---|---|---|
| 13374794 | 105 | C | T | Leu -> Phe |
| 13374795 | 220 | T | C | Leu -> Pro |
| 13374796 | 381 | A | G | Met -> Val |

**GPCR8**

There are 7 variants reported for GPCR8. As shown in Table 26, variant 13374086 is a G to C SNP at 460 bp of the nucleotide sequence that results in no change in the protein sequence (silent), variant 13374087 is a G to T SNP at 601 bp of the nucleotide sequence that results in no change in the protein sequence (silent), variant 13374088 is a T to C SNP at 743 bp of the nucleotide sequence that results in a Phe to Leu change at amino acid 239 of protein sequence, variant 13374271 is a G to C SNP at 460 bp of the nucleotide sequence that results in no change in the protein sequence (silent), variant 13374273 is a C to G SNP at 660 bp of the nucleotide sequence that results in a Pro to Arg change at amino acid 211 of protein sequence, variant 13374792 is a C to T SNP at 43 bp of the nucleotide sequence that results in no change in the protein sequence (silent), and variant 13374793 is a C to A SNP at 107 bp of the nucleotide sequence that results in a Pro to Thr change at amino acid 27 of protein

sequence. Because the nucleotide sequences of GPCR8a, 8b, 8c, and 8d vary from each other, it is important to note that the base positions of cSNPs are determined using the nucleotide sequence of GPCR8b (also known as CG50259-01)as the reference sequence.

**Table 26. cSNPs for GPCR8**

| Name of Variant | Base Position of cSNP | Wild Type | Variant | Amino Acid Change |
|---|---|---|---|---|
| 13374086 | 460 | G | C | None |
| 13374087 | 601 | G | T | None |
| 13374088 | 743 | T | C | Phe -> Leu |
| 13374271 | 460 | G | C | None |
| 13374273 | 660 | C | G | Pro -> Arg |
| 13374792 | 43 | C | T | None |
| 13374793 | 107 | C | A | Pro -> Thr |

**Example 4. Other SNPs**

One or more consensus positions (Cons. Pos.) of the nucleotide sequence have been identified as SNPs for the following GPCR clones. "Depth" represents the number of clones covering the region of the SNP. The Putative Allele Frequency (Putative Allele Freq.) is the fraction of all the clones containing the SNP. A dash ("-"), when shown, means that a base is not present. The sign ">" means "is changed to".

**GPCR4b**

```
Cons.Pos.:  85  Depth:  76  Change: G > A    Putative Allele Freq.:  0.039
Cons.Pos.: 287  Depth:  76  Change: A > G    Putative Allele Freq.:  0.026
Cons.Pos.: 314  Depth:  78  Change: G > A    Putative Allele Freq.:  0.192
Cons.Pos.: 346  Depth:  83  Change: T > C    Putative Allele Freq.:  0.024
Cons.Pos.: 376  Depth:  92  Change: T > C    Putative Allele Freq.:  0.022
Cons.Pos.: 410  Depth: 102  Change: C > T    Putative Allele Freq.:  0.020
Cons.Pos.: 442  Depth: 106  Change: T > A    Putative Allele Freq.:  0.057
Cons.Pos.: 464  Depth: 111  Change: A > G    Putative Allele Freq.:  0.189
Cons.Pos.: 544  Depth: 107  Change: T > C    Putative Allele Freq.:  0.178
Cons.Pos.: 569  Depth:  94  Change: C > T    Putative Allele Freq.:  0.117
Cons.Pos.: 603  Depth:  82  Change: A > G    Putative Allele Freq.:  0.024
Cons.Pos.: 609  Depth:  79  Change: T > C    Putative Allele Freq.:  0.038
Cons.Pos.: 630  Depth:  74  Change: T > C    Putative Allele Freq.:  0.027
Cons.Pos.: 768  Depth:  74  Change: T > C    Putative Allele Freq.:  0.027
Cons.Pos.: 772  Depth:  74  Change: C > T    Putative Allele Freq.:  0.027
Cons.Pos.: 906  Depth:  50  Change: A > G    Putative Allele Freq.:  0.180
Cons.Pos.: 907  Depth:  50  Change: C > T    Putative Allele Freq.:  0.060
Cons.Pos.: 924  Depth:  48  Change: A > G    Putative Allele Freq.:  0.062
```

**GPCR5a**

```
Cons.Pos.: 525 Depth: 17 Change: C > T   Putative Allele Freq.: 0.118
```

```
Cons.Pos.: 868 Depth: 16 Change: T > C   Putative Allele Freq.: 0.375
Cons.Pos.: 889 Depth: 16 Change: T > C   Putative Allele Freq.: 0.375
```

**GPCR6a**

```
Cons.Pos.:  34  Depth:  15  Change: A > -   Putative Allele Freq.:  0.133
  -> 131443392(-,i)  unrev. Fpos: 778
  -> 131449019(-,i)  unrev. Fpos: 804

Cons.Pos.:  89  Depth:  34  Change: T > -   Putative Allele Freq.:  0.147
  -> 132684332(+,i)  unrev. Fpos:  90
  -> 132684478(+,i)  unrev. Fpos:  97
  -> 132721601(+,i)  unrev. Fpos:  90
  -> 132721626(+,i)  unrev. Fpos:  65
  -> 132721669(+,i)  unrev. Fpos:  78

Cons.Pos.: 110  Depth:  34  Change: T > C   Putative Allele Freq.:  0.059
  -> 131443532(+,i)  unrev. Fpos: 173
  -> 131449125(-,i)  unrev. Fpos: 730

Cons.Pos.: 142  Depth:  33  Change: T > -   Putative Allele Freq.:  0.182
  -> 132684332(+,i)  unrev. Fpos: 140
  -> 132684478(+,i)  unrev. Fpos: 147
  -> 132721601(+,i)  unrev. Fpos: 140
  -> 132721626(+,i)  unrev. Fpos: 115
  -> 132721669(+,i)  unrev. Fpos: 128
  -> 132721682(+,i)  unrev. Fpos: 120

Cons.Pos.: 409  Depth:  41  Change: G > A   Putative Allele Freq.:  0.049
  -> 131443502(-,i)  unrev. Fpos: 420
  -> 131449096(+,i)  unrev. Fpos: 493

Cons.Pos.: 702  Depth:  34  Change: - > A   Putative Allele Freq.:  0.206
  -> 131443455(-,i)  unrev. Fpos: 131
  -> 132684371(-,i)  unrev. Fpos: 385
  -> 132684452(-,i)  unrev. Fpos: 381
  -> 132721492(-,i)  unrev. Fpos: 367
  -> 132721587(-,i)  unrev. Fpos: 374
  -> 132721658(-,i)  unrev. Fpos: 378
  -> 132721708(-,i)  unrev. Fpos: 382

Cons.Pos.: 710  Depth:  33  Change: - > A   Putative Allele Freq.:  0.152
  -> 131443407(-,i)  unrev. Fpos: 126
  -> 131449064(-,i)  unrev. Fpos: 143
  -> 132684494(-,i)  unrev. Fpos: 401
  -> 132721639(-,i)  unrev. Fpos: 357
  -> 132721695(-,i)  unrev. Fpos: 355
```

**GPCR7a**

```
Cons.Pos.: 465  Depth:   6  Change: T > C   Putative Allele Freq.:  0.333
```

**GPCR8b**

```
Cons.Pos.: 229  Depth:  13  Change: T > C
Cons.Pos.: 463  Depth:  12  Change: G > C
Cons.Pos.: 604  Depth:  13  Change: G > T
Cons.Pos.: 663  Depth:  14  Change: C > G
Cons.Pos.: 746  Depth:  14  Change: T > C
```

**GPCR8c**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Cons.Pos.: | 238 | Depth: | 8 | Change: G > C | Putative Allele Freq.: | 0.250 |
| Cons.Pos.: | 379 | Depth: | 8 | Change: T > G | Putative Allele Freq.: | 0.250 |
| Cons.Pos.: | 438 | Depth: | 9 | Change: C > G | Putative Allele Freq.: | 0.444 |

## EQUIVALENTS

Although particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims, which follow. In particular, it is contemplated by the inventors that various substitutions, alterations, and modifications may be made to the invention without departing from the spirit and scope of the invention as defined by the claims. The choice of nucleic acid starting material, clone of interest, or library type is believed to be a matter of routine for a person of ordinary skill in the art with knowledge of the embodiments described herein. Other aspects, advantages, and modifications considered to be within the scope of the following claims.

# SEQUENCE LISTING

<110> CuraGen Corporation

<120> Novel Proteins and Nucleic Acids Encoding Same

<130> 435-MLC114

<140> Not Yet Known (based upon PCT/US01/22637)
<141> 2001-07-18

<150> 60/218,903
<151> 2000-07-18

<150> 60/259,072
<151> 2000-12-29

<150> 60/264,124
<151> 2001-01-25

<150> 60/260,327
<151> 2001-01-08

<150> 60/261,030
<151> 2001-01-11

<150> 60/262,158
<151> 2001-01-17

<150> 60/266,084
<151> 2001-02-02

<150> 60/266,108
<151> 2001-02-02

<150> 60/274,858
<151> 2001-03-09

<150> 60/278,917
<151> 2001-03-26

<150> 09/908,006
<151> 2001-08-18

<160> 109

<170> PatentIn Ver. 2.1

<210> 1
<211> 968
<212> DNA
<213> Homo sapiens

<400> 1
ggatgggaaa accaggcaga gtgaaccaaa ccactgtttc agacttcctc cttctaggac 60
tctctgagtg gccagaggag cagcctcttc tgtttggcat cttccttggc atgtacctgg 120
tcaccatggt ggggaacctg ctcattatcc tggccatcag ctctgaccca cacctccata 180

ctcccatgta cttctttctg gccaacctgt cattaactga tgcctgtttc acttctgcct 240
ccatccccaa aatgctggcc aacattcata cccagagtca gatcatctcg tattctgggt 300
gtcttgcaca gctatatttc ctccttatgt ttggtggcct tgacaactgc ctgctggctg 360
tgatggcata tgaccgctat gtggccatct gccaaccact ccattacagc acatctatga 420
gtccccagct ctgtgcacta atgctgggtg tgtgctgggt gctaaccaac tgtcctgccc 480
tgatgcacac actgttgctg acccgcgtgg ctttctgtgc ccagaaagcc atccctcatt 540
tctattgtga tcctagtgct ctcctgaagc ttgcctgctc agatacccat gtaaacgagc 600
tgatgatcat caccatgggc ttgctgttcc tcactgttcc cctcctgctg atcgtcttct 660
cctatgtccg cattttctgg gctgtgtttg tcatctcatc tcctggaggg agatggaagg 720
ccttctctac ctgtggttct catctcacgg tggttctgct cttctatggg tctcttatgg 780
gtgtgtattt acttcctcca tcaacttact ctacagagag ggaaagtagg gctgctgttc 840
tctatatggt gattattccc acgctaaacc cattcattta tagcttgagg aacagagaca 900
tgaaggaggc tttgggtaaa ctttttgtca gtggaaaaac attcttttta tgattagaca 960
tctagacg 968

<210> 2
<211> 316
<212> PRT
<213> Homo sapiens

<400> 2
Met Gly Lys Pro Gly Arg Val Asn Gln Thr Thr Val Ser Asp Phe Leu
1 5 10 15

Leu Leu Gly Leu Ser Glu Trp Pro Glu Glu Gln Pro Leu Leu Phe Gly
20 25 30

Ile Phe Leu Gly Met Tyr Leu Val Thr Met Val Gly Asn Leu Leu Ile
35 40 45

Ile Leu Ala Ile Ser Ser Asp Pro His Leu His Thr Pro Met Tyr Phe
50 55 60

Phe Leu Ala Asn Leu Ser Leu Thr Asp Ala Cys Phe Thr Ser Ala Ser
65 70 75 80

Ile Pro Lys Met Leu Ala Asn Ile His Thr Gln Ser Gln Ile Ile Ser
85 90 95

Tyr Ser Gly Cys Leu Ala Gln Leu Tyr Phe Leu Leu Met Phe Gly Gly
100 105 110

Leu Asp Asn Cys Leu Leu Ala Val Met Ala Tyr Asp Arg Tyr Val Ala
115 120 125

Ile Cys Gln Pro Leu His Tyr Ser Thr Ser Met Ser Pro Gln Leu Cys
130 135 140

Ala Leu Met Leu Gly Val Cys Trp Val Leu Thr Asn Cys Pro Ala Leu
145 150 155 160

Met His Thr Leu Leu Leu Thr Arg Val Ala Phe Cys Ala Gln Lys Ala
165 170 175

Ile Pro His Phe Tyr Cys Asp Pro Ser Ala Leu Leu Lys Leu Ala Cys
180 185 190

Ser Asp Thr His Val Asn Glu Leu Met Ile Ile Thr Met Gly Leu Leu
195 200 205

Phe Leu Thr Val Pro Leu Leu Leu Ile Val Phe Ser Tyr Val Arg Ile
210 215 220

Phe Trp Ala Val Phe Val Ile Ser Ser Pro Gly Gly Arg Trp Lys Ala
225 230 235 240

Phe Ser Thr Cys Gly Ser His Leu Thr Val Val Leu Leu Phe Tyr Gly
245 250 255

Ser Leu Met Gly Val Tyr Leu Leu Pro Pro Ser Thr Tyr Ser Thr Glu
260 265 270

Arg Glu Ser Arg Ala Ala Val Leu Tyr Met Val Ile Ile Pro Thr Leu
275 280 285

Asn Pro Phe Ile Tyr Ser Leu Arg Asn Arg Asp Met Lys Glu Ala Leu
290 295 300

Gly Lys Leu Phe Val Ser Gly Lys Thr Phe Phe Leu
305 310 315

<210> 3
<211> 1034
<212> DNA
<213> Homo sapiens

<400> 3
gtgaacccac aactatggga agaaataacc taacaagacc ctctgaattc atcctccttg 60
gactctcctc tcgacctgag gatcagaagc cgctctttgc tgtgttcctc cccatctacc 120
ttatcacagt gataggaaac ctgcttatca tcctggccat ccgctcagac actcgtctcc 180
agacgcccat gtacttcttt ctaagcatcc tgtcttttgt tgacatttgc tatgtgacag 240
tcattatccc taagatgctg gtgaacttct tatcagagac aaagaccatc tcttacagtg 300
agtgtctgac ccagatgtac tttttcttag cctttggaaa cacagacagt tacctgctag 360
cagccatggc cattgaccgc tatgtggcca tatgtaatcc cttccactac atcaccatta 420
tgagtcacag atgctgtgtc ctgcttctgg ttctctcctt ctgcattcca cattttcact 480
ccctcctgca cattcttctg actaatcagc tcatcttctg tgcctcaaat gtcatccatc 540
acttttctg cgatgatcaa ccagtgctaa aattgtcctg ttcctcccat tttgtcaaag 600
aaatcacagt aatgacagaa ggcttggctg tcataatgac cccgttttca tgcatcatca 660
tctcttattt aagaatcctc atcactgttc tgaagattcc ttcagctgct ggaaagcgta 720
aagcattttc tacctgtggc tctcatctca cagtggtgac cctgttttat ggaagcatta 780
gctatctcta ttttcagccc ctgtccaact atactgtcaa ggatcaaata gcaacaatta 840
tctacaccgt actgactcct atgctaaatc catttatcta tagtctgagg aacaaagaca 900
tgaagcaggg tttggcaaag ttgatgcaca ggatgaaatg tcagtaaaag acctaaggtc 960
ttaagagaat accacagatc tcttgccctg gactataggt tattaatggg tatgtgattc 1020
tgaaatgatt atta 1034

<210> 4
<211> 310
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Gly Arg Asn Asn Leu Thr Arg Pro Ser Glu Phe Ile Leu Leu Gly
  1               5                  10                  15

Leu Ser Ser Arg Pro Glu Asp Gln Lys Pro Leu Phe Ala Val Phe Leu
            20                  25                  30

Pro Ile Tyr Leu Ile Thr Val Ile Gly Asn Leu Leu Ile Ile Leu Ala
        35                  40                  45

Ile Arg Ser Asp Thr Arg Leu Gln Thr Pro Met Tyr Phe Phe Leu Ser
    50                  55                  60

Ile Leu Ser Phe Val Asp Ile Cys Tyr Val Thr Val Ile Ile Pro Lys
65                  70                  75                  80

Met Leu Val Asn Phe Leu Ser Glu Thr Lys Thr Ile Ser Tyr Ser Glu
                85                  90                  95

Cys Leu Thr Gln Met Tyr Phe Phe Leu Ala Phe Gly Asn Thr Asp Ser
            100                 105                 110

Tyr Leu Leu Ala Ala Met Ala Ile Asp Arg Tyr Val Ala Ile Cys Asn
        115                 120                 125

Pro Phe His Tyr Ile Thr Ile Met Ser His Arg Cys Cys Val Leu Leu
    130                 135                 140

Leu Val Leu Ser Phe Cys Ile Pro His Phe His Ser Leu Leu His Ile
145                 150                 155                 160

Leu Leu Thr Asn Gln Leu Ile Phe Cys Ala Ser Asn Val Ile His His
                165                 170                 175

Phe Phe Cys Asp Asp Gln Pro Val Leu Lys Leu Ser Cys Ser Ser His
            180                 185                 190

Phe Val Lys Glu Ile Thr Val Met Thr Glu Gly Leu Ala Val Ile Met
        195                 200                 205

Thr Pro Phe Ser Cys Ile Ile Ile Ser Tyr Leu Arg Ile Leu Ile Thr
    210                 215                 220

Val Leu Lys Ile Pro Ser Ala Ala Gly Lys Arg Lys Ala Phe Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Thr Val Val Thr Leu Phe Tyr Gly Ser Ile Ser
                245                 250                 255

Tyr Leu Tyr Phe Gln Pro Leu Ser Asn Tyr Thr Val Lys Asp Gln Ile
            260                 265                 270

Ala Thr Ile Ile Tyr Thr Val Leu Thr Pro Met Leu Asn Pro Phe Ile
        275                 280                 285

