

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-500098

(P2004-500098A)

(43) 公表日 平成16年1月8日(2004.1.8)

(51) Int. Cl. ⁷		F I		テーマコード (参考)	
C 1 2 N	15/09	C 1 2 N	15/00	Z N A A	2 G O 4 5
A 6 1 K	38/00	A 6 1 K	39/395	N	4 B O 2 4
A 6 1 K	39/395	A 6 1 P	1/04		4 B O 6 3
A 6 1 P	1/04	A 6 1 P	1/14		4 B O 6 4
A 6 1 P	1/14	A 6 1 P	3/00		4 B O 6 5
		審査請求	未請求	予備審査請求	有 (全 275 頁) 最終頁に続く

(21) 出願番号	特願2001-564362 (P2001-564362)	(71) 出願人	500056297
(86) (22) 出願日	平成13年2月28日 (2001.2.28)		キュラジェン コーポレイション
(85) 翻訳文提出日	平成14年8月27日 (2002.8.27)		アメリカ合衆国 コネチカット州
(86) 国際出願番号	PCT/US2001/006474		11 ニュー ヘイブン ロング ウォー
(87) 国際公開番号	W02001/064879		フ ドライブ 5 5 5
(87) 国際公開日	平成13年9月7日 (2001.9.7)	(74) 代理人	100107489
(31) 優先権主張番号	60/185,674		弁理士 大塩 竹志
(32) 優先日	平成12年2月28日 (2000.2.28)	(72) 発明者	パディガル, マラリダラ
(33) 優先権主張国	米国 (US)		アメリカ合衆国 ニューヨーク
(31) 優先権主張番号	60/185,535		1, ブロンクス, ラインランダー
(32) 優先日	平成12年2月28日 (2000.2.28)		アベニュー 1 5 7 9, アpartment
(33) 優先権主張国	米国 (US)		6 イー
(31) 優先権主張番号	60/186,585		
(32) 優先日	平成12年3月3日 (2000.3.3)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 Gタンパク質結合レセプター関連ポリペプチド

(57) 【要約】

Gタンパク質結合レセプター関連ポリペプチドをコードする核酸配列が、本明細書中で開示される。これらの核酸配列によってコードされるポリペプチド、およびこのポリペプチドに免疫特異的に結合する抗体、ならびにこれらのポリペプチド、ポリヌクレオチドまたは抗体のフラグメントの誘導体、改変体、変異体もまた、開示される。本発明はさらに、これらの新規なヒト核酸およびタンパク質のいずれか1つに関与する障害を診断、処置、および予防するための、治療方法、診断方法、および研究方法を開示する。

【特許請求の範囲】

【請求項 1】

単離されたポリペプチドであって、以下：

- (a) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列の成熟形態；
- (b) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列の成熟形態の改変体であって、ここで、該改変体中の 1 以上のアミノ酸残基が、該成熟形態のアミノ酸配列と異なり、ただし、該改変体は、該成熟形態のアミノ酸配列と、そのアミノ酸残基のうちの 15% 以下において異なる、改変体；
- (c) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列；ならびに
- (d) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列の改変体であって、ここで、該改変体中の 1 以上のアミノ酸残基が、該成熟形態のアミノ酸配列と異なり、ただし、該改変体は、該アミノ酸配列と、そのアミノ酸残基のうちの 15% 以下において異なる、改変体；

10

からなる群より選択されるアミノ酸配列を含む、ポリペプチド。

【請求項 2】

請求項 1 に記載のポリペプチドであって、該ポリペプチドが、配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列の天然に存在する対立遺伝子改変体のアミノ酸配列を含む、ポリペプチド。

20

【請求項 3】

請求項 2 に記載のポリペプチドであって、ここで、前記対立遺伝子改変体が、配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 からなる群から選択される核酸配列と単一ヌクレオチドで異なる核酸配列の翻訳であるアミノ酸配列を含む、ポリペプチド。

【請求項 4】

前記改変体のアミノ酸配列が、保存的アミノ酸置換を含む、請求項 1 に記載のポリペプチド。

30

【請求項 5】

単離された核酸分子であって、以下：

- (a) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群から選択されるアミノ酸配列の成熟形態；
- (b) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列の成熟形態の改変体であって、ここで、該改変体の 1 以上のアミノ酸残基が、該成熟形態のアミノ酸配列と異なり、ただし、該改変体は、該成熟形態のアミノ酸配列と、そのアミノ酸残基のうちの 15% 以下において異なる、改変体；
- (c) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列；
- (d) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列の改変体であって、ここで、該改変体の 1 以上のアミノ酸残基が、該成熟形態のアミノ酸配列と異なり、ただし、該改変体は、該アミノ酸配列と、そのアミノ酸残基のうちの 15% 以下において異なる、改変体；
- (e) 配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 からなる群より選択されるアミノ酸配列を含むポリペプチドまたは該ポリペプチドの改変体の、少なくとも一部をコードする核酸フラグメント

40

50

であって、ここで、該改変体の 1 以上のアミノ酸残基が、該成熟形態のアミノ酸配列と異なり、ただし、該改変体は、該アミノ酸配列と、そのアミノ酸残基のうちの 15% 以下において異なる、核酸フラグメント；ならびに

(f) (a)、(b)、(c)、(d) または (e) の相補体を含む、核酸分子、
からなる群から選択されるアミノ酸配列を含むポリペプチドをコードする核酸配列を含む、核酸分子。

【請求項 6】

請求項 5 に記載の核酸分子であって、ここで、該核酸分子が天然に存在する対立遺伝子核酸改変体のヌクレオチド配列を含む、核酸分子。

【請求項 7】

請求項 5 に記載の核酸分子であって、ここで、該核酸分子が、天然に存在するポリペプチド改変体のアミノ酸配列を含むポリペプチドをコードする、核酸分子。

【請求項 8】

請求項 5 に記載の核酸分子であって、ここで、該核酸分子が、配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 からなる群から選択される核酸配列と単一ヌクレオチドで異なる、核酸分子。

【請求項 9】

請求項 5 に記載の核酸分子であって、ここで、該核酸分子が以下：

(a) 配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 からなる群より選択されるヌクレオチド配列；

(b) 配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 からなる群より選択されるヌクレオチド配列と 1 つ以上のヌクレオチドで異なる、ヌクレオチド配列であって、ただし、そのヌクレオチドの 20% 以下が、該ヌクレオチド配列と異なる、ヌクレオチド配列；

(c) (a) の核酸フラグメント；ならびに

(d) (b) の核酸フラグメント；

からなる群より選択されるヌクレオチド配列を含む、核酸分子。

【請求項 10】

請求項 5 に記載の核酸分子であって、ここで、該核酸分子が、配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 からなる群より選択されるヌクレオチド配列または該ヌクレオチド配列の相補鎖にストリンジェントな条件下でハイブリダイズする、核酸分子。

【請求項 11】

請求項 5 に記載の核酸分子であって、該核酸分子は、以下：

(a) 第 1 ヌクレオチド配列であって、前記アミノ酸配列をコードするコード配列と 1 以上のヌクレオチド配列で異なるコード配列を含み、ただし、該第 1 ヌクレオチド配列中のコード配列におけるヌクレオチドの 20% 以下が、該コード配列と異なる、第 1 ヌクレオチド配列；

(b) 該第 1 ポリヌクレオチドの相補体である、単離された第 2 ポリヌクレオチド；ならびに

(c) (a) または (b) の核酸フラグメント；

からなる群から選択されるヌクレオチド配列を含む、核酸分子。

【請求項 12】

請求項 11 に記載の核酸分子を含む、ベクター。

【請求項 13】

前記核酸分子に作動可能に連結されたプロモーターをさらに含む、請求項 12 に記載のベクター。

【請求項 14】

請求項 12 に記載のベクターを含む、細胞。

【請求項 15】

10

20

30

40

50

請求項 1 に記載のポリペプチドに免疫特異的に結合する、抗体。

【請求項 1 6】

前記抗体がモノクローナル抗体である、請求項 1 5 に記載の抗体。

【請求項 1 7】

前記抗体がヒト化抗体である、請求項 1 5 に記載の抗体。

【請求項 1 8】

サンプル中の請求項 1 に記載のポリペプチドの存在または量を決定するための方法であって、該方法は、以下：

(a) 該サンプルを提供する工程；

(b) 該サンプルを、該ポリペプチドに免疫特異的に結合する抗体と接触させる工程；および

(c) 該ポリペプチドに結合した抗体の存在または量を決定して、それによって該サンプル中のポリペプチドの存在または量を決定する工程、
を包含する、方法。

【請求項 1 9】

サンプル中の請求項 5 に記載の核酸分子の存在または量を決定するための方法であって、該方法は、以下：

(a) 該サンプルを提供する工程；

(b) 該サンプルを、該核酸分子に結合するプローブと接触させる工程；および

(c) 該核酸分子に結合した該プローブの存在または量を決定して、それによって該サンプル中の該核酸分子の存在または量を決定する工程、
を包含する、方法。

【請求項 2 0】

請求項 1 9 に記載の方法であって、前記核酸分子の存在または量が、細胞型または組織型についてのマーカーとして使用される、方法。

【請求項 2 1】

前記細胞型または組織型が癌性である、請求項 2 0 に記載の方法。

【請求項 2 2】

請求項 1 に記載のポリペプチドに結合する因子を同定する方法であって、該方法は、以下：

(a) 該ポリペプチドを該因子と接触させる工程；および

(b) 該因子が該ポリペプチドに結合するか否かを決定する工程；
を包含する、方法。

【請求項 2 3】

前記因子が、細胞レセプターまたは下流エフェクターである、請求項 2 2 に記載の方法。

【請求項 2 4】

請求項 1 のポリペプチドの発現または活性を調節する因子を同定するための方法であって、該方法は、以下：

(a) 該ポリペプチドを発現する細胞を提供する工程；

(b) 該因子と該細胞を接触させる工程；および

(c) 該因子が、該ポリペプチドの発現または活性を調節するか否かを決定する工程、
を包含し、それによって、該ペプチドの発現または活性における変化が、該ポリペプチドの発現または活性を調節する因子を示す、方法。

【請求項 2 5】

請求項 1 に記載のポリペプチドの活性を調節するための方法であって、該方法は、請求項 1 に記載のポリペプチドを発現する細胞サンプルを、該ポリペプチドの活性を調節するのに十分な量で、該ポリペプチドに結合する化合物とを接触させる工程を包含する、方法。

【請求項 2 6】

GPCR 関連障害を処置または予防する方法であって、該方法は、そのような処置または予防が所望される被験体に、該被験体における該 GPCR 関連障害を処置または予防

10

20

30

40

50

するに十分な量で、請求項 1 に記載のポリペプチドを投与する工程を包含する、方法。

【請求項 27】

前記障害が、心筋症およびアテローム硬化症からなる群から選択される、請求項 26 に記載の方法。

【請求項 28】

前記障害が、細胞シグナルのプロセッシングおよび代謝経路の調節に関連する、請求項 26 に記載の方法。

【請求項 29】

前記被験体がヒトである、請求項 26 に記載の方法。

【請求項 30】

GPCR 関連障害を処置または予防する方法であって、該方法は、そのような処置または予防が所望される被験体に、該被験体における該 GPCR 関連障害を処置または予防するに十分な量で、請求項 5 に記載の核酸を投与する工程を包含する、方法。

【請求項 31】

前記障害が、心筋症およびアテローム硬化症からなる群から選択される、請求項 30 に記載の方法。

【請求項 32】

前記障害が、細胞シグナルのプロセッシングおよび代謝経路の調節に関連する、請求項 30 に記載の方法。

【請求項 33】

前記被験体がヒトである、請求項 30 に記載の方法。

【請求項 34】

GPCR 関連障害を処置または予防する方法であって、該方法は、そのような処置または予防が所望される被験体に、該被験体における該 GPCR 関連障害を処置または予防するに十分な量で、請求項 15 に記載の抗体を投与する工程を包含する、方法。

【請求項 35】

前記障害が糖尿病である、請求項 34 に記載の方法。

【請求項 36】

前記障害が、細胞シグナルのプロセッシングおよび代謝経路の調節に関連する、請求項 34 に記載の方法。

【請求項 37】

前記被験体がヒトである、請求項 34 に記載の方法。

【請求項 38】

請求項 1 に記載のポリペプチドおよび薬学的に受容可能なキャリアを含有する、薬学的組成物。

【請求項 39】

請求項 5 に記載の核酸分子および薬学的に受容可能なキャリアを含有する、薬学的組成物。

【請求項 40】

請求項 15 に記載の抗体および薬学的に受容可能なキャリアを含有する、薬学的組成物。

【請求項 41】

請求項 38 に記載の薬学的組成物を 1 つ以上の容器に含む、キット。

【請求項 42】

請求項 39 に記載の薬学的組成物を 1 つ以上の容器に含む、キット。

【請求項 43】

請求項 40 に記載の薬学的組成物を 1 つ以上の容器に含む、キット。

【請求項 44】

第 1 の哺乳動物被験体における請求項 1 に記載のポリペプチドの変化したレベルと関連する疾患の存在または該疾患に対する素因を決定するための方法であって、該方法は、以下：

10

20

30

40

50

(a) 該第1の哺乳動物被験体由来のサンプル中の該ポリペプチドの発現レベルを測定する工程；および

(b) 工程(a)のサンプル中の該ポリペプチドの量を、該疾患を有さないことが既知であるかまたは該疾患にかかりにくいことが既知の、第2の哺乳動物被験体由来のコントロールサンプル中に存在する該ポリペプチドの量と比較する工程、
を包含し、ここで、該コントロールサンプルと比較した場合の該第1の被験体における該ポリペプチドの発現レベルにおける変化が、該疾患の存在または該疾患に対する素因を示す、方法。

【請求項45】

前記素因が癌に対するものである、請求項44に記載の方法。

10

【請求項46】

第1の哺乳動物被験体における請求項5に記載の核酸分子の変化したレベルと関連する疾患の存在または該疾患に対する素因を決定するための方法であって、該方法は、以下：

(a) 該第1の哺乳動物被験体由来のサンプル中の該核酸の量を測定する工程；および

(b) 工程(a)のサンプル中の該核酸の量を、該疾患を有さないことが既知であるかまたは該疾患にかかりにくいことが既知の、第2の哺乳動物被験体由来のコントロールサンプル中に存在する該核酸の量と比較する工程、

を包含し、ここで、該コントロールサンプルと比較した場合の該第1の被験体における該核酸のレベルにおける変化が、該疾患の存在または該疾患に対する素因を示す、方法。

【請求項47】

前記素因が癌に対するものである、請求項46に記載の方法。

20

【請求項48】

哺乳動物において病理学的状態を処置する方法であって、該方法は、該病理状態を緩和するに十分な量でポリペプチドを該哺乳動物に投与する工程を包含し、ここで、該ポリペプチドが、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のうちの少なくとも1つのアミノ酸配列を含むポリペプチドまたはそれらの生物学的に活性なフラグメントと、少なくとも95%同一なアミノ酸配列を有するポリペプチドである、方法。

【請求項49】

哺乳動物において病理学的状態を処置する方法であって、該方法は、該病理学的状態を緩和するに十分な量で請求項15に記載の抗体を該哺乳動物に投与する工程を包含する、方法。

30

【請求項50】

配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83からなる群から選択される嗅覚レセプターポリペプチド、またはそのフラグメントもしくは改変体と相互作用する候補物質をスクリーニングするための方法であって、該方法は、以下の工程：

a) 配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83の配列からなる群から選択されるポリペプチド、あるいはそのペプチドフラグメントまたは改変体を提供する工程；

40

b) 候補物質を得る工程；

c) 該ポリペプチドを、該候補物質と接触させる工程；ならびに

d) 該ポリペプチドと該候補物質との間で形成した複合体を検出する工程、
を包含する、方法。

【請求項51】

配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83からなる群から選択される嗅覚レセプターポリペプチドと相互作用するリガンド分子をスクリーニングするための方法であって、該方法は、以下の工程：

a) 配列番号2、4、6、8、10、12、17、19、21、23、25、29、31

50

、 33、35、37、および83のアミノ酸配列を含むポリペプチドからなる群から選択されるポリペプチドをコードする核酸を含む組換え真核生物宿主細胞を提供する工程；
b) 該組換え真核生物宿主細胞の膜抽出物を調製する工程；
c) 工程b)で調製した膜抽出物を、選択したリガンド分子と接触させる工程；ならびに
d) セカンドメッセンジャー代謝産物の産生レベルを検出する工程、
を包含する、方法。

【請求項52】

配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83からなる群から選択される嗅覚レセプターポリペプチドと相互作用するリガンド分子をスクリーニングするための方法であって、該方法は、以下：

10

a) 配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のアミノ酸配列を含むポリペプチドからなる群から選択されるポリペプチドをコードする核酸を含むアデノウイルスを提供する工程；
b) 嗅覚上皮に該アデノウイルスを感染させる工程；
c) 工程b)の嗅覚上皮を、選択したリガンド分子と接触させる工程；ならびに
d) 該リガンド分子に対する応答の増加を検出する工程、
を包含する、方法。

【発明の詳細な説明】

【0001】

(発明の背景)

20

本発明は、一般に、核酸およびポリペプチドに関する。より詳細には、本発明は、新規なGタンパク質結合レセプター(GPCR)ポリペプチドをコードする核酸、ならびにこれらの核酸およびポリペプチドを産生するためのベクター、宿主細胞、抗体および組換え方法に関する。

【0002】

(発明の要旨)

本発明は、新規ポリペプチドをコードする核酸配列の発見に一部基づく。これらの新規核酸およびポリペプチドは、本明細書中でGPCRX、あるいはGPCR1、GPCR2、GPCR3、GPCR4、GPCR5、GPCR6、GPCR7、GPCR8、GPCR9およびGPCR10核酸およびポリペプチドと呼ばれる。これらの核酸およびポリペプチド、ならびにそれらの誘導体、ホモログ、アナログおよびフラグメントは、本明細書中

30

【0003】

1つの局面において、本発明は、GPCRXポリペプチドをコードする単離されたGPCRX核酸分子を提供し、これは、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36および38に開示される核酸に対して同一性を有する核酸配列を含む。いくつかの実施形態において、このGPCRX核酸分子は、GPCRX核酸配列のタンパク質コード配列を含む核酸分子に相補的な核酸配列に、ストリンジェントな条件下でハイブリダイズする。本発明はまた、GPCRXポリペプチド、あるいはそのフラグメント、ホモログ、アナログまたは誘導体をコードする、単離された核酸を包含する。例えば、この核酸は、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37および83のアミノ酸配列を含むポリペプチドに対して少なくとも80%同一であるポリペプチドをコードし得る。この核酸は、例えば、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36および38のいずれかの核酸配列を含む、ゲノムDNAフラグメントまたはcDNA分子であり得る。

40

【0004】

オリゴヌクレオチド(例えば、GPCRX核酸(例えば、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36および38)の少なくとも6個連続するヌクレオチドを含むオリゴヌクレオチドまたはそのオリゴヌ

50

クレオチドの相補体)もまた、本発明に含まれる。

【0005】

実質的に精製されたGPCRXポリペプチド(配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37および83)もまた、本発明に含まれる。特定の実施形態において、このGPCRXポリペプチドは、ヒトGPCRXポリペプチドのアミノ酸配列に実質的に同一であるアミノ酸配列を含む。

【0006】

本発明はまた、GPCRXポリペプチド、あるいはそれらのフラグメント、ホモログ、アナログまたは誘導体に免疫選択的に結合する抗体を特徴とする。

【0007】

別の局面において、本発明は、治療有効量または予防有効量の治療剤および薬学的に受容可能なキャリアを含む、薬学的組成物を包含する。この治療剤は、例えば、GPCRX核酸、GPCRXポリペプチド、またはGPCRXポリペプチドに特異的な抗体であり得る。さらなる局面において、本発明は、1以上の容器に、この薬学的組成物の治療有効量または予防有効量を含む。

【0008】

さらなる局面において、本発明は、ポリペプチドを産生する方法を包含し、この方法は、GPCRX核酸を含む細胞を、このDNAによってコードされるGPCRXポリペプチドの発現を可能にする条件下で培養することによる。所望の場合、次いで、このGPCRXポリペプチドを回収し得る。

【0009】

別の局面において、本発明は、サンプル中のGPCRXポリペプチドの存在を検出する方法を包含する。この方法において、サンプルを、このポリペプチドに選択的に結合する化合物と、このポリペプチドとこの化合物との間での複合体の形成を可能にする条件下で接触させる。この複合体は、存在するならば、検出され、それによって、このサンプル中のGPCRXポリペプチドを同定する。

【0010】

本発明はまた、GPCRXの発現に基づいて特定の細胞型または組織型を同定するための方法を包含する。

【0011】

サンプル中のGPCRX核酸分子の存在を検出する方法もまた本発明に包含され、この方法は、このサンプルに、GPCRX核酸プローブまたはプライマーを接触させ、そしてこの核酸プローブまたはプライマーが、このサンプル中のGPCRX核酸分子に結合したか否かを検出するによる。

【0012】

さらなる局面において、本発明は、GPCRXポリペプチドの活性を調節するための方法を提供し、この方法は、このGPCRXポリペプチドを含む細胞サンプルに、このGPCRXポリペプチドに結合する化合物を、このポリペプチドの活性を調節するに十分な量で接触させることによる。この化合物は、例えば、本明細書中にさらに記載されるような、小分子(例えば、核酸、ペプチド、ポリペプチド、ペプチド模倣物、糖質、脂質または他の有機分子(炭素含有)または無機分子)であり得る。

【0013】

例えば、以下を含む障害または症候群を処置または予防するための医薬の製造における治療剤の使用もまた、本発明の範囲内にある：糖尿病、肥満に関連する代謝障害、代謝症候群X、食欲不振、るいそう障害(慢性疾患、代謝障害、糖尿病、肥満、感染性疾患、食欲不振、癌関連悪液質、癌、神経変性障害、アルツハイマー病、パーキンソン病、免疫障害および造血障害に関連する)、または細胞シグナル伝達プロセッシングおよび代謝経路調節に関連する他の障害。治療剤は、例えば、GPCRX核酸、GPCRXポリペプチド、またはGPCRX特異的な抗体、あるいはそれらの生物学的に活性な誘導体またはフラグメントであり得る。

10

20

30

40

50

【0014】

例えば、本発明の組成物は、以下に罹患している患者の処置に対する効力を有する：発生疾患、MHC II および I I I 疾患（免疫疾患）、味覚および嗅覚の検出能の障害、パーキットリンパ腫、皮質神経性（c o r t i c o n e u r o g e n i c）疾患、シグナル伝達経路障害、網膜疾患（光受容性に関連する疾患を含む）、細胞増殖速度障害；細胞形状障害、摂食障害；摂食の制御；過食に起因する潜在的肥満；飢餓（食欲の損失）に起因する潜在的障害、非インスリン依存性糖尿病（N I D D M 1）、細菌、真菌、原生生物およびウイルス感染（特に、H I V - 1 または H I V - 2 によって引き起こされる感染）、疼痛、癌（新生物；腺癌；リンパ腫；前立腺癌；子宮癌を含むが、これらに限定されない）、食欲不振、過食症、ぜん息、パーキンソン病、急性心不全、低血圧、高血圧、残尿、骨粗しょう症、クローン病；多発性硬化症；オールブライト遺伝性骨形成異常症、狭心症、心筋梗塞、潰瘍、ぜん息、アレルギー、良性前立腺肥大、ならびに精神学的および神経学的障害（不安、精神分裂病、躁鬱、せん妄、痴呆、重篤な精神遅滞、歯状核赤核淡蒼球ルイ体萎縮症（D R P L A）、低リン酸血症性くる病、常染色体優性（2）先端脳梁（A c r o c a l l o s a l）症候群およびジスキネジー（例えば、ハンティングトン病またはジル・ド・ラ・ツレット症候群）を含む）ならびに／または他の病態および障害など。