Tyr Ser Leu Arg Asn Lys Asp Met Lys Gln Gly Leu Ala Lys Leu Met
    290                 295                 300
```

His Arg Met Lys Cys Gln
305 310

<210> 5
<211> 1011
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (996)
<223> Wherein N is G or A or T or C

<400> 5
ttatctttac gtgaacccac aactatggga agaaataacc taacaagacc ctctgaattc 60
atcctccttg gactctcctc tcgacctgag gatcagaagc cgctctttgc tgtgttcctc 120
cccatctacc ttatcacagt gataggaaac ctgcttatca tcctggccat ccgctcagac 180
actcgtctcc agacgcccat gtacttcttt ctaagcatcc tgtcttttgt tgacatttgc 240
tatgtgacag tcattatccc taagatgctg gtgaacttct tatcagagac aaagaccatc 300
tcttacagtg agtgtctgac ccagatgtac ttttttcttag cctttggaaa cacagacagt 360
tacctgctag cagccatggc cattgaccgc tatgtggcca tatgtaatcc cttccactac 420
atcaccatta tgagtcacag atgctgtgtc ctgcttctgg ttctctcctt ctgcattcca 480
cattttcact ccctcctgca cattcttctg actaatcagc tcatcttctg tgcctcaaat 540
gtcatccatc acttttttctg cgatgatcaa ccagtgctaa aattgtcctg ttcctcccat 600
tttgtcaaag aaatcacagt aatgacagaa ggcttggctg tcataatgac cccgttttca 660
tgcgtcatca tctcttattt aagaatcctc atcactgttc tgaagattcc ttcagctgct 720
ggaaagcgta aagcattttc tacctgtggc tctcatctca cagtggtgac cctgttttat 780
ggaagcatta gctatctcta ttttcagccc ctgtccaact atactgtcaa ggatcaaata 840
gcaacaatta tctacaccgt actgactcct atgctaaatc catttatcta tagtctgagg 900
aacaaagaca tgaagcaggg tttggcaaag ttgatgcaca ggatgaaatg tcagtaaaag 960
acctaaggtc ttaagagaat accacagatc tcttgncctg gactataggt t 1011

<210> 6
<211> 310
<212> PRT
<213> Homo sapiens

<400> 6
Met Gly Arg Asn Asn Leu Thr Arg Pro Ser Glu Phe Ile Leu Leu Gly
1 5 10 15

Leu Ser Ser Arg Pro Glu Asp Gln Lys Pro Leu Phe Ala Val Phe Leu
20 25 30

Pro Ile Tyr Leu Ile Thr Val Ile Gly Asn Leu Leu Ile Ile Leu Ala
35 40 45

Ile Arg Ser Asp Thr Arg Leu Gln Thr Pro Met Tyr Phe Phe Leu Ser
50 55 60

Ile Leu Ser Phe Val Asp Ile Cys Tyr Val Thr Val Ile Ile Pro Lys
65 70 75 80

Met Leu Val Asn Phe Leu Ser Glu Thr Lys Thr Ile Ser Tyr Ser Glu

85 90 95

Cys Leu Thr Gln Met Tyr Phe Phe Leu Ala Phe Gly Asn Thr Asp Ser
100 105 110

Tyr Leu Leu Ala Ala Met Ala Ile Asp Arg Tyr Val Ala Ile Cys Asn
115 120 125

Pro Phe His Tyr Ile Thr Ile Met Ser His Arg Cys Cys Val Leu Leu
130 135 140

Leu Val Leu Ser Phe Cys Ile Pro His Phe His Ser Leu Leu His Ile
145 150 155 160

Leu Leu Thr Asn Gln Leu Ile Phe Cys Ala Ser Asn Val Ile His His
165 170 175

Phe Phe Cys Asp Asp Gln Pro Val Leu Lys Leu Ser Cys Ser Ser His
180 185 190

Phe Val Lys Glu Ile Thr Val Met Thr Glu Gly Leu Ala Val Ile Met
195 200 205

Thr Pro Phe Ser Cys Val Ile Ile Ser Tyr Leu Arg Ile Leu Ile Thr
210 215 220

Val Leu Lys Ile Pro Ser Ala Ala Gly Lys Arg Lys Ala Phe Ser Thr
225 230 235 240

Cys Gly Ser His Leu Thr Val Val Thr Leu Phe Tyr Gly Ser Ile Ser
245 250 255

Tyr Leu Tyr Phe Gln Pro Leu Ser Asn Tyr Thr Val Lys Asp Gln Ile
260 265 270

Ala Thr Ile Ile Tyr Thr Val Leu Thr Pro Met Leu Asn Pro Phe Ile
275 280 285

Tyr Ser Leu Arg Asn Lys Asp Met Lys Gln Gly Leu Ala Lys Leu Met
290 295 300

His Arg Met Lys Cys Gln
305 310

<210> 7
<211> 981
<212> DNA
<213> Homo sapiens

<400> 7
gctgactgtc actcatgatg agctttgccc ctaatgcttc acactctccg gtttttttgc 60
tccttgggtt ctcgagagct aacatctcct acactctcct cttcttcctg ttcctggcta 120
tttacctgac caccatactg gggaatgtga cactggtgct gctcatctcc tgggactcca 180
gactgcactc acccatgtat tatctgcttc gtggcctctc tgtgatagac atggggctat 240
ccacagttac actgccccag ttgctggccc atttggtctc tcattaccca accattcctg 300
ctgcccgctg cttggctcag ttctttttct tctatgcatt tggggttaca gatacacttg 360

tcattgctgt catggctctg gatcgctatg tggccatctg tgacccccctg cactatgctt 420
tggtaatgaa tcaccaacgg tgtgcctgct tactagcctt gagctgggtg gtgtccatac 480
tgcacaccat gttgcgtgtg ggactcgtcc tgcctctttg ctggactggg gatgctgggg 540
gcaacgttaa ccttcctcac ttcttttgtg accaccggcc acttctgcga gcctcttgtt 600
ctgacataca ttctaatgag ctggccatat tctttgaggg tggcttcctt atgctgggcc 660
cctgtgccct cattgtactc tcttatgtcc gaattggggc cgctattcta cgtttgcctt 720
cagctgctgg tcgccgccga gcagtctcca cctgtggatc ccacctcacc atggttggtt 780
tcctctacgg caccatcatt tgtgtctact tccagcctcc cttccagaac tctcagtatc 840
aggacatggt ggcttcagta atgtatactg ccattacacc tttggccaac ccatttgtgt 900
atagcctcca caataaggat gtcaagggtg cactctgcag gctgcttgaa tgggtgaagg 960
tagacccctg attagcctgc t 981

<210> 8
<211> 318
<212> PRT
<213> Homo sapiens

<400> 8
Met Met Ser Phe Ala Pro Asn Ala Ser His Ser Pro Val Phe Leu Leu
1 5 10 15

Leu Gly Phe Ser Arg Ala Asn Ile Ser Tyr Thr Leu Leu Phe Phe Leu
20 25 30

Phe Leu Ala Ile Tyr Leu Thr Thr Ile Leu Gly Asn Val Thr Leu Val
35 40 45

Leu Leu Ile Ser Trp Asp Ser Arg Leu His Ser Pro Met Tyr Tyr Leu
50 55 60

Leu Arg Gly Leu Ser Val Ile Asp Met Gly Leu Ser Thr Val Thr Leu
65 70 75 80

Pro Gln Leu Leu Ala His Leu Val Ser His Tyr Pro Thr Ile Pro Ala
85 90 95

Ala Arg Cys Leu Ala Gln Phe Phe Phe Phe Tyr Ala Phe Gly Val Thr
100 105 110

Asp Thr Leu Val Ile Ala Val Met Ala Leu Asp Arg Tyr Val Ala Ile
115 120 125

Cys Asp Pro Leu His Tyr Ala Leu Val Met Asn His Gln Arg Cys Ala
130 135 140

Cys Leu Leu Ala Leu Ser Trp Val Val Ser Ile Leu His Thr Met Leu
145 150 155 160

Arg Val Gly Leu Val Leu Pro Leu Cys Trp Thr Gly Asp Ala Gly Gly
165 170 175

Asn Val Asn Leu Pro His Phe Phe Cys Asp His Arg Pro Leu Leu Arg
180 185 190

Ala Ser Cys Ser Asp Ile His Ser Asn Glu Leu Ala Ile Phe Phe Glu
195 200 205

```
Gly Gly Phe Leu Met Leu Gly Pro Cys Ala Leu Ile Val Leu Ser Tyr
    210                 215                 220

Val Arg Ile Gly Ala Ala Ile Leu Arg Leu Pro Ser Ala Ala Gly Arg
225                 230                 235                 240

Arg Arg Ala Val Ser Thr Cys Gly Ser His Leu Thr Met Val Gly Phe
                245                 250                 255

Leu Tyr Gly Thr Ile Ile Cys Val Tyr Phe Gln Pro Pro Phe Gln Asn
            260                 265                 270

Ser Gln Tyr Gln Asp Met Val Ala Ser Val Met Tyr Thr Ala Ile Thr
        275                 280                 285

Pro Leu Ala Asn Pro Phe Val Tyr Ser Leu His Asn Lys Asp Val Lys
    290                 295                 300

Gly Ala Leu Cys Arg Leu Leu Glu Trp Val Lys Val Asp Pro
305                 310                 315
```

<210> 9
<211> 1021
<212> DNA
<213> Homo sapiens

<400> 9

```
agcattctaa ctgtcttccc tcagctccag atgctgcaga gtcatggaaa accaatccag 60
catttctgaa tttttcctcc gaggaatatc agcgtctcca gagcaacagc agtccctctt 120
cggaattttc ctgtgtatgt atcttgtcac cttgactggg aacctgctca tcatcctggc 180
cattggctct gacctgcacc tccacacccc catgtacttt ttcttggcca acctgtcttt 240
tgttgacatg ggtttaacgt cctccacagt taccaagatg ctggtgaata tacagactcg 300
gcatcacacc atctcctata cgggttgcct cacgcaaatg tatttctttc tgatgtttgg 360
tgatctagac agcttcttcc tggctgccat ggcgtatgac cgctatgtgg ccatttgcca 420
ccccctctgc tactccacag tcatgaggcc ccaagtctgt gccctaatgc ttgcattgtg 480
ctgggtcctc accaatatcg ttgccctgac tcacacgttc ctcatggctc ggttgtcctt 540
ctgtgtgact ggggaaattg ctcacttttt ctgtgacatc actcctgtcc tgaagctgtc 600
atgttctgac acccacatca acgagatgat ggtttttgtc ttgggaggca ccgtactcat 660
cgtccccttt ttatgcattg tcacctccta catccacatt gtgccagcta tcctgagggt 720
ccgaacccgt ggtggggtgg gcaaggcctt ttccacctgc agttcccacc tctgcgttgt 780
ttgtgtgttc tatgggaccc tcttcagtgc ctacctgtgt cctccctcca ttgcctctga 840
agagaaggac attgcagcag ctgcaatgta caccatagtg actcccatgt tgaacccctt 900
tatctatagc ctaaggaaca aggacatgaa gggggcccta aagaggctct tcagtcacag 960
gagtattgtt tcctcttaga tgtggtgaca gcaacattta atgaaaagac ataggcttgg 1020
a                                                                  1021
```

<210> 10
<211> 311
<212> PRT
<213> Homo sapiens

<400> 10

```
Met Glu Asn Gln Ser Ser Ile Ser Glu Phe Phe Leu Arg Gly Ile Ser
  1                 5                  10                  15
```

```
Ala Ser Pro Glu Gln Gln Gln Ser Leu Phe Gly Ile Phe Leu Cys Met
                20                  25                  30

Tyr Leu Val Thr Leu Thr Gly Asn Leu Leu Ile Ile Leu Ala Ile Gly
            35                  40                  45

Ser Asp Leu His Leu His Thr Pro Met Tyr Phe Phe Leu Ala Asn Leu
        50                  55                  60

Ser Phe Val Asp Met Gly Leu Thr Ser Ser Thr Val Thr Lys Met Leu
65                  70                  75                  80

Val Asn Ile Gln Thr Arg His His Thr Ile Ser Tyr Thr Gly Cys Leu
                85                  90                  95

Thr Gln Met Tyr Phe Phe Leu Met Phe Gly Asp Leu Asp Ser Phe Phe
            100                 105                 110

Leu Ala Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His Pro Leu
        115                 120                 125

Cys Tyr Ser Thr Val Met Arg Pro Gln Val Cys Ala Leu Met Leu Ala
    130                 135                 140

Leu Cys Trp Val Leu Thr Asn Ile Val Ala Leu Thr His Thr Phe Leu
145                 150                 155                 160

Met Ala Arg Leu Ser Phe Cys Val Thr Gly Glu Ile Ala His Phe Phe
                165                 170                 175

Cys Asp Ile Thr Pro Val Leu Lys Leu Ser Cys Ser Asp Thr His Ile
            180                 185                 190

Asn Glu Met Met Val Phe Val Leu Gly Gly Thr Val Leu Ile Val Pro
        195                 200                 205

Phe Leu Cys Ile Val Thr Ser Tyr Ile His Ile Val Pro Ala Ile Leu
    210                 215                 220

Arg Val Arg Thr Arg Gly Gly Val Gly Lys Ala Phe Ser Thr Cys Ser
225                 230                 235                 240

Ser His Leu Cys Val Val Cys Val Phe Tyr Gly Thr Leu Phe Ser Ala
                245                 250                 255

Tyr Leu Cys Pro Pro Ser Ile Ala Ser Glu Glu Lys Asp Ile Ala Ala
            260                 265                 270

Ala Ala Met Tyr Thr Ile Val Thr Pro Met Leu Asn Pro Phe Ile Tyr
        275                 280                 285

Ser Leu Arg Asn Lys Asp Met Lys Gly Ala Leu Lys Arg Leu Phe Ser
    290                 295                 300

His Arg Ser Ile Val Ser Ser
305                 310
```

<210> 11
<211> 1021
<212> DNA
<213> Homo sapiens

<400> 11
agcattctaa ctgtcttccc tcagctccag atgctgcaga gtcatggaaa accaatccag 60
catttctgaa tttttcctcc gaggaatatc agcgtctcca gagcaacagc agtccctctt 120
cggaattttc ctgtgtatgt atcttgtcac cttgactggg aacctgctca tcatcctggc 180
cattggctct gacctgcacc tccacacccc catgtacttt ttcttggcca acctgtcttt 240
tgttgacatg ggtttaacgt cctccacagt taccaagatg ctggtgaata tacagactcg 300
gcatcacacc atctcctata cgggttgcct cacgcaaatg tatttctttc tgatgtttgg 360
tgatctagac agcttcttcc tggctgccat ggcgtatgac cgctatgtgg ccatttgcca 420
ccccctctac tactccacag tcatgaggcc ccaagtctgt gccctaatgc ttgcattgtg 480
ctgggtcctc accaatatcg ttgccctgac tcacacgttc ctcatggctc ggttgtcctt 540
ctgtgtgact ggggaaattg ctcacttttt ctgtgacatc actcctgtcc tgaagctgtc 600
atgttctgac acccacatca acgagatgat ggtttttgtc ttgggaggca ccgtactcat 660
cgtccccttt ttatgcattg tcacctccta catccacatt gtgccagcta tcctgagggt 720
ccgaacccgt ggtggggtgg gcaaggcctt ttccacctgc agttcccacc tctgcgttgt 780
ttgtgtgttc tatgggaccc tcttcagtgc ctacctgtgt cctccctcca ttgcctctga 840
agagaaggac attgcagcag ctgcaatgta caccatagtg actcccatgt tgaacccctt 900
tatctatagc ctaaggaaca aggacatgaa gggggcccta aagaggctct tcagtcacag 960
gagtattgtt tcctcttaga tgtggtgaca gcaacattta atgaaaagac ataggcttgg 1020
a 1021

<210> 12
<211> 311
<212> PRT
<213> Homo sapiens

<400> 12
Met Glu Asn Gln Ser Ser Ile Ser Glu Phe Phe Leu Arg Gly Ile Ser
1 5 10 15

Ala Ser Pro Glu Gln Gln Gln Ser Leu Phe Gly Ile Phe Leu Cys Met
20 25 30

Tyr Leu Val Thr Leu Thr Gly Asn Leu Leu Ile Ile Leu Ala Ile Gly
35 40 45

Ser Asp Leu His Leu His Thr Pro Met Tyr Phe Phe Leu Ala Asn Leu
50 55 60

Ser Phe Val Asp Met Gly Leu Thr Ser Ser Thr Val Thr Lys Met Leu
65 70 75 80

Val Asn Ile Gln Thr Arg His His Thr Ile Ser Tyr Thr Gly Cys Leu
85 90 95

Thr Gln Met Tyr Phe Phe Leu Met Phe Gly Asp Leu Asp Ser Phe Phe
100 105 110

Leu Ala Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His Pro Leu
115 120 125

Tyr Tyr Ser Thr Val Met Arg Pro Gln Val Cys Ala Leu Met Leu Ala
130 135 140

Leu Cys Trp Val Leu Thr Asn Ile Val Ala Leu Thr His Thr Phe Leu
145 150 155 160

Met Ala Arg Leu Ser Phe Cys Val Thr Gly Glu Ile Ala His Phe Phe
165 170 175

Cys Asp Ile Thr Pro Val Leu Lys Leu Ser Cys Ser Asp Thr His Ile
180 185 190

Asn Glu Met Met Val Phe Val Leu Gly Gly Thr Val Leu Ile Val Pro
195 200 205

Phe Leu Cys Ile Val Thr Ser Tyr Ile His Ile Val Pro Ala Ile Leu
210 215 220

Arg Val Arg Thr Arg Gly Gly Val Gly Lys Ala Phe Ser Thr Cys Ser
225 230 235 240

Ser His Leu Cys Val Val Cys Val Phe Tyr Gly Thr Leu Phe Ser Ala
245 250 255

Tyr Leu Cys Pro Pro Ser Ile Ala Ser Glu Glu Lys Asp Ile Ala Ala
260 265 270

Ala Ala Met Tyr Thr Ile Val Thr Pro Met Leu Asn Pro Phe Ile Tyr
275 280 285

Ser Leu Arg Asn Lys Asp Met Lys Gly Ala Leu Lys Arg Leu Phe Ser
290 295 300

His Arg Ser Ile Val Ser Ser
305 310

<210> 13
<211> 940
<212> DNA
<213> Homo sapiens

<400> 13
tcatggaaaa ccaatccagc atttctgaat ttttcctccg aggaatatca gcgcctccag 60
agcaacagca gtccctcttc ggaattttcc tgtgtatgta tcttgtcacc ttgactggga 120
acctgctcat catcctggcc attggctctg acctgcacct ccacaccccc atgtactttt 180
tcttggccaa cctgtctttt gttgacatgg gtttaacgtc ctccacagtt accaagatgc 240
tggtgaatat acagactcgg catcacacca tctcctatac gggttgcctc acgcaaatgt 300
atttctttct gatgtttggt gatctagaca gcttcttcct ggctgccatg gcgtatgacc 360
gctatgtggc catttgccac cccctctgct actccacagt catgaggccc caagtctgtg 420
ccctaatgct tgcattgtgc tgggtcctca ccaatatcgt tgccctgact cacacgttcc 480
tcatggctcg gttgtccttc tgtgtgactg gggaaattgc tcacttttтc tgtgacatca 540
ctcctgtcct gaagctgtca tgttctgaca cccacatcaa cgagatgatg gtttttgtct 600
tgggaggcac cgtactcatc gtccccttтt tatgcattgt cacctcctac atccacattg 660
tgccagctat cctgagggtc cgaacccgtg gtggggtggg caaggccttt tccacctgca 720
gttcccacct ctgcgttgtt tgtgtgttct atgggaccct cttcagtgcc tacctgtgtc 780

ctccctccat tgcctctgaa gagaaggaca ttgcagcagc tgcaatgtac accatagtga 840
ctcccatgtt gaaccccttt atctatagcc taaggaacaa ggacatgaag ggggccctaa 900
agaggctctt cagtcacagg agtattgttt cctcttagat 940

<210> 14
<211> 311
<212> PRT
<213> Homo sapiens

<400> 14
Met Glu Asn Gln Ser Ser Ile Ser Glu Phe Phe Leu Arg Gly Ile Ser
1 5 10 15

Ala Pro Pro Glu Gln Gln Gln Ser Leu Phe Gly Ile Phe Leu Cys Met
20 25 30

Tyr Leu Val Thr Leu Thr Gly Asn Leu Leu Ile Ile Leu Ala Ile Gly
35 40 45

Ser Asp Leu His Leu His Thr Pro Met Tyr Phe Phe Leu Ala Asn Leu
50 55 60

Ser Phe Val Asp Met Gly Leu Thr Ser Ser Thr Val Thr Lys Met Leu
65 70 75 80

Val Asn Ile Gln Thr Arg His His Thr Ile Ser Tyr Thr Gly Cys Leu
85 90 95

Thr Gln Met Tyr Phe Phe Leu Met Phe Gly Asp Leu Asp Ser Phe Phe
100 105 110

Leu Ala Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His Pro Leu
115 120 125

Cys Tyr Ser Thr Val Met Arg Pro Gln Val Cys Ala Leu Met Leu Ala
130 135 140

Leu Cys Trp Val Leu Thr Asn Ile Val Ala Leu Thr His Thr Phe Leu
145 150 155 160

Met Ala Arg Leu Ser Phe Cys Val Thr Gly Glu Ile Ala His Phe Phe
165 170 175

Cys Asp Ile Thr Pro Val Leu Lys Leu Ser Cys Ser Asp Thr His Ile
180 185 190

Asn Glu Met Met Val Phe Val Leu Gly Gly Thr Val Leu Ile Val Pro
195 200 205

Phe Leu Cys Ile Val Thr Ser Tyr Ile His Ile Val Pro Ala Ile Leu
210 215 220

Arg Val Arg Thr Arg Gly Gly Val Gly Lys Ala Phe Ser Thr Cys Ser
225 230 235 240

Ser His Leu Cys Val Val Cys Val Phe Tyr Gly Thr Leu Phe Ser Ala

245 250 255

Tyr Leu Cys Pro Pro Ser Ile Ala Ser Glu Glu Lys Asp Ile Ala Ala
260 265 270

Ala Ala Met Tyr Thr Ile Val Thr Pro Met Leu Asn Pro Phe Ile Tyr
275 280 285

Ser Leu Arg Asn Lys Asp Met Lys Gly Ala Leu Lys Arg Leu Phe Ser
290 295 300

His Arg Ser Ile Val Ser Ser
305 310

<210> 15
<211> 1012
<212> DNA
<213> Homo sapiens

<400> 15
ggtttgtttc tgccttttc aatgtccctc tatttccagc agagagaaga ctgtcagcat 60
gaagagggag aatcagagca gtgtgtctga gttcctcctc ctggacctcc ccatctggcc 120
agagcagcag gctgtgttct tcgccctgtt cttgggcatg tgcctgatca cggtgctggg 180
gaacctgctc atcatcctgc tcatccggct ggactctcac cttcacaccc ccatgttctt 240
cttcctcagc cacttggctc tcactgacat ctccctttca tctgtcactg tcccaaagat 300
gttattaagc atgcaaactc aggatcaatc cattctttat gcagggtgtg taactcagat 360
gtatttttc atatttttca ctgatctaga caatttcctt ctcactacaa tggcatacga 420
tcggtatgtg gccatctgtc acccccctccg ctacaccact atcatgaaag agggactgtg 480
taacttacta gtcactgtgt cctggatcct ctcctgtacc aatgccctgt ctcacactct 540
cctcctggcc cagctgtcct tttgtgctga caacaccatc ccccatttct tctgtgatct 600
tgttgcccta ctcaagctct catgctcaga catctccctc aatgagctgg tcattttcac 660
agtgggacag gcagtcatta ctctaccact aatatgcatc ttgatctctt atggccacat 720
tggggtcacc atcctcaagg ctccatctac taagggcatc ttcaaagctt tgtccacctg 780
tggctctcac ctctctgtgg tgtctctgta ttatggcaca attattggac tgtattttct 840
cccctcatcc agtgcctcca gtgacaagga cgtaattgcc tctgtgatgt acacggtgat 900
caccccattg ctgaatccct tcatttatag cctaaggaac agggacataa agggagcccc 960
tggagagact cttcaacagg gcaacagtct tatctcaatg tgatttactc tt 1012

<210> 16
<211> 314
<212> PRT
<213> Homo sapiens

<400> 16
Met Lys Arg Glu Asn Gln Ser Ser Val Ser Glu Phe Leu Leu Leu Asp
1 5 10 15

Leu Pro Ile Trp Pro Glu Gln Gln Ala Val Phe Phe Ala Leu Phe Leu
20 25 30

Gly Met Cys Leu Ile Thr Val Leu Gly Asn Leu Leu Ile Ile Leu Leu
35 40 45

Ile Arg Leu Asp Ser His Leu His Thr Pro Met Phe Phe Phe Leu Ser
50 55 60

```
His Leu Ala Leu Thr Asp Ile Ser Leu Ser Ser Val Thr Val Pro Lys
 65                  70                  75                  80

Met Leu Leu Ser Met Gln Thr Gln Asp Gln Ser Ile Leu Tyr Ala Gly
                85                  90                  95

Cys Val Thr Gln Met Tyr Phe Phe Ile Phe Phe Thr Asp Leu Asp Asn
            100                 105                 110

Phe Leu Leu Thr Thr Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His
        115                 120                 125

Pro Leu Arg Tyr Thr Thr Ile Met Lys Glu Gly Leu Cys Asn Leu Leu
    130                 135                 140

Val Thr Val Ser Trp Ile Leu Ser Cys Thr Asn Ala Leu Ser His Thr
145                 150                 155                 160

Leu Leu Leu Ala Gln Leu Ser Phe Cys Ala Asp Asn Thr Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Leu Val Ala Leu Leu Lys Leu Ser Cys Ser Asp Ile
            180                 185                 190

Ser Leu Asn Glu Leu Val Ile Phe Thr Val Gly Gln Ala Val Ile Thr
        195                 200                 205

Leu Pro Leu Ile Cys Ile Leu Ile Ser Tyr Gly His Ile Gly Val Thr
    210                 215                 220

Ile Leu Lys Ala Pro Ser Thr Lys Gly Ile Phe Lys Ala Leu Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Tyr Tyr Gly Thr Ile Ile
                245                 250                 255

Gly Leu Tyr Phe Leu Pro Ser Ser Ser Ala Ser Ser Asp Lys Asp Val
            260                 265                 270

Ile Ala Ser Val Met Tyr Thr Val Ile Thr Pro Leu Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Arg Asp Ile Lys Gly Ala Pro Gly Glu Thr
    290                 295                 300

Leu Gln Gln Gly Asn Ser Leu Ile Ser Met
305                 310
```

<210> 17
<211> 971
<212> DNA
<213> Homo sapiens

<400> 17
catgaagagg gagaatcaga gcagtgtgtc tgagttcctc ctcctggacc tccccatctg 60

gccagagcag caggctgtgt tcttcaccct gttcttgggc atgtacctga tcacggtgct 120
ggggaacctg ctcatcatcc tgctcatccg gctggactct caccttcaca cccccatgtt 180
cttcttcctc agccacttgg ctctcactga catctccctt tcatctgtca ctgtcccaaa 240
gatgttatta agcatgcaaa ctcaggatca atccattctt tatgcagggt gtgtaactca 300
gatgtatttt ttcatatttt tcactgatct agacaatttc cttctcactt caatggcata 360
cgatcggtat gtggccatct gtcacccccc ccgctacacc actatcatga aagagggact 420
gtgtaactta ctagtcactg tgtcctggat cctctcctgt accaatgccc tgtctcacac 480
tctcctcctg gcccagctgt ccttttgtgc tgacaacacc atccccccatt tcttctgtga 540
tcttgttgcc ctactcaagc tctcatgctc agacatctcc ctcaatgagc tggtcatttt 600
cacagtggga caggcagtca ttactctacc actaatatgc atcttgatct cttatggcca 660
cattggggtc accatcctca aggctccatc tactaagggc atcttcaaag ctttgtccac 720
ctgtggctct cacctctctg tggtgtctct gtattatggc acaattattg gactgtattt 780
tctcccctca tccagtgcct ccagtgacaa ggacgtaatt gcctctgtga tgtacacggt 840
gatcacccca ttgctgaatc ccttcattta tagcctaagg aacagggaca taaagggagc 900
cctggagaga ctcttcaaca gggcaacagt cttatctcaa tgacatttac tcttctttat 960
aacagacata t 971

<210> 18
<211> 313
<212> PRT
<213> Homo sapiens

<400> 18
Met Lys Arg Glu Asn Gln Ser Ser Val Ser Glu Phe Leu Leu Leu Asp
1 5 10 15

Leu Pro Ile Trp Pro Glu Gln Gln Ala Val Phe Phe Thr Leu Phe Leu
20 25 30

Gly Met Tyr Leu Ile Thr Val Leu Gly Asn Leu Leu Ile Ile Leu Leu
35 40 45

Ile Arg Leu Asp Ser His Leu His Thr Pro Met Phe Phe Phe Leu Ser
50 55 60

His Leu Ala Leu Thr Asp Ile Ser Leu Ser Ser Val Thr Val Pro Lys
65 70 75 80

Met Leu Leu Ser Met Gln Thr Gln Asp Gln Ser Ile Leu Tyr Ala Gly
85 90 95

Cys Val Thr Gln Met Tyr Phe Phe Ile Phe Phe Thr Asp Leu Asp Asn
100 105 110

Phe Leu Leu Thr Ser Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His
115 120 125

Pro Leu Arg Tyr Thr Thr Ile Met Lys Glu Gly Leu Cys Asn Leu Leu
130 135 140

Val Thr Val Ser Trp Ile Leu Ser Cys Thr Asn Ala Leu Ser His Thr
145 150 155 160

Leu Leu Leu Ala Gln Leu Ser Phe Cys Ala Asp Asn Thr Ile Pro His
165 170 175

Phe Phe Cys Asp Leu Val Ala Leu Leu Lys Leu Ser Cys Ser Asp Ile
180 185 190

Ser Leu Asn Glu Leu Val Ile Phe Thr Val Gly Gln Ala Val Ile Thr
195 200 205

Leu Pro Leu Ile Cys Ile Leu Ile Ser Tyr Gly His Ile Gly Val Thr
210 215 220

Ile Leu Lys Ala Pro Ser Thr Lys Gly Ile Phe Lys Ala Leu Ser Thr
225 230 235 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Tyr Tyr Gly Thr Ile Ile
245 250 255

Gly Leu Tyr Phe Leu Pro Ser Ser Ser Ala Ser Ser Asp Lys Asp Val
260 265 270

Ile Ala Ser Val Met Tyr Thr Val Ile Thr Pro Leu Leu Asn Pro Phe
275 280 285

Ile Tyr Ser Leu Arg Asn Arg Asp Ile Lys Gly Ala Leu Glu Arg Leu
290 295 300

Phe Asn Arg Ala Thr Val Leu Ser Gln
305 310

<210> 19
<211> 996
<212> DNA
<213> Homo sapiens

<400> 19
tatgagccct gagaaccaga gcagcgtgtc cgagttcctc cttctgggcc tccccatccg 60
gccagagcag caggctgtgt tcttcaccct gttcctgggc atgtacctga ccacggtgct 120
ggggaacctg ctcatcatgc tgctcatcca gctggactct caccttcaca cccccatgta 180
cttcttcctc agccacttgg ctctcactga catctccttt tcatctgtca ctgtccctaa 240
gatgctgatg gacatgcgga ctaagtacaa atcgatcctc tatgaggaat gcatttctca 300
gatgtatttt tttatatttt ttactgacct ggacagcttc cttattacat caatggcata 360
tgaccgatat gttgccatat gtcacccctct ccactacact gtcatcatga gggaagagct 420
ctgtgtcttc ttagtggctg tatcttggat tctgtcttgt gccagctccc tctctcacac 480
ccttctcctg acccggctgt ctttctgtgc tgcgaacacc atcccccatg tcttctgtga 540
ccttgctgcc ctgctcaagc tgtcctgctc agatatcttc ctcaatgagc tggtcatgtt 600
cacagtaggg gtggtggtca ttaccctgcc attcatgtgt atcctggtat catatggcta 660
cattggggcc accatcctga gggtcccttc aaccaaaggg atccacaaag cattgtccac 720
atgtggctcc catctctctg tggtgtctct ctattatggg tcaatatttg gccagtacct 780
tttcccgact gtaagcagtt ctattgacaa ggatgtcatt gtggctctca tgtacacggt 840
ggtcacaccc atgttgaacc cctttatcta cagccttagg aacaggggca tgaaagaggc 900
ccttgggaaa ctcttcagta gagcaacatt tttctcttgg tgacatctga ctttttaaaa 960
aattagaatc tcattttggt ttatctcatg tatttg 996

<210> 20
<211> 313
<212> PRT
<213> Homo sapiens

<400> 20

```
Met Ser Pro Glu Asn Gln Ser Ser Val Ser Glu Phe Leu Leu Leu Gly
  1               5                  10                  15

Leu Pro Ile Arg Pro Glu Gln Gln Ala Val Phe Phe Thr Leu Phe Leu
             20                  25                  30

Gly Met Tyr Leu Thr Thr Val Leu Gly Asn Leu Leu Ile Met Leu Leu
         35                  40                  45

Ile Gln Leu Asp Ser His Leu His Thr Pro Met Tyr Phe Phe Leu Ser
     50                  55                  60

His Leu Ala Leu Thr Asp Ile Ser Phe Ser Ser Val Thr Val Pro Lys
 65                  70                  75                  80

Met Leu Met Asp Met Arg Thr Lys Tyr Lys Ser Ile Leu Tyr Glu Glu
                 85                  90                  95

Cys Ile Ser Gln Met Tyr Phe Phe Ile Phe Phe Thr Asp Leu Asp Ser
            100                 105                 110

Phe Leu Ile Thr Ser Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His
        115                 120                 125

Pro Leu His Tyr Thr Val Ile Met Arg Glu Glu Leu Cys Val Phe Leu
    130                 135                 140

Val Ala Val Ser Trp Ile Leu Ser Cys Ala Ser Ser Leu Ser His Thr
145                 150                 155                 160

Leu Leu Leu Thr Arg Leu Ser Phe Cys Ala Ala Asn Thr Ile Pro His
                165                 170                 175

Val Phe Cys Asp Leu Ala Ala Leu Leu Lys Leu Ser Cys Ser Asp Ile
            180                 185                 190

Phe Leu Asn Glu Leu Val Met Phe Thr Val Gly Val Val Val Ile Thr
        195                 200                 205

Leu Pro Phe Met Cys Ile Leu Val Ser Tyr Gly Tyr Ile Gly Ala Thr
    210                 215                 220

Ile Leu Arg Val Pro Ser Thr Lys Gly Ile His Lys Ala Leu Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Tyr Tyr Gly Ser Ile Phe
                245                 250                 255

Gly Gln Tyr Leu Phe Pro Thr Val Ser Ser Ser Ile Asp Lys Asp Val
            260                 265                 270

Ile Val Ala Leu Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Arg Gly Met Lys Glu Ala Leu Gly Lys Leu
```

290 295 300

Phe Ser Arg Ala Thr Phe Phe Ser Trp
305 310

<210> 21
<211> 978
<212> DNA
<213> Homo sapiens

<400> 21
tatgagccct gagaaccaga gcagcgtgtc cgagttcctc cttctgggcc tccccatccg 60
gccagagcag caggctgtgt tcttcaccct gttcctgggc atgtacctga ccacggtgct 120
ggggaacctg ctcatcatgc tgctcatcca gctggactct caccttcaca cccccatgta 180
cttcttcctc agccacttgg ctctcactga catctccttt tcatctgtca ctgtccctaa 240
gatgctgatg gacatgcgga ctaagtacaa atcgatcctc tatgaggaat gcatttctca 300
gatgtatttt tttatatttt ttactgacct ggacagcttc cttattacat caatggcata 360
tgaccgatat gttgccatat gtcaccctct ccactacact gtcatcatga gggaagagct 420
ctgtgtcttc ttagtggctg tatcttggat tctgtcttgt gccagctccc tctctcacac 480
ccttctcctg acccggctgt ctttctgtgc tgcgaacacc atcccccatg tcttctgtga 540
ccttgctgcc ctgctcaagc tgtcctgctc agatatcttc ctcaatgagc tggtcatgtt 600
cacagtaggg gtggtggtca ttaccctgcc attcatgtgt atcctggtat catatggcta 660
cattggggcc accatcctga gggtcccttc aaccaaaggg atccacaaag cattgtccac 720
atgtggctcc catctctctg tggtgtctct ctattatggg tcaatatttg gccagtacct 780
tttcccgact gtaagcagtt ctattgacaa ggatgtcatt gtggctctca tgtacacggt 840
ggtcacaccc atgttgaacc cctttatcta cagccttagg aacagggaca tgaaagaggc 900
ccttgggaaa ctcttcagta gagcaacatt tttctcttgg tgacatctga ctttttaaaa 960
aattagaatc tcattttg 978

<210> 22
<211> 313
<212> PRT
<213> Homo sapiens

<400> 22
Met Ser Pro Glu Asn Gln Ser Ser Val Ser Glu Phe Leu Leu Leu Gly
1 5 10 15

Leu Pro Ile Arg Pro Glu Gln Gln Ala Val Phe Phe Thr Leu Phe Leu
20 25 30

Gly Met Tyr Leu Thr Thr Val Leu Gly Asn Leu Leu Ile Met Leu Leu
35 40 45

Ile Gln Leu Asp Ser His Leu His Thr Pro Met Tyr Phe Phe Leu Ser
50 55 60

His Leu Ala Leu Thr Asp Ile Ser Phe Ser Ser Val Thr Val Pro Lys
65 70 75 80

Met Leu Met Asp Met Arg Thr Lys Tyr Lys Ser Ile Leu Tyr Glu Glu
85 90 95

Cys Ile Ser Gln Met Tyr Phe Phe Ile Phe Phe Thr Asp Leu Asp Ser
100 105 110

```
Phe Leu Ile Thr Ser Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His
        115                 120                 125

Pro Leu His Tyr Thr Val Ile Met Arg Glu Glu Leu Cys Val Phe Leu
    130                 135                 140

Val Ala Val Ser Trp Ile Leu Ser Cys Ala Ser Ser Leu Ser His Thr
145                 150                 155                 160

Leu Leu Leu Thr Arg Leu Ser Phe Cys Ala Ala Asn Thr Ile Pro His
                165                 170                 175

Val Phe Cys Asp Leu Ala Ala Leu Leu Lys Leu Ser Cys Ser Asp Ile
            180                 185                 190

Phe Leu Asn Glu Leu Val Met Phe Thr Val Gly Val Val Val Ile Thr
        195                 200                 205

Leu Pro Phe Met Cys Ile Leu Val Ser Tyr Gly Tyr Ile Gly Ala Thr
    210                 215                 220

Ile Leu Arg Val Pro Ser Thr Lys Gly Ile His Lys Ala Leu Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Tyr Tyr Gly Ser Ile Phe
                245                 250                 255

Gly Gln Tyr Leu Phe Pro Thr Val Ser Ser Ser Ile Asp Lys Asp Val
            260                 265                 270

Ile Val Ala Leu Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Arg Asp Met Lys Glu Ala Leu Gly Lys Leu
    290                 295                 300