【0015】

ポリペプチドは、本発明に特異的な抗体を産生するための免疫原として、およびワクチンとして使用され得る。これらはまた、潜在的なアゴニストおよびアンタゴニスト化合物についてスクリーニングするために使用され得る。例えば、G P C R X をコードする c D N A は、遺伝子治療において有用であり得、そして G P C R X は、それを必要とする被験体に投与された場合に有用であり得る。非限定的な例として、本発明の組成物は、以下に罹患している患者の処置に対する効力を有する：細菌、真菌、原生生物およびウイルス感染（特に、H I V - 1 または H I V - 2 によって引き起こされる感染）、疼痛、癌（新生物；腺癌；リンパ腫；前立腺癌；子宮癌を含むが、これらに限定されない）、食欲不振、過食症、ぜん息、パーキンソン病、急性心不全、低血圧、高血圧、残尿、骨粗しょう症、クローン病；多発性硬化症；オールブライト遺伝性骨形成異常症、狭心症、心筋梗塞、潰瘍、ぜん息、アレルギー、良性前立腺肥大、ならびに精神学的および神経学的障害（不安、精神分裂病、躁鬱、せん妄、痴呆、重篤な精神遅滞およびジスキネジー（例えば、ハンティングトン病またはジル・ド・ラ・ツレット症候群））、ならびに／または他の病態および障害。

【0016】

本発明はさらに、例えば、以下を含む障害または症候群のモジュレーターについてスクリーニングするための方法を包含する：糖尿病、肥満に関連する代謝障害、代謝症候群 X、食欲不振、るいそう障害（慢性疾患、代謝障害、糖尿病、肥満、感染性疾患、食欲不振、癌関連悪液質、癌、神経変性障害、アルツハイマー病、パーキンソン病、免疫障害および造血障害に関連する）、細胞シグナル伝達プロセッシングおよび代謝経路調節に関連する他の障害。この方法は、試験化合物に G P C R X ポリペプチドを接触させる工程、およびこの試験化合物がこの G P C R X ポリペプチドに結合するか否かを決定する工程を包含する。G P C R X ポリペプチドへの試験化合物の結合は、この試験化合物が、前述の障害または症候群に対する活性あるいは前述の障害または症候群に対する潜在性または素因のモジュレーターであることを示す。

【0017】

また、本発明の範囲には、以下を含む障害もしくは症候群の活性のモジュレーター、または以下を含む障害もしくは症候群の潜伏もしくは素因のモジュレーターについてスクリーニングする方法が含まれる：例えば、糖尿病、肥満に関連する代謝障害、代謝異常症候群 X（m e t a b o l i c s y n d r o m e X）、食欲不振、慢性病に関連する消耗障害、代謝障害、糖尿病、肥満、感染性疾患、食欲不振、癌に関連した悪液質、癌、神経変性障害、アルツハイマー病、パーキンソン障害、免疫障害、および造血障害、または、細胞のシグナルプロセッシングおよび代謝経路調節に関連した他の障害。このスクリーニング

方法は、上述の障害または症候群についての危険性の増大した試験動物に対して、試験化合物を投与することによる。試験動物は、G P C R X 核酸によってコードされる組換えポリペプチドを発現する。次いで、G P C R X ポリペプチドの発現または活性を、この試験動物において測定し、同様に、コントロール動物（これは、G P C R X ポリペプチドを組換え的に発現し、かつ上記の障害についての危険性も上記の症候群についても危険性も増大していない）におけるこのタンパク質の発現または活性を測定する。次いで、試験動物およびコントロール動物の両方におけるG P C R X ポリペプチドの発現を比較する。コントロール動物と比較した、試験動物におけるG P C R X ポリペプチドの活性の変化は、この試験化合物が、上記の障害または症候群の潜伏のモジュレーターであることを示す。

【0018】

10

なお別の局面では、本発明は、被験体（例えば、ヒト被験体）におけるG P C R X ポリペプチド、G P C R X 核酸またはその両方のレベルの改変に関連した疾患の存在または素因を決定する方法を含む。この方法は、この被験体由来の試験サンプル中におけるG P C R X ポリペプチドの量を測定する工程、およびコントロールサンプル中において存在するG P C R X ポリペプチドの量に対して、この試験サンプル中におけるこのポリペプチドの量を比較する工程、を包含する。コントロールサンプルと比較した場合の試験サンプルにおけるG P C R X ポリペプチドレベルの変化は、この被験体における疾患の存在または素因を示す。好ましくは、この素因としては、以下の素因が挙げられる：例えば、糖尿病、肥満に関連する代謝障害、代謝異常症候群X、食欲不振、慢性病に関連する消耗障害、代謝障害、糖尿病、肥満、感染性疾患、食欲不振、癌に関連した悪液質、癌、神経変性障害、アルツハイマー病、パーキンソン障害、免疫障害、および造血障害。また、本発明の新規なポリペプチドの発現レベルは、種々の癌についてスクリーニングするための方法および癌の病期を決定するための方法において使用され得る。

20

【0019】

さらなる局面では、本発明は、病理学的状態を緩和または予防するに十分な量において、被験体（例えば、ヒト被験体）にG P C R X ポリペプチド、G P C R X 核酸、またはG P C R X 特異的抗体を投与することによって、哺乳動物における障害に関連した病理学的状態を処置または予防する方法を包含する。好ましい実施形態では、この障害としては、以下が挙げられる：例えば、糖尿病、肥満に関連する代謝障害、代謝異常症候群X、食欲不振、慢性病に関連する消耗障害、代謝障害、糖尿病、肥満、感染性疾患、食欲不振、癌に関連した悪液質、癌、神経変性障害、アルツハイマー病、パーキンソン障害、免疫障害、および造血障害。

30

【0020】

なお別の局面では、本発明は、当該分野において一般的に使用される多数の技術のうちのいずれか1つによって、本発明の細胞性レセプターおよび下流エフェクターを同定する方法において使用され得る。これらの技術としては、ツーハイブリッド系、アフィニティ精製、抗体または他の特異的な相互作用分子を用いる共沈が挙げられるが、これらに限定されない。

【0021】

他に規定されない限り、本明細書中において使用されるすべての専門用語および科学技術用語は、本発明が属する分野の当業者によって通常理解されるのと同じ意味を有する。本明細書中に記載された方法および物質と類似または等価な方法および物質が、本発明の実施または試験において使用され得るが、適切な方法および物質を以下に記載する。本明細書中で引用されたすべての刊行物、特許出願、特許および他の参考文献が、それらの全体において、本明細書中において参考として援用される。矛盾する場合には、定義を含む本明細書中が制御する。さらに、物質、方法および例は、例示のみであり、限定されることを意図されない。

40

【0022】

本発明の他の特徴および利点が、以下の詳細な説明および特許請求の範囲から明らかである。

50

【0023】

(発明の詳細な説明)

本発明は、一部、新規なポリペプチドをコードする新規な核酸配列の発見に基づく。この新規な核酸およびそれによりコードされるポリペプチドを、個々に、GPCR1、GPCR2、GPCR3、GPCR4、GPCR5、GPCR6、GPCR7、GPCR8、GPCR9およびGPCR10という。これらの核酸およびこれらによってコードされるポリペプチドを、本明細書中では集合的に、「GPCRX」と称する。

【0024】

本発明の新規なGPCRX核酸は、表1A、1D、1G、2A、3A、4A、4C、4G、5A、5C、5G、6A、7A、8A、9A、10A、10C、および10Fまでのすべてを含んだ中(「表1A~10F」)から配列が提供される核酸、それらのフラグメント、誘導体、アナログまたはホモログを含む。本発明の新規なGPCRXタンパク質は、表1B、1E、1H、2B、3B、4B、4H、5B、5D、5H、6B、7B、8B、9B、10B、10D、および10Gまでのすべてを含んだ中(「表1B~10G」)から配列が提供されるタンパク質フラグメントを含む。個々のGPCRX核酸およびタンパク質を、以下に記載する。細胞のシグナル伝達または代謝経路調節に関連した障害の処置または予防において、これらの核酸およびペプチドを使用する方法は、本発明の範囲内である。

10

【0025】

Gタンパク質共役型レセプタータンパク質(GPCR)は、多数の種における、Gタンパク質共役型レセプターの大きなファミリーとして同定された。これらのレセプターは、多くの神経伝達物質レセプターおよびホルモンレセプターと、7回膜貫通ドメイン構造を共有し、そして種々のシグナルの認識およびGタンパク質媒介性変換の根底をなすようである。ヒトGPCRは、一般的に、イントロンを含まず、そして高い配列可変性を示す4つの異なる遺伝子サブファミリーに属する。これらの遺伝子は、優先的に、嗅覚上皮において発現される。例えば、Ben-Arieら、Hum. Mol. Genet. 1994 3:229-235; および Online Mendelian Inheritance in Man (OMIM) エントリー番号164342 (<http://www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?>)を参照のこと。

20

30

嗅覚レセプター(OR)遺伝子ファミリーは、最大のGPCR多重遺伝子ファミリーの1つを構成し、そしてヒトゲノムにおける多くの染色体部位の中に分散されている。Rouquierら、Hum. Mol. Genet. 7(9):1337-45(1998); Malnicら、Cell 96:713-23(1999)を参照のこと。嗅覚レセプターは、Gタンパク質共役型レセプターの中でも最大のファミリーを構成し、推定上1000までのメンバーを有する。Vanderhaeghenら、Genomics 39(3):239-46(1997); Xieら、Mamm. Genome 11(12):1070-78(2000); Issel-Tarverら、Proc. Natl. Acad. Sci. USA 93(20):10897-902(1996)を参照のこと。嗅覚レセプターによる臭気物質の認識は、臭気の識別における最初の段階である。Krautwurstら、Cell 95(7):917-26(1998); Buckら、Cell 65(1):175-87(1991)を参照のこと。多くのORは、いくつかの特徴的な配列モチーフを共有し、そして推定リガンド結合部位に対応する中心可変領域を有する。Issel-Tarverら、Proc. Natl. Acad. Sci. USA 93:10897-902(1996)を参照のこと。

40

【0026】

GPCRに関連した7回膜を貫通するタンパク質の他の例は、化学受容体である。Thomasら、Gene 178(1-2):1-5(1996)を参照のこと。化学受容体は、味覚組織、嗅覚組織、および雄性生殖組織において同定された。同上; Walenskyら、J. Biol. Chem. 273(16):9378-87(1998); Pa

50

rmentierら、Nature 355 (6359) : 453-55 (1992) ; Asaiら、Biochem. Biophys. Res. Commun. 221 (2) : 240-47 (1996) を参照のこと。

【0027】

(GPCR1)

GPCR1は、以下に開示される、3つの新規なGタンパク質共役型レセプター (GPCR) タンパク質のファミリーを含む。開示されたタンパク質には、GPCR1a、GPCR1bおよびGPCR1cと命名した。これらのタンパク質は、嗅覚レセプターに関連する。

【0028】

(GPCR1a)

1019ヌクレオチドの開示されたGPCR1a核酸を、表1Aに示す。開示されたGPCR1aのオープンリーディングフレーム (「ORF」) は、表1Aにおいてボールドで示された、ヌクレオチド27~29のATG開始コドンで始まる。コードされるポリペプチドは、本明細書中では、GPCR1aまたはb a 1 1 3 a 1 0 _ Bとして代替的に言及される。開示されたGPCR1aのORFは、ヌクレオチド984~986のTAGコドンで終結する。表1Aに示されるように、開始コドンに対して5'側および停止コドンに対して3'側の推定非翻訳領域に下線を付す。そして開始コドンおよび停止コドンはボールド字体である。

【0029】

【表1】

表1A. GPCR1a ヌクレオチド配列 (配列番号 1).

```

CCTTCAGTTGACAGAGGAGATACACTATGGTAAGTGCCAATCAGACAGCCTCTGTGACCGAGTTTATTC
TCCTGGGCCTCTCTGCCCACCCAAAGCTGGAGAAAACGTTCTTTGTGCTCATCTGCTGATGTACCTGG
TGATCCTACTGGGCAATGGGGTCCCTCATCCTGATGACTGTGTCCAACCTCCCACCTGCACATGCCCATGT
ACTTCTTCTCTGGGGAACCTCTCCTTCTCTGGACATCTGCTATACACATCCTCAGTCCCCCTCATCCTTG
ACAGCTTCTTGACCCCGAGGAAAACCATCTCCTTCTCAGCCTGTGCAGTGCAGATGTTCTCTCCTTTG
CCATGGGAGCCACAGAGTGTGTTCTCCTGAGCATGATGGCGTTTGATCGCTACGTGGCCATCTGCAACC
CCCTTAGGTACCTGTGGTCAATGAGCAAGGCTGCCTACATGCCCCATAAGGCTGCCGGCTCCTGGGTAG
CTGGAAGCACTGCTTCCATGGTGCAGACATCCCTTGCAATGAGGCTGCCCTTCTGTGGAGACAACATCA
TCAACCACTTACCTGTGAGATTCTGGCTGTCTGAAGTTGGCCTGTGCTGATATCTCTGTCAATGTGA
TCAGTATGGGAGTGACCAATGTGATCTTCTGGGGGTCCCGGTTCTGTTTCATCTCTTTCTCCTATGTCT
TCATCATTGCCACCATCCTGAGGATCCCTCAGCTGAGGGGAGGAAAAGGCCCTTCTCCACCTGCTCTG
CCCACCTCAGAGTCGTGGTCATCTTCTATGGGACCATCCTCTTCATGTATGGGAAGCCCAAGTCTAAGG
ACCCGCTGGGGGCAGACAAGCAAGACCTTGACAGAAAACCTATTTCCTTTTCTATGGGGTGGTGACCC
CCATGCTCAACCCCATCATCTACAGCCTGAGGAACAAGGATGTAAGGCTGCTGTGAGGGACTTGATAT
TTCAGAAATGCTTTGCCTAGTGATGTTTGGGGGAACAGATGTCCTCATAGCTC

```

開示された、コードされるGPCR1aタンパク質は、319アミノ酸残基を有し、GPCR1aタンパク質と呼ばれる。GPCR1aタンパク質を、シグナルペプチドの予測および細胞の局在化について分析した。SignalPの結果では、GPCR1aが、配列番号2の44位と55位との間 (すなわち、アミノ酸配列GNG-VLにおけるスラッシュ) において切断されると予測している。PsortおよびHydropathyプロファイルはまた、GPCR1が、シグナルペプチドを含み、かつ形質膜に局在化される可能性がある (信頼度0.6000) ことを予測している。開示されたGPCR1ポリペプチド配列は、一文字アミノ酸コードを使用して、表1Bに示される。

【0030】

【表2】

10

20

30

40

表1B. コードされたGPCR1aタンパク質自己配列(自己配列番号2)

```

MVSANQTASVTEFILLGLSAHPKLEKTFVLLILLMYLVILLGNG/VLIILMTVSNHLMHPMYFFLGNLS
FLDICYTTSSVPLILOSFLTFRKTISSACAVQMFLSFAMGATECVLLSMMAFDRIYVAICNPLRYFVVM
SKRAYMPHKAAGSWAGSTASMVQTSIAMRLPFCGDNIINHFTCEILAVLKLACADISVNVISMVGTNV
IFLGVEVLFISSYVFIIATILRIPSAEGRKKAFSTCSAHLTVVVIIFYGTILEMYGKPKSKDPLGADKQ
DLADKLISLFYGVVTFMLNPIIYSLRNKDVKAARLDLIFQKCF

```

GPCR1aを、GenBankによって利用可能にされたGenomic Daily Files 10
 に対して実行したGタンパク質結合レセプタープローブまたは相同体につい
 ての専有配列ファイルのTblastN分析によって、染色体9上に初めに同定した。専
 有ソフトウェアプログラム(GenScanTM)を使用して、核酸配列およびエキソンの
 選択をさらに予測した。得られる配列を、BLAST検索を用いて類似性によってさら
 に改変した。次いでこの配列を、見かけの不一致について手動で補正し、それによって全
 長タンパク質をコードする配列を得た。

【0031】

GPCR1a核酸配列の領域は、 3.6×10^{-160} のE値を有する、部分的Musculus 20
 嗅覚レセプターmRNA(1731bp)(GENBANK登録番号:Aj133427)をコードする配列に対して同一な1016塩基のうち873塩基(85%)
)を有する。本明細書中における全てのBLASTアラインメントにおいて、「E値」また
 は「期待値」の値は、確率の数値表示であり、整列された配列は、検索されたデータベ
 ース内で、単なる偶然にBLAST照会配列に対してそれらの類似性を達し得る。例えば
 、GPCR1a BLAST分析から検索された対象(「Subject」)(例えば、Mus
 musculus嗅覚レセプター)が、Query GPCR1a配列と偶然に純粋
 に一致する確率は、 3.6×10^{-160} である。

【0032】

BLASTX検索を、公共のタンパク質データベースに対して行った。本発明のタンパク 30
 質の全長アミノ配列は、Musculus由来の319アミノ酸嗅覚レセプター
 タンパク質(ptnr:TREMBLNEW-ACC:CAB55592、E=1.8e-142)に対して同一な316アミノ酸残基のうち274残基(86%)、およびこの
 319アミノ酸嗅覚レセプタータンパク質に対して陽性な316残基のうち295残基(93%)
)を有することを見出した。開示されたGPCR1aタンパク質(配列番号2)は
 、嗅覚レセプタータンパク質の数と良好な同一性を有する。例えば、GPCR1aは、3
 19アミノ酸Musculus嗅覚レセプター37aタンパク質に対して同一な
 256/316(81%)アミノ酸、およびMusculus由来の嗅覚レセプ
 ター37aに対して同一な376/316(87%)アミノ酸を有する(期待値=e-1
 28、gi|11276075|ref|NP_062346.1|)。開示されたタン
 パク質はまた、表1Cに開示した嗅覚タンパク質と類似している。

【0033】

【表3】

表1C. GPCR1aのBLAST結果					
遺伝子インデックス/ 識別名	タンパク質/生物体	長さ (アミノ酸)	同一性 (%)	類似性 (%)	期待値
gi 11464981 ref NP_062349.1	嗅覚 レセプター (OR) 37a <i>Mus musculus</i>	319	255/318 (80%)	275/318 (86%)	e-127
gi 11276077 ref NP_062347.1	OR 37b <i>Mus musculus</i>	318	248/317 (78%)	276/317 (86%)	e-127
gi 11276079 ref NP_062348.1	OR 37c <i>Mus musculus</i>	318	251/314 (79%)	274/314 (86%)	e-127
gi 10092669 ref NP_063950.1	ORファミリー2, サブファミリーS, メンバー2 <i>Homo sapiens</i>	309	238/306 (77%)	261/306 (84%)	e-118
gi 3769624 gb AAC64588.1 (AF091565)	OR <i>Rattus norvegicus</i>	227	206/227 (90%)	217/227 (94%)	e-102

10

(GPCR1b)

開示された1015ヌクレオチドのGPCR1b (ba32713_Aとも呼ばれる) 核酸を表1Dに示す。オープンリーディングフレームを、ヌクレオチド17~19のATG開始コドンで開始し、そしてヌクレオチド971~973のTAGコドンで終止して同定した。表1Dにおいて、開始コドンからの推定非翻訳領域の上流および停止コドンからの下流に下線を付し、そして開始コドンおよび停止コドンは太字である。

【0034】

20

【表4】

表1D. GPCR1b 核酸配列 (配列番号3)
ACAGAGGAGATACACTATGGTAAGTGCCAATCAGACAGCCCTCTGTGACCGAGTTTATTCTCCTGGGCCTC TCTGCCACCCAAAGCTGGAGAAACAGTTCTTTGTGCTCATCTGCTGATGATCTGCTGATCTACTGG GCAATGGGGTCTCTCATCTGATGACTGTGTCCAACTCCACCTGCACATGCCCATGTACTTCTTCTGGG GAACCTCTCTCTCTCTGACATCTGCTATACAACATCCTCAGTCCCCCTCATCCTTGACAGCTTCTTGACC CCCAGAAAACCATCTCTTCTCAGCCCTGTGCAGTGCAGATGTCTCTCTCTTGGCATGGGAGCCACAG AGTGTGTTCTCTGAGCATGATGGCGTTTGATCGCTACGTGGCCATCTGCAACCCCTTAGGTACCCCTGT GGTCATGAGCAAGGCTGCCATACATGCCATAGCTGCCGGCTCCTGGGTAGCTGGAAGCACTGTCTCCATG GTGCAGACATCCCTTGCAATGAGGCTGCCCTTCTGTGGAGACAACATCATCAACCACTTACCTGTGAGA TTCTGGCTGTCTCTGAAGTTGGCTGTGCTGATATCTCTGTCAATGTGATCAGTATGGGAGTGACCAATGT GATCTTCTCTGGGGTCCCGGTTCTGTTTCATCTCTTCTCTCTATGTCTTTCATCATTTGCCACATCCTGAGG ATCCCTCTCAGCTGAGGGGAGGAAAGGCTTCTCCACCTGCTCTGCCACCTCAGCTGCTGTGCTCATCT TCTATGGGACCATCTCTTCTATGTATGGGAGCCCAAGTCTAAGGACCCGCTGGGGGAGACAAGCAAGA CCTTGACAGCAAACTCATTTCCCTTTCTATGGGGTGGTGACCCCATGCTCAACCCCATCATCTACAGC CTGAGGAACAAGGATGTAAGGCTGCTGTGAGGGACTTGATATTTCAGAAATGCTTTGCCCTAGTGATGT TGGGGGAACAGATGTCCTCATAGCTCTTTCCTCT

30

配列データベースの検索において、例えば、核酸配列が、1731bp *Mus musculus* OR 37d 偽遺伝子 (GENBANK登録番号: MMU133427 | acc: AJ133427、E=5.5e-161) に対して同一な1015塩基のうち869塩基 (85%) を有することを見出した。核酸が、嗅覚レセプタータンパク質 (GENBANK登録番号: HSOLFMP | acc: Y14442、E=9.5e-49) について、部分的ヒトmRNAに対して同一な791塩基のうち505塩基 (63%) を有することも見出した。

40

【0035】

318アミノ酸残基を有するコードされたタンパク質を、表1Eにおいて一文字コードを用いて示す。本発明のタンパク質の全長アミノ酸配列を、*Mus musculus* 由来の319アミノ酸残基嗅覚レセプタータンパク質 (ptnr: TREMBLNEW-ACC: CAB55592、E=6.3e-144) に対して同一な315アミノ酸残基のうち274残基 (86%)、およびこの319アミノ酸嗅覚レセプタータンパク質に対して陽性な315残基のうち296残基 (93%) を有することを見出した。

【0036】

開示されたGPCR1bタンパク質は、2つの位置においてのみ、開示されたGPCR1

50

aタンパク質と異なる。145および146の位置において、GPCR1aはHKを有し、一方GPCR1bは欠失(Δ)およびIを有する。

【0037】

【表5】

表1E. コドされたGPCR1bタンパク質配列 (配列番号4)	
MVSAHQASVTEFILLGLSAHPKLEKTFEVLILLMYLVILLGNQ/VLILMTVSNSHLHMPMYFPLGNLSPL	
DICYTTSSVPLILDSFLTFRKTISSFSACAVQMFLSFANGATECVLLSMMAFDRYVAICNPLRYPVVMSKA	
AYNPIAAGSNVAGSTASMVQTSIAMRLFFCGDNIINHFTCHILAVLKLACADISVNVISMGVTVIPLGV	
PVLFISPSYVFIIATILRIPSAEQRKAFSTCSAHLTVVVIFYGTILFMYGKPKSKDPLGADKQDLADKL	
ISLFYGVVTPMLNFIITYSLRNKDVKAARDLIFQKCPA	