Phe Ser Arg Ala Thr Phe Phe Ser Trp
305                 310
```

<210> 23
<211> 925
<212> DNA
<213> Homo sapiens

<400> 23
tgtccgagtt cctccttctg ggcctcccca tccggccaga gcagcaggct gtgttcttca 60
ccctgttcct gggcatgtac ctgaccacgg tgctggggaa cctgctcatc atgctgctca 120
tccagctgga ctctcacctt cacacccccа tgtacttctt cctcagccac ttggctctca 180
ctgacatctc cttttcatct gtcactgtcc ctaagatgct gatggacatg cggactaagt 240
acaaatcgat cctctatgag gaatgcattt ctcagatgta ttttttttata ttttttactg 300
acctggacag cttccttatt acatcaatga catatgaccg atatgttgcc atatgtcacc 360
ctctccacta cactgtcatc atgagggaag agctctgtgt cttcttagtg gctgtatctt 420
ggattctgtc ttgtgccagc tccctctctc acacccttct cctgacccgg ctgtctttct 480
gtgctgcgaa caccatcccc catgtcttct gtgaccttgc tgccctgctc aagctgtcct 540
gctcagatat cttcctcaat gagctggtca tgttcacagt aggggtggtg gtcattaccc 600

agccattcat gtgtatcctg gtatcatatg gctacattgg ggccaccatc ctgagggtcc 660
cttcaaccaa agggatccac aaagcattgt ccacatgtgg ctcccatctc tctgtggtgt 720
ctctctatta tgggtcaata tttggccagt acctttccc gactgtaagc agttctattg 780
acaaggatgt cattgtggct ctcatgtaca cggtggtcac acccatgttg aacccccttta 840
tctacagcct taggaacagg gacatgaaag aggcccttgg gaaactcttc agtagagcaa 900
cattttctc ttggtgacat ctgac 925

<210> 24
<211> 304
<212> PRT
<213> Homo sapiens

<400> 24
Ser Glu Phe Leu Leu Leu Gly Leu Pro Ile Arg Pro Glu Gln Gln Ala
1 5 10 15

Val Phe Phe Thr Leu Phe Leu Gly Met Tyr Leu Thr Thr Val Leu Gly
20 25 30

Asn Leu Leu Ile Met Leu Leu Ile Gln Leu Asp Ser His Leu His Thr
35 40 45

Pro Met Tyr Phe Phe Leu Ser His Leu Ala Leu Thr Asp Ile Ser Phe
50 55 60

Ser Ser Val Thr Val Pro Lys Met Leu Met Asp Met Arg Thr Lys Tyr
65 70 75 80

Lys Ser Ile Leu Tyr Glu Glu Cys Ile Ser Gln Met Tyr Phe Phe Ile
85 90 95

Phe Phe Thr Asp Leu Asp Ser Phe Leu Ile Thr Ser Met Thr Tyr Asp
100 105 110

Arg Tyr Val Ala Ile Cys His Pro Leu His Tyr Thr Val Ile Met Arg
115 120 125

Glu Glu Leu Cys Val Phe Leu Val Ala Val Ser Trp Ile Leu Ser Cys
130 135 140

Ala Ser Ser Leu Ser His Thr Leu Leu Leu Thr Arg Leu Ser Phe Cys
145 150 155 160

Ala Ala Asn Thr Ile Pro His Val Phe Cys Asp Leu Ala Ala Leu Leu
165 170 175

Lys Leu Ser Cys Ser Asp Ile Phe Leu Asn Glu Leu Val Met Phe Thr
180 185 190

Val Gly Val Val Val Ile Thr Gln Pro Phe Met Cys Ile Leu Val Ser
195 200 205

Tyr Gly Tyr Ile Gly Ala Thr Ile Leu Arg Val Pro Ser Thr Lys Gly
210 215 220

Ile His Lys Ala Leu Ser Thr Cys Gly Ser His Leu Ser Val Val Ser

225 230 235 240

Leu Tyr Tyr Gly Ser Ile Phe Gly Gln Tyr Leu Phe Pro Thr Val Ser
245 250 255

Ser Ser Ile Asp Lys Asp Val Ile Val Ala Leu Met Tyr Thr Val Val
260 265 270

Thr Pro Met Leu Asn Pro Phe Ile Tyr Ser Leu Arg Asn Arg Asp Met
275 280 285

Lys Glu Ala Leu Gly Lys Leu Phe Ser Arg Ala Thr Phe Phe Ser Trp
290 295 300

<210> 25
<211> 936
<212> DNA
<213> Homo sapiens

<400> 25
tccatgggaa tgtccaacct gacaagactc tctgaattta ttctcttggg actctcctct 60
cggtctgaag accagaggcc actctttgcc ctctttctta tcatatacct ggtcactttg 120
atgggaaatc tgctcatcat cttggctatc cactctgatc ctcgacttca aaaccctatg 180
tatttttcc taagcatctt gtcctttgct gatatttgct acacaacagt catagtccca 240
aagatgctcg tgaacttctt atcagagaaa aagaccattt cctatgctga atgtctggca 300
cagatgtatt tcttcctggt ttttggaaac atagatagtt atctcctggc ggctatggcc 360
atcaaccgct gtgtagccat ttgtaaccca ttccattatg tcactgttat gaaccgcaga 420
tgctgtgtgt tgctactagc attccccatc actttctcct atttccactc cctcctacat 480
gtcctcctgg tgaatcggct caccttttgt acatcaaatg ttatccatca ttttttttgt 540
gatgtcaacc ctgtgctgaa actgtcctgc tcctccacct ttgtcaatga aattgtggcc 600
atgacagaag ggctggcctc tgtgatggct ccatttgtct gtatcatcat ctcttatcta 660
agaattctca tcgctgttct caagattccc tcagcagctg gaaaacacaa agccttctcc 720
acctgcagct cccatctcac tgtggtgatt ctgttttatg ggagtattag ctatgtctat 780
ttgcagcctt tgtccagcta tactgtcaag gaccgaatag caacaatcaa ctacactgtg 840
ttgacatcag tgttgaaccc atttatctac agtttaagaa acaaagacat gaaacggggc 900
ttacagaaat tgataaacaa gattaagtct caatga 936

<210> 26
<211> 310
<212> PRT
<213> Homo sapiens

<400> 26
Met Gly Met Ser Asn Leu Thr Arg Leu Ser Glu Phe Ile Leu Leu Gly
1 5 10 15

Leu Ser Ser Arg Ser Glu Asp Gln Arg Pro Leu Phe Ala Leu Phe Leu
20 25 30

Ile Ile Tyr Leu Val Thr Leu Met Gly Asn Leu Leu Ile Ile Leu Ala
35 40 45

```
Ile His Ser Asp Pro Arg Leu Gln Asn Pro Met Tyr Phe Phe Leu Ser
     50                  55                  60

Ile Leu Ser Phe Ala Asp Ile Cys Tyr Thr Thr Val Ile Val Pro Lys
 65                  70                  75                  80

Met Leu Val Asn Phe Leu Ser Glu Lys Lys Thr Ile Ser Tyr Ala Glu
                 85                  90                  95

Cys Leu Ala Gln Met Tyr Phe Phe Leu Val Phe Gly Asn Ile Asp Ser
            100                 105                 110

Tyr Leu Leu Ala Ala Met Ala Ile Asn Arg Cys Val Ala Ile Cys Asn
        115                 120                 125

Pro Phe His Tyr Val Thr Val Met Asn Arg Arg Cys Cys Val Leu Leu
    130                 135                 140

Leu Ala Phe Pro Ile Thr Phe Ser Tyr Phe His Ser Leu Leu His Val
145                 150                 155                 160

Leu Leu Val Asn Arg Leu Thr Phe Cys Thr Ser Asn Val Ile His His
                165                 170                 175

Phe Phe Cys Asp Val Asn Pro Val Leu Lys Leu Ser Cys Ser Ser Thr
            180                 185                 190

Phe Val Asn Glu Ile Val Ala Met Thr Glu Gly Leu Ala Ser Val Met
        195                 200                 205

Ala Pro Phe Val Cys Ile Ile Ile Ser Tyr Leu Arg Ile Leu Ile Ala
    210                 215                 220

Val Leu Lys Ile Pro Ser Ala Ala Gly Lys His Lys Ala Phe Ser Thr
225                 230                 235                 240

Cys Ser Ser His Leu Thr Val Val Ile Leu Phe Tyr Gly Ser Ile Ser
                245                 250                 255

Tyr Val Tyr Leu Gln Pro Leu Ser Ser Tyr Thr Val Lys Asp Arg Ile
            260                 265                 270

Ala Thr Ile Asn Tyr Thr Val Leu Thr Ser Val Leu Asn Pro Phe Ile
        275                 280                 285

Tyr Ser Leu Arg Asn Lys Asp Met Lys Arg Gly Leu Gln Lys Leu Ile
    290                 295                 300

Asn Lys Ile Lys Ser Gln
305                 310
```

<210> 27
<211> 936
<212> DNA
<213> Homo sapiens

<400> 27
tccatgggaa tgtccaacct gacaagactc tctgaattta ttctcttggg actctcctct 60
cggtctgaag accagaggcc actctttgcc ctctttctta tcatatacct ggtcactttg 120
atgggaaatc tgctcatcat cttggctatc cactctgatc ctcgacttca aaaccctatg 180
tatttttcc taagcatctt gtcctttgct gatatttgct acacaacagt catagtccca 240
aagatgctcg tgaacttctt atcagagaaa aagaccattt cctatgctga atgtctggca 300
cagatgtatt tcttcctggt ttttggaaac atagatagtt atctcctggc ggctatggcc 360
atcaaccgct gtgtagccat ttgtaaccca ttccattatg tcactgttat gaaccgcaga 420
tgctgtgtgt tgctactagc attccccatc actttctcct atttccactc tctcctacat 480
gtcctcctgg tgaatcggct caccttttgt acatcaaatg ttatccatca ttttttttgt 540
gatgtcaacc ctgtgctgaa actgtcctgc tcctccacct ttgtcaatga aattgtggcc 600
atgacagaag ggctggcctc tgtgatggct ccatttgtct gtatcatcat ctcttatcta 660
agaattctca tcgctgttct caagattccc tcagcagctg gaaaacacaa agccttctcc 720
acctgcagct cccatctcac tgtggtgatt ctgttttatg ggagtattag ctatgtctat 780
ttgcagcctt tgtccagcta tactgtcaag gaccgaatag caacaatcaa ctacactgtg 840
ttgacatcag tgttgaaccc atttatctac agtttaagaa acaaagacat gaaacggggc 900
ttacagaaat tgataaacaa gattaagtct caatga 936

<210> 28
<211> 310
<212> PRT
<213> Homo sapiens

<400> 28
Met Gly Met Ser Asn Leu Thr Arg Leu Ser Glu Phe Ile Leu Leu Gly
1 5 10 15

Leu Ser Ser Arg Ser Glu Asp Gln Arg Pro Leu Phe Ala Leu Phe Leu
20 25 30

Ile Ile Tyr Leu Val Thr Leu Met Gly Asn Leu Leu Ile Ile Leu Ala
35 40 45

Ile His Ser Asp Pro Arg Leu Gln Asn Pro Met Tyr Phe Phe Leu Ser
50 55 60

Ile Leu Ser Phe Ala Asp Ile Cys Tyr Thr Thr Val Ile Val Pro Lys
65 70 75 80

Met Leu Val Asn Phe Leu Ser Glu Lys Lys Thr Ile Ser Tyr Ala Glu
85 90 95

Cys Leu Ala Gln Met Tyr Phe Phe Leu Val Phe Gly Asn Ile Asp Ser
100 105 110

Tyr Leu Leu Ala Ala Met Ala Ile Asn Arg Cys Val Ala Ile Cys Asn
115 120 125

Pro Phe His Tyr Val Thr Val Met Asn Arg Arg Cys Cys Val Leu Leu
130 135 140

Leu Ala Phe Pro Ile Thr Phe Ser Tyr Phe His Ser Leu Leu His Val
145 150 155 160

Leu Leu Val Asn Arg Leu Thr Phe Cys Thr Ser Asn Val Ile His His
165 170 175

Phe Phe Cys Asp Val Asn Pro Val Leu Lys Leu Ser Cys Ser Ser Thr
180 185 190

Phe Val Asn Glu Ile Val Ala Met Thr Glu Gly Leu Ala Ser Val Met
195 200 205

Ala Pro Phe Val Cys Ile Ile Ile Ser Tyr Leu Arg Ile Leu Ile Ala
210 215 220

Val Leu Lys Ile Pro Ser Ala Ala Gly Lys His Lys Ala Phe Ser Thr
225 230 235 240

Cys Ser Ser His Leu Thr Val Val Ile Leu Phe Tyr Gly Ser Ile Ser
245 250 255

Tyr Val Tyr Leu Gln Pro Leu Ser Ser Tyr Thr Val Lys Asp Arg Ile
260 265 270

Ala Thr Ile Asn Tyr Thr Val Leu Thr Ser Val Leu Asn Pro Phe Ile
275 280 285

Tyr Ser Leu Arg Asn Lys Asp Met Lys Arg Gly Leu Gln Lys Leu Ile
290 295 300

Asn Lys Ile Lys Ser Gln
305 310

<210> 29
<211> 1042
<212> DNA
<213> Homo sapiens

<400> 29
atgacaaaac ttcttgttta taactgagcc caagtcaatg gaaagaatca accacaccag 60
cagtgtctcc gagtttatcc tcctgggact ctcctcccgg cctgaggacc aaaagacact 120
ctttgttctc ttcctcatcg tgtacctggt caccataaca gggaacctgc tcatcatcct 180
ggccattcgc ttcaaccccc atcttcagac ccctatgtat ttcttcttga gtttttgtc 240
tctcactgat atttgcttta caacaagcgt tgtccccaag atgctgatga acttcctgtc 300
agaaaagaag accatctcct atgctgggtg tctgacacag atgtattttc tctatgcctt 360
gggcaacagt gacagctgcc ttctggcaat catggccttt gaccgctatg tggccgtctg 420
tgacccttc cactatgtca ccaccatgag ccaccaccac tgtgtcctgc tggtggcctt 480
ctcctgctca tttcctcacc tccactcact cctgcacaca cttctgctga atcgtctcac 540
cttctgtgac tccaatgtta tccaccactt tctctgtgac ctcagccctg tgctgaaatt 600
gtcctgctcg tccatatttg tcaatgaaat tgtgcagatg acagaagcac ctattgtttt 660
ggtgactcct tttctctgca ttgctttctc ttatatacga atcctcacta cagttctcaa 720
gattccctct acttctggga aacgcaaagc cttctccacc tgtggttttt acctcaccgt 780
ggtgacgctc ttttatggaa gcatcttctg tgtctattta cagcccccat ccacctacgc 840
tgtcaaggac cacgtggcaa caattgttta cacagttttg tcatccatgc tcaatccttt 900
tatctacagc ctgagaaaca aagacctgaa acagggcctg aggaagctta tgagcaagag 960
atcctaggaa gcaccctctt gaaaaactcg taagtggaat ctgctcaact tggacgtgtt 1020
ttctactggt ttctggtgaa ca 1042

<210> 30
<211> 309

<212> PRT
<213> Homo sapiens

<400> 30

```
Met Glu Arg Ile Asn His Thr Ser Ser Val Ser Glu Phe Ile Leu Leu
  1               5                  10                  15

Gly Leu Ser Ser Arg Pro Glu Asp Gln Lys Thr Leu Phe Val Leu Phe
             20                  25                  30

Leu Ile Val Tyr Leu Val Thr Ile Thr Gly Asn Leu Leu Ile Ile Leu
         35                  40                  45

Ala Ile Arg Phe Asn Pro His Leu Gln Thr Pro Met Tyr Phe Phe Leu
     50                  55                  60

Ser Phe Leu Ser Leu Thr Asp Ile Cys Phe Thr Thr Ser Val Val Pro
 65                  70                  75                  80

Lys Met Leu Met Asn Phe Leu Ser Glu Lys Lys Thr Ile Ser Tyr Ala
                 85                  90                  95

Gly Cys Leu Thr Gln Met Tyr Phe Leu Tyr Ala Leu Gly Asn Ser Asp
            100                 105                 110

Ser Cys Leu Leu Ala Ile Met Ala Phe Asp Arg Tyr Val Ala Val Cys
        115                 120                 125

Asp Pro Phe His Tyr Val Thr Thr Met Ser His His His Cys Val Leu
    130                 135                 140

Leu Val Ala Phe Ser Cys Ser Phe Pro His Leu His Ser Leu Leu His
145                 150                 155                 160

Thr Leu Leu Leu Asn Arg Leu Thr Phe Cys Asp Ser Asn Val Ile His
                165                 170                 175

His Phe Leu Cys Asp Leu Ser Pro Val Leu Lys Leu Ser Cys Ser Ser
            180                 185                 190

Ile Phe Val Asn Glu Ile Val Gln Met Thr Glu Ala Pro Ile Val Leu
        195                 200                 205

Val Thr Pro Phe Leu Cys Ile Ala Phe Ser Tyr Ile Arg Ile Leu Thr
    210                 215                 220

Thr Val Leu Lys Ile Pro Ser Thr Ser Gly Lys Arg Lys Ala Phe Ser
225                 230                 235                 240

Thr Cys Gly Phe Tyr Leu Thr Val Val Thr Leu Phe Tyr Gly Ser Ile
                245                 250                 255

Phe Cys Val Tyr Leu Gln Pro Pro Ser Thr Tyr Ala Val Lys Asp His
            260                 265                 270

Val Ala Thr Ile Val Tyr Thr Val Leu Ser Ser Met Leu Asn Pro Phe
        275                 280                 285
```

Ile Tyr Ser Leu Arg Asn Lys Asp Leu Lys Gln Gly Leu Arg Lys Leu
290 295 300

Met Ser Lys Arg Ser
305

<210> 31
<211> 1033
<212> DNA
<213> Homo sapiens

<400> 31
cttcttgttt ataactgagc ccaagtcaat ggaaagaatc aaccatacca gcagtgtctc 60
cgagtttatc ctcctgggac tctcctcccg gcctgaggac caaaagccac tctttgttct 120
cttcctcatc gtgtacctgg tcaccataac agggaaccag ctcatcatcc tggccattcg 180
cttcaacccc catcttcaga cccctatgta tttcttcttg agtttttttgt ctctcactga 240
tatttgcttt acaacaagcg ttgtccccaa gatgctgatg aacttcctgt cagaaaagaa 300
gaccatctcc tatgctgggt gtctgacaca gatgtatttt ctctatgcct tgggcaacag 360
tgacagctgc cttctggcaa tcatggcctt tgaccgctat gtggccgtct gtgacccttt 420
ccactatgtc accaccatga gccaccacca ctgtgtcctg ctggtggcct tctcctgctc 480
atttcctcac ctccactcac tcctgcacac acttctgctg aatcgtctca ccttctgtga 540
ctccaatgtt atccaccact ttctctgtga cctcagccct gtgctgaaat tgtcctgctc 600
gtccatattt gtcaatgaaa ttgtgcagat gacagaagca cctattgttt tggtgactcc 660
ttttctctgc attgctttct cttatatacg aatcctcact acagttctca agattccctc 720
tacttctggg aaacgcaaag ccttctccac ctgtggtttt tacctcaccg tggtgacgct 780
cttttatgga agcatcttct gtgtctattt acagccccca tccacctacg ctgtcaagga 840
ccacgtggca acaattgttt acacagtttt gtcatccatg ctcaatcctt ttatctacag 900
cctgagaaac aaagacctga aacagggcct gaggaagctt atgagcaaga gatcctagga 960
agcaccctct tgaaaaactc gtaagtggaa tctgctcaac ttggacgtgt tttctactgg 1020
tttctggtga aca 1033

<210> 32
<211> 309
<212> PRT
<213> Homo sapiens

<400> 32
Met Glu Arg Ile Asn His Thr Ser Ser Val Ser Glu Phe Ile Leu Leu
1 5 10 15

Gly Leu Ser Ser Arg Pro Glu Asp Gln Lys Pro Leu Phe Val Leu Phe
20 25 30

Leu Ile Val Tyr Leu Val Thr Ile Thr Gly Asn Gln Leu Ile Ile Leu
35 40 45

Ala Ile Arg Phe Asn Pro His Leu Gln Thr Pro Met Tyr Phe Phe Leu
50 55 60

Ser Phe Leu Ser Leu Thr Asp Ile Cys Phe Thr Thr Ser Val Val Pro
65 70 75 80

Lys Met Leu Met Asn Phe Leu Ser Glu Lys Lys Thr Ile Ser Tyr Ala
85 90 95

```
Gly Cys Leu Thr Gln Met Tyr Phe Leu Tyr Ala Leu Gly Asn Ser Asp
            100                 105                 110

Ser Cys Leu Leu Ala Ile Met Ala Phe Asp Arg Tyr Val Ala Val Cys
        115                 120                 125

Asp Pro Phe His Tyr Val Thr Thr Met Ser His His His Cys Val Leu
    130                 135                 140

Leu Val Ala Phe Ser Cys Ser Phe Pro His Leu His Ser Leu Leu His
145                 150                 155                 160

Thr Leu Leu Leu Asn Arg Leu Thr Phe Cys Asp Ser Asn Val Ile His
                165                 170                 175

His Phe Leu Cys Asp Leu Ser Pro Val Leu Lys Leu Ser Cys Ser Ser
            180                 185                 190

Ile Phe Val Asn Glu Ile Val Gln Met Thr Glu Ala Pro Ile Val Leu
        195                 200                 205

Val Thr Pro Phe Leu Cys Ile Ala Phe Ser Tyr Ile Arg Ile Leu Thr
    210                 215                 220

Thr Val Leu Lys Ile Pro Ser Thr Ser Gly Lys Arg Lys Ala Phe Ser
225                 230                 235                 240

Thr Cys Gly Phe Tyr Leu Thr Val Val Thr Leu Phe Tyr Gly Ser Ile
                245                 250                 255

Phe Cys Val Tyr Leu Gln Pro Pro Ser Thr Tyr Ala Val Lys Asp His
            260                 265                 270

Val Ala Thr Ile Val Tyr Thr Val Leu Ser Ser Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Lys Asp Leu Lys Gln Gly Leu Arg Lys Leu
    290                 295                 300

Met Ser Lys Arg Ser
305
```

<210> 33
<211> 969
<212> DNA
<213> Homo sapiens

<400> 33

```
agaatcaacc acaccagcag tgtctccgag tttatcctcc tgggactctc ctcccggcct 60
gaggaccaaa agacactctt tgttctcttc ctcatcgtgt acctggtcac cataacaggg 120
aacctgctca tcatcctggc cattcgcttc aacccccatc ttcagacccc tatgtatttc 180
ttcttgagtt ttctgtctct cactgatatt tgctttacaa caagcgttgt ccccaagatg 240
ctgatgaact tcctgtcaga aaagaagacc atctcctatg ctgggtgtct gacacagatg 300
tattttctct atgccttggg caacagtgac agctgccttc tggcagtcat ggcctttgac 360
cgctatgtgg ccgtctgtga ccctttccac tatgtcacca ccatgagcca ccaccactgt 420
```

gtcctgctgg tggccttctc ctgctcattt cctcacctcc actcactcct gcacacactt 480
ctgctgaatc gtctcacctt ctgtgactcc aatgttatcc accactttct ctgtgacctc 540
agccctgtgc tgaaattgtc ctgctcttcc atatttgtca atgaaattgt gcagatgaca 600
gaagcaccta ttgttttggt gactcgtttt ctctgcattg ctttctctta tatacgaatc 660
ctcactacag ttctcaagat tccctctact tctgggaaac gcaaagcctt ctccacctgt 720
ggtttttacc tcaccgtggt gacgctcttt tatggaagca tcttctgtgt ctatttacag 780
cccccatcca cctacgctgt caaggaccac gtggcaacaa ttgtttacac agttttgtca 840
tccatgctca atccttttat ctacagcctg agaaacaaag acctgaaaca gggcctgagg 900
aagcttatga gcaagagatc ctaggaagca ccctcttgaa aaactcgtaa gtggaatctg 960
ctcaacttg 969

<210> 34
<211> 307
<212> PRT
<213> Homo sapiens

<400> 34
Arg Ile Asn His Thr Ser Ser Val Ser Glu Phe Ile Leu Leu Gly Leu
1 5 10 15

Ser Ser Arg Pro Glu Asp Gln Lys Thr Leu Phe Val Leu Phe Leu Ile
20 25 30

Val Tyr Leu Val Thr Ile Thr Gly Asn Leu Leu Ile Ile Leu Ala Ile
35 40 45

Arg Phe Asn Pro His Leu Gln Thr Pro Met Tyr Phe Phe Leu Ser Phe
50 55 60

Leu Ser Leu Thr Asp Ile Cys Phe Thr Thr Ser Val Val Pro Lys Met
65 70 75 80

Leu Met Asn Phe Leu Ser Glu Lys Lys Thr Ile Ser Tyr Ala Gly Cys
85 90 95

Leu Thr Gln Met Tyr Phe Leu Tyr Ala Leu Gly Asn Ser Asp Ser Cys
100 105 110

Leu Leu Ala Val Met Ala Phe Asp Arg Tyr Val Ala Val Cys Asp Pro
115 120 125

Phe His Tyr Val Thr Thr Met Ser His His His Cys Val Leu Leu Val
130 135 140

Ala Phe Ser Cys Ser Phe Pro His Leu His Ser Leu Leu His Thr Leu
145 150 155 160

Leu Leu Asn Arg Leu Thr Phe Cys Asp Ser Asn Val Ile His His Phe
165 170 175

Leu Cys Asp Leu Ser Pro Val Leu Lys Leu Ser Cys Ser Ser Ile Phe
180 185 190

Val Asn Glu Ile Val Gln Met Thr Glu Ala Pro Ile Val Leu Val Thr
195 200 205

Arg Phe Leu Cys Ile Ala Phe Ser Tyr Ile Arg Ile Leu Thr Thr Val
210 215 220

Leu Lys Ile Pro Ser Thr Ser Gly Lys Arg Lys Ala Phe Ser Thr Cys
225 230 235 240

Gly Phe Tyr Leu Thr Val Val Thr Leu Phe Tyr Gly Ser Ile Phe Cys
245 250 255

Val Tyr Leu Gln Pro Pro Ser Thr Tyr Ala Val Lys Asp His Val Ala
260 265 270

Thr Ile Val Tyr Thr Val Leu Ser Ser Met Leu Asn Pro Phe Ile Tyr
275 280 285

Ser Leu Arg Asn Lys Asp Leu Lys Gln Gly Leu Arg Lys Leu Met Ser
290 295 300

Lys Arg Ser
305

<210> 35
<211> 955
<212> DNA
<213> Homo sapiens

<400> 35
tcaatggaaa gaatcaacca caccagcagt gtctccgagt ttatcctcct gggactctcc 60
tcccggcctg aggaccaaaa gacactcttt gttctcttcc tcatcgtgta cctggtcacc 120
ataacaggga acctgctcat catcctggcc attcgcttca acccccatct tcagacccct 180
atgtatttct tcttgagttt tctgtctctc actgatattt gctttacaac aagcgttgtc 240
cccaagatgc tgatgaactt cctgtcagaa aagaagacca tctcctatgc tgggtgtctg 300
acacagatgt attttctcta tgccttgggc aacagtgaca gctgccttct ggcagtcatg 360
gcctttgacc gctatgtggc cgtctgtgac cctttccact atgtcaccac catgagccac 420
caccactgtg tcctgctggt ggccttctcc tgctcatttc ctcacctcca ctcactcctg 480
cacacacttc tgctgaatcg tctcaccttc tgtgactcca atgttatcca ccactttctc 540
tgtgacctca gccctgtgct gaaattgtcc tgctcttcca tatttgtcaa tgaaattgtg 600
cagatgacag aagcacctat tgttttggtg actcgttttc tctgcattgc tttctcttat 660
atacgaatcc tcactacagt tctcaagatt ccctctactt ctgggaaacg caaagccttc 720
tccacctgtg gtttttacct caccgtggtg acgctctttt atggaagcat cttctgtgtc 780
tatttacagc ccccatccac ctacgctgtc aaggaccacg tggcaacaat tgtttacaca 840
gttttgtcat ccatgctcaa tccttttatc tacagcctga gaaacaaaga cctgaaacag 900
ggcctgagga agcttatgag caagagatcc taggaagcac cctcttgaaa aactc 955

<210> 36
<211> 309
<212> PRT
<213> Homo sapiens

<400> 36
Met Glu Arg Ile Asn His Thr Ser Ser Val Ser Glu Phe Ile Leu Leu
1 5 10 15

Gly Leu Ser Ser Arg Pro Glu Asp Gln Lys Thr Leu Phe Val Leu Phe
20 25 30

```
Leu Ile Val Tyr Leu Val Thr Ile Thr Gly Asn Leu Leu Ile Ile Leu
          35                  40                  45

Ala Ile Arg Phe Asn Pro His Leu Gln Thr Pro Met Tyr Phe Phe Leu
      50                  55                  60

Ser Phe Leu Ser Leu Thr Asp Ile Cys Phe Thr Thr Ser Val Val Pro
 65                  70                  75                  80

Lys Met Leu Met Asn Phe Leu Ser Glu Lys Lys Thr Ile Ser Tyr Ala
                 85                  90                  95

Gly Cys Leu Thr Gln Met Tyr Phe Leu Tyr Ala Leu Gly Asn Ser Asp
            100                 105                 110

Ser Cys Leu Leu Ala Val Met Ala Phe Asp Arg Tyr Val Ala Val Cys
        115                 120                 125

Asp Pro Phe His Tyr Val Thr Thr Met Ser His His His Cys Val Leu
    130                 135                 140

Leu Val Ala Phe Ser Cys Ser Phe Pro His Leu His Ser Leu Leu His
145                 150                 155                 160

Thr Leu Leu Leu Asn Arg Leu Thr Phe Cys Asp Ser Asn Val Ile His
                165                 170                 175

His Phe Leu Cys Asp Leu Ser Pro Val Leu Lys Leu Ser Cys Ser Ser
            180                 185                 190

Ile Phe Val Asn Glu Ile Val Gln Met Thr Glu Ala Pro Ile Val Leu
        195                 200                 205

Val Thr Arg Phe Leu Cys Ile Ala Phe Ser Tyr Ile Arg Ile Leu Thr
    210                 215                 220

Thr Val Leu Lys Ile Pro Ser Thr Ser Gly Lys Arg Lys Ala Phe Ser
225                 230                 235                 240

Thr Cys Gly Phe Tyr Leu Thr Val Val Thr Leu Phe Tyr Gly Ser Ile
                245                 250                 255

Phe Cys Val Tyr Leu Gln Pro Pro Ser Thr Tyr Ala Val Lys Asp His
            260                 265                 270

Val Ala Thr Ile Val Tyr Thr Val Leu Ser Ser Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Lys Asp Leu Lys Gln Gly Leu Arg Lys Leu
    290                 295                 300

Met Ser Lys Arg Ser
305
```

<210> 37

<211> 958
<212> DNA
<213> Homo sapiens

<400> 37
atggacaaca gcaactggac cagtgtgtcc cattttgttc tcttgggcat ttccacccac 60
ccagaagagc aaatcccact cttccttgtt ttctcactca tgtacgcaat caatatttct 120
ggcaacttgg ccatcatcac actgattctc tctgctccac gcctccacat ccccatgtac 180
atcttcctca gtaacttggc cttgacagac atctgcttca cctccaccac ggtccccaag 240
atgctgcaga ttattttctc ccctacaaag gtaatttcct acacaggctg tttagcccaa 300
acttatttct tcatttgctt cgccgtcatg gaaaacttca tcctggctgt gatggcctat 360
gacaggtaca ttgccatctg ccacccctttc cactacacta tgatcctgac tagaatgctg 420
tgtgtgaaga tggtggtcat gtgccatgct ctctcccacc ttcatgccat gctgcatacc 480
tttctcatag gccaactaat cttctgtgca gataacagaa tcccccactt cttctgtgac 540
ctctacgctc tgatgaagat ctcctgcacc agcacctacc tcaacaccct tatgattcac 600
acagaaggtg ctgttgtaat cagtggagct ctggccttca ttactgcctc ctatgcctgc 660
atcatcctgg tggtcctccg gatcccctca gccaagggca ggtggaaaac cttttctacc 720
tgcggctccc acctcactgt ggtggccata ttctatggca ccctcagttg ggtctacttc 780
cggccccttt ccagctattc agtgaccaag ggtcgcatta taacagtcgt gtacacagtg 840
gtgactccca tgctgaaccc cttcatctac agcctgagga atggggatgt caagggaggc 900
ttcatgaaat ggatgagcag aatgcagact tttttcttta gataaaaccc caaacaca 958

<210> 38
<211> 314
<212> PRT
<213> Homo sapiens

<400> 38
Met Asp Asn Ser Asn Trp Thr Ser Val Ser His Phe Val Leu Leu Gly
1 5 10 15

Ile Ser Thr His Pro Glu Glu Gln Ile Pro Leu Phe Leu Val Phe Ser
20 25 30

Leu Met Tyr Ala Ile Asn Ile Ser Gly Asn Leu Ala Ile Ile Thr Leu
35 40 45

Ile Leu Ser Ala Pro Arg Leu His Ile Pro Met Tyr Ile Phe Leu Ser
50 55 60

Asn Leu Ala Leu Thr Asp Ile Cys Phe Thr Ser Thr Thr Val Pro Lys
65 70 75 80

Met Leu Gln Ile Ile Phe Ser Pro Thr Lys Val Ile Ser Tyr Thr Gly
85 90 95

Cys Leu Ala Gln Thr Tyr Phe Phe Ile Cys Phe Ala Val Met Glu Asn
100 105 110

Phe Ile Leu Ala Val Met Ala Tyr Asp Arg Tyr Ile Ala Ile Cys His
115 120 125

Pro Phe His Tyr Thr Met Ile Leu Thr Arg Met Leu Cys Val Lys Met
130 135 140

Val Val Met Cys His Ala Leu Ser His Leu His Ala Met Leu His Thr

```
145                 150                 155                 160

Phe Leu Ile Gly Gln Leu Ile Phe Cys Ala Asp Asn Arg Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Leu Tyr Ala Leu Met Lys Ile Ser Cys Thr Ser Thr
            180                 185                 190

Tyr Leu Asn Thr Leu Met Ile His Thr Glu Gly Ala Val Val Ile Ser
        195                 200                 205

Gly Ala Leu Ala Phe Ile Thr Ala Ser Tyr Ala Cys Ile Ile Leu Val
    210                 215                 220

Val Leu Arg Ile Pro Ser Ala Lys Gly Arg Trp Lys Thr Phe Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Thr Val Val Ala Ile Phe Tyr Gly Thr Leu Ser
                245                 250                 255

Trp Val Tyr Phe Arg Pro Leu Ser Ser Tyr Ser Val Thr Lys Gly Arg
            260                 265                 270

Ile Ile Thr Val Val Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Gly Asp Val Lys Gly Gly Phe Met Lys Trp
    290                 295                 300

Met Ser Arg Met Gln Thr Phe Phe Phe Arg
305                 310
```

<210> 39
<211> 254
<212> PRT
<213> Homo sapiens

<400> 39

```
Gly Asn Leu Leu Val Ile Leu Val Ile Leu Arg Thr Lys Lys Leu Arg
  1               5                  10                  15

Thr Pro Thr Asn Ile Phe Leu Leu Asn Leu Ala Val Ala Asp Leu Leu
            20                  25                  30

Phe Leu Leu Thr Leu Pro Pro Trp Ala Leu Tyr Tyr Leu Val Gly Gly
        35                  40                  45

Asp Trp Val Phe Gly Asp Ala Leu Cys Lys Leu Val Gly Ala Leu Phe
    50                  55                  60

Val Val Asn Gly Tyr Ala Ser Ile Leu Leu Leu Thr Ala Ile Ser Ile
65                  70                  75                  80

Asp Arg Tyr Leu Ala Ile Val His Pro Leu Arg Tyr Arg Arg Ile Arg
                85                  90                  95
```

```
Thr Pro Arg Arg Ala Lys Val Leu Ile Leu Leu Val Trp Val Leu Ala
            100                 105                 110

Leu Leu Leu Ser Leu Pro Pro Leu Leu Phe Ser Trp Leu Arg Thr Val
        115                 120                 125

Glu Glu Gly Asn Thr Thr Val Cys Leu Ile Asp Phe Pro Glu Glu Ser
    130                 135                 140

Val Lys Arg Ser Tyr Val Leu Leu Ser Thr Leu Val Gly Phe Val Leu
145                 150                 155                 160

Pro Leu Leu Val Ile Leu Val Cys Tyr Thr Arg Ile Leu Arg Thr Leu
                165                 170                 175

Arg Lys Arg Ala Arg Ser Gln Arg Ser Leu Lys Arg Arg Ser Ser Ser
            180                 185                 190

Glu Arg Lys Ala Ala Lys Met Leu Leu Val Val Val Val Val Phe Val
        195                 200                 205

Leu Cys Trp Leu Pro Tyr His Ile Val Leu Leu Leu Asp Ser Leu Cys
    210                 215                 220

Leu Leu Ser Ile Trp Arg Val Leu Pro Thr Ala Leu Leu Ile Thr Leu
225                 230                 235                 240

Trp Leu Ala Tyr Val Asn Ser Cys Leu Asn Pro Ile Ile Tyr
                245                 250
```

<210> 40
<211> 17
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:consensus sequence

<220>
<221> VARIANT
<222> (1)
<223> Wherein Xaa is Gly or Ser or Thr or Ala or Leu or Ile or Val or Met or Phe or Tyr or Trp or Cys

<220>
<221> VARIANT
<222> (2)
<223> Wherein Xaa is Gly or Ser or Thr or Ala or Asn or Cys or Pro or Asp or Glu

<220>
<221> VARIANT
<222> (3)
<223> Wherein Xaa is not Glu or Asp or Pro or Lys or Arg or His

<220>
<221> VARIANT
<222> (4)
<223> Wherein Xaa is any amino acid as defined in the specification

<220>
<221> VARIANT
<222> (5)
<223> Wherein Xaa is any amino acid as defined in the specification

<220>
<221> VARIANT
<222> (6)
<223> Wherein Xaa is Leu or Ile or Val or Met or Asn or Gln or Gly or Ala

<220>
<221> VARIANT
<222> (7)
<223> Wherein Xaa is any amino acid as defined in the specification

<220>
<221> VARIANT
<222> (8)
<223> Wherein Xaa is any amino acid as defined in the specification

<220>
<221> VARIANT
<222> (9)
<223> Wherein Xaa is Leu or Ile or Val or Met or Phe or Thr

<220>
<221> VARIANT
<222> (10)
<223> Wherein Xaa is Gly or Ser or Thr or Ala or Asn or Cys

<220>
<221> VARIANT
<222> (11)
<223> Wherein Xaa is Leu or Ile or Val or Met or Phe or Tyr or Trp or Ser or Thr or Ala or Cys

<220>
<221> VARIANT
<222> (12)
<223> Wherein Xaa is Asp or Glu or Asn or His

<220>
<221> VARIANT
<222> (14)

<223> Wherein Xaa is Phe or Tyr or Trp or Cys or Ser or His

<220>
<221> VARIANT
<222> (15)
<223> Wherein Xaa is any amino acid as defined in the specification

<220>
<221> VARIANT
<222> (16)
<223> Wherein Xaa is any amino acid as defined in the specification

<220>
<221> VARIANT
<222> (17)
<223> Wherein Xaa is Leu or Ile or Val or Met

<400> 40
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Xaa Xaa Xaa
1 5 10 15

Xaa

<210> 41
<211> 311
<212> PRT
<213> Rattus norvegicus

<400> 41
Met Glu Asn Gln Ser Ser Val Ser Glu Phe Phe Leu Gln Gly Ile Ser
1 5 10 15

Gly Phe Pro Glu Gln Gln Gln Leu Leu Tyr Gly Leu Phe Leu Cys Met
20 25 30

Tyr Leu Val Thr Leu Thr Gly Asn Val Leu Ile Ile Met Ala Ile Gly
35 40 45

Ser Asp Pro His Leu His Thr Pro Met Tyr Phe Phe Leu Ala Asn Leu
50 55 60

Ser Phe Ala Asp Met Gly Leu Ile Ser Ser Thr Val Thr Gln Met Leu
65 70 75 80

Phe Asn Val Gln Thr Gln Arg His Thr Ile Ser Tyr Thr Gly Cys Leu
85 90 95

Thr Gln Met Tyr Phe Phe Leu Met Phe Gly Asp Leu Asp Ser Phe Phe
100 105 110

Leu Ala Val Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His Pro Leu
115 120 125

```
His Tyr Ser Thr Ile Met Arg Ala Lys Val Cys Val Leu Met Leu Ala
    130                 135                 140

Leu Cys Trp Val Leu Thr Asn Ile Val Ala Leu Thr His Thr Leu Leu
145                 150                 155                 160

Met Ala Arg Leu Ser Phe Cys Val Val Gly Glu Ile Ala His Phe Phe
                165                 170                 175

Cys Asp Ile Thr Pro Val Leu Lys Leu Ser Cys Ser Asp Thr Tyr Val
            180                 185                 190

Asn Glu Leu Met Val Phe Ala Leu Gly Gly Thr Val Leu Met Leu Pro
        195                 200                 205

Phe Ile Cys Ile Val Ile Ser Tyr Ile His Ile Val Phe Ala Ile Leu
    210                 215                 220

Arg Val Arg Thr Pro Gly Gly Gly Thr Lys Ala Phe Ser Thr Cys Ser
225                 230                 235                 240

Ser His Leu Cys Val Val Cys Val Phe Tyr Gly Thr Leu Phe Ser Ala
                245                 250                 255

Tyr Leu Ser Pro Pro Ser Val Val Ser Thr Glu Lys Asp Ile Ala Ala
            260                 265                 270

Ala Ala Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Tyr
        275                 280                 285

Ser Leu Arg Asn Lys Asp Met Lys Gly Ala Leu Lys Arg Leu Leu Phe
    290                 295                 300

His Arg Arg Ile Leu Ser Ser
305                 310


<210> 42
<211> 313
<212> PRT
<213> Rattus norvegicus

<400> 42
Met Ser Ser Thr Asn Gln Ser Ser Val Thr Glu Phe Leu Leu Leu Gly
  1                 5                  10                  15

Leu Ser Arg Gln Pro Gln Gln Gln Gln Leu Leu Phe Leu Leu Phe Leu
             20                  25                  30

Ile Met Tyr Leu Ala Thr Val Leu Gly Asn Leu Leu Ile Ile Leu Ala
         35                  40                  45

Ile Gly Thr Asp Ser Arg Leu His Thr Pro Met Tyr Phe Phe Leu Ser
     50                  55                  60