10

P S O R T分析は、ba32713_Aタンパク質(GPCR1b)が、0.6000の確実性を有するか、またはミトコンドリアの内膜、ミトコンドリアの内膜空間もしくはGolgi体においてより低い確実性を有する、原形質膜に局在下されると予測する。このタンパク質が、残基44と45(表1Eにおいて斜線で示したGNG-VL)との間の大部分の切断部位を有する単一ペプチドを有することもまた予測される。

【0038】

B L A S T X検索を、公共のタンパク質データベースに対して行った。本発明のタンパク質の全長アミノ配列は、Mus musculus由来の319アミノ酸嗅覚レセプタータンパク質(p t n r : T R E M B L N E W - A C C : C A B 5 5 5 9 2、E = 6 . 3 e - 1 4 4)に対して同一な315アミノ酸残基のうち274残基(86%)、およびこの319アミノ酸嗅覚レセプタータンパク質に対して陽性な315残基のうち296残基(93%)を有することを見出した。開示されたGPCR1bタンパク質(配列番号4)は、表1Fに示すように、嗅覚レセプタータンパク質の数と良好な同一性を有する。

20

【0039】

【表6】

表1F. GPCR1bのBLAST結果					
遺伝子/タンパク質/生物体	長さ(アミノ酸)	同一性(%)	類似性(%)	期待値	特徴値
gi 11276075 ref NP_062346.1	OR 37A Mus musculus	319	256/315 (81%)	277/315 (87%)	e-129
gi 11464981 ref NP_062349.1	(OR)37B Mus musculus	319	255/317 (80%)	276/317 (86%)	e-128
gi 11276077 ref NP_062347.1	OR 37b Mus musculus	318	248/316 (78%)	277/316 (87%)	e-128
gi 11276079 ref NP_062348.1	OR 37C Mus musculus	318	251/313 (80%)	275/313 (87%)	e-128
gi 10092669 ref NP_063950.1	ORファミリー2, サブファミリーS, メンバー2 Homo sapiens	309	238/305 (78%)	262/305 (85%)	e-119

30

(GPCR1c)

開示された1003ヌクレオチドのGPCR1c(ba113a10_Cとも呼ばれる)核酸を、表1Gに示す。オープンリーディングフレームを、ヌクレオチド26~28のATG開始コドンで開始し、そしてヌクレオチド974~976のTGAコドンで終止して同定した。開始コドンに対する5'および停止コドンに対する3'の推定非翻訳領域は、表1Gにおいて下線を付し、そして開始コドンおよび停止コドンは太字である。

40

【0040】

配列データベースの検索において、例えば、核酸配列が、Mus musculus GPCR mRNA(GENBANK登録番号:AJ133428、E=1.9e-120)に対して同一な899塩基のうち719塩基(79%)を有することを見出した。

【0041】

【表7】

表1G. GPCR1c核酸配列 (配列番号5)

```

TTTGTACAAAGTGACATAGAAACACCATGGTCAGTTCC
AATCAGACCTCCCTGTGCTGGGGTTCCTTCTCCTGGGGCTCTCTGCCCATCCAAAGCTGGAGAAGACAT
TCTTCGTGCTCATCTGCTGATGTACCTGGTGATCCTACTGGGCAATGGGGTCTCATCTGGTGACCAT
CCTTGACTCCCGCCTGGACACACCATGTACTTCTTCTGGGGAACTCTCCTTCTGGACATCTGCTAT
ACAACCTCCTCATCCTTGACAGCTTCCCTGACCCCGAGGAAAACCATCTCCTTCTCAGCCTGTGACGTAC
AGATGTTCTCTCCCTTGGCATGGGAGCCACAGAGTGTGTCTCCTGAGCATGATGGCGTTGATCGCTA
CGTGGCCATCTGCAACCCCTTTGGTACCTGAAGTCATGAACAAAGCTACTTATGTGCCCATGGCTGCT
GGCTCCTGGGTAGCTGGAAGCCTCACTGCCATGGTGACAGACCCCTTGCATTGAGGCTGCCCTTCTGTG
GAGACAACATCATCAATCACTTCACTGTGAGATTCTGGCTGTCTGAAAGTTGGCCTGTGCTGATATCTC
TGTCAATGTGATCAGTATGGGAGTGGCCAATGTGATCTTCTGGGGTCCCTGTCTGTCTCATCTCTTC
TCCTATGTCTTCATCATTTGCCACCATCTGAGGATCCCTCAGCTGAGGGGAGGAAAAGGCCTTCTCCA
CCTGCTCTGCCACCTCACTGTGCTGATCGTCTTCTACGGGACCATCCTTTCATGTACGGGAAGCCCAA
GTCTAAGGACCCACTGGGAGCAGACAAACAGACCTTGACAGCAAACTCATTCCCTTTTCTATGGGTG
GTGACCCCATGCTCAACCCCATCATCTACAGCCTGAGGAACAAGGAAGTGAAGGCTGCTGTGAGGAACC
TGGTATTTGAGAAACGCTTCTGCACTGATGGTGGAGGGTCTGTATGGCTCTGTG

```

10

316のアミノ酸残基を有する、開示されたGPCR1cタンパク質を、表1Hにおいて一文字コードを用いて示す。P S O R Tプログラムを使用する分析は、b a l l 3 a 1 0 _ Cタンパク質が、0.6400の確実性を有する原形質膜中に局在化すると予測する；それはまた、中程度の確実性を有するG o l g i体中にも局在化される。このタンパク質は、大部分の切断部位が、残基44と45（表1Hにおいて斜線で示したG N G - V L）との間にある単一ペプチドを有することもまた予測される。

【0042】

【表8】

20

表1H. 1ドメインGPCR1cタンパク質自己配列 (配列番号6)

```

MVSSNQTSPVLGFLLLGLSARPKLEKTFVLLILLMYLVILLGNG/VLILVTILDSRLDTPMYFFLG
LSFLDICYTTSSSLTASLTAPRTISFSACAVQMFLSLAMGATECVLLSMMAFDYVAICNPLWYREVMNK
ATYVPMAGSWVAGSLTAMVQTPALRLPFCGDNIIINHFTCEILAVLKACADISVNVISMGVANVIFLG
VPVLFISFSYVFIIATILRIPSAEGRKKAFTSCSAHLTVVIVFYGTILFMYGKPKSKDPLGADKQDLADK
LISLFYGVVTEMLNPPIIYSLRNKEVKAAVRNLVFQKRLQ

```

B L A S T X 検索を、公共のタンパク質データベースに対して行った。本発明のタンパク質の全長アミノ配列は、Mus musculus由来の319アミノ酸残基OR (p t n r : T R E M B L N E W - A C C : C A B 5 5 5 9 6、E = 5.2e-138) に対して同一な317アミノ酸残基のうち270残基(85%)、およびこの319アミノ酸残基ORに対して陽性な317残基のうち287残基(90%)を有することを見出した。開示されたGPCR1bタンパク質(配列番号6)は、表1Iに示すように、嗅覚レセプタータンパク質の数と良好な同一性を有する。

【0043】

【表9】

30

表1I. GPCR1cのBLAST結果

遺伝子インデックス/ 識別名	タンパク質/生物体	長さ (アミノ酸)	同一性 (%)	ポジティブ (%)	E値
gi 11464981 ref NP_062349.1	OR 37a マウス	319	242/317 (76%)	257/317 (80%)	e-120
gi 11276075 ref NP_062346.1	OR 37a マウス	319	241/319 (75%)	256/319 (80%)	e-119
gi 11276077 ref NP_062347.1	OR 37b マウス	318	236/319 (74%)	255/319 (80%)	e-116
gi 11276079 ref NP_062348.1	OR 37c マウス	318	235/308 (76%)	250/308 (81%)	e-116
gi 10092669 ref NP_063950.1	OR ファミリー2, サブファミリー3, メンバー2 ヒト	309	220/309 (71%)	242/309 (78%)	e-108

40

50

3つのGPCR1タンパク質間のアミノ酸相違を、表1Jに示す。欠失を、デルタ(Δ)によって標識する。この3つのタンパク質間の相違は、少しの異なる領域に局在化されるためと考えられる。従って、これらのタンパク質は、類似の機能性(例えば、嗅覚またはケモカインレセプター(以下を参照のこと)としての役目を果たす)を有し得る。

【0044】

【表10】

表1J. GPCR1タンパク質の相違													
位置	4	8	9	11	12	14	49	51	52	53	55	57	58
GPCR1a	A	A	S	T	E	I	M	V	S	N	H	H	M
GPCR1b	A	A	S	T	E	I	M	V	S	N	H	H	M
GPCR1c	S	S	P	L	G	L	V	I	L	D	R	D	T
位置	79	80	81	82	84	85	86	106	132	135	138	141	143
GPCR1a	V	P	L	I	D	S	F	F	R	V	S	A	M
GPCR1b	V	P	L	I	D	S	F	F	R	V	S	A	M
GPCR1c	S	Δ	Δ	Δ	T	A	S	L	W	E	N	T	V
位置	145	146	156	157	158	163	166	204	251	252	304	311	313
GPCR1a	H	K	T	A	S	S	M	T	V	I	D	D	I
GPCR1b	Δ	I	T	A	S	S	M	T	V	I	D	D	I
GPCR1c	A	M	L	T	A	P	L	A	I	V	E	N	V
位置	317	319	320										
GPCR1a	C	A	Δ										
GPCR1b	C	A	Δ										
GPCR1c	R	L	Q										

10

20

本発明の開示されたタンパク質と関連したORタンパク質配列とを比較するClustalW分析を、表1Kで与え、この表で、GPCR1aは1行目に示し、そしてGPCR1cは2行目に示す。

【0045】

GPCR1aタンパク質のClustalWアラインメント、および本明細書中のClustalW分析において、黒で輪郭をとられているアミノ酸残基は、保存配列領域(つまり、構造的特性または機能的特性を保存するように要求され得る領域)を示す。それに対して、強調されていないアミノ酸残基は、ほとんど保存されず、そしてタンパク質の構造または機能を変えることなしに、非常に広範な範囲まで潜在的に変異され得る。GPCR1a、GPCR1bまたはGPCR1cとして特にアドレスを取らない限り、GPCR1に対する任意の参考は、全ての変異体を含むと推定される。本明細書中における任意のGPCR X変異体配列間の残基相違は、「a」変異体における残基および「a」変異体についての残基位置を示すために描かれる。個々のGPCR変異体の間で異なる、全ての以下の配列アラインメントにおけるGPCR残基を、本明細書中の全てのアラインメントにおける変異体残基上で、囲みで強調しており、そして(o)記号で標識する。例えば、表1Kの1行目に示すタンパク質は、GPCR1aに対する配列を示し、そしてGPCR1bが異なる位置を、(o)記号で標識し、そして囲みで強調している。全てのGPCR1タンパク質は、嗅覚レセプター(OR)タンパク質(Mus musculus由来の37a、37b、37eおよび37c)に対して有意な相同性を有し、そしてファミリー2、サブファミリーS由来のヒトORタンパク質メンバー2(表1C、1Fおよび1Iもまた参照のこと)に対して有意な相同性を有する。

30

40

【0046】

【表11】

表 1K. GPCR1 の ClustalW 分析

- 1) Novel GPCR1a (配列番号5)
- 2) Novel GPCR1c (配列番号6)
- 3) gi|11276075|ref|NP_062346.1| olfactory receptor 37a *Mus musculus* (配列番号3)
- 4) gi|11276077|ref|NP_062347.1| olfactory receptor 37b *Mus musculus* (配列番号4)
- 5) gi|11464981|ref|NP_062349.1| olfactory receptor 37c *Mus musculus* (配列番号1)
- 6) gi|11276079|ref|NP_062348.1| olfactory receptor 37c *Mus musculus* (配列番号2)
- 7) gi|10092669|ref|NP_063950.1| OR, fam. 2, subfam. S, mem. 2 *H. sapiens* (配列番号4)

	10	20	30	40	50	60
GPCR1a	VS	QT	AS	TE	LL	LS
GPCR1c	VS	QT	SP	TE	LL	LS
gi 11276075	DR	QT	AP	SG	LL	LS
gi 11276077	RG	QT	AP	SG	LL	LS
gi 11464981	ER	KT	TP	SS	LL	LS
gi 11276079	GV	QT	TE	LL	LS	AS
gi 10092669	VS	QT	SP	TE	LL	LS
	70	80	90	100	110	120
GPCR1a	YF	FL	CH	LS	FL	DI
GPCR1c	YF	FL	CH	LS	FL	DI
gi 11276075	YF	FL	CH	LS	FL	DI
gi 11276077	YF	FL	CH	LS	FL	DI
gi 11464981	YF	FL	CH	LS	FL	DI
gi 11276079	YF	FL	CH	LS	FL	DI
gi 10092669	YF	FL	CH	LS	FL	DI
	130	140	150	160	170	180
GPCR1a	FR	YV	AI	CH	PL	RY
GPCR1c	FR	YV	AI	CH	PL	RY
gi 11276075	FR	YV	AI	CH	PL	RY
gi 11276077	FR	YV	AI	CH	PL	RY
gi 11464981	FR	YV	AI	CH	PL	RY
gi 11276079	FR	YV	AI	CH	PL	RY
gi 10092669	FR	YV	AI	CH	PL	RY
	190	200	210	220	230	240
GPCR1a	CE	LL	AV	KL	AC	AD
GPCR1c	CE	LL	AV	KL	AC	AD
gi 11276075	CE	LL	AV	KL	AC	AD
gi 11276077	CE	LL	AV	KL	AC	AD
gi 11464981	CE	LL	AV	KL	AC	AD
gi 11276079	CE	LL	AV	KL	AC	AD
gi 10092669	CE	LL	AV	KL	AC	AD
	250	260	270	280	290	300
GPCR1a	ST	CS	AH	LT	VS	VE
GPCR1c	ST	CS	AH	LT	VS	VE
gi 11276075	ST	CS	AH	LT	VS	VE
gi 11276077	ST	CS	AH	LT	VS	VE
gi 11464981	ST	CS	AH	LT	VS	VE
gi 11276079	ST	CS	AH	LT	VS	VE
gi 10092669	ST	CS	AH	LT	VS	VE
	310	320				
GPCR1a	RR	KK	VA	AV	RI	CF
GPCR1c	RR	KK	VA	AV	RI	CF
gi 11276075	RR	KK	VA	AV	RI	CF
gi 11276077	RR	KK	VA	AV	RI	CF
gi 11464981	RR	KK	VA	AV	RI	CF
gi 11276079	RR	KK	VA	AV	RI	CF
gi 10092669	RR	KK	VA	AV	RI	CF

10

20

30

この GPCR1 タンパク質はまた、表 1 L (配列番号 44) のライン 3 に示されるように、321 アミノ酸ヒト膜貫通レセプター (gi|6691937|emb|CAB65797.1|bA150A6.2、新規な 7 膜貫通レセプター (ロドプシンファミリー)、嗅覚レセプター様タンパク質 (hs6M1-21)) と、同一性の領域を有する。

【0047】

【表 12】

40

表 1L. GPCR1a の保守性と配列番号442の整列

	10	20	30	40	50	60
GPCR1a						
GPCR1c						
gi 6691937						
	70	80	90	100	110	120
GPCR1a						
GPCR1c						
gi 6691937						
	130	140	150	160	170	180
GPCR1a						
GPCR1c						
gi 6691937						
	190	200	210	220	230	240
GPCR1a						
GPCR1c						
gi 6691937						
	250	260	270	280	290	300
GPCR1a						
GPCR1c						
gi 6691937						
	310	320				
GPCR1a						
GPCR1c						
gi 6691937						

10

20

GPCR1、ならびに全ての他のGPCRXタンパク質における同定可能なドメインの存在を、ソフトウェアアルゴリズム（例えば、PROSITE、DOMAIN、Blocks、Pfam、ProDomain、およびPrints）を使用する検索によって決定し、次いでInterpro番号をInterproウェブサイト（<http://www.ebi.ac.uk/interpro>）を使用するドメイン適合（または数）に横線を引くことによって決定した。例えば、表1Mに開示されるようなGPCR1aについてのDOMAINの結果は、Reverse Position Specific BLAST分析を用いて、Conserved Domain Database (CDD) から収集した。このBLAST分析ソフトウェアサンプルは、SmartおよびPfam収集において見出されたドメインをサンプリングする。表1Mおよび全ての連続したDOMAIN配列整列において、完全に保存された単一の残基が、黒い影によって示され、そして「強い」半保存残基が灰色の影によって示される。この保存されたアミノ酸残基の「強い」群は、アミノ酸の以下の群のいずれか1つであり得る：STA、NEQK、NHQK、NDEQ、QHRK、MILV、MILF、HY、FYW。

30

40

【0048】

表1Mは、GPCR1aに対するDOMAIN分析結果からのドメイン記述を列挙する。アミノ酸残基53～239（配列番号2）の領域は、最もおそらく（ $E = 2 \times 10^{-19}$ ）、「7膜貫通レセプター（ロドプシンファミリー）フラグメント」ドメインを含み、これは、本明細書中でPfamデータベースの7tm1entry（TM7、配列番号45、完全な配列については、表1Nを参照のこと）の残基12～180と整列されている。このことは、GPCR1配列が、このドメインを含むことが知られている他のタンパク質の特性、ならびに377個のアミノ酸の7tmドメイン自体の特性に類似する特性を有することを示す。GPCR1aはまた、TM7タンパク質の別の領域に対する同一性を有する。アミノ酸残基226～298（配列番号2の）の領域が、TM7のアミノ酸残基310～377と整列する（ $E = 3 \times 10^{-4}$ ）。GPCR1bおよびGPCR1cはまた、この

50

ドメインと整列し；GPCR1bの残基53～238は、TM7の残基12～180と整列し（ $E = 6e^{-19}$ ）、そしてGPCR1bの残基225～297は、TM7の残基310～377と整列し（ $E = 3e^{-4}$ ）；GPCR1cの残基55～235は、TM7の残基14～180と整列し（ $E = 2e^{-12}$ ）、そしてGPCR1cの残基222～294は、TM7の残基310～377と整列する（ $E = 2e^{-4}$ ）。

【0049】

【表13】

表 1M. GPCR1a ドメイン分析	
Sbjct: 7 transmembrane receptor (rhodopsin family) fragment	
Gn1 Pfam pfam00001; Length = 377	
Score = 89.7 bits (221), Expect = 2e-19	
GPCR1a	NSHLMPMYHFGNLSFLCYTSSYLKDSFTPRKTIFFSAARQ
TM7	GNVLVCMVAVSREKALQTTNLLVSLAVADLLVATLVMPWVYLEVVGWKFSTRHCDIF
GPCR1a	VFSSFANGATECULSMZAFDRYVACNIRPVVMKAAYPHKAGSNVAGSTASMVQ
TM7	MTLDVMMCTASILNLCAISIDRYTAVAMPMLYNTRYSSKRRRTVMIAIVVWVLSFTISCPM
GPCR1a	TSLAMRLPFCGDNIINHFTSLAVLKLACADISVNIEMGMINVFLGVIVVFISFSN
TM7	LFGLNNTDQNE-----IIANP-----FVYYSIVSFY-----FVITLLVY
GPCR1a	FIIATL-----IPSAEGEKKA-----
TM7	KVIYVRRRRKRVNTKSSRAFRANLKAPLKGNCTHPEDMKLCTVIMKSNGSFPVNRVRV

10

20

この7膜貫通レセプタードメインファミリーの代表的なメンバーは、Bos taurus (SWISSPROT: locus D2DR_BOVIN, accession P20288; gene index 118205) 由来のD2ドーパミンである。このD2レセプターは、一体化膜タンパク質であり、そしてGタンパク質結合レセプターのFamily 1に属する。このD2レセプターの活性は、アデニルイルシクラーゼを阻害するGタンパク質によって媒介される。Chioら、Nature 343: 255-269 (1990)。この444個のアミノ酸タンパク質の残基51～427は、表1

30

【0050】

【表14】

表 1N. TM7に対する アミノ酸配列 (配列番号: 45)
GNVLVCMVAVSREKALQTTNLLVSLAVADLLVATLVMPWVYLEVVGWKFSTRHCDIF
VTLDMVMCTASILNLCAISIDRYTAVAMPMLYNTRYSSKRRRTVMIAIVVWVLSFTISCPM
LFGLNNTDQNECIIANPAFVYSSIVSFYVPFIVTLLVYIKIYIVLRRRRKRVNTKRSSR
AFRANLKAPLKGNCTHPEDMKLCTVIMKSNGSFPVNRVRVEAARRAQELEMELSSSTSPF
ERTRYSPIPPSSHQLTLPLDPSSHGLHSTPDSPAKPEKNGHAKTVNPKIAKIFEIQSMPNG
KRTSLKTMSSRRKLSQQKEKKATQMLAIVLGVFIICWLPFFITHIILNIHCDNCNIPFVLYS
AFTWLGYNVSAVNPIIY

40

この7膜貫通レセプターファミリーは、多くの異なるタンパク質（例えば、セロトニンレセプター、ドーパミンレセプター、ヒスタミンレセプター、アドレナリン作動性レセプター、カンナビノイドレセプター、アンギオテンシンIIレセプター、ケモカインレセプター、オピオイドレセプター、Gタンパク質結合レセプター（GPCR）タンパク質、嗅覚レセプター（OR）などを含む）を含む。いくつかのタンパク質およびProtein Data Base Ids/gene indexesは、例えば、以下を含む：ロドプシン (129209)；5-ヒドロキシトリプタミンレセプター；(112821、8

50

4 8 8 9 6 0、1 1 2 8 0 5、2 3 1 4 5 4、1 1 6 8 2 2 1、3 9 8 9 7 1、1 1 2 8 0 6) ; Gタンパク質結合レセプター (1 1 9 1 3 0、5 4 3 8 2 3、1 7 3 0 1 4 3、1 3 2 2 0 6、1 3 7 1 5 9、6 1 3 6 1 5 3、4 1 6 9 2 6、1 1 6 9 8 8 1、1 3 6 8 8 2、1 3 4 0 7 9) ; 味覚レセプター (5 4 4 4 6 3、4 6 2 2 0 8) ; c-x-c ケモカインレセプター (4 1 6 7 1 8、1 2 8 9 9 9、4 1 6 8 0 2、5 4 8 7 0 3、1 3 5 2 3 3 5) ; オプシン (1 2 9 1 9 3、1 2 9 1 9 7、1 2 9 2 0 3) ; および聴覚レセプター様タンパク質 (1 2 9 0 9 1、1 1 7 1 8 9 3、4 0 0 6 7 2、5 4 8 4 1 7) 。

【0051】

G P C R X R N A についての発現情報は、組織供給源 (実施例に記載されるような、商品名データベース供給源、公的な E S T 供給源、文献供給源、および/または R A C E 供給源があげられるがこれらに限定されない) を使用して、誘導された。 10

【0052】

G P C R 1 の核酸およびタンパク質は、種々の G P C R または嗅覚レセプター (O R) 関連の症状および/または障害に関係した潜在的な治療適用において有用である。例えば、Gタンパク質結合レセプター様タンパク質をコードする c D N A は、遺伝子治療において有用であり得、そして Gタンパク質結合レセプター様タンパク質は、その必要性において被験体に投与される場合、有用であり得る。G P C R 1 タンパク質、またそのフラグメントをコードする新規な核酸は、診断適用においてさらに有用であり得、ここで、核酸またはタンパク質の存在または量が、評価される。これらの物質は、治療方法または診断方法における使用のための、本発明の新規な物質に免疫特異的に結合する抗体の産生においてさらに有用である。この G P C R 1 核酸およびタンパク質は、以下に記載される種々の疾患および障害および/または他の症状に関係する潜在的な診断適用および治療適用において有用である。例えば、本発明の組成物は、以下を患う患者の処置のための効力を有する：心筋症、アテローム性動脈硬化症、高血圧、先天性心疾患、大動脈狭窄症、心房中隔欠損症 (A S D)、房室 (A-V) 管欠陥、動脈管、肺動脈弁狭窄症、大動脈弁下部狭窄症、心室中隔欠損症 (V S D)、弁疾患、結節硬化症、強皮症、肥満症、移植、副腎脳白質ジストロフィー、先天性副腎過形成、受胎能、血友病、凝固亢進、特発性血小板減少性紫斑病、免疫不全、対宿主性移植片病、気管支ぜん息、ならびに類似の他の疾患、障害、および状態。非限定的な例によって、本発明の組成物は、以下を患う患者の処置のための効力を有する：新生物、腺癌、リンパ腫、前立腺癌、子宮癌、免疫応答、A I D S、喘息、クローン病、多発性硬化症、およびオールブライト遺伝性骨形成異常症。さらなる G P C R 関連疾患および障害は、本明細書を通して言及される。 20 30