Asn Leu Ser Phe Val Asp Val Cys Phe Ser Ser Thr Thr Val Pro Lys
```

65 70 75 80

Val Leu Ala Asn His Ile Leu Gly Ser Gln Ala Ile Ser Phe Ser Gly
85 90 95

Cys Leu Thr Gln Leu Tyr Phe Leu Ala Val Phe Gly Asn Met Asp Asn
100 105 110

Phe Leu Leu Ala Val Met Ser Tyr Asp Arg Phe Val Ala Ile Cys His
115 120 125

Pro Leu His Tyr Thr Thr Lys Met Thr Arg Gln Leu Cys Val Leu Leu
130 135 140

Val Val Gly Ser Trp Val Val Ala Asn Met Asn Cys Leu Leu His Ile
145 150 155 160

Leu Leu Met Ala Arg Leu Ser Phe Cys Ala Asp Asn Met Ile Pro His
165 170 175

Phe Phe Cys Asp Gly Thr Pro Leu Leu Lys Leu Ser Cys Ser Asp Thr
180 185 190

His Leu Asn Glu Leu Met Ile Leu Thr Glu Gly Ala Val Val Met Val
195 200 205

Thr Pro Phe Val Cys Ile Leu Ile Ser Tyr Ile His Ile Thr Cys Ala
210 215 220

Val Leu Arg Val Ser Ser Pro Arg Gly Gly Trp Lys Ser Phe Ser Thr
225 230 235 240

Cys Gly Ser His Leu Ala Val Val Cys Leu Phe Tyr Gly Thr Val Ile
245 250 255

Ala Val Tyr Phe Asn Pro Ser Ser Ser His Leu Ala Gly Arg Asp Met
260 265 270

Ala Ala Ala Val Met Tyr Ala Val Val Thr Pro Met Leu Asn Pro Phe
275 280 285

Ile Tyr Ser Leu Arg Asn Ser Asp Met Lys Ala Ala Leu Arg Lys Val
290 295 300

Leu Ala Met Arg Phe Pro Ser Lys Gln
305 310

<210> 43
<211> 311
<212> PRT
<213> Rattus norvegicus

<400> 43
Met Glu Asn Gln Ser Ser Val Ser Glu Phe Phe Leu Arg Gly Ile Ser
1 5 10 15

```
Gly Phe Pro Glu Gln Gln Gln Leu Leu Tyr Gly Leu Phe Leu Cys Met
                20                  25                  30

Tyr Leu Val Thr Leu Thr Gly Asn Val Leu Ile Ile Leu Ala Ile Gly
            35                  40                  45

Ser Asp Pro His Leu His Thr Pro Met Tyr Phe Phe Leu Ala Asn Leu
        50                  55                  60

Ser Phe Ala Asp Met Gly Leu Ile Ser Ser Thr Val Thr Lys Met Leu
65                  70                  75                  80

Phe Asn Val Gln Thr Gln Cys His Thr Ile Ser Tyr Thr Gly Cys Leu
                85                  90                  95

Thr Gln Met Tyr Leu Phe Met Met Phe Gly Asp Leu Asp Ser Phe Phe
            100                 105                 110

Leu Ala Val Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His Pro Leu
        115                 120                 125

His Tyr Ser Thr Ile Met Asn Ala Arg Ile Cys Val Leu Met Leu Ile
    130                 135                 140

Leu Cys Trp Ile Leu Thr Asn Val Val Ala Leu Thr His Thr Leu Leu
145                 150                 155                 160

Met Ala Arg Leu Ser Phe Cys Val Val Gly Glu Ile Ala His Phe Phe
                165                 170                 175

Cys Asp Val Thr Ser Val Met Lys Leu Ser Cys Ser Asp Thr His Val
            180                 185                 190

Asn Glu Leu Val Leu Ser Gly Phe Gly Gly Thr Val Leu Met Val Pro
        195                 200                 205

Phe Val Ser Ile Val Ile Ser Tyr Val His Ile Val Phe Ala Val Leu
    210                 215                 220

Arg Ile Gln Ser Ser Gly Gly Ser Ser Lys Ala Phe Ser Thr Cys Ser
225                 230                 235                 240

Ser His Leu Cys Val Val Cys Val Phe Tyr Gly Thr Leu Phe Ser Val
                245                 250                 255

Tyr Leu Phe Pro Ser Ser Val Glu Thr Thr Glu Lys Asp Val Ala Ala
            260                 265                 270

Ala Ala Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Tyr
        275                 280                 285

Ser Leu Arg Asn Lys Asp Ile Lys Gly Ala Leu Lys Arg Leu Leu Ser
    290                 295                 300

His Arg Arg Ile Leu Ser Ser
305                 310
```

<210> 44
<211> 311
<212> PRT
<213> Pan troglodytes

<400> 44
Met Glu Asn Gln Ser Ser Val Ser Glu Phe Phe Leu Gln Gly Ile Ser
1 5 10 15

Gly Phe Pro Glu Gln Gln Gln Leu Leu Tyr Gly Leu Phe Leu Cys Met
20 25 30

Tyr Leu Val Thr Leu Thr Gly Asn Val Leu Ile Ile Met Ala Ile Gly
35 40 45

Ser Asp Pro His Leu His Thr Pro Met Tyr Phe Phe Leu Ala Asn Leu
50 55 60

Ser Phe Ala Asp Met Gly Leu Ile Ser Ser Thr Val Thr Gln Met Leu
65 70 75 80

Phe Asn Val Gln Thr Gln Arg His Thr Ile Ser Tyr Thr Gly Cys Leu
85 90 95

Thr Gln Met Tyr Phe Phe Leu Met Phe Gly Asp Leu Asp Ser Phe Phe
100 105 110

Leu Ala Val Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His Pro Leu
115 120 125

His Tyr Ser Thr Ile Met Arg Ala Lys Val Cys Val Leu Met Leu Ala
130 135 140

Leu Cys Trp Val Leu Thr Asn Ile Val Ala Leu Thr His Thr Leu Leu
145 150 155 160

Met Ala Arg Leu Ser Phe Cys Val Val Gly Glu Ile Ala His Phe Phe
165 170 175

Cys Asp Ile Thr Pro Val Leu Lys Leu Ser Cys Ser Asp Thr Tyr Val
180 185 190

Asn Glu Leu Met Val Phe Ala Leu Gly Gly Thr Val Leu Met Leu Pro
195 200 205

Phe Ile Cys Ile Val Ile Ser Tyr Ile His Ile Val Phe Ala Ile Leu
210 215 220

Arg Val Arg Thr Pro Gly Gly Gly Thr Lys Ala Phe Ser Thr Cys Ser
225 230 235 240

Ser His Leu Cys Val Val Cys Val Phe Tyr Gly Thr Leu Phe Ser Ala
245 250 255

Tyr Leu Ser Pro Pro Ser Val Val Ser Thr Glu Lys Asp Ile Ala Ala
260 265 270

```
Ala Ala Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Tyr
        275                 280                 285

Ser Leu Arg Asn Lys Asp Met Lys Gly Ala Leu Lys Arg Leu Leu Phe
    290                 295                 300

His Arg Arg Ile Leu Ser Ser
305                 310
```

<210> 45
<211> 314
<212> PRT
<213> Gorilla gorilla

<400> 45

```
Met Met Gly Gln Asn Gln Thr Ser Ile Ser Asp Phe Leu Leu Leu Gly
  1               5                  10                  15

Leu Pro Ile Gln Pro Glu Gln Gln Asn Leu Cys Tyr Ala Leu Phe Leu
             20                  25                  30

Ala Met Tyr Leu Thr Thr Leu Leu Gly Asn Leu Leu Ile Ile Val Leu
         35                  40                  45

Ile Arg Leu Asp Ser His Leu His Thr Pro Met Tyr Leu Phe Leu Ser
     50                  55                  60

Asn Leu Ser Phe Ser Asp Leu Cys Phe Ser Ser Val Thr Ile Pro Lys
 65                  70                  75                  80

Leu Leu Gln Asn Met Gln Asn Gln Asp Pro Ser Ile Pro Tyr Ala Asp
                 85                  90                  95

Cys Leu Thr Gln Met Tyr Phe Phe Leu Leu Phe Gly Asp Leu Glu Ser
            100                 105                 110

Phe Leu Leu Val Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys Phe
        115                 120                 125

Pro Leu His Tyr Thr Ala Ile Met Ser Pro Met Leu Cys Leu Ser Leu
    130                 135                 140

Val Ala Leu Ser Trp Val Leu Thr Thr Phe His Ala Met Leu His Thr
145                 150                 155                 160

Leu Leu Met Ala Arg Leu Cys Phe Cys Ala Asp Asn Val Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Met Ser Ala Leu Leu Lys Leu Ala Cys Ser Asp Thr
            180                 185                 190

Arg Val Asn Glu Trp Val Ile Phe Ile Met Gly Gly Leu Ile Val Val
        195                 200                 205

Ile Pro Phe Leu Leu Ile Leu Gly Ser Tyr Ala Arg Ile Val Ser Ser
```

210 215 220

Ile Leu Lys Val Pro Ser Ser Lys Gly Ile Cys Lys Ala Phe Ser Thr
225 230 235 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Phe Tyr Gly Thr Ile Ile
245 250 255

Gly Leu Tyr Leu Cys Pro Ser Ala Asn Ser Ser Thr Leu Lys Glu Thr
260 265 270

Val Met Ala Met Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
275 280 285

Ile Tyr Ser Leu Arg Asn Arg Asp Met Lys Gly Ala Leu Glu Arg Val
290 295 300

Ile Cys Lys Arg Lys Asn Pro Phe Leu Leu
305 310

<210> 46
<211> 314
<212> PRT
<213> Pan troglodytes

<400> 46
Met Met Gly Gln Asn Gln Thr Ser Ile Ser Asp Phe Leu Leu Leu Gly
1 5 10 15

Leu Pro Ile Gln Pro Glu Gln Gln Asn Leu Cys Tyr Ala Leu Phe Leu
20 25 30

Ala Met Tyr Leu Thr Thr Leu Leu Gly Asn Leu Leu Ile Ile Val Leu
35 40 45

Ile Arg Leu Asp Ser His Leu His Thr Pro Met Tyr Leu Phe Leu Ser
50 55 60

Asn Leu Ser Phe Ser Asp Leu Cys Phe Ser Ser Val Thr Ile Pro Lys
65 70 75 80

Leu Leu Gln Asn Met Gln Asn Gln Asp Pro Ser Ile Pro Tyr Ala Asp
85 90 95

Cys Leu Thr Gln Met Tyr Phe Phe Leu Leu Phe Gly Asp Leu Glu Ser
100 105 110

Phe Leu Leu Val Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys Phe
115 120 125

Pro Leu His Tyr Thr Ala Ile Met Ser Pro Met Leu Cys Leu Ser Leu
130 135 140

Val Ala Leu Ser Trp Val Leu Thr Thr Phe His Ala Met Leu His Thr
145 150 155 160

```
Leu Leu Met Ala Arg Leu Cys Phe Cys Ala Asp Asn Val Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Met Ser Ala Leu Leu Lys Leu Ala Cys Ser Asp Thr
            180                 185                 190

Arg Val Asn Glu Trp Val Ile Phe Ile Met Gly Gly Leu Ile Val Val
        195                 200                 205

Ile Pro Phe Leu Leu Ile Leu Gly Ser Tyr Ala Arg Ile Val Ser Ser
    210                 215                 220

Ile Leu Lys Val Pro Ser Ser Lys Gly Ile Cys Lys Ala Phe Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Phe Tyr Gly Thr Ile Ile
                245                 250                 255

Gly Leu Tyr Leu Cys Pro Ser Ala Asn Ser Ser Thr Leu Lys Glu Thr
            260                 265                 270

Val Met Ala Met Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Arg Asp Met Lys Gly Ala Leu Glu Arg Val
    290                 295                 300

Ile Xaa Lys Arg Lys Asn Pro Phe Leu Leu
305                 310
```

<210> 47
<211> 312
<212> PRT
<213> Homo sapiens

<400> 47

```
Met Ser Gly Thr Asn Gln Ser Ser Val Ser Glu Phe Leu Leu Leu Gly
  1               5                  10                  15

Leu Ser Arg Gln Pro Gln Gln Gln His Leu Leu Phe Val Phe Phe Leu
            20                  25                  30

Ser Met Tyr Leu Ala Thr Val Leu Gly Asn Leu Leu Ile Ile Leu Ser
        35                  40                  45

Val Ser Ile Asp Ser Cys Leu His Thr Pro Met Tyr Phe Phe Leu Ser
    50                  55                  60

Asn Leu Ser Phe Val Asp Ile Cys Phe Ser Phe Thr Thr Val Pro Lys
65                  70                  75                  80

Met Leu Ala Asn His Ile Leu Glu Thr Gln Thr Ile Ser Phe Cys Gly
                85                  90                  95

Cys Leu Thr Gln Met Tyr Phe Val Phe Met Phe Val Asp Met Asp Asn
            100                 105                 110
```

```
Phe Leu Leu Ala Val Met Ala Tyr Asp His Phe Val Ala Val Cys His
        115                 120                 125

Pro Leu His Tyr Thr Ala Lys Met Thr His Gln Leu Cys Ala Leu Leu
    130                 135                 140

Val Ala Gly Leu Trp Val Val Ala Asn Leu Asn Val Leu Leu His Thr
145                 150                 155                 160

Leu Leu Met Ala Pro Leu Ser Phe Cys Ala Asp Asn Ala Ile Thr His
                165                 170                 175

Phe Phe Cys Asp Val Thr Pro Leu Leu Lys Leu Ser Cys Ser Asn Thr
            180                 185                 190

His Leu Asn Glu Val Ile Ile Leu Ser Glu Gly Ala Leu Val Met Ile
        195                 200                 205

Thr Pro Phe Leu Cys Ile Leu Ala Ser Tyr Met His Ile Thr Cys Thr
    210                 215                 220

Val Leu Lys Val Pro Ser Thr Lys Gly Arg Trp Lys Ala Phe Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ala Val Val Leu Leu Phe Tyr Ser Thr Ile Ile
                245                 250                 255

Ala Val Tyr Phe Asn Pro Leu Ser Ser His Ser Ala Glu Lys Asp Thr
            260                 265                 270

Met Ala Thr Val Leu Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Arg Tyr Leu Lys Gly Ala Leu Lys Lys Val
    290                 295                 300

Ile Gly Arg Val Val Phe Ser Val
305                 310


<210> 48
<211> 313
<212> PRT
<213> Gorilla gorilla

<400> 48
Met Glu Gly Lys Asn Leu Thr Ser Ile Ser Glu Phe Phe Leu Leu Gly
  1               5                  10                  15

Phe Ser Glu Gln Leu Glu Glu Gln Lys Ala Leu Phe Gly Ser Phe Leu
            20                  25                  30

Phe Met Tyr Leu Val Thr Val Ala Gly Asn Leu Leu Ile Ile Leu Val
        35                  40                  45

Ile Ile Thr Asp Thr Gln Leu His Thr Pro Met Tyr Phe Phe Leu Ala
```

```
     50                  55                  60

Asn Leu Ser Leu Ala Asp Ala Cys Phe Val Ser Thr Thr Val Pro Lys
 65                  70                  75                  80

Met Leu Ala Asn Ile Gln Ile Gln Ser Gln Ala Ile Ser Tyr Ser Gly
                 85                  90                  95

Cys Leu Leu Gln Leu Tyr Phe Phe Met Leu Phe Val Met Leu Glu Ala
            100                 105                 110

Phe Leu Leu Ala Val Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His
        115                 120                 125

Pro Leu His Tyr Ile Leu Ile Met Ser Pro Gly Leu Cys Val Phe Leu
    130                 135                 140

Val Ser Ala Ser Trp Ile Met Asn Ala Leu His Ser Leu Leu His Thr
145                 150                 155                 160

Leu Leu Met Asn Ser Leu Ser Phe Cys Ala Asn His Glu Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Ile Asp Pro Leu Leu Ser Leu Ser Cys Thr Asp Pro
            180                 185                 190

Phe Thr Asn Glu Leu Val Ile Phe Ile Thr Gly Gly Leu Thr Gly Leu
        195                 200                 205

Val Cys Val Leu Cys Leu Ile Ile Ser Tyr Thr Asn Ile Phe Ser Thr
    210                 215                 220

Ile Leu Lys Ile Pro Ser Ala Gln Gly Lys Arg Lys Ala Phe Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Phe Phe Gly Thr Ser Phe
                245                 250                 255

Cys Val Tyr Phe Ile Pro Pro Ser Thr Arg Ser Ala Gln Lys Asp Thr
            260                 265                 270

Val Ala Ser Val Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Gln Glu Ile Lys Ser Ser Leu Arg Lys Leu
    290                 295                 300

Ile Trp Val Arg Glu Ile His Ser Pro
305                 310
```

<210> 49
<211> 309
<212> PRT
<213> Homo sapiens

<400> 49

```
Met Arg Glu Asn Asn Gln Ser Ser Thr Leu Glu Phe Ile Leu Leu Gly
 1               5                  10                  15

Val Thr Gly Gln Gln Glu Gln Glu Asp Phe Phe Tyr Ile Leu Phe Leu
            20                  25                  30

Phe Ile Tyr Pro Ile Thr Leu Ile Gly Asn Leu Leu Ile Val Leu Ala
        35                  40                  45

Ile Cys Ser Asp Val Arg Leu His Asn Pro Met Tyr Phe Leu Leu Ala
    50                  55                  60

Asn Leu Ser Leu Val Asp Ile Phe Phe Ser Ser Val Thr Ile Pro Lys
65                  70                  75                  80

Met Leu Ala Asn His Leu Leu Gly Ser Lys Ser Ile Ser Phe Gly Gly
                85                  90                  95

Cys Leu Thr Gln Met Tyr Phe Met Ile Ala Leu Gly Asn Thr Asp Ser
            100                 105                 110

Tyr Ile Leu Ala Ala Met Ala Tyr Asp Arg Ala Val Ala Ile Ser His
        115                 120                 125

Pro Leu His Tyr Thr Thr Ile Met Ser Pro Arg Ser Cys Ile Trp Leu
    130                 135                 140

Ile Ala Gly Ser Trp Val Ile Gly Asn Ala Asn Ala Leu Pro His Thr
145                 150                 155                 160

Leu Leu Thr Ala Ser Leu Ser Phe Cys Gly Asn Gln Glu Val Ala Asn
                165                 170                 175

Phe Tyr Cys Asp Ile Thr Pro Leu Leu Lys Leu Ser Cys Ser Asp Ile
            180                 185                 190

His Phe His Val Lys Met Met Tyr Leu Gly Val Gly Ile Phe Ser Val
        195                 200                 205

Pro Leu Leu Cys Ile Ile Val Ser Tyr Ile Arg Val Phe Ser Thr Val
    210                 215                 220

Phe Gln Val Pro Ser Thr Lys Gly Val Leu Lys Ala Phe Ser Thr Cys
225                 230                 235                 240

Gly Ser His Leu Thr Val Val Ser Leu Tyr Tyr Gly Thr Val Met Gly
                245                 250                 255

Thr Tyr Phe Arg Pro Leu Thr Asn Tyr Ser Leu Lys Asp Ala Val Ile
            260                 265                 270

Thr Val Met Tyr Thr Ala Val Thr Pro Met Leu Asn Pro Phe Ile Tyr
        275                 280                 285

Ser Leu Arg Asn Arg Asp Met Lys Ala Ala Leu Arg Lys Leu Phe Asn
    290                 295                 300
```

Lys Arg Ile Ser Ser
305

<210> 50
<211> 313
<212> PRT
<213> Canis familiaris

<400> 50
Met Thr Glu Lys Asn Gln Thr Val Val Ser Glu Phe Val Leu Leu Gly
1 5 10 15

Leu Pro Ile Asp Pro Asp Gln Arg Asp Leu Phe Tyr Ala Leu Phe Leu
20 25 30

Ala Met Tyr Val Thr Thr Ile Leu Gly Asn Leu Leu Ile Ile Val Leu
35 40 45

Ile Gln Leu Asp Ser His Leu His Thr Pro Met Tyr Leu Phe Leu Ser
50 55 60

Asn Leu Ser Phe Ser Asp Leu Cys Phe Ser Ser Val Thr Met Pro Lys
65 70 75 80

Leu Leu Gln Asn Met Gln Ser Gln Val Pro Ser Ile Pro Tyr Ala Gly
85 90 95

Cys Leu Thr Gln Met Tyr Phe Phe Leu Phe Phe Gly Asp Leu Glu Ser
100 105 110

Phe Leu Leu Val Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys Phe
115 120 125

Pro Leu His Tyr Thr Thr Ile Met Ser Pro Lys Leu Cys Phe Ser Leu
130 135 140

Leu Val Leu Ser Trp Val Leu Thr Met Phe His Ala Val Leu His Thr
145 150 155 160

Leu Leu Met Ala Arg Leu Cys Phe Cys Ala Asn Thr Ile Pro His Phe
165 170 175

Phe Cys Asp Met Ser Ala Leu Leu Lys Leu Ala Cys Ser Asp Thr Gln
180 185 190

Val Asn Glu Leu Val Ile Phe Ile Met Gly Gly Leu Ile Leu Val Ile
195 200 205

Pro Phe Leu Leu Ile Ile Thr Ser Tyr Ala Arg Ile Val Ser Ser Ile
210 215 220

Leu Lys Val Pro Ser Ala Ile Gly Ile Cys Lys Val Phe Ser Thr Cys
225 230 235 240

Gly Ser His Leu Ser Val Val Ser Leu Phe Tyr Gly Thr Val Ile Gly
245 250 255

```
Leu Tyr Leu Cys Pro Ser Ala Asn Asn Ser Thr Val Lys Glu Thr Ile
            260                 265                 270

Met Ala Met Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile
        275                 280                 285

Tyr Ser Leu Arg Asn Lys Asp Met Lys Gly Ala Leu Arg Arg Val Ile
    290                 295                 300

Cys Arg Lys Lys Ile Thr Phe Ser Val
305                 310
```

```
<210> 51
<211> 314
<212> PRT
<213> Rattus norvegicus

<400> 51
Met Thr Glu Glu Asn Gln Thr Val Ile Ser Gln Phe Leu Leu Leu Phe
  1               5                  10                  15

Leu Pro Ile Pro Ser Glu His Gln His Val Phe Tyr Ala Leu Phe Leu
             20                  25                  30

Ser Met Tyr Leu Thr Thr Val Leu Gly Asn Leu Ile Ile Ile Ile Leu
         35                  40                  45

Ile His Leu Asp Ser His Leu His Thr Pro Met Tyr Leu Phe Leu Ser
     50                  55                  60

Asn Leu Ser Phe Ser Asp Leu Cys Phe Ser Ser Val Thr Met Pro Lys
 65                  70                  75                  80

Leu Leu Gln Asn Met Gln Ser Gln Val Pro Ser Ile Pro Phe Ala Gly
                 85                  90                  95

Cys Leu Thr Gln Leu Tyr Phe Tyr Leu Tyr Phe Ala Asp Leu Glu Ser
            100                 105                 110

Phe Leu Leu Val Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys Phe
        115                 120                 125

Pro Leu His Tyr Met Ser Ile Met Ser Pro Lys Leu Cys Val Ser Leu
    130                 135                 140

Val Val Leu Ser Trp Val Leu Thr Thr Phe His Ala Met Leu His Thr
145                 150                 155                 160

Leu Leu Met Ala Arg Leu Ser Phe Cys Ala Asp Asn Met Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Ile Ser Pro Leu Leu Lys Leu Ser Cys Ser Asp Thr
            180                 185                 190

His Val Asn Glu Leu Val Ile Phe Val Met Gly Gly Leu Val Ile Val
```

195 200 205

Ile Pro Phe Val Leu Ile Ile Val Ser Tyr Ala Arg Val Val Ala Ser
210 215 220

Ile Leu Lys Val Pro Ser Val Arg Gly Ile His Lys Ile Phe Ser Thr
225 230 235 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Phe Tyr Gly Thr Ile Ile
245 250 255

Gly Leu Tyr Leu Cys Pro Ser Ala Asn Asn Ser Thr Val Lys Glu Thr
260 265 270

Val Met Ala Met Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
275 280 285

Ile Tyr Ser Leu Arg Asn Arg Asp Met Lys Glu Ala Leu Ile Arg Val
290 295 300

Leu Cys Lys Lys Lys Ile Thr Phe Cys Leu
305 310

<210> 52
<211> 216
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 52
Phe Val Asp Met Gly Leu Thr Ser Ser Thr Val Thr Lys Met Leu Val
1 5 10 15

Asn Ile Gln Thr Arg His His Thr Ile Thr Tyr Thr Gly Cys Leu Thr
20 25 30

Gln Met Tyr Phe Phe Leu Met Phe Gly Asp Leu Asp Ser Phe Phe Leu
35 40 45

Ala Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His Pro Leu Cys
50 55 60

Tyr Ser Thr Val Met Arg Pro Gln Val Cys Ala Leu Met Leu Ala Leu
65 70 75 80

Cys Trp Val Leu Thr Asn Ile Val Ala Leu Thr His Thr Phe Leu Met
85 90 95

Ala Arg Leu Ser Phe Cys Val Thr Gly Glu Ile Ala His Phe Phe Cys
100 105 110

Asp Ile Thr Pro Val Leu Lys Leu Ser Cys Ser Asp Thr His Ile Asn
115 120 125

```
Glu Met Met Val Phe Val Leu Gly Gly Thr Val Leu Ile Val Pro Phe
    130                 135                 140

Leu Cys Ile Val Thr Ser Tyr Ile His Ile Val Pro Ala Ile Leu Arg
145                 150                 155                 160

Val Arg Thr Arg Gly Gly Val Gly Lys Ala Phe Ser Thr Cys Ser Ser
                165                 170                 175

His Leu Cys Val Val Cys Val Phe Tyr Gly Thr Leu Phe Ser Ala Tyr
            180                 185                 190

Leu Cys Pro Pro Ser Ile Ala Ser Glu Glu Lys Asp Ile Ala Ala Ala
        195                 200                 205

Ala Met Tyr Thr Ile Val Thr Pro
    210                 215


<210> 53
<211> 338
<212> PRT
<213> Gorilla gorilla

<400> 53
Met Gly Leu Thr Gln Asp Phe Phe Pro Pro Thr Ser Asp Leu Leu Glu
  1               5                  10                  15

Gly Gly Asn Gln Thr Ser Thr Phe Glu Phe Leu Leu Ser Gly Leu Ser
            20                  25                  30

Asp Gln Pro Gln Gln Gln His Ile Leu Phe Leu Leu Phe Leu Trp Met
        35                  40                  45

Tyr Val Val Thr Val Ala Gly Asn Leu Leu Ile Val Leu Ala Ile Gly
    50                  55                  60

Thr Asp Thr His Leu His Thr Pro Met Tyr Phe Phe Leu Ala Ser Leu
65                  70                  75                  80

Ser Cys Ala Asp Ile Phe Phe Thr Ser Thr Thr Val Pro Lys Ala Leu
                85                  90                  95

Val Asn Ile Gln Thr Gln Ser Arg Ser Ile Ser Tyr Ala Gly Cys Leu
            100                 105                 110

Ala Gln Leu Tyr Phe Phe Leu Thr Phe Gly Asp Met Asp Ile Phe Leu
        115                 120                 125

Leu Ala Thr Met Ala Tyr Asp Arg Tyr Met Ala Ile Cys His Pro Leu
    130                 135                 140

His Tyr Met Met Ile Met Ser Leu His Arg Cys Ala Leu Leu Val Thr
145                 150                 155                 160

Ala Cys Trp Thr Leu Thr Ser Leu Val Ala Met Thr His Thr Phe Leu
```

```
                165                 170                 175

Ile Phe Arg Leu Ser Phe Cys Ser Lys Ile Ile Pro Asp Phe Phe Cys
            180                 185                 190

Asp Leu Gly Pro Leu Met Lys Val Ser Cys Ser Asp Thr Gln Val Ser
        195                 200                 205

Glu Leu Val Leu Leu Phe Leu Gly Gly Ala Val Ile Leu Ile Pro Phe
    210                 215                 220

Met Leu Ile Leu Val Ser Tyr Ile Arg Ile Val Ser Ala Ile Leu Arg
225                 230                 235                 240

Ala Pro Ser Ala Gln Gly Arg Arg Lys Ala Phe Ser Thr Cys Gly Ser
                245                 250                 255

His Leu Val Val Val Ala Leu Phe Phe Gly Thr Val Ile Arg Ala Tyr
            260                 265                 270

Leu Cys Pro Ser Ser Ser Ser Ser Asn Ser Val Asp Glu Asp Thr Ala
        275                 280                 285

Ala Ala Val Met Tyr Thr Val Val Thr Pro Leu Leu Asn Pro Phe Ile
    290                 295                 300

Tyr Ser Leu Arg Asn Lys Asp Met Lys Ala Ala Val Val Arg Leu Leu
305                 310                 315                 320

Lys Gly Arg Val Ser Phe Ser Gln Gly Gln Gly Leu Leu Leu Arg Asn
                325                 330                 335

Pro Arg
```

<210> 54
<211> 314
<212> PRT
<213> Pan troglodytes

<400> 54

```
Met Met Glu Gln Asn Gln Thr Ser Thr Ser Asp Phe Leu Leu Leu Gly
  1               5                  10                  15

Leu Pro Ile Gln Pro Glu Gln Gln Asn Leu Cys Tyr Ala Leu Phe Leu
             20                  25                  30

Ala Met Tyr Leu Thr Thr Leu Leu Gly Asn Leu Leu Ile Ile Val Leu
         35                  40                  45

Ile Arg Leu Asp Ser His Leu His Thr Pro Val Tyr Leu Phe Leu Ser
     50                  55                  60

Asn Leu Ser Phe Ser Asp Leu Cys Phe Ser Ser Val Thr Met Pro Lys
 65                  70                  75                  80
```

```
Leu Leu Gln Asn Met Gln Ser Gln Asn Pro Ser Ile Pro Tyr Ala Asp
                  85                  90                  95

Cys Leu Thr Gln Met Tyr Phe Phe Leu Tyr Phe Ser Asp Leu Glu Ser
             100                 105                 110

Phe Leu Leu Val Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys Phe
         115                 120                 125

Pro Leu His Tyr Thr Ala Ile Met Ser Pro Met Leu Cys Leu Ser Val
     130                 135                 140

Val Ala Leu Ser Trp Val Leu Thr Thr Phe His Ala Met Leu His Thr
145                 150                 155                 160

Leu Leu Met Ala Arg Leu Cys Phe Cys Ala Asp Asn Val Ile Pro His
                 165                 170                 175

Phe Phe Cys Asp Met Ser Ala Leu Leu Lys Leu Ala Cys Ser Asp Thr
             180                 185                 190

Arg Val Asn Glu Trp Val Ile Phe Ile Met Gly Gly Leu Ile Val Val
         195                 200                 205

Ile Pro Phe Leu Leu Ile Leu Gly Ser Tyr Ala Arg Ile Val Ser Ser
     210                 215                 220

Ile Leu Lys Val Pro Ser Ser Lys Gly Ile Cys Lys Ala Leu Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Phe Tyr Gly Thr Val Ile
                 245                 250                 255

Gly Leu Tyr Leu Cys Pro Ser Ala Asn Ser Ser Thr Leu Lys Asp Thr
             260                 265                 270

Val Met Ala Met Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
         275                 280                 285

Ile Tyr Ser Leu Arg Asn Arg Asp Met Lys Gly Ala Leu Glu Arg Val
     290                 295                 300

Ile Cys Lys Arg Lys Asn Pro Phe Leu Ile
305                 310


<210> 55
<211> 314
<212> PRT
<213> Pan troglodytes

<400> 55
Met Met Gly Gln Asn Gln Thr Ser Ile Ser Asp Phe Leu Leu Leu Gly
  1               5                  10                  15

Leu Pro Ile Gln Pro Glu Gln Gln Asn Leu Cys Tyr Ala Leu Phe Leu
             20                  25                  30
```

```
Ala Met Tyr Leu Thr Thr Leu Leu Gly Asn Leu Leu Ile Ile Val Leu
          35                  40                  45

Ile Arg Leu Asp Ser His Leu His Thr Pro Met Tyr Leu Phe Leu Ser
      50                  55                  60

Asn Leu Ser Phe Ser Asp Leu Cys Phe Ser Ser Val Thr Ile Pro Lys
 65                  70                  75                  80

Leu Leu Gln Asn Met Gln Asn Gln Asp Pro Ser Ile Pro Tyr Ala Asp
                 85                  90                  95

Cys Leu Thr Gln Met His Phe Phe Leu Leu Phe Gly Asp Leu Glu Ser
            100                 105                 110

Phe Leu Leu Val Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys Phe
        115                 120                 125

Pro Leu His Tyr Thr Ala Ile Met Ser Pro Met Leu Cys Leu Ser Val
    130                 135                 140

Val Ala Leu Ser Trp Val Leu Thr Thr Phe His Ala Met Leu His Thr
145                 150                 155                 160

Leu Leu Met Ala Arg Leu Cys Phe Cys Ala Asp Asn Val Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Met Ser Ala Leu Leu Lys Leu Ala Cys Ser Asp Thr
            180                 185                 190

Arg Val Asn Glu Trp Val Ile Phe Ile Met Gly Gly Leu Ile Val Val
        195                 200                 205

Ile Pro Phe Leu Leu Ile Leu Gly Ser Tyr Ala Arg Ile Val Ser Ser
    210                 215                 220

Ile Leu Lys Val Pro Ser Phe Lys Gly Ile Cys Lys Ala Leu Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Phe Tyr Gly Thr Val Ile
                245                 250                 255

Gly Leu Tyr Leu Cys Pro Ser Ala Asn Ser Ser Thr Leu Lys Asp Thr
            260                 265                 270

Val Met Ala Met Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Arg Asp Met Lys Gly Ala Leu Glu Arg Val
    290                 295                 300

Ile Cys Lys Arg Lys Asn Pro Phe Leu Leu
305                 310
```

<210> 56

<211> 313
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<220>
<221> VARIANT
<222> (1)..(313)
<223> Wherein Xaa is any amino acid as defined in the specification

<400> 56

```
Met Asp Gly Gly Asn Gln Ser Glu Gly Ser Glu Phe Leu Leu Leu Gly
  1               5                  10                  15

Met Ser Glu Ser Pro Glu Gln Gln Arg Ile Leu Phe Trp Met Phe Leu
             20                  25                  30

Ser Met Tyr Leu Val Thr Val Leu Gly Asn Val Leu Ile Ile Leu Ala
         35                  40                  45

Ile Ser Ser Asp Ser Arg Leu His Thr Pro Met Tyr Phe Phe Leu Ala
     50                  55                  60

Asn Leu Ser Phe Thr Asp Leu Phe Phe Val Thr Asn Thr Ile Pro Lys
 65                  70                  75                  80

Met Leu Val Asn Leu Gln Ser Gln Asp Lys Ala Ile Ser Tyr Ala Gly
                 85                  90                  95

Cys Leu Thr Gln Leu Tyr Phe Leu Leu Ser Leu Val Thr Leu Asp Asn
            100                 105                 110

Leu Ile Leu Ala Val Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys Cys
        115                 120                 125

Pro Leu His Tyr Val Thr Ala Met Ser Pro Arg Leu Cys Ile Leu Leu
    130                 135                 140

Leu Ser Leu Cys Trp Val Phe Ser Val Leu Tyr Gly Leu Ile His Thr
145                 150                 155                 160

Leu Leu Met Thr Arg Val Thr Phe Cys Gly Ser Arg Lys Ile His Tyr
                165                 170                 175

Leu Phe Cys Glu Met Tyr Phe Leu Leu Arg Leu Ala Cys Ser Asn Ile
            180                 185                 190

Gln Ile Asn His Thr Val Leu Xaa Ala Thr Gly Cys Phe Ile Phe Leu
        195                 200                 205

Ile Pro Leu Gly Phe Met Ile Xaa Ser Tyr Ala Arg Ile Val Arg Ala
    210                 215                 220
```

```
Ile Leu Arg Ile Pro Ser Ala Thr Gly Lys Tyr Lys Ala Phe Ser Thr
225                 230                 235                 240

Cys Ala Ser His Leu Ala Val Val Ser Leu Phe Tyr Gly Thr Leu Gly
                245                 250                 255

Met Val Tyr Leu Gln Pro Leu Gln Thr Tyr Ser Thr Lys Asp Ser Val
            260                 265                 270

Ala Thr Val Met Tyr Ala Val Val Thr Pro Met Met Asn Pro Phe Ile
        275                 280                 285

Tyr Ser Leu Arg Asn Lys Asp Ile His Gly Ala Leu Gly Arg Leu Leu
    290                 295                 300

Gln Gly Lys Ala Phe Gln Lys Leu Thr
305                 310


<210> 57
<211> 313
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
      primer

<220>
<221> VARIANT
<222> (1)..(313)
<223> Wherein Xaa is any amino acid as defined in the
      specification

<400> 57
Met Glu Gly Lys Asn Leu Thr Ser Ile Ser Glu Phe Phe Leu Leu Gly
  1                 5                  10                  15

Phe Ser Glu Gln Leu Glu Glu Gln Lys Ala Leu Phe Gly Ser Phe Leu
             20                  25                  30

Phe Met Tyr Leu Val Met Val Ala Gly Asn Leu Leu Ile Ile Leu Val
         35                  40                  45

Ile Ile Thr Asp Thr Gln Leu His Thr Pro Met Tyr Phe Phe Leu Ala
     50                  55                  60

Asn Leu Ser Leu Ala Asp Ala Cys Phe Val Ser Thr Thr Val Pro Lys
 65                  70                  75                  80

Met Leu Ala Asn Ile Gln Ile Gln Ser Gln Ala Ile Ser Tyr Ser Gly
                 85                  90                  95

Cys Leu Leu Gln Leu Tyr Phe Phe Met Leu Phe Val Met Leu Glu Ala
            100                 105                 110

Phe Leu Leu Ala Val Met Ala Tyr Asp His Tyr Val Ala Ile Cys His
```

```
        115                 120                 125

Pro Leu His Tyr Ile Leu Ile Met Ser Pro Gly Leu Cys Val Phe Leu
    130                 135                 140

Val Ser Ala Ser Trp Ile Met Asn Ala Leu Tyr Ser Leu Leu His Thr
145                 150                 155                 160

Leu Leu Met Asn Ser Leu Ser Phe Cys Ala Asn His Glu Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Ile Asp Pro Leu Leu Ser Leu Ser Cys Ala Asp Pro
            180                 185                 190

Phe Thr Asn Glu Leu Val Ile Phe Ile Thr Gly Gly Leu Thr Gly Leu
        195                 200                 205

Ile Cys Val Leu Cys Leu Ile Ile Ser Tyr Thr Asn Val Phe Ser Thr
    210                 215                 220

Ile Leu Lys Ile Pro Ser Ala Gln Gly Lys Arg Lys Ala Phe Ser Thr
225                 230                 235                 240

Cys Ser Ser His Leu Ser Val Val Ser Leu Phe Xaa Gly Thr Ser Phe
                245                 250                 255

Cys Val Tyr Phe Ser Pro Pro Ser Thr Arg Xaa Ala Gln Lys Asp Thr
            260                 265                 270

Val Ala Ser Val Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Gln Glu Ile Lys Ser Ser Leu Arg Lys Leu
    290                 295                 300

Ile Trp Val Arg Lys Ile His Ser Pro
305                 310


<210> 58
<211> 157
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
      primer

<400> 58
Phe Cys His Pro Phe His Tyr Thr Met Ile Leu Thr Arg Met Leu Cys
  1               5                  10                  15

Val Lys Met Val Val Met Cys His Ala Leu Ser His Leu His Ala Met
            20                  25                  30