【0053】

さらに、G P C R 1 に対するタンパク質類似性情報、発現パターン、およびマップ配置は、G P C R 1 が G P C R ファミリーの重要な構造的および/または生理学的機能の特徴を有し得ることを示唆する。従って、本発明の核酸およびタンパク質は、潜在的な診断適用および治療適用において有用であり、そして検索ツールとして有用である。これらは、特定のまたは選択的核酸またはタンパク質診断および/または予防マーカーとして機能すること (ここで、核酸またはタンパク質の存在または量が評価される)、ならびに以下のような潜在的な治療適用を含む：(i) タンパク質治療剤、(i i) 低分子薬物標的、(i i i) 抗体標的 (治療的、診断的、薬物標的/細胞傷害性抗体)、(i v) 遺伝子治療に有用な核酸 (遺伝子送達/遺伝子切除)、ならびに (v) インビボおよびインビトロで組織再生を促進する組成物、(v i) 生物学的防御武器。 40

【0054】

これらの物質は、さらに治療方法または診断方法における使用のための新規な G P C R 1 物質の免疫特異的に結合する抗体の生成において、さらに有用である。これらの抗体は、以下の「抗 G P C R 抗体」の節に記載されるような、疎水性チャートからの予測を使用して、当該分野で公知の方法に従って、産生され得る。

【0055】

(GPCR2)

本発明のさらなるGPCR様タンパク質（本明細書中でGPCR2と称される）は、嗅覚レセプター（「OR」）様タンパク質である。1254ヌクレオチドの新規な核酸（11612531—1、配列番号7）（これは、新規なGタンパク質結合レセプター様タンパク質をコードする）が、図2Aに示される。GPCR2についてのSeq Callingフラグメントは、胎盤由来であり、このことは、それが雌の健康について重要な組織において発現され得ることを示す。

【0056】

【表15】

10

表 2A. GPCR2 ヌクレオチド配列 (配列番号7)
ACGCGCTTCGACATCTATTCCTGGTGGGCTCTTCTCCATATCAGGTCGTTAAATAAGCTGCCAGATTCTGCCTTTACAGCCCAA GGAGCTTGTTCATGGACCATGGGCAATGGAGGGTCTTCTCCAGAACTCCACTAACTTCGTCTCCAGGCTTCATCACCCTATCCTGCCT TCCCCGGGCTTCTCTTTCGAATAGTCTTCTCCATCTTTGTGGTGGCTATAACAGCCAACTTGGTCATGATTCTGCTCATCCACATGG ACTCCCGGCTCCACAGCCCATGTACTTCTGCTCAGCCAGCTCTCCATCATGGATACCATCTACATCTGTATCACTGTCCCAAGA TGCTCCAGGACCTCTGTGTCGAAGGACAAGACCAATTCCTTCATGGGCTGTGCAGTTCAGATCTTCCTCTACCTGACCCCTGATTGGAG GGGAATTCCTCTGCTGGGCTCTCATGGGCTATGACCGCTATGTGGCTGTGTGCAACCCCTCTACGGTACCCTCTCCTCATGAACCGCA GGGTTTGCCTATTCTATGGTGGTGGGCTCTGGGTTGGTGGTCTCTTGGATGGGTTCTGCTGACTCCTGTCACTATGAGTTTCCCT TCTGTAGATCCCGAGAGATCAATCACTTTTCTGTGAGATCCAGCCGCTGTGAAGTTGTCTTGCACAGACAGTCACTCTATGAGA CCCTGATGTATCCCTGCTGCTGCTGATTATCCCTCTATCTGTCTCTGTCTTACACGCACATCCTCCTGACTGTCCACAGGA TGAATCTGCTGAGGGCCGGCCAAAGCCTTTGCTACGTTCTCTCCACATTTATGGTGGTGGGCTTTCTACGGGGCAGCTTTCT ACACCAACGTGCAGCCCACTCTACCACACTCCAGAGAAAGATAAAGTGGTGTCTGCTTCTACACCATCCTACCCCATGCTCA ACCCACTCATCTACAGCTTGAGGAATAAGATGTGGCTGCAGCTCTGAGGAAAGTACTAGGGAGATGTGGCTCCTCCAGAGCATCA GGGTGATGACTGTGTGATCAGGAAGGACTAGCAGGGACTCCAGAGTATCAGCGTGGTGAATGATCAGGAAGGACTAGTGGGGAC TCCTAGAGCATCAGGCTGGGACTGTGATCAGGAAGGACTAGCAAGGACTAGCGCAACATCTGCGGTGCTGCGGCCAATAACGCAG CTATTACAGAAATATGTTATTGGTCTGGAAGT

20

GPCR2についてのオープンリーディングフレーム（ORF）が、ヌクレオチド105～1058に同定された。配列番号7によってコードされる開示されたGPCR2ポリペプチド（配列番号8）は、318アミノ酸残基であり、そして表2Bに一文字表記を使用して提供される。GPCR2タンパク質は、シグナルペプチド予測および細胞局在化について分析された。Signal Pepの結果は、GPCR2が配列番号8の位置42と位置43との間で（すなわち、アミノ酸配列ITANLにおけるスラッシュにおいて）切断されることを予測する。PsorおよびHydropathyプロファイルはまた、GPCR2が、シグナルペプチドを含み、そして原形質膜に局在化するようであることを予測する（0.6400の確実度）。

30

【0057】

【表16】

表 2B. コードされるGPCR2 タンパク質配列 (配列番号8)
MGMEGLQNSTNFVLTGLITHPAFFGLLEFAIVSIFVVAITA/NLVMILLIHMSRLHTPMYFLLSLS IMDTIYICVTPKMLQDLSKDKTISFMGCAVQIFLYLTLLIGFEFFLLGLMAYDRYVAVCNPLRYPLEM NRRVCLFMVGVSWVGGSLDGFMLTPTMTSFFFCRSREINHFFCEIPAVLKLSCDTSLYETILMYACCVL IIPLSVISVSYTHILLTVHRMNSAEGRRKAFATCSSHIMVSVVFYGAIFYTNVQPHSYHTPEKDKVSA FYTILTEMLNPLIYSLRNKDVAALRKVLGRCGSSQSIRVMTV

40

本発明のタンパク質の全アミノ酸配列は、Homo sapiens由来の216アミノ酸残基の嗅覚レセプター（ptnr: SPTREMBL-ACC: O43869）（E=2.2e⁻⁶¹）に、216アミノ酸残基のうち151残基が同一（69%）であり、216残基のうち177残基が陽性（81%）であることが見出された。GPCR2（配列番号7）によってコードされるこのタンパク質は、嗅覚、臭気物質、および味覚の化学レセプターに有意な相同性を有し、そしてGタンパク質結合レセプター（GPCR）のファミリーに属する。この遺伝子ファミリーは、小分子薬物の標的として使用され、そしてGPCRは、原形質膜に発現され、そしてまた、タンパク質薬様の治療抗体、細胞傷害性抗体、および診断抗体に対する適切な標的である。

50

【0058】

表2Cに示されるように、BLAST分析によって、GPCR2が多数の嗅覚レセプターと有意な相同性を有することが示される。

【0059】

【表17】

表 2C. GPCR2 BLASTの結果					
高スコアのセグメント対を生成する配列:		リーディング フレーム	高スコア	最小の 可能性 P(N)	
Ptnr:SPTREMBL-ACC:Q43869	OLFACTORY RECEPTOR - HOMO SAPIENS...	927	1.8e-82	1	
Ptnr:SWISSPROT-ACC:P23275	OLFACTORY RECEPTOR 15 (OR3) - M...	712	2.7e-70	1	
Ptnr:SPTREMBL-ACC:Q90808	OLFACTORY RECEPTOR 4 - GALLUS GALLUS...	708	7.3e-70	1	
Ptnr:TREMBLNEW-ACC:CA855593	OLFACTORY RECEPTOR - Mus musculus...	706	1.2e-69	1	
Ptnr:TREMBLNEW-ACC:CA855594	OLFACTORY RECEPTOR - Mus musculus...	704	1.9e-69	1	
Ptnr:SPTREMBL-ACC:Q63394	OL1 RECEPTOR - RATTUS NORVEGICUS...	702	3.1e-69	1	
Ptnr:TREMBLNEW-ACC:CA855592	OLFACTORY RECEPTOR - Mus musculus...	698	8.3e-69	1	
Ptnr:TREMBLNEW-ACC:CA855596	OLFACTORY RECEPTOR - Mus musculus...	697	1.1e-68	1	
Ptnr:SWISSPROT-ACC:Q95156	OLFACTORY RECEPTOR-LIKE PROTEIN...	695	1.7e-68	1	
Ptnr:SWISSPROT-ACC:Q13606	OLFACTORY RECEPTOR-LIKE PROTEIN...	694	2.2e-68	1	
Ptnr:SWISSPROT-ACC:Q13607	OLFACTORY RECEPTOR-LIKE PROTEIN...	689	7.5e-68	1	
Ptnr:TREMBLNEW-ACC:AAC64376	OLFACTORY RECEPTOR-LIKE PROTEIN...	685	2.0e-67	1	

10

ClustalW分析に使用した配列を含む他のBLAST結果を、表2Dに示す。

【0060】

【表18】

20

Table 2D. GPCR2に関するBLAST結果					
遺伝子インデックス/ 識別名	タンパク質/ 生物	長さ (aa)	同一性 (%)	陽性 (%)	期待値
gi 3983382 gb AAD13319.1 (AF102527)	OR E3 Mus musculus	223	156/223 (69%)	184/223 (81%)	2e-78
gi 2921628 gb AAC39611.1 (U86215)	OR Homo sapiens	216	151/216 (69%)	177/216 (81%)	5e-75
gi 12007423 gb AAG45196.1 (AF321234)	T2 OR Mus musculus	316	155/298 (52%)	205/298 (68%)	2e-72
gi 12007424 gb AAG45197.1 (AF321234)	T3 OR Mus musculus	315	156/304 (51%)	209/304 (68%)	2e-72
gi 12007425 gb AAG45198.1 (AF321234)	T4 OR Mus musculus	319	152/304 (50%)	207/304 (68%)	9e-72
gi 12007422 gb AAG45195.1 (AF321234)	T1 OR Mus musculus	316	156/305 (51%)	207/305 (67%)	2e-67

30

この情報は、GPCR2を関連タンパク質配列と比較するClustalW分析のように、表2Eに提供する複数配列整列にグラフで表す（GPCR2を1行目に示す）。

【0061】

(表2E. ClustalWタンパク質に関する情報)

- 1) 新規GPCR2 (配列番号8)
- 2) gi|3983382|gb|AAD13319.1| olfactory receptor E3 Mus musculus (配列番号46)
- 3) gi|2921628|gb|AAC39611.1| olfactory receptor Homo sapiens (配列番号47)
- 4) gi|12007423|gb|AAG45196.1| T2 olfactory receptor Mus musculus (配列番号48)
- 5) gi|12007424|gb|AAG45197.1| T3 olfactory receptor Mus musculus (配列番号49)
- 6) gi|12007425|gb|AAG45198.1| T4 olfactory receptor Mus musculus (配列番号50)
- 7) gi|12007422|gb|AAG45195.1| T1 olfactory

50

receptor Mus musculus (配列番号51)
 【0062】
 【表19】

	10	20	30	40	50	60
GPCR2						
gi 3983382	MGMEGLLNSTN	VTFTLTHPAF	GGTFAIVF	SAVVAIT	ALLNTHLHMTS	SHHTPN
gi 2921628	-----	-----	-----	-----	-----	-----
gi 12007423	-MEFWNSTLGT	DNVETDDSGS	ELLCAITFT	AYMARI	SGGSI	VITMELRLHVR
gi 12007424	-MEVCNSTLRS	GLMGLDDND	PELLCATIT	AGGPAHS	SGGSI	VITMELRLHVR
gi 12007425	-MEFRNSTMG	NGGIIVGGLDD	SGARDLLCATIT	AYMARI	SGGSI	VITMELRLHVR
gi 12007422	-MELWNSTLES	GGIIVGGLNG	SSSELLCAIT	VTAYMARI	SGGSI	VITMELRLHVR
	70	80	90	100	110	120
GPCR2						
gi 3983382	YFLSGLS	MDTICHTV	PMQOLSK	KKISEM	GGVAF	HEFLYTLIGCEFFLGLNA
gi 2921628	-----	-----	-----	-----	-----	-----
gi 12007423	YFLSGLS	MDTICHTV	PMQOLSK	KKISEM	GGVAF	HEFLYTLIGCEFFLGLNA
gi 12007424	YFLSGLS	MDTICHTV	PMQOLSK	KKISEM	GGVAF	HEFLYTLIGCEFFLGLNA
gi 12007425	YFLSGLS	MDTICHTV	PMQOLSK	KKISEM	GGVAF	HEFLYTLIGCEFFLGLNA
gi 12007422	YFLSGLS	MDTICHTV	PMQOLSK	KKISEM	GGVAF	HEFLYTLIGCEFFLGLNA
	130	140	150	160	170	180
GPCR2						
gi 3983382	YDRYVAIC	PLNRRP	LLFVAVGS	AVGGG	DGFMIL	PVITSSFFGSSG
gi 2921628	-----	-----	-----	-----	-----	-----
gi 12007423	YDRYVAIC	PLNRRP	LLFVAVGS	AVGGG	DGFMIL	PVITSSFFGSSG
gi 12007424	YDRYVAIC	PLNRRP	LLFVAVGS	AVGGG	DGFMIL	PVITSSFFGSSG
gi 12007425	YDRYVAIC	PLNRRP	LLFVAVGS	AVGGG	DGFMIL	PVITSSFFGSSG
gi 12007422	YDRYVAIC	PLNRRP	LLFVAVGS	AVGGG	DGFMIL	PVITSSFFGSSG
	190	200	210	220	230	240
GPCR2						
gi 3983382	RIAPLKL	SDTSL	ETPA	ACCT	-----	-----
gi 2921628	-----	-----	-----	-----	-----	-----
gi 12007423	RIAPLKL	SDTSL	ETPA	ACCT	-----	-----
gi 12007424	RIAPLKL	SDTSL	ETPA	ACCT	-----	-----
gi 12007425	RIAPLKL	SDTSL	ETPA	ACCT	-----	-----
gi 12007422	RIAPLKL	SDTSL	ETPA	ACCT	-----	-----
	250	260	270	280	290	300
GPCR2						
gi 3983382	TCSSSL	TVVSG	YGAAT	ETN	-----	-----
gi 2921628	-----	-----	-----	-----	-----	-----
gi 12007423	TCSSSL	TVVSG	YGAAT	ETN	-----	-----
gi 12007424	TCSSSL	TVVSG	YGAAT	ETN	-----	-----
gi 12007425	TCSSSL	TVVSG	YGAAT	ETN	-----	-----
gi 12007422	TCSSSL	TVVSG	YGAAT	ETN	-----	-----
	310	320				
GPCR2						
gi 3983382	NRRL	SGSSQS	IRVMTV	-----	-----	-----
gi 2921628	-----	-----	-----	-----	-----	-----
gi 12007423	NRRL	SGSSQS	IRVMTV	-----	-----	-----
gi 12007424	NRRL	SGSSQS	IRVMTV	-----	-----	-----
gi 12007425	NRRL	SGSSQS	IRVMTV	-----	-----	-----
gi 12007422	NRRL	SGSSQS	IRVMTV	-----	-----	-----

10

20

30

GPCR2に関するDOMAINの結果は、Reverse Position Specific BLASTを用いてConserved Domain Database (CDD) から収集した。このBLASTサンプルドメインは、Smart and Pfamコレクションから見出された。7tm_1 (7回膜貫通レセプター (ロドプシンファミリー)) は、GPCR2に対する相同性を有する2つのセグメント有することが示された。アミノ酸残基43~238の領域は、「7回膜貫通レセプター (ロドプシンファミリー) フラグメント」ドメイン (配列番号45、E=2e-22) のアミノ酸2~181と整列し、そしてGPCR2のアミノ酸226~289は、Pfamデータベースの7tm_1 エントリー (配列番号45、E=0.08) の残基313~317と整列した。これは、GPCR2配列が、このドメインならびに7tm_1ドメイン自体を含むことが公知の他のタンパク質に類似した特性を有することを示す。

40

【0063】

開示されたGPCR2は、雌性の生殖健康に重要な組織において発現され、故に、GPC

50

R 2 は、例えば、早期分娩、子宮内膜症、およびインビトロ受精の薬物標的としてはたらしめ得る。嗅覚レセプターに対する相同性は、内因性小分子リガンドがこの遺伝子を調節し、故に内因性リガンドに構造的に類似する薬物がアゴニストおよびアンタゴニストとしてはたらい、G P C R 2 の生物学的な効果を調節し得ることを示唆する。

【0064】

本発明の核酸およびタンパク質は、以下にさらに記載されるように、G P C R 関連病理学的障害および／またはO R 関連病理学的障害に関与する潜在的な治療適用に有用である。例えば、嗅覚関連様タンパク質をコードするc D N A は、遺伝子治療に有用であり得、そして嗅覚レセプター様タンパク質は、それを必要とする被験体に投与される場合、有用であり得る。非限定の例として、本発明の組成物は、新生物、腺癌、リンパ腫、子宮癌、免疫応答、A I D S、喘息、クローン病、多発性硬化症、およびA l b r i g h t 遺伝性骨ジストロフィーを患う被験体の処置に有効性を有する。他のG P C R 関連疾患および障害が、意図される。

10

【0065】

本発明のG P C R 様タンパク質をコードする新規核酸、またはそのフラグメントは、診断適用にさらに有用であり得、ここで、この核酸またはタンパク質の存在または量が、評価される。これらの材料は、治療方法または診断方法における使用のための本発明の新規物質に免疫特異的に結合する抗体の産生においてさらに有用である。これらの抗体は、以下の「抗G P C R X 抗体」の節に記載されるように、疎水性チャートからの推測によって、当該分野で公知の方法に従って作製され得る。1つの実施形態において、意図されるG P C R 3 エピトープは、約アミノ酸105～140である。これらの新規タンパク質は、種々のヒトの障害の機能分析のためのアッセイ系に使用され得、これは、疾患の病理学の理解および種々の障害のための新規薬物標的の開発を補助する。

20

【0066】

(G P C R 3)

本発明のさらなるG P C R 様タンパク質（本明細書においてG P C R 3 と称される）は、嗅覚レセプター（「O R」）様タンパク質である。この新規核酸は、G P C R プローブもしくはホモログに関するC u r a G e n 社配列ファイルを用いたT b l a s t N によって、G e n B a n k によって利用可能なG e n o m i c D a i l y F i l e に対して実行され、第6染色体上に同定された。この核酸はさらに、プログラムG e n S c a n^{T M}によって推測された（エキソンの選択を含む）。これらはさらに、B L A S T 検索を用いて類似性によって改変された。次いでこの配列は、明らかな不一致が手動で訂正され、それによって、全長タンパク質をコードする配列を得た。新規嗅覚レセプター様タンパク質をコードする957ヌクレオチドの新規核酸（b a l 4 5 1 2 2 _ B、配列番号9）を、表3 A に示す。ヌクレオチド10～12のA T G 開始コドンで開始し、ヌクレオチド955～957のT G A コドンで終わるオープンリーディングフレーム（O R F）を同定した。開始コドンの上流および終止コドンの下流である推定非翻訳領域は、表3 A の下線であり、そして開始コドンおよび終止コドンは太字である。

30

【0067】

【表20】

40

表 3A. GPCR3 ヌクレオチド配列 (配列番号 9)

```

AGATTAGTCATGAAGGCCAACTACAGCGCAGAGGAGCGCTTCTCCTGCTGGGTTTCTCCGACTGGCCCTT
CCCTGCAGCCGGTCTCTTGCCTTGTCTCCTGTGCTACCTCCTGACCTTGACGGGCAACTCGGCGCT
GGTGCTGCTGGCGGTGCGCGACCCGCGCTGCACACGCCCATGTACTACTTCTCTGCCACCTGGCCTTG
GTAGACGCGGGCTTCACTACTAGCGTGGTGGCCGCGCTGCTGGCCAACTGCGCGGACCAGCGCTTGGC
TGCCGCGCAGCCACTGCACGCGCAGCTGTGCGCATCGCTGGCTCTGGGTTCCGCCGAATGCGTCTCCT
GGCGGTGATGGCTCTGGACCGCGCGGCCGCACTGTGCCGCCGCTGCGCTATGCGGGGCTCGTCTCCCCG
CGCCTATGTGCGACGCTGGCCAGCGCTCCTGGCTAAGCGGCCTCACCAACTCGGTTGCGCAAACCGCGC
TCCTGGCTGAGCGGCGCTGTGCGCGCCCCGCTGCTGGGCCACTTCATCTGTGAGCTGCCGGCGTTGCT
CAAGTGCGCCGCGGAGGCGACGGAGACACTACCGAGAACCAGATGTTGCGCCCGCGGTGGTCATCCTG
CTGCTGCCGTTTGCCGTCATCCTGGCCTCCTACGGTGCCGTGCCCGGAGCTGTCTGTTGTATGCGGTTCA
GCGGAGCGCGGAGGAGGGCGGTGGGCACGTGTGGTCCCACCTGACAGCGCTGTGCTGTTCTACGGCTC
GGCCATCTACACCTACCTGCAGCCCGCGCAGCGCTACAACCAAGGACGCGGGCAAGTTCGTATCGCTCTTC
TACACCGTGGTCACACCTGCTCTCAACCGCTCATCTACACCCTCAGGAATAAGAAAGTGAAGGGGGCAG
CGAGGAGGCTGCTGCGGAGTCTGGGGAGAGGCCAGGCTGGGCAGTGA

```

10

推定スプライス部位は、配列番号 9 のヌクレオチド 15 と 16 との間に位置する。1 つの実施形態において、ヌクレオチド 1 ~ 15 はエキソン 1 からであり、ヌクレオチド 16 ~ 957 はエキソン 2 からである。

【0068】

開示された ba145122_B 核酸配列は、Homo sapiens 由来の GPCR mRNA (GENBANK-ID: HUMORLMHC | acc: L35475) に対して 934 塩基のうちの 592 塩基が同一性 (63%) を有する ($E = 2.4 \times 10^{-5}$)

20

【0069】

配列番号 9 によってコードされる開示された GPCR3 ポリペプチド (配列番号 10) は、315 アミノ酸残基であり、表 3B に一文字コードを用いて表示される。開示された GPCR3 タンパク質の最初の 70 アミノ酸を、シグナルペプチド推測および細胞局在に関して分析した。SignalP 結果は、GPCR3 が配列番号 10 の 46 位と 47 位との間 (すなわち、アミノ酸配列 NSA-LV 中のスラッシュ) で切断されることを推測する。Sort および Hydrophathy プロフィールはまた、GPCR3 がシグナルペプチドを含み、おそらく原形質膜に局在することを (0.6000 の確実性) 推測する。

【0070】

30

【表 21】

表 3B. コードされる GPCR3 タンパク質配列 (配列番号 10)

```

MKANYSAEERFLLLGFSDFSLQPVLFALVLLCYLLTLTGNSA/LVLLAVRDPRLHTFMYFLCHLALVDA
GFTTSVVPILLANLRGPALWLP RSHCTAQLCASLALGSAECVLLAVMALDRAAAVCRPLRYAGLVSPRLC
RTLASASWLSGLTNSVAQTALLAERPLCAPRLLGHFICELPALLKLARGGDDTTENQMFARVVILLP
FAVILASYGAVARAVCCMRFSGGRRRAVGTGSHLTAVCLFYGSAIYTYLQPAQRYNQARGKFVSLFYTV
VTPALNPLIYTLRNKKVKGAARRLLRLSLGRGQAGQ

```

BLASTX 検索を、公共のタンパク質データベースに対して実行した。本発明のタンパク質の全長アミノ酸配列は、Mus musculus 由来の 309 アミノ酸残基の MM17M-6、7 回膜貫通レセプター (OR 様タンパク質) (ptnr: SP TREMBL - ACC: Q9WV09、配列番号 52) ($E = 1.5 \times 10^{-128}$) に、308 アミノ酸残基のうち 253 残基が同一 (82%) であり、308 残基のうち 264 残基が陽性 (85%) であることが見出された。これらのタンパク質の整列を、表 3C に示す。

40

【0071】

【表 22】

表 3C. QpWV09(配列番号52)とのGPCR3の整列

	10	20	30	40	50	60	
GPCR3							
Q9WV09	MRANYSADERFLLGFSQNPSPLOPVLFAVLVLCYLLTLTGNSALVLLAPRDPRLTPTMYV						60
							
	--ANYSASAEERFLLGFSQNPSPLOPVLFAVLVLCYLLTLTGNSALVLLAPRDPRLTPTMYV						58
	70	80	90	100	110	120	
GPCR3							
Q9WV09	FLCHLALVLAGETTSVVPPLLENERGPALWLEFSHCIAQLLASLALGSAECVLLAVMALE						120
							
	FLCHLALVLAGETTSVVPPLLENERGSMQLPPAGSAQLGSSALGSAECVLLAVMALE						118
	130	140	150	160	170	180	
GPCR3							
Q9WV09	RAAAVCRPLRYAGVSEPLCRTLASASWLSGLINNVACTALLAEFLCAPPLCHFCICEL						180
							
	RAAAVCRPLRYTSLASPLCRTLAGVSWLGLANSAAQTALLAEFLCAPPECHFCICEL						178
	190	200	210	220	230	240	
GPCR3							
Q9WV09	PALLKILARGDGDTEENNFARVVILLLEFAVCLASYGAVARAVCCRFSGGRERAVG						240
							
	PALLKILARGDGRSATERNFARVVILLLEFAVCLASYGAVARAVGWHSSSGWRFAAS						238
	250	260	270	280	290	300	
GPCR3							
Q9WV09	QCSHLTAVCLFYGSATITTYLQFAQRYNCAQKQVSLFYTVVTPALNPLIYTLRNHVKGA						300
							
	QCSHLTAVCLFYGSATITTYLQFTSYNCAQKQVSLFYTVVTPALNPLIYTLRNHVKGA						298
	310						
GPCR3							
Q9WV09	PRLLRSLSGGOAGQ						315
							
	PRLLRSLSGPE						309

開示されたGPCR3タンパク質（配列番号10）は、表3Dに示されるように多数の嗅覚レセプタータンパク質と良好な同一性を有する。

【0072】

【表23】

表 3D. GPCR3に関するBLASTの結果

遺伝子インデックス 識別名	タンパク質/生物	長さ(aa)	同一性(%)	陽性(%)	期待値
gi 5051404 emb CA B45012.1 (AL078630)	573K1.15 - mm17M1-6 (新規 7回膜貫通レセプター) <i>Mus musculus</i>	309	229/296 (77%)	240/296 (80%)	e-108
gi 12054359 emb C AC20487.1 (AJ302567) (AJ02568)	OR <i>Homo sapiens</i>	312	145/298 (48%)	191/298 (63%)	2e-69
gi 12054355 emb C AC20485.1 (AJ302565) (AJ302566) (AJ3025 69) (AJ302570)	OR <i>Homo sapiens</i>	312	144/298 (48%)	191/298 (63%)	3e-69
gi 12231029 sp Q1 5062 O2H3 HUMAN	OR 2H3 <i>Homo sapiens</i>	316	143/294 (48%)	185/294 (62%)	3e-68
gi 9798920 gb AAF 98752.1 AF211940_ 1 (AF211940)	OR <i>Homo sapiens</i>	303	142/287 (49%)	183/287 (63%)	5e-68

この情報は、GPCR3を関連タンパク質配列と比較するClustalW分析のように、表3Eに提供する複数配列整列にグラフで表す（GPCR3を1行目に示す）。

【0073】

（表3E. ClustalWタンパク質に関する情報）

- 1) 新規GPCR3（配列番号10）
- 2) gi|5051404|emb|CAB45012.1|573K1.15（mm17M1-6（新規7回膜貫通レセプター（ロドプシンファミリー）（嗅覚レセプター様）タンパク質）*Mus musculus*（配列番号53）
- 3) gi|12054359|emb|CAC20487.1|olfactory r

ceptor Homo sapiens (配列番号54)

4) gi|12054355|emb|CAC20485.1|olfactory receptor
ceptor Homo sapiens (配列番号55)

5) gi|12231029|sp|Q15062|O2H3_HUMAN OLFACTORY
RECEPTOR 2H3 (OLFACTORY RECEPTOR-LIKE
PROTEIN FAT11 (配列番号56)

6) gi|9798920|gb|AAF98752.1|AF211940_1 olfactory
receptor Homo sapiens (配列番号57)

【0074】

【表24】

10

	10	20	30	40	50	60
GPCR3						
gi 5051404	---MKAFYAEERFLLLGFESESLGPVLEAEVIECYLLTITNSAAYLLVRUPRLHTPN					
gi 12054359	---AKHSASAEERFLLLGFESESLGPVLEAEVIECYLLTITNSAAYLLVRUPRLHTPN					
gi 12054355	MLMKKAAEFEEFLLLGFSNPHLEVVLEVVDLHFLITLGLSLFLLSLYLSHHTPN					
gi 12231029	MLMKKAAEFEEFLLLGFSNPHLEVVLEVVDLHFLITLGLSLFLLSLYLSHHTPN					
gi 9798920	---MCKQSTPGFLLLGFESESLGLERTLFVVVETSLLTITGLTLILLGALDSELSI					
	70	80	90	100	110	120
GPCR3	YFICCHLLAGETTSVVEPLIANLRSEALMPPRHDAICASSLALCSAECVLIAMV					
gi 5051404	YFICCHLLAGETTSVVEPLIASIRSMICAPRAGMAICASSLALCSAECVLIAMV					
gi 12054359	YFFLSNLSFLLDLCETTSSEPLLVLNMGPKTISVAGFVOISFVLALGTAEVLIAMV					
gi 12054355	YFFLSNLSFLLDLCETTSSEPLLVLNMGPKTISVAGFVOISFVLALGTAEVLIAMV					
gi 12231029	YFFLSNLSFLLDLCETTSSEPLLVLNMGPKTISVAGFVOISFVLALGTAEVLIAMV					
gi 9798920	YFFLSNLSFLLDLCETTSSEPLLVLNMGPKTISVAGFVOISFVLALGTAEVLIAMV					

20

【0075】

【表25】

	130	140	150	160	170	180
GPCR3	LDRAAAYCPDRRAGTMSURLCPTLASASWESLITNSVASTPLLAERPLCAPPLCHFC					
gi 5051404	LDRAAAYCPDRRAGTMSURLCPTLAGVSWGGIANSAAOTALLAARPLCAPPLCHFC					
gi 12054359	LDRAAAYCPDRRAGTMSURLCPTLAGVSWGGIANSAAOTALLAARPLCAPPLCHFC					
gi 12054355	LDRAAAYCPDRRAGTMSURLCPTLAGVSWGGIANSAAOTALLAARPLCAPPLCHFC					
gi 12231029	LDRAAAYCPDRRAGTMSURLCPTLAGVSWGGIANSAAOTALLAARPLCAPPLCHFC					
gi 9798920	LDRAAAYCPDRRAGTMSURLCPTLAGVSWGGIANSAAOTALLAARPLCAPPLCHFC					
	190	200	210	220	230	240
GPCR3	EEPAALLRLRGDGGDTTENDMFAARVYLLLEDFVILLASYGATRAVCCRFECRRFV					
gi 5051404	EEPAALLRLRGDGGDTTENDMFAARVYLLLEDFVILLASYGATRAVCCRFECRRFV					
gi 12054359	EEPAALLRLRGDGGDTTENDMFAARVYLLLEDFVILLASYGATRAVCCRFECRRFV					
gi 12054355	EEPAALLRLRGDGGDTTENDMFAARVYLLLEDFVILLASYGATRAVCCRFECRRFV					
gi 12231029	EEPAALLRLRGDGGDTTENDMFAARVYLLLEDFVILLASYGATRAVCCRFECRRFV					
gi 9798920	EEPAALLRLRGDGGDTTENDMFAARVYLLLEDFVILLASYGATRAVCCRFECRRFV					
	250	260	270	280	290	300
GPCR3	GTGSSHLTAVCLFYGEANTLQEAQRVNTARGREYSLEYTVVTFALNPLIYTLRNEV					
gi 5051404	GTGSSHLTAVCLFYGEANTLQEAQRVNTARGREYSLEYTVVTFALNPLIYTLRNEV					
gi 12054359	GTGSSHLTAVCLFYGEANTLQEAQRVNTARGREYSLEYTVVTFALNPLIYTLRNEV					
gi 12054355	GTGSSHLTAVCLFYGEANTLQEAQRVNTARGREYSLEYTVVTFALNPLIYTLRNEV					
gi 12231029	GTGSSHLTAVCLFYGEANTLQEAQRVNTARGREYSLEYTVVTFALNPLIYTLRNEV					
gi 9798920	GTGSSHLTAVCLFYGEANTLQEAQRVNTARGREYSLEYTVVTFALNPLIYTLRNEV					
	310	320				
GPCR3	DAFRRLLRSLSRGQAGQ---					
gi 5051404	DAFRRLLRSLSRGQAGQ---					
gi 12054359	DAFRRLLRSLSRGQAGQ---					
gi 12054355	DAFRRLLRSLSRGQAGQ---					
gi 12231029	DAFRRLLRSLSRGQAGQ---					
gi 9798920	DAFRRLLRSLSRGQAGQ---					

30

40

GPCR3に関するDOMAINの結果は、Reverse Position Specific
BLASTを用いてConserved Domain Database
(CDD) から収集した。このBLASTサンプルドメインは、Smart and P

50

f a mコレクションから見出された。7 t m__1 (7回膜貫通レセプター (ロドプシンファミリー)) は、G P C R 3に対する有意な相同性を有する2つのセグメントを有することが示された。アミノ酸残基40~226の領域は、「7回膜貫通レセプター (ロドプシンファミリー) フラグメント」ドメイン (配列番号45、E = 1 e - 15) のアミノ酸1~167と整列し、そしてG P C R 3のアミノ酸219~290は、P f a mデータベースの7 t m__1エントリーの残基305~377と整列した (E = . 004)。これは、G P C R 3配列が、このドメインならびに7 t m__1ドメイン自体を含むことが公知の他のタンパク質に類似した特性を有することを示す。

【0076】

本発明の核酸およびタンパク質は、以下でさらに記載されるように、様々なG P C R関連の病理学的障害および／またはO R関連の病理学的障害に関する潜在的な治療的適用において有用である。例えば、嗅覚受容体様タンパク質をコードするc D N Aは、遺伝子治療において有用であり得、そしてこの嗅覚受容体様タンパク質は、これを必要とする被験体に投与される場合に有用であり得る。非限定的な例として、本発明の組成物は、新生物、腺癌、リンパ腫、前立腺癌、子宮癌、免疫応答、A I D S、喘息、クローン病、多発性硬化症、およびオールブライト遺伝性骨形成異常症に罹患している患者の処置のための効力を有する。他のG P C R関連疾患および障害が企図される。

【0077】

本発明のG P C R様タンパク質をコードする新規の核酸、またはそのフラグメントは、診断用途においてさらに有用であり得、ここで、この核酸またはタンパク質の存在または量が評価されるべきである。これらの材料は、治療方法または診断方法において使用するための本発明の新規の物質に免疫特異的に結合する抗体の作製において、さらに有用である。これらの抗体は、当該分野で公知の方法に従って、疎水性チャートによる予測を使用して、以下の「抗G P C R X抗体」の章に記載されるようにして作製され得る。これらの新規のタンパク質は、様々なヒト障害の機能分析のためのアッセイシステムにおいて使用され得、このタンパク質は、疾患の病理学の理解および様々な障害のための新規の薬物標的の開発の際に役立つ。

【0078】

(G P C R 4)

G P C R 4は、以下に開示される3つの核酸のファミリーを含む。この開示される核酸は、G P C R様タンパク質をコードする。

【0079】

(G P C R 4 a)

開示されるG P C R 4 aは、3つの異なる核酸、G P C R 4 a 1 (d j 4 0 8 b 2 0__C)、G P C R 4 a 2 (d j 4 0 8 b 2 0__C__d a 1)、およびG P C R 4 a 3 (C G 5 5 3 5 8 - 0 3) によりコードされる。第1の核酸であるd j 4 0 8 b 2 0__C (G P C R 4 a 1) は、947ヌクレオチド長 (配列番号11) である。オープンリーディングフレームは、ヌクレオチド3~5においてA T G開始コドンで始まり、そしてヌクレオチド939~941においてT G Aコドンで終結することが同定された。開始コドンから上流であり、かつ終止コドンから下流の推定非翻訳領域は、表4において下線を引かれており、そして開始コドンおよび終止コドンは、太字である。312アミノ酸残基を有するコードされたタンパク質は、表4 Bにおいて一文字のコードを使用して表される (配列番号12)。

【0080】

【表26】

表 4A. GPCR4a1 スクレオチド配列 (配列番号11)

```

GTATGGAAAACGATAATACAAGTTCTTTTGAAGGCTTCATCCTGGTGGGCTTCTCTGATCGTCCCCACCT
AGAGCTGATCGTCTTTGTGTTGTCTCATCTTTTATCTGCTGACTCTTCTTGGCAACATGACCATTGTC
TTGCTTTTCAGCTCTGGATTCCCGGCTGCACACACCAATGTATTTCTTTTGGCAAACCTCTCATTCCTGG
ACATGTGTTTCACCCACAGGTCCATCCCTCAGATGCTCTACAACCTTTGGGGTCCAGATAAGACCATCAG
CTATGTGGGTGTGCCATCCAGCTGTAATTTGCTGCTGGCCCTGGGAGGGGTGGAGTGTCTCTCTGGCT
GTCATGGCATATGACCGCTATGCTGCAGTCTGCAAACCCCTGCACTACACCATCATCATGCCACCCAGTC
TCTGTGGACAGCTGGCTTCACTGGCATGGCTGAGTGGCTTTGGCAATTTCTCTATAATGGCACCCAGAC
ATTGATGCTACCCCGCTGTGGGCACAGACGAGTTGACCACTTTCTCTGTGAGATGCCAGCACTAATTGGT
ATGGCCTGTGTAGACACCATGATGCTTGAGGCAGCTGGCTTTTGGCCTGGCAATCTTTATCATCCTGGCAC
CACTCATCCTCATTCTCATTCTTTATGTTACGTTGGAGAACAGTCTTAGGATCAAGTCAGCTGCTGG
GCGAAAGAAAGCCTTCAACACTTGCAGCTCGCATCTAATTGTGTCTCTCTCTTATGGTACAAATCATA
TACATGTACCTCCAGCCAGCAAATACCTTATCCAGGACAGGGCAAGTTTCTTACCCTTTTCTACACAA
TTGTCACTCCAGTGTAAACCCCTGATCTATACACTAAGAAACAAAGATGTTAAAGAGGCCATGAAGAA
GGTGTAGGGAAGGGGAGTGCAGAAATATAGTAAGG

```

10

20

30

開示される核酸 GPCR4a1 配列は、939bp の HOMO sapiens 嗅覚受容体様タンパク質 (OR2C1) 遺伝子 (GENBANK 登録番号: AF098664) ($E = 3.8 \times 10^{-72}$) に対して 624/931 (67%) 同一の (624/931 の陽性、67%) 塩基を有する。配列データベースの検索において、部分的な一致点がまた同定され、例えば、ヌクレオチド 719~947 のマイナス鎖は、1320bp の合成 GPCR mRNA (患者番号: T72050) (染色体 6p2.1 の HH 領域におけるマーカー 2B8 の周辺の配列) に対して 229/229 (100%) 同一な塩基を有し、そして同じ領域の GPCR4a1 は、合成 GPCR mRNA (患者番号: T72050) (これもまた染色体 6p2.1 の HH 領域中のマーカー 2B8 の周辺の配列) に対して 229/229 (100%) 同一な塩基を有した (両方の場合の E 値は、 9.6×10^{-47} である)。

【0081】

配列番号 11 によってコードされる GPCR4A ポリペプチド (配列番号 12) は、表 4B において、一文字のアミノ酸コードを使用して表される。GPCR4A の Sort プロフィールは、この配列がシグナルペプチドを有し、そして 0.6000 の確実度を有する形質膜に位置するようであることを示す。GPCR4a ペプチドの最も可能な切断部位は、シグナル P 結果に基づいて、アミノ酸 41 と 42 との間、すなわち、アミノ酸配列 L L G - N M におけるスラッシュである。

【0082】

【表 27】

表 4B. GPCR4a タンパク質配列 (配列番号12)

```

MENDNTSSFEGFIVGVSDRPHLEIVFVVVLI FYLLTL LG/NMTIVLLSALDSRLHTPMYFFLANLSF
LDMCFTTGSIFQMLYNLWGPDKTISYVGCAIQLYFVLALGGVECVLLAVMAYDRYA AVCKPLHYTIIMH
PRLCGQLASVAVLSGFGNSLIMAPQTLMLPRCGHRRVDHFLCEMPALIGMACVDTMLEALAFALAI FI
ILAPLILILISYGVGGTVLRIKSAAGRKKAFNTCSSHLIVVSLFYGTIIYMYLQFANTYSQDQGKELT
LFYTIIVTPSVNPLIYTLRNKDVKEMKKVLGKSAEI

```

40

上記の推定 GPCR4a1 配列を、この配列を確認するためにエキソン結合プロセスに供した。PCR プライマーは、順方向プライマーについて使用可能な最も上流の配列、および逆方向プライマーについて利用可能な最も下流の配列において開始することによって設計された。各場合において、独特であるかまたは高度に選択性であるかのいずれかの適切な配列に遭遇するまで、または逆方向プライマーの場合、終止コドンに到達するまで、それぞれの末端からコード配列まで内向きに走査して、この配列を試験した。次いで、このような適切な配列を、広範な cDNA ライブラリに基づいて、PCR 増幅において順方向プライマーおよび逆方向プライマーとして使用した。得られた単位複製配列を、実施例に記載されるように、ゲル精製し、クローン化し、そして高重複性になるまで配列決定した。

50

【0083】

クローン化した配列は、本明細書中でGPCR4a2と称され、そして表4Cで報告される、GPCR4a2（配列番号13）の代替の実施形態として開示される。あるいは、この945ヌクレオチド配列（GPCR4a2）は、本明細書中でdj408b20__C__da1と称される。この核酸は、5' UTRにおいて、GPCR4a1より短い2個のヌクレオチドである。しかし、GPCR4a2は、同じ312アミノ酸タンパク質（GPCR4a、配列番号12）をコードする。

【0084】

【表28】

10

表4C. GPCR4a2ヌクレオチド配列（配列番号13）	
GTATGGAAACGATAATACAAGTTCTTTTGAAGGCTTCATCCTGGTGGGCTTCTCTGATCGTCCCCACCTAGAGCTGATC GTCTTTGTGGTTGTCTCATCTTTTATCTGCTGACTCTTTGGCAACATGACCATTTGCTTGTCTTTTCAGCTCTGGATTC CCGGCTGCACACCAATGTATTCTTTTGGCAAACCTCTCATTCTGGACATGTGTTTCACCACAGGTTCCATCCCTC AGATGCTCTACAACTTTGGGGTCCAGATAAGACCATCAGCTATGTGGGTTGTGCCATCCAGCTGTACTTTGTCTGGCC CTGGGAGGGGTGCAGTGTCTCTCTGGCTGTCTGGCATATGACCGCTATGCTGCAGTCTGCAAACCCCTGCACTACAC CATCATCATGCACCCAGCTCTCTGTGGACAGCTGGCTTCAGTGGCATGGCTGAGTGGCTTTGGCAATCTCTCATATATGG CACCCAGACATTGATGCTACCCCGCTGTGGGCACAGACGAGTTGACCACTTTCTCTGTGAGATGCCAGCACTAATTGGT ATGGCTGTGTAGACACCATGATGCTTGAGGCACCTGGCTTTTGGCCTGGCAATCTTTATCATCCTGGCACCCTCATCCT CATTTCTATTCTTTATGTTTACGTTGGAGGAACAGTGCCTTAGGATCAAGTCAGCTGCTGGGCGAAAGAAAGCCTTCAACA CTGTCAGCTCGCATCTAATTGTTGTCTCTCTCTATGGTACAATCATATACATGTACCTCCAGCCAGCAATACTTAT TCCCAGGACCAGGGCAAGTTTCTTACCTTTTCTACACAATGTGCTACTCCAGTGTAAACCCCTGATCTATACACTAAG	

開示されるGPCR4aポリペプチドの全アミノ酸配列は、Homo sapiensの新規の7膜貫通レセプター（ロドプシンファミリー、嗅覚受容体様タンパク質HS6M1-15、ptnr: SPTREMBL-ACC: Q9Y3N9 DJ88J8.1、E=5.0e⁻¹⁰⁸）由来の320アミノ酸残基に対して197/305（64%）同一なアミノ酸残基、およびこれに対して242/305（79%）陽性の残基を有する。

【0085】

GPCR4aとの整列の最良のヒットのBLASTP（非重複複合データベース）分析を表4Dに示す。

【0086】

【表29】

30

表4D. GPCR4aのBLASTP結果					
送付インデックス 識別子	クイック検索/生物体	長さ (aa)	同一性 (%)	陽性 (%)	期待値
ACC:Q9Y3N9 DJ88J8.1	新規の7膜貫通レセプター (ロドプシンファミリー、OR様) Ha6m1-15 Homo Sapiens	320	197/305 (64%)	242/305 (79%)	9.4e- 109
ACC:P23275	嗅覚受容体 15 (OR3) Mus musculus	312	190/308 (61%)	241/308 (78%)	5.5e- 104
ACC:Q76001 DJ80119.7	OR様タンパク質 (HS6M1-3) Homo sapiens	311	193/302 (63%)	244/302 (80%)	1.5e- 103