Leu His Thr Phe Leu Met Gly Gln Leu Ile Phe Cys Ala Asp Asn Arg
        35                  40                  45
```

Ile Pro His Phe Phe Cys Asp Leu Tyr Ala Leu Met Lys Ile Ser Cys
50 55 60

Thr Ser Thr Tyr Leu Asn Thr Leu Met Ile His Thr Glu Gly Ala Val
65 70 75 80

Val Ile Ser Gly Ala Leu Ala Phe Ile Thr Ala Ser Tyr Ala Cys Ile
85 90 95

Ile Leu Val Val Leu Arg Ile Pro Ser Ala Lys Gly Arg Trp Lys Thr
100 105 110

Phe Ser Thr Cys Gly Ser His Leu Thr Val Val Ala Ile Phe Tyr Gly
115 120 125

Thr Leu Ser Trp Val Tyr Phe Arg Pro Leu Ser Ser Tyr Ser Val Thr
130 135 140

Lys Gly Arg Ile Ile Thr Val Val Tyr Thr Val Val Thr
145 150 155

<210> 59
<211> 314
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 59
Met Met Gly Gln Asn Gln Thr Ser Ile Ser Asp Phe Leu Leu Leu Gly
1 5 10 15

Leu Pro Ile Gln Pro Glu Gln Gln Asn Leu Cys Tyr Ala Leu Phe Leu
20 25 30

Ala Met Tyr Leu Thr Thr Leu Leu Gly Asn Leu Leu Ile Ile Val Leu
35 40 45

Ile Arg Leu Asp Ser His Leu His Thr Pro Met Tyr Leu Phe Leu Ser
50 55 60

Asn Leu Ser Phe Ser Asp Leu Cys Phe Ser Ser Val Thr Ile Pro Lys
65 70 75 80

Leu Leu Gln Asn Met Gln Asn Gln Asp Pro Ser Ile Pro Tyr Ala Asp
85 90 95

Cys Leu Thr Gln Met Tyr Phe Phe Leu Leu Phe Gly Asp Leu Glu Ser
100 105 110

Phe Leu Leu Val Ala Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys Phe
115 120 125

Ala Leu His Tyr Thr Ala Ile Met Ser Pro Met Leu Cys Leu Ser Leu
130 135 140

Val Ala Leu Ser Trp Val Leu Thr Thr Phe His Ala Met Leu His Thr
145 150 155 160

Leu Leu Met Ala Arg Leu Cys Phe Cys Ala Asp Asn Val Ile Pro His
165 170 175

Phe Phe Cys Asp Met Ser Ala Leu Leu Lys Leu Ala Cys Ser Asp Thr
180 185 190

Arg Val Asn Glu Trp Val Ile Phe Ile Met Gly Gly Leu Ile Val Val
195 200 205

Ile Pro Phe Leu Leu Ile Leu Gly Ser Tyr Ala Arg Ile Val Ser Ser
210 215 220

Ile Leu Lys Val Pro Ser Ser Lys Gly Ile Cys Lys Ala Phe Ser Thr
225 230 235 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Phe Tyr Gly Thr Val Ile
245 250 255

Gly Leu Tyr Leu Cys Pro Ser Ala Asn Ser Ser Thr Leu Lys Glu Thr
260 265 270

Val Met Ala Met Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe
275 280 285

Ile Tyr Ser Leu Arg Asn Gly Asp Met Lys Gly Ala Leu Ser Arg Val
290 295 300

Ile His Gln Lys Lys Thr Phe Phe Ser Leu
305 310

<210> 60
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 60
ggagaaattt tggacagatt gtatcaga 28

<210> 61
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide

primer

<400> 61
aatcatttca gaatcacata cccattaata acc 33

<210> 62
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 62
ctgaattttt cctccgagga atatc 25

<210> 63
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 63
tctaagagga aacaatactc ctgtgac 27

<210> 64
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 64
gtttgtttct gccttttca atgtcc 26

<210> 65
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 65
aagagtaaat gtcattgaga taagactgtt gc 32

<210> 66
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
primer

<400> 66
tatgagccct gagaaccaga gcag 24

<210> 67
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
primer

<400> 67
tcaaatacat gagataaacc aaaatgagat tc 32

<210> 68
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
primer

<400> 68
tttttgtctc tcactgatat ttgc 24

<210> 69
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
primer

<400> 69
tgttcaccag aaaccagtag aaaacac 27

<210> 70
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 70
cttcttgttt ataactgagc ccaagtc 27

<210> 71
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 71
aaccagtaga aaacacgtcc aagttg 26

<210> 72
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 72
agccatccct catttctatt gt 22

<210> 73
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 73
tcctagtgct ctcctgaagc ttgcct 26

<210> 74
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 74
tcagctcgtt tacatgggta tc 22

<210> 75
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 75
atccctaaga tgctggtgaa ct 22

<210> 76
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 76
ccatctctta cagtgagtgt ctgaccca 28

<210> 77
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 77
tgtgtttcca aaggctaaga aa 22

<210> 78
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 78
ggtctctcat tacccaacca tt 22

<210> 79

<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
primer

<400> 79
cccgctgctt ggctcagttc tttt 24

<210> 80
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
primer

<400> 80
tgtaacccca aatgcataga ag 22

<210> 81
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
primer

<400> 81
ccaacctgtc ttttgttgac at 22

<210> 82
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
primer

<400> 82
cgtcctccac agttaccaag atgctg 26

<210> 83
<211> 22
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 83
ccgtatagga gatggtgtga tg 22

<210> 84
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 84
gacgtaattg cctctgtgat gt 22

<210> 85
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 85
accccattgc tgaatccctt cattta 26

<210> 86
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 86
ccctttatgt ccctgttcct ta 22

<210> 87
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 87
tcactgacat ctccctttca tc 22

<210> 88
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 88
tgtcactgtc ccaaagatgt tattaagca 29

<210> 89
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 89
gaatggattg atcctgagtt tg 22

<210> 90
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 90
ctaagggcat cttcaaagct tt 22

<210> 91
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 91
ctgtggctct cacctctctg tggtgt 26

<210> 92
<211> 22
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 92
gaggggagaa aatacagtcc aa 22

<210> 93
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 93
ggacatgcgg actaagtaca aa 22

<210> 94
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 94
cgatcctcta tgaggaatgc atttctca 28

<210> 95
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 95
gaagctgtcc aggtcagtaa aa 22

<210> 96
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 96
aggctgtaga taaaggggtt ca 22

<210> 97
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide primer

<400> 97
tgagagccac aatgacatcc ttgtca 26

<210> 98
<211> 22
<212> DNA
<213> Homo sapiens

<400> 98
ttcccgactg taagcagttc ta 22

<210> 99
<211> 22
<212> DNA
<213> Homo sapiens

<400> 99
actctcctct cggtctgaag ac 22

<210> 100
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide primer

<400> 100
agaggccact ctttgccctc tttctt 26

<210> 101
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide primer

<400> 101
tagccaagat gatgagcaga tt 22

<210> 102
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide primer

<400> 102
tccaatgtta tccaccactt tc 22

<210> 103
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide primer

<400> 103
ccctgtgctg aaattgtcct gctctt 26

<210> 104
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide primer

<400> 104
tctgcacaat ttcattgaca aa 22

<210> 105
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide primer

<400> 105
gccatcatca cactgattct ct 22

<210> 106
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide
primer

<400> 106
ccacatcccc atgtacatct tcctca 26

<210> 107
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide
primer

<400> 107
gatgtctgtc aaggccaagt ta 22

<210> 108
<211> 1051
<212> DNA
<213> Homo sapiens

<400> 108
tctatttcca gcagagagaa gactgtcagc atgaagaggg agaatcagag cagtgtgtct 60
gagttcctcc tcctggacct ccccatctgg ccagagcagc aggctgtgtt cttcaccctg 120
ttcttgggca tgtacctgat cacggtgctg gggaacctgc tcatcatcct gctcatccgg 180
ctggactctc accttcacac ccccatgttc ttcttcctca gccacttggc tctcactgac 240
atctcccttt catctgtcac tgtcccaaag atgttattaa gcatgcaaac tcaggatcaa 300
tccattcttt atgcagggtg tgtaactcag atgtattttt tcatattttt cactgatcta 360
gacaatttcc ttctcacttc aatggcatac gatcggtatg tggccatctg tcacccccctc 420
cgctacacca ctatcatgaa agagggactg tgtaacttac tagtcactgt gtcctggatc 480
ctctcctgta ccaatgccct gtctcacact ctcctcctgg cccagctgtc cttttgtgct 540
gacaacacca tcccccattt cttctgtgat cttgttgccc tactcaagct ctcatgctca 600
gacatctccc tcaatgagct ggtcattttc acagtgggac aggcagtcat tactctacca 660
ctaatatgca tcttgatctc ttatggccac attggggtca ccatcctcaa ggctccatct 720
actaagggca tcttcaaagc tttgtccacc tgtggctctc acctctctgt ggtgtctctg 780
tattatggca caattattgg actgtatttt ctcccctcat ccagtgcctc cagtgacaag 840
gacgtaattg cctctgtgat gtacacggtg atcaccccat tgctgaatcc cttcatttat 900
agcctaagga acagggacat aaagggagcc ctggagagac tcttcaacag ggcaacagtc 960
ttatctcaat gacatttact cttctttata acagacatat gtactgacct atttccagat 1020
catagatcct tacttctgat cccagcaagg g 1051

<210> 109
<211> 313
<212> PRT
<213> Homo sapiens

<400> 109

```
Met Lys Arg Glu Asn Gln Ser Ser Val Ser Glu Phe Leu Leu Leu Asp
  1               5                  10                  15

Leu Pro Ile Trp Pro Glu Gln Gln Ala Val Phe Phe Thr Leu Phe Leu
             20                  25                  30

Gly Met Tyr Leu Ile Thr Val Leu Gly Asn Leu Leu Ile Ile Leu Leu
         35                  40                  45

Ile Arg Leu Asp Ser His Leu His Thr Pro Met Phe Phe Phe Leu Ser
     50                  55                  60

His Leu Ala Leu Thr Asp Ile Ser Leu Ser Ser Val Thr Val Pro Lys
 65                  70                  75                  80

Met Leu Leu Ser Met Gln Thr Gln Asp Gln Ser Ile Leu Tyr Ala Gly
                 85                  90                  95

Cys Val Thr Gln Met Tyr Phe Phe Ile Phe Phe Thr Asp Leu Asp Asn
            100                 105                 110

Phe Leu Leu Thr Ser Met Ala Tyr Asp Arg Tyr Val Ala Ile Cys His
        115                 120                 125

Pro Leu Arg Tyr Thr Thr Ile Met Lys Glu Gly Leu Cys Asn Leu Leu
    130                 135                 140

Val Thr Val Ser Trp Ile Leu Ser Cys Thr Asn Ala Leu Ser His Thr
145                 150                 155                 160

Leu Leu Leu Ala Gln Leu Ser Phe Cys Ala Asp Asn Thr Ile Pro His
                165                 170                 175

Phe Phe Cys Asp Leu Val Ala Leu Leu Lys Leu Ser Cys Ser Asp Ile
            180                 185                 190

Ser Leu Asn Glu Leu Val Ile Phe Thr Val Gly Gln Ala Val Ile Thr
        195                 200                 205

Leu Pro Leu Ile Cys Ile Leu Ile Ser Tyr Gly His Ile Gly Val Thr
    210                 215                 220

Ile Leu Lys Ala Pro Ser Thr Lys Gly Ile Phe Lys Ala Leu Ser Thr
225                 230                 235                 240

Cys Gly Ser His Leu Ser Val Val Ser Leu Tyr Tyr Gly Thr Ile Ile
                245                 250                 255

Gly Leu Tyr Phe Leu Pro Ser Ser Ser Ala Ser Ser Asp Lys Asp Val
            260                 265                 270

Ile Ala Ser Val Met Tyr Thr Val Ile Thr Pro Leu Leu Asn Pro Phe
        275                 280                 285

Ile Tyr Ser Leu Arg Asn Arg Asp Ile Lys Gly Ala Leu Glu Arg Leu
    290                 295                 300
```

Phe Asn Arg Ala Thr Val Leu Ser Gln
305 310

**WHAT IS CLAIMED IS:**

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38;

(b) a variant of a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of the amino acid residues from the amino acid sequence of said mature form;

(c) an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38; and

(d) a variant of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence.

2 The polypeptide of claim 1, wherein said polypeptide comprises the amino acid sequence of a naturally-occurring allelic variant of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38.

3. The polypeptide of claim 2, wherein said allelic variant comprises an amino acid sequence that is the translation of a nucleic acid sequence differing by a single nucleotide from a nucleic acid sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37.

4. The polypeptide of claim 1, wherein the amino acid sequence of said variant comprises a conservative amino acid substitution.

5. An isolated nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38;

(b) a variant of a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of the amino acid residues from the amino acid sequence of said mature form;

(c) an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38;

(d) a variant of an amino acid sequence selected from the group consisting SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence;

(e) a nucleic acid fragment encoding at least a portion of a polypeptide comprising an amino acid sequence chosen from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, or a variant of said polypeptide, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence; and

(f) a nucleic acid molecule comprising the complement of (a), (b), (c), (d) or (e).

6. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule comprises the nucleotide sequence of a naturally-occurring allelic nucleic acid variant.

7. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule encodes a polypeptide comprising the amino acid sequence of a naturally-occurring polypeptide variant.

8. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule differs by a single nucleotide from a nucleic acid sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37.

9. The nucleic acid molecule of claim 5, wherein said nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37;

(b) a nucleotide sequence differing by one or more nucleotides from a nucleotide sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, provided that no more than 20% of the nucleotides differ from said nucleotide sequence;

(c) a nucleic acid fragment of (a); and

(d) a nucleic acid fragment of (b).

10. The nucleic acid molecule of claim 5, wherein said nucleic acid molecule hybridizes under stringent conditions to a nucleotide sequence chosen from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 and 37, or a complement of said nucleotide sequence.

11. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of:

(a) a first nucleotide sequence comprising a coding sequence differing by one or more nucleotide sequences from a coding sequence encoding said amino acid sequence, provided that no more than 20% of the nucleotides in the coding sequence in said first nucleotide sequence differ from said coding sequence;

(b) an isolated second polynucleotide that is a complement of the first polynucleotide; and

(c) a nucleic acid fragment of (a) or (b).

12. A vector comprising the nucleic acid molecule of claim 11.

13. The vector of claim 12, further comprising a promoter operably-linked to said nucleic acid molecule.

14. A cell comprising the vector of claim 12.

15. An antibody that binds immunospecifically to the polypeptide of claim 1.

16. The antibody of claim 15, wherein said antibody is a monoclonal antibody.

17. The antibody of claim 15, wherein the antibody is a humanized antibody.

18. A method for determining the presence or amount of the polypeptide of claim 1 in a sample, the method comprising:

(a) providing the sample;

(b) contacting the sample with an antibody that binds immunospecifically to the polypeptide; and

(c) determining the presence or amount of antibody bound to said polypeptide,

thereby determining the presence or amount of polypeptide in said sample.

19. A method for determining the presence or amount of the nucleic acid molecule of claim 5 in a sample, the method comprising:

(a) providing the sample;

(b) contacting the sample with a probe that binds to said nucleic acid molecule; and

(c) determining the presence or amount of the probe bound to said nucleic acid molecule,

thereby determining the presence or amount of the nucleic acid molecule in said sample.

20. The method of claim 19 wherein presence or amount of the nucleic acid molecule is used as a marker for cell or tissue type.

21. The method of claim 20 wherein the cell or tissue type is cancerous.

22. A method of identifying an agent that binds to a polypeptide of claim 1, the method comprising:

(a) contacting said polypeptide with said agent; and

(b) determining whether said agent binds to said polypeptide.

23. The method of claim 22 wherein the agent is a cellular receptor or a downstream effector.

24. A method for identifying an agent that modulates the expression or activity of the polypeptide of claim 1, the method comprising:

(a) providing a cell expressing said polypeptide;

(b) contacting the cell with said agent, and

(c) determining whether the agent modulates expression or activity of said polypeptide, whereby an alteration in expression or activity of said peptide indicates said agent modulates expression or activity of said polypeptide.

25. A method for modulating the activity of the polypeptide of claim 1, the method comprising contacting a cell sample expressing the polypeptide of said claim with a compound that binds to said polypeptide in an amount sufficient to modulate the activity of the polypeptide.

26. A method of treating or preventing a GPCRX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the polypeptide of claim 1 in an amount sufficient to treat or prevent said GPCRX-associated disorder in said subject.

27. The method of claim 26 wherein the disorder is selected from the group consisting of cardiomyopathy and atherosclerosis.

28. The method of claim 26 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

29. The method of claim 26, wherein said subject is a human.

30. A method of treating or preventing a GPCRX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the nucleic acid of claim 5 in an amount sufficient to treat or prevent said GPCRX-associated disorder in said subject.

31. The method of claim 30 wherein the disorder is selected from the group consisting of cardiomyopathy and atherosclerosis.

32. The method of claim 30 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

33. The method of claim 30, wherein said subject is a human.

34. A method of treating or preventing a GPCRX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the antibody of claim 15 in an amount sufficient to treat or prevent said GPCRX-associated disorder in said subject.

35. The method of claim 34 wherein the disorder is diabetes.

36. The method of claim 34 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

37. The method of claim 34, wherein the subject is a human.

38. A pharmaceutical composition comprising the polypeptide of claim 1 and a pharmaceutically-acceptable carrier.

39. A pharmaceutical composition comprising the nucleic acid molecule of claim 5 and a pharmaceutically-acceptable carrier.

40. A pharmaceutical composition comprising the antibody of claim 15 and a pharmaceutically-acceptable carrier.

41. A kit comprising in one or more containers, the pharmaceutical composition of claim 38.

42. A kit comprising in one or more containers, the pharmaceutical composition of claim 39.

43. A kit comprising in one or more containers, the pharmaceutical composition of claim 40.

44. A method for determining the presence of or predisposition to a disease associated with altered levels of the polypeptide of claim 1 in a first mammalian subject, the method comprising:

(a) measuring the level of expression of the polypeptide in a sample from the first mammalian subject; and

(b) comparing the amount of said polypeptide in the sample of step (a) to the amount of the polypeptide present in a control sample from a second mammalian subject known not to have, or not to be predisposed to, said disease;

wherein an alteration in the expression level of the polypeptide in the first subject as compared to the control sample indicates the presence of or predisposition to said disease.

45. The method of claim 44 wherein the predisposition is to cancers.

46. A method for determining the presence of or predisposition to a disease associated with altered levels of the nucleic acid molecule of claim 5 in a first mammalian subject, the method comprising:

(a) measuring the amount of the nucleic acid in a sample from the first mammalian subject; and

(b) comparing the amount of said nucleic acid in the sample of step (a) to the amount of the nucleic acid present in a control sample from a second mammalian subject known not to have or not be predisposed to, the disease;

wherein an alteration in the level of the nucleic acid in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.

47. The method of claim 46 wherein the predisposition is to a cancer.

48. A method of treating a pathological state in a mammal, the method comprising administering to the mammal a polypeptide in an amount that is sufficient to alleviate the pathological state, wherein the polypeptide is a polypeptide having an amino acid sequence at least 95% identical to a polypeptide comprising an amino acid sequence of at least one of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, or a biologically active fragment thereof.

49. A method of treating a pathological state in a mammal, the method comprising administering to the mammal the antibody of claim 15 in an amount sufficient to alleviate the pathological state.

50. A method for the screening of a candidate substance interacting with an olfactory receptor polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, or fragments or variants thereof, comprises the following steps:

a) providing a polypeptide selected from the group consisting of the sequences of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, or a peptide fragment or a variant thereof;

b) obtaining a candidate substance;

c) bringing into contact said polypeptide with said candidate substance; and

d) detecting the complexes formed between said polypeptide and said candidate substance.

51. A method for the screening of ligand molecules interacting with an olfactory receptor polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, wherein said method comprises:

a) providing a recombinant eukaryotic host cell containing a nucleic acid encoding a polypeptide selected from the group consisting of the polypeptides comprising the amino acid sequences SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38;

b) preparing membrane extracts of said recombinant eukaryotic host cell;

c) bringing into contact the membrane extracts prepared at step b) with a selected ligand molecule; and

d) detecting the production level of second messengers metabolites.

52. A method for the screening of ligand molecules interacting with an olfactory receptor polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38, wherein said method comprises:

a) providing an adenovirus containing a nucleic acid encoding a polypeptide selected from the group consisting of polypeptides comprising the amino acid sequences SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36 and 38;

b) infecting an olfactory epithelium with said adenovirus;

c) bringing into contact the olfactory epithelium b) with a selected ligand molecule; and

d) detecting the increase of the response to said ligand molecule.