BLASTXをまた行って、GPCR4aと有意な同一性を有するタンパク質を決定した。BLASTXの結果を表4Eに示す。

【0087】

【表30】

40

表4E. GPCR4aのBLASTX結果					
高スコアセグメントを産生する配列 -		ハイディング フレーム	高スコア	最小P値 P(N)	N
Ptnr:SPTRMBL-ACC:Q9Y3N9	DJ88J8.1 (NOVEL 7 TRANSMEMBR...	+3	1076	5.6e-108	1
Ptnr:SWISSPROT-ACC:P23275	OLFACTORY RECEPTOR 15 (OR3) ...	+3	1031	3.3e-103	1
Ptnr:SPTRMBL-ACC:O76001	DJ80I19.7 (OLFACTORY RECEPTOR...	+3	1027	8.7e-103	1
Ptnr:SPTRMBL-ACC:O95371	OLFACTORY RECEPTOR-LIKE PROTE...	+3	992	4.5e-99	1
Ptnr:SPTRMBL-ACC:O95918	DJ271M21.2 (HS6M1-12 (7 TRANS...	+3	988	1.2e-98	1
Ptnr:SPTRMBL-ACC:Q9WV11	573K1.8 (MM17M1-2 (NOVEL 7 TR...	+3	984	3.2e-98	1
Ptnr:SPTRMBL-ACC:Q9WV14	573K1.2 (MM17M1-3 (NOVEL 7 TR...	+3	981	6.6e-98	1

10

G P C R 4 a 2 について見出された可能な S N P を表 4 F に列挙する。

【 0 0 8 8 】

【 表 3 1 】

表 4F: SNPs		
塩基 位置	塩基 前	塩基 後
44	T	C(20)
147	T	C(2)
220	T	C(2)
271	T	C(3)
432	A	G(2)
452	C	T(2)
493	A	G(2)
771	T	C(3)

20

(G P C R 4 b)

d j 4 0 8 b 2 0 _ C (G P C R 4 a 1) として同定された標的配列を、再び、エキソン結合プロセスに供して、配列を確認した。順方向プライマーについて、利用可能な最も上流の配列、および逆方向プライマーについて利用可能な最も下流の配列で開始することによって、P C R プライマーを設計した。この配列をコードする c D N A を、以下のプライマーを使用するポリメラーゼ連鎖反応 (P C R) によってクローン化した：

30

以下を含むヒト c D N A のプール 1 上の A T A C A A G T T C T T T C G A A G G C T T C A T C C (配列番号 1 4) および

C C C T T A C T A T A T T T C T G C A C T C C C C T T (配列番号 1 5) ；

副腎、骨髄、脳－扁桃、脳－小脳、脳－海馬、脳－黒質、脳－視床、脳－全脳、胎児脳、胎児腎臓、胎児肝臓、胎児肺、心臓、腎臓、リンパ腫－ラージ、乳腺、脾臓、下垂体、胎盤、前立腺、唾液腺、骨格筋、小腸、脊髄、脾臓、胃、精巣、甲状腺、気管、子宮。

【 0 0 8 9 】

40

プライマーを、本発明の D N A / タンパク質配列の全長またはその一部 (1 つ以上のエキソン) のインシリコ (i n s i l i c o) 予測に基づいて、または密接に関連したヒト配列または他の種の配列に対する推定エキソンの翻訳相同性によって、設計した。通常、複数のクローンを配列決定してこの配列を駆動し、次いでこれを、S e q C a l l i n g プロセスと同様に構築した。さらに、配列トレースを、手で評価し、そして適切な場合、補正について編集した。エキソン結合によって駆動された P C R 産物を、I n v i t r o g e n 製の p C R 2 . 1 ベクターにクローン化した。細菌性クローン 1 1 5 8 4 3 : : D J 4 0 8 B 2 0 _ C . 6 9 8 3 2 2 . D 1 0 は、I n v i t r o g e n 製の P C R 2 . 1 ヘクローン化されたオープンリーディングフレーム全体を網羅する挿入物を有する。

【 0 0 9 0 】

50

各場合において、独特であるかまたは高度に選択性であるかのいずれかの適切な配列に遭遇するまで、または逆方向プライマーの場合、終止コドンに到達するまで、それぞれの末端からコード配列まで内向きに走査して、この配列を試験した。このようなプライマーは、本発明のDNA／タンパク質配列の全長またはその一部（1つ以上のエキソン）のインシリコ予測に基づいて、または密接に関連したヒト配列または他の種の配列に対する推定エキソンの翻訳相同性によって、設計した。代表的には、実施例に記載されるように、得られた単位複製配列をゲル精製し、クローン化し、そして高重複性になるまで配列決定した。全てのクローンから得られる配列を、それ自体を用い、L u r a G e n C o r p o r a t i o nのデータベースの他のフラグメントおよび公開されたESTを用いて構築した。アセンブリの別の成分とのそれらの同一性の程度が、50bpにわたって少なくとも95%である場合、フラグメントおよびESTは、このアセンブリの成分として含まれた。さらに、この配列トレースを、手で評価し、そして適切な場合は、補正のために編集した。

【0091】

これらの手順は、GPCR4タンパク質をコードする第3の核酸を提供する。この核酸は、GPCR4a3またはCG55358-03といわれる。この核酸は、932ヌクレオチド長（配列番号16、表4G）であり、そして、GPCR2a1よりも、5'UTRにおいて、16ヌクレオチド短い。ヌクレオチド3～5におけるアミノ酸トレオニンについてコードするACAコドンで始まり、そして、ヌクレオチド924～926におけるTAGコドンで終わる成熟タンパク質のオープンリーディングフレームが、同定された。存在する場合、推定の非翻訳領域は、開始コドンから上流に、そして終結コドンから下流に見出される。1つの沈黙塩基置換が存在する：C767は、GPCR4a3中のT752である。GPCR4b（GPCR4a3によってコードされるタンパク質）は、表4H（配列番号17）において示されるように、N末端の5アミノ酸欠失を除いて、GPCR4aと同一である。

【0092】

本明細書中に開示されるGPCR嗅覚レセプターは、少なくとも以下の組織において発現される：網膜色素上皮の先端の微絨毛、動脈（大動脈）、基底前脳、脳、パーキットリンパ腫細胞株、脳梁、心（房および室）、尾状核、CNSおよび末梢組織、小脳、大脳皮質、結腸、皮質神経細胞、内皮（冠状動脈および臍静脈）細胞、口蓋上皮、眼、新生児眼、前頭皮質、胎児造血細胞、心臓、海馬、視床下部、白血球、肝臓、胎児肝臓、肺、肺リンパ腫細胞株、胎児リンパ様組織、成人リンパ様組織、MHCIおよびIIを発現する組織、神経組織、髄質、視床下核、卵巣、膵臓、下垂体、胎盤、橋、前立腺、被殻、血清、骨格筋、小腸、平滑筋（大動脈中の冠状動脈）、脊髄、脾臓、胃、舌の味覚レセプター細胞、精巣、視床、および胸腺組織。この情報は、Seq Calling供給源、公のEST供給源、文献供給源、および／またはRACE供給源を含むがこれらに限定されない、本発明に含まれる配列の組織供給源を決定することによって、引き出される。

【0093】

配列データベースの研究において、例えば、開示されたGPCR4a3核酸配列は、gb：GENBANK-ID：AF098664 | acc：AF098664.1 mRNA from Homo sapiens (Homo sapiens olfactory receptor-like protein (OR2C1) gene, complete cds, E=2.1e-71) に対して、913塩基のうち614（67%）が同一である。

【0094】

【表32】

表 4 G. GPCR 4a3 ヌクレオチド配列 (配列番号 16)

ATACAAAGTTCTTTTCGAAGGCTTCATCCTGGTGGGCTTCTCTGATCGTCCACCTAGAGC	60
TGATCGTCTTTGTGGTTGTCTCATCTTTATCTGCTGACTCTTCTGGCAACATGACCA	120
TTGTCTTGCTTTCAAGTCTGGATTCCCGGCTGCACACCAATGTATTTCTTTTGGCAA	180
ACCTCTCATTCTGGACATGTGTTTACCACAGGTTCCATCCCTCAGATGCTCTACAACC	240
TTTGGGGTCCAGATAAGACCATCAGCTATGTGGGTTGTGCCATCCAGCTGTACTTTGTCC	300
TGGCCCTGGGAGGGGTGGAGTGTGCTCCTCCTGGCTGTGATGGCATATGACCGCTATGCTG	360
CAGTCTGCAAAACCCCTGCCTACACCATCATCATGCCACCCAGTCTCTGTGGACAGGTGG	420
CTTCAGTGGCATGGCTGAGTGGCTTTGGCAATTCTCTCATAATGGCAGCCAGACATTGA	480
TGCTACCCCGCTGTGGGCACAGACGAGTTGACCACTTCTCTGTGAGATGCCAGCACTAA	540
TTGGTATGGCCTGTGTAGACACCATGATGCTTGAGGCACTGGCTTTTGGCCCTGGCAATCT	600
TTATCATCCTGGCACCAGTCTCATCTCATTCTCATTTCTTATGGTTAGGTTGGAGGAACAG	660
TGCTTAGGATCAAGTCAGCTGCTGGCGAAAGAAAGCCTTCAACACTTGACGCTCGCATC	720
TAATTGTTGTCTCTCTCTCTTCTATGGTACAAATATATACATGTACCTCCAGCCAGCAAATA	780
CTTATTTCCAGGACAGGGCAAGTTCTTACCCTTTTCTACACAATGTCACTCCAGTGG	840
TTAACCCCTGATCTATACACTAAGAAACAAAGATGTTAAAGAGGCCATGAAGAAGGTGC	900
TAGGGAGGGGAGTGCAGAAATATAGTAAGG	932

10

20

開示された GPCR 4b 嗅覚レセプター様タンパク質についての、シグナル P、P 種類および／またはヒドロパシープロフィールは、この配列が、この配列がシグナルペプチドを有し、そして、0.6000 の確実性で、原形質膜に局在し得るようであるということを推定する。シグナル P は、表 4 H において LLG / NM の間のスラッシュによって示されるように、シグナル配列が、アミノ酸 36 と 37 との間で、切断部位を有する最初の 36 アミノ酸中にコードされることを示す。

【0095】

【表 33】

表 4 H. GPCR 4b アミノ酸配列 (配列番号 17)

TSSFEGFILVGFSDRPHLELVFVVLIFVLLTLG/NMTIVLLSALDSRLHTPMYFFLAN	60
LSFLDMCFITTSIPQMLYNLWGPDKTISYVGCAIQLYFVLALGGVECVLLAVMAYDRYAA	120
VCKPLHYTIIMHPRLCGLASVAWLSGFGNSLIMAPQTLMLPRCGHRRVDHFLCEMPALI	180
GMACVDTMMLLEALAFALAFIILAPLILILISYGVGGTVLRIKSAAGRKKAFNTCSSHL	240
IVVSLFYGTIIYMYLQPANTYSQDQKFLTLFYITIVTPSVNPLIYTLRNKDVKEAMKKVL	300
GKGSAEI	307

30

本発明の開示された GPCR 4b タンパク質の全長アミノ酸配列は、320 アミノ酸残基 p t n r : S P T R E M B L - A C C : H o m o s a p i e n s 由来の Q 9 Y 3 N 9 タンパク質 (D J 8 8 J 8 . 1、NOVEL 7 TRANSMEMBRANE RECEPTOR (RHODOPSIN FAMILY) (OLFACTORY RECEPTOR LIKE) PROTEIN) (H S 6 M 1 - 1 5), $E = 1.6e-107$) に 299 アミノ酸残基の 195 (65%) が同一であることが見出され、そして 299 アミノ酸残基のうち 239 (79%) が類似していることが見出された。

【0096】

以下の位置において、GPCR 4a3 ヌクレオチド配列の 1 つ以上のコンセンサス位置 (Cons. Pos.) が、SNP として同定されている。「深さ」は、SNP の領域をカバーするクローンの数を示す。推定対立遺伝子頻度 (Putative Allele Freq.) は、SNP を含む全てのクローンの画分である。示される場合、ダッシュ (「-」) は、塩基が存在しないことを意味する。記号「>」は、「に変更される」を意味する: Cons. Pos. : 422 深さ: 18 変化: T>C、Putative Allele Freq. : 0.333; Cons. Pos. : 546 深さ: 15 変化: T>C、Putative Allele Freq. : 0.133; Cons. Pos. : 753 深さ: 8 変化: C>T、Putative Allele Freq. : 0.250。

【0097】

50

G P C R 4 a または G P C R 4 b として特定にいわれない限り、G P C R 4 に対する任意の参照は、全ての改変体を含むことが推定される。本明細書中の任意の G P C R X 改変体配列の間の残基の差異は、「1つの」改変体における残基および「1つの」改変体に関する残基の位置を示すために記載される。全ての以下の配列において、G P C R 4 a タンパク質配列が使用される。

【0098】

開示された G P C R 4 タンパク質（配列番号 12）はまた、多数の嗅覚レセプタータンパク質と良好な同一性を有する。C l u s t a l W 分析のために使用される同一性情報を表 4 に示す。

【0099】

【表 34】

10

表 4I. GPCR4 についての BLAST 結果					
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
GI 6679170 ref NP_032788.1	OR 15 (OR3) <i>Mus musculus</i>	312	174/308 (56%)	216/308 (69%)	1e-93
GI 4826521 emb CAC42853.1 (AL035402, dJ88J8.1) (AJ302594-99) (AJ302600-01)	novel 7 tm receptor protein (rhodopsin fam., OR-like) (hs6M1-15) <i>Homo sapiens</i>	320	174/305 (57%)	213/305 (69%)	2e-93
GI 12054431 emb CAC20523.1 (AJ302603)	OR <i>Homo sapiens</i>	320	173/305 (56%)	213/305 (69%)	3e-93
GI 12054429 emb CAC20522.1 (AJ302602)	OR <i>Homo sapiens</i>	320	173/305 (56%)	213/305 (69%)	7e-93
GI 12054347 emb CAC20478.1 (AJ302558)	OR <i>Homo sapiens</i>	311	170/302 (56%)	211/302 (69%)	6e-90

20

この情報は、G P C R 4 と関連する O R 配列とを比較する C l u s t a l W 分析として、表 4（G P C R 4 は、1 行目に示される）において与えられる多重配列整列においてグラフによって示される。

【0100】

【表 35】

30

表4J ClustalW タンパク質についての情報

- 1) GPCR4 (配列番号12)
- 2) gi|6679170|ref|NP_032788.1| olfactory receptor 15 *Mus musculus* (配列番号58)
- 3) gi|4826521|emb|CAB42853.1| dJ88J8.1 (novel 7 transmembrane receptor (rhodopsin family) (OR like) protein) (hs6M1-15)) *Homo sapiens* (配列番号59)
- 4) gi|12054431|emb|CAC20523.1| olfactory receptor *Homo sapiens* (配列番号60)
- 5) gi|12054429|emb|CAC20522.1| olfactory receptor *Homo sapiens* (配列番号61)
- 6) gi|12054347|emb|CAC20478.1| olfactory receptor *Homo sapiens* (配列番号62)

	10	20	30	40	50	60
GPCR4						
GI 6679170						
GI 4826521						
GI 12054431						
GI 12054429						
GI 12054347						
	70	80	90	100	110	120
GPCR4						
GI 6679170						
GI 4826521						
GI 12054431						
GI 12054429						
GI 12054347						
	130	140	150	160	170	180
GPCR4						
GI 6679170						
GI 4826521						
GI 12054431						
GI 12054429						
GI 12054347						
	190	200	210	220	230	240
GPCR4						
GI 6679170						
GI 4826521						
GI 12054431						
GI 12054429						
GI 12054347						
	250	260	270	280	290	300
GPCR4						
GI 6679170						
GI 4826521						
GI 12054431						
GI 12054429						
GI 12054347						
	310	320				
GPCR4						
GI 6679170						
GI 4826521						
GI 12054431						
GI 12054429						
GI 12054347						

10

20

30

GPCR4についてのDOMAIN結果を、Reverse Position Specific BLASTを用いてConserved Domain Database (CDD) から回収した。このBLASTサンプルドメインは、Smart and Pfam回収物において見出された。GPCR4の2つの領域は、上記のように、377アミノ酸の7TMに対して同一性を有する。7tm_1 (上の7つの膜貫通レセプター (ロドプシンファミリー) (配列番号45)) は、GPCR4に相同性を有することが示された。GPCR4の残基41~159と整列した7TMの残基1-120 ($E = 7e-22$ 、表2Kにおいて示される) およびGPCR4の残基224~290は、7TMの残基310~377 ($E = 2e-04$) と同一性を有する。

40

【0101】

【表36】

表 4K: GPOR4と7TMとの間のフィンライメント

GPCR4	S	K	T	I	V	E	L	S	L	S	R	H	I	P	M	F	T	R	A	N	F	L	E	C	F	T	G	S	L	I	O	M	E	N	L	N	G	E	I	K	T	I	G	S	G	A	L	I	Q
TM7	S	N	I	L	V	C	A	V	E	R	K	A	S	I	T	T	N	L	L	V	S	E	V	A	R	L	V	A	L	V	M	A	N	V	L	E	V	G	E	K	T	F	R	H	G	E	D	F	
GPCR	V	F	V	E	A	G	G	V	E	C	L	L	A	V	E	R	D	R	A	V	C	K	S	H	I	T	I	M	-	H	P	L	C	G	G	H	A	S	A	N	S	G	G	N	G	L	I		
TM7	V	T	L	D	M	C	T	A	S	I	N	E	C	A	S	I	D	R	T	A	V	A	M	L	N	T	I	R	Y	S	S	K	E	V	T	V	T	H	I	V	E	L	S	L	T	I	C	P	S

20

30

40

梗塞、潰瘍、喘息、アレルギー、良性前立腺肥大、ならびに精神病性障害および神経性障害（不安、精神分裂病、躁うつ病、せん妄、痴呆、重篤な精神遅滞、ならびに運動異常（例えば、ハンティングトン病もしくは Gilles de la Tourette 症候群）。他の GPCR-4 疾患および障害が意図される。

【0103】

OR 様タンパク質をコードする新規な核酸、および本発明の OR 様タンパク質、またはそれらのフラグメントは、診断適用において、さらに有用であり、ここで、核酸またはタンパク質の存在または量が、評価される。これらの材料は、治療方法または診断方法、および他の疾患、障害および症状などにおける使用のために、本発明の新規な物質に対して免疫特異的に結合する抗体の産生において、さらに有用であり得る。これらの材料は、治療または診断方法での使用のために、本発明の新規な物質に免疫特異的に結合する抗体の産生においてさらに有用である。これらの抗体は、以下の節の「抗-GPCR X 抗体」において記載されるように、疎水性チャートからの予測値を使用して、当該分野で公知の方法に従って産生され得る。1 実施形態において、意図される GPCR 4 エピトープは、約 65 から 85 のアミノ酸由来である。別の局面において、GPCR 4 エピトープは、約 115 から 130 のアミノ酸由来である。さらなる実施形態において、GPCR 4 は、155 から 175、215 から 240、250 ~ 275 および 280 から 310 のアミノ酸由来である。これらの新規なタンパク質をまた使用して、機能分析に関するアッセイシステムを開発し得る。

【0104】

(GPCR 5)

GPCR 5 は、以下に記載されるような、3 種の類似の核酸および 3 種の類似のタンパク質のファミリーを含む。開示された核酸は、GPCR、OR 様タンパク質をコードする。

【0105】

(GPCR 5a)

開示された新規の、1003 ヌクレオチドの GPCR 5a 核酸（これはまた、115-a-12-A と呼ばれる）は、表 5A に示される。ORF は、ヌクレオチド 6 ~ 8 の ATG 開始コドンで開始し、そしてヌクレオチド 999 ~ 1001 の TAA コドンで終了する。推定非翻訳領域の、開始コドンの上流および末端コドンの下流を、表 5A において下線を引き、そして開始コドンおよび終止コドンは、太字である。

【0106】

【表 37】

表 5A. GPCR 5a ヌクレオチド 配列 (配列番号 18)

```

AGACAATGAGTCTGATGGGAACACAGTAGTGATCCAACAGAGTTCGTCCTGGCAGGGCTCCCAATCT
CAACAGCGCAAGAGTGGAAATATTTCTGTGTTCTTCTGTCTATCTCTGAATCTGACAGGCAATGTG
TTGATTGTGGGGGTGGTAAGGGCTGATACTCGACTACAGACCCCTATGTACTTCTTCTGGGTAACTGT
CCTGCCTAGAGATACTGCTCACTTCTGTCTATTCCTCAAGATGCTGAGCAATTCCTCTCAAGGCAACA
CACTATTTCTTTGCTGCTATGATCAACCAATCTATTTCTACTTCTTCTCGGGGCTCCGAGTTCTTA
CTGTTGGCTGTCTGCTGCTGCGGATCGCTACCTGGCCATCTGTCTATCCTCTGCGCTACCCCTTGCTCATGA
GTGGGGCTGTGTGCTTTCTGCTGTCCTTGGCCTGCTGGGTGGGGGACTCGTCCCTGTGCTTGGTCCCAC
AGTGGCTGTGGCCTTGCTTCTTTCTGTAAGCAGGGTGCTGTGGTACAGCACTTCTTCTGCGACAGTGGC
CCACTGCTCCGCTGGCTTGACCAACACCAAGAGCTGGAGGAGACTGACTTTGTCTGGCCTCCCTCG
TCATTGTATCTTCTTCTGCTGATCACTGCTGTGCTCTACGGCCTCATTGTGCTGGCAGTCTGAGCATCCC
CTGTGCTTCAAGCCGTGAGAGGCTTCTCTACCTGTACCTCCCACTTGATAGTGGTGACCCCTTCTCTAT
GGAGTGCCATTTTCTCTATGTGCGGCCATCGCAGAGTGGTTCTGTGGACACTAACTGGGCAGTGACAG
TAATAACGACATTTGTGACACCACTGTTGAATCCATTCTATGCCTTACGTAATGAGCAAGTCAAGGA
AGCTTTGAAGGACATGTTTAGGAAGGTAGTGCGAGCGGTTTAGGGAATCTTTACTTGATAAATGTCTC
AGTGAGAAAGCAGTAAAGTAAAA

```

配列番号 18 によってコードされた GPCR 5a タンパク質は、331 個のアミノ酸残基を有し、そして表 5B に 1 文字のコードを使用して示されている。GPCR 5a の P s o r t プロフィールにより、この配列が単一のペプチドを有し、そして 0.0600 の確実度で原形質膜において局在化され、これはまた、Golgi 体に局在化されると推定す

る。最も起こりそうなペプチドの切断部位は、アミノ酸の54と55との間、すなわち、Signal P結果に基づくアミノ酸配列VRADT（表5Bにおいてスラッシュで示してある）である。

【0107】

【表38】

表5B. ヌードネタ GPCR 5a タンパク質 配列 (配列番号 19)

```
MSPDGNHSSDPTFVLGLPLNLSARVELFSVFLVYLLNLTGNVLIVGVVRA/DTRLQTPMYFF
LGNLSCLEILLTSVIIPKMLSNFLSRQHTISFAACITQFYFYFFLGASEFLLAVMSADRYLAIC
HPLRYPELLMSGAVCFRVALACWVGGLVPVLGPTVAVALLPFCCKQGAHVQHFFCDSGPELLRLACTN
TKKLEETDFVLASLVIVSSLLITAVSYGLIVLAVLSIPSASGRQKAFSTCTSHLIVVTLFYGSAI
FLYVRPSQSGSVDTNNAVTVITTFVTPLLNPFYIYALRNEQVKEALKDMFRKVVAGVLGNLLLDKC
LSEKAVK
```

10

GPCR 5の開示された核酸配列は、Rattus norvegicus 嗅覚レセプタータンパク質mRNA (936bp) (遺伝子バンク: RATOLFPROD | acc: M64378) ($E = 1.1 \times 10^{-4.5}$) と同一な934塩基のうち604 (64%) を有し、そしてそれに陽性な934塩基のうち604 (64%) を有する。

【0108】

完全GPCR 5アミノ酸配列は、Mus musculus 由来の313アミノ酸残基嗅覚レセプター (ptnr: SPTREMBL-ACC: Q9Z1V0) ($E = 2.6 \times 10^{-7.4}$) と同一な304アミノ酸残基のうち149 (49%) を有し、そしてそれに陽性な304残基のうち201 (66%) を有する。

20

【0109】

(GPCR 5b)

GPCR 5a (115-a-12-A) を、エキソン連結プロセスに供し、おの配列を確認する。PCRプライマーを、前方向プライマーに対して最も上流の利用可能な配列、および逆方向プライマーに対して最も下流の配列で開始することによって設計した。各末端からコード配列の方へ内向きに、独特である適切な配列または高度に選択された適切な配列のいずれかが遭遇するまで、または逆方向プライマーの場合において、終止コドンが到達するまで、各場合において配列を試験した。次いで、このような適切な配列を、広範なcDNAライブラリーに基づくPCR増幅において、前方向プライマーおよび逆方向プライマーとして使用した。得られたアンプリコンを、ゲル精製し、クローン化し、そして高度な重複性まで配列決定し、GPCR 5bを提供し、これをまた、115-a-12-Bと言及する。

30

【0110】

GPCR 5bのヌクレオチド配列 (1004bp、配列番号20) は、表5Cに存在する。このヌクレオチド配列は、A5とA6との間にTを付加すること、および6個のヌクレオチド (GPCR 5aに対して番号付けされている) (C131>T; T186>C; G472>A; T579>A; A687>T; C799>T) の変化により、GPCR 5aとは異なる。

40

【0111】

【表39】

表5C. GPCR 5b マルチド配列 (配列番号 20)

AGACATATGAGTCCTGATGGGAACACAGTAGTGATCCACAGAGTTCGTCTGGCAGGGCTCCCAA
 ATCTCAACAGCGCAAGAGTGGAATTATTTTCTGTGTTCTTCTGTCTATCTCTGAATCTGATAGG
 CAATGTGTTGATTGTGGGGGTGGTAAGGGCTGATACTCGACTACAGACCCCATGTACTTCTTTCTG
 GGTAACCTGTCCTGCGCTAGAGATACTGCTCACTTCTGTCTCATTCACAAAGATGCTGAGCAATTTCC
 TCTCAAGGCAACACACTATTTCTTTGCTGCTATCACCCTTCTATTTCTACTTCTTTCTCGG
 GGCCTCCGAGTTCTTACTGTTGGCTGTCTGTGCGGATCGCTACCTGGCCATCTGTCTCCTCTG
 CGCTACCCCTTGTCTCATGAGTGGGGCTGTGTGCTTTCTGTGGCCTTGGCCTGCTGGGTGGGGGAC
 TCATCCCTGTGCTTGGTCCCACAGTGGCTGTGGCCTTGCTTCTTCTGTAAGCAGGGTGTGTGGT
 ACAGCACTTCTTCTGCGACAGTGGCCCTGCTCCGCTGGCATGCACCAACACCAAGAAGCTGGAG
 GAGACTGACTTTGTCCTGGCCTCCCTCGTCATTGTATCTTCTTGTGCTGATCACTGCTGTGCTACG
 GCCTCATTGTGCTGGCTGTCTGAGCATCCCTCTGCTTCAGGCCGTGAGAAGGCCTTCTCTACCTG
 TACCTCCCACTTGATAGTGGTGACCTCTTCTATGGAAGTGCCATTTTCTCTATGTGCGGTCATCG
 CAGAGTGGTTCTGTGGACACTAAGTGGGCGTGCAGTAATAACGACATTTGTGACACCACTGTTGA
 ATCCATTCTATGCTTACGTAATGAGCAAGTCAAGGAAGCTTCAAGGACATGTTTAGGAAGGT
 AGTGGCAGGCGTTTATAGGAATCTTTACTTGATAAATGTCTCAGTGAGAAAGCAGTAAAGTAAAA

10

このコードされたGPCR 5bタンパク質を、表5Dに示す。この開示したタンパク質は、331アミノ酸長であり、そして配列番号21と命名する。GPCR 5bは、3個のアミノ酸残基 (T42>I; V151>I; P265>S) でGPCR 5aとは異なる。GPCR 5aと同様に、GPCR 5bのSortプロフィールにより、この配列が単一のペプチドを有し、そして0.600の確実度で原形質膜において局在化されるようであることを予測する。ペプチドの最も起こりそうな切断部位は、アミノ酸54と55との間、すなわちSignalP結果に基づくアミノ酸配列VRADT (表5Dにおいてスラッシュで示してある) である。

20

【0112】

【表40】

表 5D. コードされたGPCR5b タンパク質配列 (配列番号 21)

MSPDGNHSSDPTEFVLGLPNLNSARVELFSVFLVYLLNLIGNVLIVGVVRA/DTRLQTPMYFFLGN
 LSCLEILLTSVVIIPKMLSNFLSRQHTISFAACITQFYFFFLGASEFLLAVMSADRYLAICHPLRYP
 LLMSGAVCFRVALACWVGGLIPVLGPTVAVALLPFCQKQAVVQHFFCDSGPLRLACTNTKKLEETDF
 VLASLVIVSSLLITAVSYGLIVLAVLSIPSASGRQKAFSTCTSHLIVVTLFYGSAIFLYVRSSQSGSV
 DTNWAIVTITTFVTPLLNPFYIALRNEQVKEALKDMFRKVVAGVLGNLLDKCLSEKAVK

30

GPCR 5bを有するアライメントについての最高のヒットのBLASTP (非冗長複合体データベース) 分析は、表5Eに列挙される。

【0113】

【表41】

表 5E. GPCR5bについてのBLASTPの結果

遺伝子名 / タンパク質名	タンパク質 / 生物	長さ (aa)	同一性 (%)	陽性 (%)	期待値
SPTREMBL-ACC:Q921V0	OLFACTORY RECEPTOR C6 <i>Mus musculus</i>	313	150/304 (49%)	203/304 (66%)	9.4e-76
SWISSPROT-ACC:P23267	OLFACTORY RECEPTOR-LIKE PROTEIN F6 <i>Rattus norvegicus</i>	311	152/301 (50%)	199/301 (66%)	5.2e-75
SPTREMBL-ACC:Q9Y3P5 DJ994E9.5	HS6M1-17 (NOVEL 7 TM (RHODOPSIN FAMILY) (OR LIKE PROTEIN) <i>Homo sapiens</i>	306	145/301 (48%)	197/301 (65%)	1.1e-67

40

GPCR 4aと有意な同一性を有するタンパク質を決定するためにBLASTXも行った。

50

【0114】

【表42】

表 5F. GPCR5b についての BLASTX の結果					
配列生成ハイスコアセグメント	識別	HA スコア	最小値 確率 P(N)	N	
ptnr:SPTREMBL-ACC:Q921V0	OLFACTORY RECEPTOR C6 - Mus m...	+1	764	5.6e-75	1
ptnr:SWISSPROT-ACC:P23267	OLFACTORY RECEPTOR-LIKE PROT...	+1	757	3.1e-74	1
ptnr:SWISSPROT-ACC:P23270	OLFACTORY RECEPTOR-LIKE PROT...	+1	701	2.7e-68	1
ptnr:SPTREMBL-ACC:Q9Y3P5	DJ994E9.5 (HS6M1-17 (NOVEL 7 ...	+1	688	6.3e-67	1
ptnr:SPTREMBL-ACC:O70271	OLFACTORY RECEPTOR-LIKE PROTE...	+1	688	6.3e-67	1
ptnr:SPTREMBL-ACC:O95007	WUGSC:H DJ0669B10.3 PROTEIN -...	+1	680	4.5e-66	1
ptnr:SPTREMBL-ACC:O13036	CHICK OLFACTORY RECEPTOR 7 - ...	+1	675	1.5e-65	1
ptnr:SPTREMBL-ACC:O95222	OLFACTORY RECEPTOR - Homo sap...	+1	672	3.1e-65	1
ptnr:SPTREMBL-ACC:O70269	OLFACTORY RECEPTOR-LIKE PROTE...	+1	670	5.1e-65	1
ptnr:SPTREMBL-ACC:O57597	CHICK OLFACTORY RECEPTOR 7 - ...	+1	669	6.5e-65	1
ptnr:SPTREMBL-ACC:O70270	OLFACTORY RECEPTOR-LIKE PROTE...	+1	668	8.4e-65	1
ptnr:TREMBLNEW-ACC:AAF65461	OLFACTORY RECEPTOR P2 - Mu...	+1	666	1.4e-64	1
ptnr:SPTREMBL-ACC:Q9WU86	ODORANT RECEPTOR S1 - Mus mus...	+1	665	1.7e-64	1
ptnr:SPTREMBL-ACC:Q90808	OLFACTORY RECEPTOR 4 - Gallus...	+1	665	1.7e-64	1
ptnr:SWISSPROT-ACC:P37071	OLFACTORY RECEPTOR-LIKE PROT...	+1	646	1.8e-62	1

10

(GPCR5c)

別のヌクレオチド配列は、GPCR5a (115-a-12A) がこの配列を確認するためにエキソン連結プロセスに供される場合、生じた。PCRプライマーを、順方向プライマーについて利用可能な最も上流の配列で、そして逆方向プライマーについて利用可能な最も下流の配列で開始することによって設計した。それぞれの場合において、独特であるかまたは高く選択的であるかのいずれかである適切な配列に遭遇したか、逆方向プライマーの場合において、停止コドンに達するまで、それぞれの末端からコード配列に向かって内側に、ウォーキングしてこの配列を試験した。次いで、このような適切な配列を、幅広いcDNAライブラリーに基づいて、PCR増幅における順方向プライマーおよび逆方向プライマーとして使用した。得られたアンプリコンをゲル精製し、クローンし、そして高い冗長性に対して配列決定して、以下に報告された配列（登録番号115__A__12__A__da1、またはGPCR5cとして示される）を提供した。

20

【0115】

30

クロンのSeqCalling Fragmentを提供するヒト組織は、Pool One: 副腎、骨髄、脳-扁桃、脳-小脳、脳-海馬、脳-黒質、脳-視床、脳-全体、胎児脳、胎児腎臓、胎児肝臓、胎児肺、心臓、腎臓、リンパ腫-Raji、乳腺、脾臓、下垂体、胎盤、前立腺、唾液腺、骨格筋、小腸、脊髄、脾臓、胃、精巣、甲状腺、気管、子宮を含む。クロンの組織源は、RACE (asm: 126603384) である。

【0116】

GPCR5c (1005bp、配列番号22) についてのヌクレオチド配列は、表5Gに提示される。GPCR5cヌクレオチド配列は、5'末端においてエキストラT、3'末端においてA、6ヌクレオチドの変化((GPCR5aに関する数) T123>C; C131>T; T186>C; G472>A; T579>A; A687>T)を有することによってGPCR5aと異なる。

40

【0117】

【表43】

表 5G. GPCR5c ナクレオチド配列 (配列番号 22)

TAGACAATGAGTCTGATGGGAACACAGTAGTGATCCAACAGAGTTCGTCTCGGCAGGGCTCCCAATC
TCAACAGCGCAAGAGTGGAAATATTTCTGTGTTTCTTCTGTCTATCTCCGAATCTGATAGGCAATGT
GTTGATTGTGGGGTGGTAAGGGCTGATACTCGACTACAGACCCCATGTACTTCTTTCTGGGTAACCTG
TCCTGCCTAGAGATACTGCTCACTTCTGTATCATTCCAAAGATGCTGAGCAATTTCTCTCAAGGCAAC
ACACTATTTCCTTTGTGTCATGTATCACCAATTCTATTCTACTTCTTTCTCGGGGCTCCGAGTTCTT
ACTGTTGGCTGTCTGTGCGGATCGCTACCTGGCCATCTGTATCCTCTGCGCTACCCCTTGCTCATG
AGTGGGGCTGTGTGCTTTCGTGTGGCTTGGCCTGCTGGGTGGGGGACTCATCCCTGTGCTTGGTCCCA
CAGTGGCTGTGGCCTTGCTTCTTTCTGTAAGCAGGGTGTGTGTACAGCACTTCTTCTGCGACAGTGG
CCCACTGCTCCGCTGGCATGCACCAACACCAAGAAGCTGGAGGAGACTGACTTTGTCTGGCCTCCCTC
GTCATTGTATCTTCTTGTGTCATGCTGTGTCTACGGCCTCATTTGTGCTGGCTGTCTGAGCATCC
CCTCTGCTTTCAGGCCGTGAGAAGGCCTTCTTACCTGTACCTCCCACTTGATAGTGGTGACCTCTTCTA
TGGAGTGCCATTTTCTCTATGTGCGGCCATCGCAGAGTGGTTCTGTGGACACTAACTGGGCAGTGACA
GTAATAACGACATTTGTGACACCACTGTTGAATCCATTCTATGCCTTACGTAATGAGCAAGTCAAGG
AAGCTTTGAAGGACATGTTTAGGAAGGTAGTGGCAGGCGTTTTAGGGAATCTTTACTTGATAAATGTCT
CAGTGAGAAAGCAGTAAAGTAAAA

10

GPCR5cのコード領域は、ヌクレオチド7～1000であり、表5Hに提示されるよ
うなコードされたGPCR5cタンパク質を与える。開示されたタンパク質は、331個
のアミノ酸の長さであり、配列番号23によって示される。GPCR5cは、3アミノ酸
残基(L39>P; T42>I; およびV151>I)、GPCR5aと異なる。GPC
R5aのように、GPCR5cについてのSortプロフィールは、この配列が、シグ
ナルペプチドを有し、そして0.6000の確実度を有して原形質膜に局在化するよう
であること示す。ペプチドについての最も可能性のある切断部位は、アミノ酸54と55と
の間(すなわち、SignalPの結果の基づくアミノ酸配列VRADTにおけるスラ
ッシュ(表5Hにおいてスラッシュで示される))である。

20

【0118】

【表44】

表 5H. クドにある GPCR5b タンパク質配列 (配列番号 23)

MSPDGNHSSDPTEFVLGLPNLNSARVELFVFLVLYLPNLIGNVLIVGVVRA/DTRLQTPMYFFLGNLSC
LEILLTSVITPKMLSNFLSRQHTISEAACITQEFYFFFLGASEFLLAVMSADRYLAICHPLRYFLLMSG
AVCFRVALACWVGGLIPVLGPTVAVALLPFCQKQAVVQHFTCDSGPLRLACTNTKKLEETDFVLASLVI
VSSLLITAVSYGLIVLAVLSIPSASGRQKAFSTCTSHLIVVTLFYGSALFLYVRPSQSGSVDTNWAFTVI
TTFVTPLLPFFIYALRNEQVKEALKDMFRKVVAGVLGNLLLDKCLSEKAVK

30

開示されたGPCR5タンパク質(配列番号19)は、多数の嗅覚レセプタータンパク質
と良好な同一性を有する。ClustalW分析について使用される同一性の情報は、表
5Iに提示される。GPCR5a、GPCR5bまたはGPCR5cとして具体的には取
り組まないが、GPCR5に対する任意の言及は、全ての改変体を含むとみなされる。本
明細書中において任意のGPCR5X改変体配列間の残基の差異は、「a」改変体におけ
る残基および「a」改変体に関する残基位置を示すために書かれる。個々のGPCR改
変体間で異なる全ての以下の配列のアライメントにおけるGPCR残基は、囲み内で強調表
示され、そして本明細書中において改変体残基の上に(o)シンボルでマークされる。例
えば、表5Jのライン1に示されるタンパク質は、GPCR5aについての配列を示し、
そしてGPCR5bまたはGPCR5cが異なる位置は、(o)シンボルでマークされ、
そして囲みとともに強調表示される。全てのGPCR5タンパク質が、嗅覚レセプター(
OR)タンパク質に対して有意な相同性を有する。

40

【0119】

【表45】

表 5I. GPCR5 についての BLAST の結果					
遺伝子番号 / 識別子	タンパク質 / 生物	長さ (aa)	同一性 (%)	陽性 (%)	期待値
GI 129091 sp P23267 OLF6_RAT	OR 15 (OR3) <i>Rattus norvegicus</i>	311	144/301 (47%)	189/301 (61%)	7e-67
GI 6754932 ref NP_0 35121.1 (AF102523)	OR 49, ORC6 <i>Mus musculus</i>	313	145/305 (47%)	192/305 (62%)	3e-66
GI 7242165 ref NP_0 35113.1 (AF106007) (AF321233)	OR 41, ORI7 <i>Mus musculus</i>	234	136/304 (44%)	181/304 (58%)	9e-61
GI 7363437 ref NP_0 39229.1	OR, family 10, subfamily C, member 1 <i>Homo sapiens</i>	306	139/301 (46%)	186/301 (61%)	2e-59
GI 12007431 gb AAG4 5202.1 AF321236_1 (AF321236)	M50 OR <i>Mus musculus</i>	316	129/303 (42%)	181/303 (59%)	4e-59

10

この情報は、GPCR5 を関連のタンパク質配列と比較する ClustalW 分析として表 5 J に与えられる複数配列アライメントにおいてグラフィカルに提示される (GPCR5 がライン 1 に示される)。

【0120】

(表 5 J ClustalW タンパク質のについての情報)

20

【0121】

【表 46】

- 1) GPCR5 (配列番号 19)
- 2) gi|129091|sp|P23267|OLF6_RAT OLFACTORY RECEPTOR-LIKE PROTEIN F6 (配列番号 63)
- 3) gi|6754932|ref|NP_035121.1| olfactory receptor 49 Mus musculus (配列番号 64)
- 4) gi|7242165|ref|NP_035113.1| olfactory receptor 41 Mus musculus (配列番号 65)
- 5) gi|7363437|ref|NP_039229.1| olfactory receptor, family 10, subfamily C, member 1 Homo sapiens (配列番号 66)
- 6) gi|12007431|gb|AAG45202.1|AF321236_1 m50 olfactory receptor Mus musculus (配列番号 67)

	10	20	30	40	50	60
GPCR5						
gi 129091	MAWSTGQNLSTPGPPLGPPGPRSEIGLELEFIMMELTVGQDAISGAHRCQ					
gi 6754932	MAWSTGQNLSTPGPPLGPPGPRSEIGLELEFIMMELTVGQDAISGAHRCQ					
gi 7242165	MAWSTGQNLSTPGPPLGPPGPRSEIGLELEFIMMELTVGQDAISGAHRCQ					
gi 7363437	MAWSTGQNLSTPGPPLGPPGPRSEIGLELEFIMMELTVGQDAISGAHRCQ					
gi 12007431	MAWSTGQNLSTPGPPLGPPGPRSEIGLELEFIMMELTVGQDAISGAHRCQ					
	70	80	90	100	110	120
GPCR5						
gi 129091	PMYFELGHSCELLISNTEKTESNHSRQH---TTFRAALNNTPTFTGASFL					
gi 6754932	PMYFELGHSCELLISNTEKTESNHSRQH---TTFRAALNNTPTFTGASFL					
gi 7242165	PMYFELGHSCELLISNTEKTESNHSRQH---TTFRAALNNTPTFTGASFL					
gi 7363437	PMYFELGHSCELLISNTEKTESNHSRQH---TTFRAALNNTPTFTGASFL					
gi 12007431	PMYFELGHSCELLISNTEKTESNHSRQH---TTFRAALNNTPTFTGASFL					
	130	140	150	160	170	180
GPCR5						
gi 129091	LLAVFSAQRVFAICHLRYDHEGATFVRLACVGLVPTVPTVAVLIPCKQGA					
gi 6754932	LLAVFSAQRVFAICHLRYDHEGATFVRLACVGLVPTVPTVAVLIPCKQGA					
gi 7242165	LLAVFSAQRVFAICHLRYDHEGATFVRLACVGLVPTVPTVAVLIPCKQGA					
gi 7363437	LLAVFSAQRVFAICHLRYDHEGATFVRLACVGLVPTVPTVAVLIPCKQGA					
gi 12007431	LLAVFSAQRVFAICHLRYDHEGATFVRLACVGLVPTVPTVAVLIPCKQGA					
	190	200	210	220	230	240
GPCR5						
gi 129091	LNCHFCVSGEIRKATVKKLLETDLVSLVSSLLAVSYLLVLAHSPSAS					
gi 6754932	LNCHFCVSGEIRKATVKKLLETDLVSLVSSLLAVSYLLVLAHSPSAS					
gi 7242165	LNCHFCVSGEIRKATVKKLLETDLVSLVSSLLAVSYLLVLAHSPSAS					
gi 7363437	LNCHFCVSGEIRKATVKKLLETDLVSLVSSLLAVSYLLVLAHSPSAS					
gi 12007431	LNCHFCVSGEIRKATVKKLLETDLVSLVSSLLAVSYLLVLAHSPSAS					
	250	260	270	280	290	300
GPCR5						
gi 129091	SPRAFTSTTHHIVTETVYSATVAVRISQSGVITVAVATTFVTTHLHFTVA					
gi 6754932	SPRAFTSTTHHIVTETVYSATVAVRISQSGVITVAVATTFVTTHLHFTVA					
gi 7242165	SPRAFTSTTHHIVTETVYSATVAVRISQSGVITVAVATTFVTTHLHFTVA					
gi 7363437	SPRAFTSTTHHIVTETVYSATVAVRISQSGVITVAVATTFVTTHLHFTVA					
gi 12007431	SPRAFTSTTHHIVTETVYSATVAVRISQSGVITVAVATTFVTTHLHFTVA					
	310	320	330			
GPCR5						
gi 129091	LRNDEVAEALDKMFRKVVAGVLGNLLDKCLSEKAVK					
gi 6754932	LRNDEVAEALDKMFRKVVAGVLGNLLDKCLSEKAVK					
gi 7242165	LRNDEVAEALDKMFRKVVAGVLGNLLDKCLSEKAVK					
gi 7363437	LRNDEVAEALDKMFRKVVAGVLGNLLDKCLSEKAVK					
gi 12007431	LRNDEVAEALDKMFRKVVAGVLGNLLDKCLSEKAVK					

10

20

30

GPCR5についてのDOMAINの結果は、Reverse Position Specific BLASTを用いてConserved Domain Database (CDD) から集めた。このBLASTサンプルドメインは、SmartおよびPfamコレクションにおいて見出された。この結果は、統計およびドメイン記載とともに表 40 Kに列挙される。

【0122】

7tm_1 (配列番号45) の残基1~115は、表5KにおいてGPCR5 43-156 ($E=1e-20$) とともに整列される。7tm_1の残基314~377はまた、GPCR5の残基231~293 ($E=2e-05$) と同一性を有する。

【0123】

【表47】

10

20

20

30

40

50

する。

【0125】

(GPCR6)

948ヌクレオチドの開示された新規なGPCR6核酸(6-L-19-Cと呼ばれる)は、表6Aに示される。オープンリーディングは、ヌクレオチド7~9においてATG開始コドンで始まり、そしてヌクレオチド940~941においてTAGで終わる。開始コドンから上流および停止コドンから下流の推定の非翻訳領域は、表6Aにおいて下線を引かれ、そして開始コドンおよび停止コドンは、太字である。

【0126】

【表48】

10

表 6A. GPCR6 ナクレオチド配列 (配列番号 24)
GGAGACATGGGCAAGGAAAACATGCACCACTGTGGCTGAGTTCATTCTCCTTGGACTATCAGATGTCCCTG AGTTGAGAGTCTGCCTCTTCTGCTGTTCCTTCTCATCTATGGAGTCACGTTGTTAGCCAACCTGGGCAT GATTGCACTGATTGAGGTGAGCTCTCGGCTCCACACCCCACTGACTTTTTCTCAGCCACTGTCTCTCT GTAGATTTCTGCTACTCCTCAATAATGTGCCAAAATGTGGCTAATATCTTTAACAAGGACAAAGCCA TCTCCTTCTAGGGTGCATGGTGCAATTCTACTTGTGTTTGCACCTGTGTGGTCACTGAGGTCTTCTGCT GGCCGTGATGGCCTATGACCGCTTTGTGGCCATCTGTAACCCCTTGTCTATACACAGTCACCATGTCTTGG AAGGTGCGTGTGGAGCTGGCTTCTTGCTGCTACTTCTGTGGGACGGTGTGTTCTCTGATTCAATTTGTGCT TAGCTCTTAGGATCCCTTCTATAGATCTAATGTGATTAACCACTTTTCTGTGATCTACCTCTGTCTT AAGTCTTGCTTGTCTGATATCACTGTGAATGAGACACTGCTGTTCTGGTGGCCACTTTGAATGAGAGT GTTACCATCATGATCATCTCACCTCTACCTGCTAATTCACCACTCATCTGAAGATGGGCTCTGCAG AGGGCAGGCACAAAGCCTTCTCCACCTGTGCTTCCCACCTCACAGCTATCACTGTCTTCCATGGAACAGT CCTTTCCATTTATTGTCAGGCCCAAGTTCAGGCAATAGTGGAGATGCTGACAAAGTGGCCACCGTGTCTAC ACAGTCGTGATTCTATGCTGAACCTCTGTGATCTACAGCCTGAGAAATAAAGATGTGAAAGAAGCTCTCA GAAAGTGATGGGCTCCAAAATTCATCTCCTAGGGAAGA

20

開示された核酸配列は、G. gallus cor4嗅覚レセプター4DNA (GENBANK-1D:GGCOR4GEN|acc:X94744)に対する同一性915塩基のうち617(67%)を有する(E値=8.7e⁻⁶⁵)。

【0127】

配列番号24によってコードされるGPCR6タンパク質は、318アミノ酸残基を有し、そして表6(配列番号25)において1文字コードを使用して提示される。GPCR6についてのSignalP、Sortおよび/またはHydropathyプロファイルは、GPCR6がシグナルペプチドを有し、そして0.6000の確実度で、原形質膜に局在化するのである。SignalPは、シグナル配列が最初の41アミノ酸でコードされる(すなわち、アミノ酸40と41との間で配列TLL-ANにおけるスラッシュで切断部位を有する)これは、このタイプの膜タンパク質で代表的である。

30

【0128】

【表49】

表 6B. コドンの GPCR6 アミノ酸配列 (配列番号 25)
MGKENCTTVAEFILLGLSDVPELRVCLFLLFLLIYGVTL/ANLGMIALIQVSSRLHTPMYFFLSHLSSVO FCYSSIIIVPKMLANIFNKDKAISFLGCMVQFYLFCTCVVTEVFLLAVMAYDRFVAICNPLLYTVTMSWKV RVELASCCYFCGTVCSLIHLCLALRIPFYRSNVINHHFCDLPPVLSLACSDITVNETLLFLVATLNEST IMIILTSYLLILTTILKMGSAGEGRHKAFSTCASHLTAITVFHGTVLSIYCRPSSGNSGADQKVATVFYTV VIPMLNSVIYSLRNKDVKEALRKVMGSKIHS

40

本発明のタンパク質の全長アミノ酸配列は、314アミノ酸残基ヒト嗅覚レセプター様タンパク質OLF1(ptnr:SWISSPROT-ACC:Q13606)(E値=5.8e⁻⁸⁶)に対して、307アミノ酸残基のうち同一な166アミノ酸残基(54%)、および307アミノ酸残基のうち陽性の217アミノ酸残基(70%)を有することが見出された。

【0129】

50

以前に同定された GPCR6 標的配列 (6 __ L __ 19 __ C) を、この配列を確認するためにエキソン連結プロセスに供した。PCR プライマーを、順方向プライマーについて利用可能な最も上流の配列で、そして逆方向プライマーについて利用可能な最も下流の配列で開始することによって設計した。それぞれの場合において、独特であるかまたは高く選択的であるかのいずれかである適切な配列に遭遇したか、逆方向プライマーの場合において、停止コドンに達するまで、それぞれの末端からコード配列に向かって内側に、ウォーキングしてこの配列を試験した。

【0130】

この配列をコードする cDNA を、以下のプライマー：C A C T G T G G C T G A G T T C A T T C T C C T T (配列番号 26) および T C T T C C C T A G G A G T G A A T T T T G G A G C (該列番号 27) を使用して、以下のヒト cDNA のプール (副腎、骨髄、脳一扁桃、脳一小脳、脳一海馬、脳一黒質、脳一視床、脳一全体、胎児脳、胎児腎臓、胎児肝臓、胎児肺、心臓、腎臓、リンパ腫-Raji、乳腺、脾臓、下垂体、胎盤、前立腺、唾液腺、骨格筋、小腸、脊髄、脾臓、胃、精巣、甲状腺、気管、子宮) でポリメラーゼ連鎖反応によってクローン化した。プライマーを、本発明の DNA/タンパク質配列の全長または部分 (1 つ以上のエキソン) についてのシリコ (s i l i c o) 予測に基づいて、あるいは密接に関連するヒト配列または他の種由来の配列に対する推定されるエキソンの翻訳された相同性によって設計した。通常、複数のクローンは、この配列を誘導するために配列決定され、次いで、この配列は、Seq Calling プロセスと同様に組み立てられ得た。さらに、配列のトレースを、手動で評価し、そして適切であれば訂正のために編集した。 10 20

【0131】

エキソン結合によって誘導された PCR 産物を、Invitrogen からの pCR2.1 ベクターにクローニングした。細菌クローン 55446::6 __ L __ 19 __ C. 698018. M1 は、Invitrogen からの pCR2.1 ベクターにクローニングしたオープンリーディングフレーム全体をカバーするインサートを有する。

【0132】

通常、得られたアンプリコンをゲル精製し、クローニングし、そして高い重複性まで配列決定した。全てのクローンから得られた配列を、互いに、CuraGen Corporation のデータベースにおける他のフラグメントと、および公共の EST と構成した。構築のための別の成分とのフラグメントおよび EST の同一性の程度が 50 bp にわたって少なくとも 95% である場合には、これらのフラグメントおよび EST を、構築のための成分として含めた。さらに、配列の残りを手動で評価し、そして適切である場合には、補正のために編集した。これらの手順は、以下に報告する配列を提供し、これは、登録番号 CG50383__01 と指定され、これは、アミノ酸もヌクレオチド配列も、GPCR6 (6 __ L __ 19 __ C) と異なる。 30

【0133】

開示される GPCR6-Olfactory Receptor 様タンパク質は、少なくとも以下の組織において発現される：網膜色素上皮の尖微小柔毛、動脈 (大動脈)、基底前脳、脳、パーキットリンパ腫細胞株、脳染、心臓 (房および室)、尾状核、CNS および末梢組織、小脳、小脳皮質、結腸、皮質性神経原性細胞、内皮 (冠状動脈および臍静脈) 細胞、血小板上皮、眼、新生児の眼、前頭皮質、胎児造血細胞、心臓、海馬、視床下部、白血球、肝臓、胎児肝臓、肺、肺リンパ腫細胞株、胎児リンパ組織、成人リンパ組織、MHC II および III を発現する組織、神経組織、髄質、視床下核、卵巣、脾臓、下垂体、胎盤、橋、前立腺、被殻、血清、骨格筋、小腸、平滑筋 (大動脈における冠状動脈) 脊髄、脾臓、胃、舌の味覚受容細胞、精巣、視床、および胸腺組織。この情報は、本発明に含まれた配列の組織供給源 (Seq Calling 供給源、Public EST 供給源、Literature 供給源、および/または RACE 供給源が挙げられるが、これらに限定されない) を決定することによって誘導された。 40

【0134】

開示されるGPCR6タンパク質（配列番号25）は、多数の嗅覚受容タンパク質との良好な同一性を有する。ClustalW分析のために使用した同一性情報を、表6Cに与える。GPCR6タンパク質は、嗅覚受容（OR）タンパク質に対して有意な同一性を有する：

【0135】

【表50】

表 6C. GPCR6 についてのBLAST結果					
遺伝子アクセス/ 識別名	タンパク質/生物	長さ (aa)	同一性 (%)	陽性 (%)	期待値
GI 5729960 ref NP_006628.1	OR fam. 5, subfam. I, mem 1 <i>Homo sapiens</i>	314	154/307 (50%)	200/307 (64%)	2e-72
GI 2495054 sp Q95155 OLF2 CANFA	OR-like prt OLF2 <i>Canis familiaris</i>	311	152/308 (49%)	207/308 (66%)	8e-70
GI 11692519 gb AAG39856.1 AF282271.1 (AF282271)	OR 41, K11 <i>Mus musculus</i>	314	150/308 (48%)	199/308 (63%)	8e-69
GI 3746443 gb AAC63969.1 (AF045577)	OR, OR93ch <i>Pan troglodytes</i>	314	151/306 (49%)	199/306 (64%)	1e-68
GI 3746448 gb AAC63971.1 (AF045580)	OR OR93Gib <i>Hylabates lar</i>	313	149/305 (48%)	198/305 (64%)	3e-68

10

この情報を、関連するタンパク質配列とGPCR6とを比較するClustalW分析として、表6Dに与えられる複数配列アライメントとして、図式的に提供する（GPCR6が1行目に示されている）。

【0136】

【表51】

20

表 6D ClustalW タンパク質についての情報

- 1) GPCR6 (配列番号 25)
- 2) gi|5729960|ref|NP_006628.1| OR, family 5, subfamily 1, member 1 Homo sapiens (配列番号 25)
- 3) gi|2495054|sp|Q95155|OLF2_CANFA OR-LIKE PROTEIN OLF2 (配列番号 25)
- 4) gi|11692519|gb|AAG39856.1|AF282271_1 odorant receptor K11 Mus musculus (配列番号 25)
- 5) gi|3746443|gb|AAC63969.1| olfactory receptor OR93Ch [Pan troglodytes] (配列番号 25)
- 6) gi|3746448|gb|AAC63971.1| olfactory receptor OR93Gib [Hylobates lar] (配列番号 25)

	10	20	30	40	50	60
GPCR6						
gi 5729960	-----	-----	-----	-----	-----	-----
gi 2495054	-----	-----	-----	-----	-----	-----
gi 11692519	-----	-----	-----	-----	-----	-----
gi 3746443	-----	-----	-----	-----	-----	-----
gi 3746448	-----	-----	-----	-----	-----	-----
	70	80	90	100	110	120
GPCR6						
gi 5729960	-----	-----	-----	-----	-----	-----
gi 2495054	-----	-----	-----	-----	-----	-----
gi 11692519	-----	-----	-----	-----	-----	-----
gi 3746443	-----	-----	-----	-----	-----	-----
gi 3746448	-----	-----	-----	-----	-----	-----
	130	140	150	160	170	180
GPCR6						
gi 5729960	-----	-----	-----	-----	-----	-----
gi 2495054	-----	-----	-----	-----	-----	-----
gi 11692519	-----	-----	-----	-----	-----	-----
gi 3746443	-----	-----	-----	-----	-----	-----
gi 3746448	-----	-----	-----	-----	-----	-----
	190	200	210	220	230	240
GPCR6						
gi 5729960	-----	-----	-----	-----	-----	-----
gi 2495054	-----	-----	-----	-----	-----	-----
gi 11692519	-----	-----	-----	-----	-----	-----
gi 3746443	-----	-----	-----	-----	-----	-----
gi 3746448	-----	-----	-----	-----	-----	-----
	250	260	270	280	290	300
GPCR6						
gi 5729960	-----	-----	-----	-----	-----	-----
gi 2495054	-----	-----	-----	-----	-----	-----
gi 11692519	-----	-----	-----	-----	-----	-----
gi 3746443	-----	-----	-----	-----	-----	-----
gi 3746448	-----	-----	-----	-----	-----	-----
	310					
GPCR6						
gi 5729960	-----	-----	-----	-----	-----	-----
gi 2495054	-----	-----	-----	-----	-----	-----
gi 11692519	-----	-----	-----	-----	-----	-----
gi 3746443	-----	-----	-----	-----	-----	-----
gi 3746448	-----	-----	-----	-----	-----	-----

10

20

30

GPCR6における同定可能なドメインの存在を、PROSITE、Blocks、Pfam、ProDomain、Printsのようなアルゴリズムを使用する検索によって決定し、次いで、Interproウェブサイト (<http://www.ebi.ac.uk/interpro/>) を使用してドメインのマッチ (または数) を交差させることによって、Interpro数を決定した。

【0137】

GPCR6についてのDOMAIN結果を、Reverse Position Specific BLASTでのConserved Domain Database (CDD) から収集した。このBLASTは、SmartおよびPfam収集において見出されたドメインをサンプリングする。これらの結果を、統計およびドメインの説明とともに表 6Eに列挙する。これらの結果は、このタンパク質が7tm_1 (InterPro) 7回膜貫通レセプター (ロドプシンファミリー) (InterProによって定義されるような) を、残基42~203 (これは、7TMDドメインの残基2~158と整列する) において含むことを示す。このことは、GPCR6の配列がこの/これらのドメインを含

50

【表 5 2】

[illegible]

10

20

30

40

50

ポリペプチドを免疫原として使用して、本発明に特異的な抗体を産生し得る。そしてワクチンとして使用し得る。これらはまた、潜在的なアゴニスト化合物およびアンタゴニスト化合物についてスクリーニングするために、使用され得る。例えば、OR様タンパク質をコードするcDNAは、遺伝子治療において有用であり得、そしてOR様タンパク質は、投与の必要がある被験体に投与される場合に、有用であり得る。非限定的な例として、本発明の組成物は、細菌感染、真菌感染、原虫感染、およびウイルス感染（特に、HIV-1またはHIV-2によって引き起こされる感染）、疼痛、癌（新生物；腺癌；リンパ腫；前立腺癌；子宮癌が挙げられるがこれらに限定されない）、食欲不振、過食症、喘息、パーキンソン病、急性心不全、低血圧、高血圧、尿停留、骨粗しょう症、クローン病；多発性硬化症；ならびに以下の処置：オールブライト遺伝性骨形成異常症、狭心症、心筋梗塞、潰瘍、喘息、アレルギー、良性前立腺肥大、ならびに精神病性および神経学的な障害（不安、精神分裂病、躁うつ病、せん妄、痴呆、重篤な精神遅滞を含む）、およびジスキネジー（例えば、ハンティングトン病またはジル・ド・ラ・ツレット症候群）ならびに／または他の病理および障害を罹患する患者の処置のために、効力を有する。

10

【0141】

OR様タンパク質をコードする新規核酸、および本発明のOR様タンパク質またはそのフラグメントは、この核酸またはタンパク質の存在または量が評価される診断用途において、さらに有用であり得る。

【0142】

これらの物質は、治療方法または診断方法において使用するための新規なGPCR6基質に免疫特異的に結合する抗体の生成において、さらに有用である。これらの抗体は、以下の「抗GPCRX抗体」の節において記載されるような、疎水性チャートからの予測を使用して、当該分野において公知の方法によって生成し得る。1つの実施形態において、意図されるGPCR6エピトープは、約225～240アミノ酸である。別の実施形態において、GPCR6エピトープは、約255～275アミノ酸である。さらなる実施形態において、GPCR6エピトープは、280～310アミノ酸である。

20

【0143】

（GPCR7）

新規GPCR核酸が、TblastNによって、GPCRプローブおよびホモログに関するCuraGen Corporationの配列ファイルを使用し、そしてGenBankによって利用可能にされるGenomic Daily Filesに対して実行されて、同定された。この核酸をさらに、エキソンの選択を含むプログラムGenScanTMによってさらに予測した。これらを、BLAST検索を使用する類似性によって、さらに改変した。次いで、これらの配列を、明らかな不一致に関して手動で補正し、これによって、全長タンパク質をコードする配列を得た。開示される1013ヌクレオチドの新規GPCR7核酸（dj313i6_Dともまた称される）を、表7Aに示す。オープンリーディングフレームは、ATG開始コドンでヌクレオチド5～7において開始し、そしてTAGコドンでヌクレオチド997～999において終結する。開始コドンの上流および終止コドンの下流の推定の未翻訳領域は、表7Aにおいて下線を引かれており、そして開始コドンおよび終止コドンは、太字である。

30

40

【0144】

【表53】

表 7A. GPCR7 のヌクレオチド配列 (配列番号 28)

TACAATGGAAAGAGCTAACGACAGCACCTTCTCTGGATTATCCTCCTGGGCTTCTCCAACAGGCCTCAGCTGGAAAC
 AGCTCTCTTTGTGGTCATCTTGATCATCTACTTTCTGAGCTTTCTGGGCAATGGCACCATTATACTTTATCCATTGT
 AGATCCTCGCCTCCATACCCCTATGTATTTCTCCTCTCCAATCTCTCTTTATGGATCTTTGTTGACCACCTGTAC
 TGTCCCTCAGACACTGGTCAACTTAAGGGGAAGGACAAGACCATCACCTATGGTGGCTGCGTGACCCAGCTATTCAT
 TGCCTTGGGACTCGGGGGGGAGTGGAGTGTGTCTTATTGTCTGCCATGGCCTATGACCGCTATGCAGCCGTCTGCCG
 CCCACTCCACTACATGGT GAGCATGCATCCCCAACTTTGCTTGCAGTTGGTTGTAACCACTTGGCTCACAGGGTTTGG
 CAATCTGTGATACAGACAGCATTGACCATGACTCTCCCCCTCTGTGATAAAAACCAAGTGGATCATTTCTTCTGTGA
 AGTTCCAGTGATGCTGAAACTGTCCTGCACCAACACCTCCATCAACGAGGCTGAAATCTTTGCTGTCAGTGTCTTCTT
 CTTGGTGGTGCCTCTCTCACTCATCTTAGCATCCTATGGTCACATTACTCATGCAGTCTCTGAAGATAAAGTCAGCTCA
 AGGGAGGCAGAGGCTTTTGGAACTGTGGTTCTCACCTCCTGGTAGTGATCATTTTCTTTGGGACACTCATCTCCAT
 GTACCTCCAGCCTCCCTCCAGTTATTACAGGATGTGAACAAAAGCATTGCACCTCTTCTATACTCTGGTGACTCCTCT
 ACTGAATCCCCTAATTACACTCTGAGGAACAAGGAAGTCAAGGGGCAACTAAGAAGACTAGTGGGAGCACCATAG
 ATGCATGAGAAGTTAACGAGGGTTTGCAGTTCCAAACATTTGTGCACTAGAAGACTGCTGAGAAGCTACAACTA

10

開示された核酸配列は、Homo sapiens 嗅覚レセプター様タンパク質 (OR2C1) 遺伝子 (GENBANK-ID: AF098664) と 939 塩基のうち 615 塩基 (65%) 同一である (E 値 = 1.7×10^{-6})。

【0145】

配列番号 28 にコードされる GPCR7 タンパク質は、327 アミノ酸残基を有し、そして表 7B において 1 文字コードを用いて示される (配列番号 19)。GPCR7 についてのこの Signal P、Sort および/または疎水性親水性指標プロフィールによって、GPCR7 がシグナルペプチドを有すること、そして 0.6000 の確実性で原形質膜に局在するらしいことが予期される。この Signal P は、シグナル配列が、最初の 44 アミノ酸 (アミノ酸 43 と 44 との間) に、すなわち、配列 GNG-TI におけるスラッシュで切断部位とともに、コードされていることを示す。これは、このタイプの膜タンパク質の代表である。

20

【0146】

【表 54】

表 7B. コードされた GPCR7 タンパク質配列 (SEQ ID NO: 29).

MERANDSTFSGFILLGFSNRPQLETALFVILIIYFLSFLNG/TIILLSIVDPRLHTPMYFFLSNL
 SFMDLCLTCTVPQTLVNFKGKDKTITYGGCVTQLFIALGLGGGVECVLLSAMAYDRYAAVCRPLHY
 MVSMPQLCLQLVVTWLTGFGNSVIQTALTMTPLCDKNQVDHFFCEVPVMLKLSCTNTSINEAEI
 FAVSVFFLVVPLSLILASYGHIHAVLKIKAQGRQKAFGTGSHLLVVIIFGTLISMYLQPPSSY
 SQOVNKSIALFYTLVTPLLNPLIYTLRNKEVKGATKKTSGEDHRCMRKLTQGLQFQTFVH

30

本発明のタンパク質の全長アミノ酸配列は、Rattus norvegicus (ptnr: SPTREMBL-ACC: Q63394) 由来の 313 アミノ酸残基の OL1 レセプタータンパク質に対して、304 アミノ酸残基のうち 183 (60%) 同一であり、そして 304 残基のうち 229 残基 (75%) 陽性であることが見出された (E 値 = 2.6×10^{-9})。さらなる BLAST 分析によって、表 7C に列挙される重要な結果を得た。開示された GPCR7 タンパク質 (配列番号 29) は、嗅覚レセプタータンパク質の数と良好な同一性を有する。

40

【0147】

【表 55】

表 7C. GPCR7配列のBLASTの結果

遺伝子名/配列名	タンパク質 名	長さ(残基)	同一性(%)	陽性(%)	E値
gi 11177906 ref NP_068632.1 (L34074)	OR Rattus norvegicus	313	168/304 (55%)	209/304 (68%)	2e-86
gi 10944516 emb CAC14158.1 (AL133267) dJ408B20.2	novel 7 TM -OR family (hS6M1- 32) Homo sapiens	313	169/304 (55%)	207/304 (67%)	2e-85
gi 12054411 emb CAC20513.1 (AJ302593)	OR 41, K11 Homo sapiens	357	166/304 (54%)	206/304 (67%)	1e-83
gi 12054393 emb CAC20504.1 (AJ302584 - 592)	OR Homo sapiens	357	165/304 (54%)	206/304 (67%)	4e-83
gi 3080467 emb CAB11427.1 (Z98744)	OR Homo sapiens	310	165/304 (54%)	206/304 (67%)	4e-83

10

この情報は、関連タンパク質配列とGPCR7を比較するClustalW分析として、表7D（GPCR7は1行目に示されている）において与えられた複数の配列アラインメントにおいて図示されている。

【0148】

表7D. ClustalWタンパク質についての情報：

20

- 1) GPCR7 (配列番号29)
- 2) gi|11177906|ref|NP_068632.1|Olfactory receptor Rattus norvegicus (配列番号73)
- 3) gi|10944516|emb|CAC14158.1|dJ408B20.2 (novel 7 TM receptor (olfactory family) (hS6M1-32) Homosapiens (配列番号74)
- 4) gi|12054411|emb|CAC20513.1|olfactory receptor Homo sapiens (配列番号75)
- 5) gi|12054393|emb|CAC20504.1|olfactory receptor Homo sapiens (配列番号76)
- 6) gi|3080467|emb|CAB11427.1|olfactory receptor Homo sapiens (配列番号77)

30

【0149】

【表56】

	10	20	30	40	50	60
GPCR7						
gi 11177906						
gi 10944516						
gi 12054411						
gi 12054393						
gi 3080467						
	70	80	90	100	110	120
GPCR7						
gi 11177906						
gi 10944516						
gi 12054411						
gi 12054393						
gi 3080467						
	130	140	150	160	170	180
GPCR7						
gi 11177906						
gi 10944516						
gi 12054411						
gi 12054393						
gi 3080467						
	190	200	210	220	230	240
GPCR7						
gi 11177906						
gi 10944516						
gi 12054411						
gi 12054393						
gi 3080467						
	250	260	270	280	290	300
GPCR7						
gi 11177906						
gi 10944516						
gi 12054411						
gi 12054393						
gi 3080467						
	310	320	330	340	350	
GPCR7						
gi 11177906						
gi 10944516						
gi 12054411						
gi 12054393						
gi 3080467						

10

20

GPCR7における同定可能なドメインの存在は、PROSITE、Blocks、Pfam、ProDomain、Printsのようなアルゴリズムを用いて検索すること、
次いで、Interproウェブサイト (<http://www.ebi.ac.uk/interpro/>) を用いてドメイン適合 (または数) を掛け合わせるによってInterpro数を決定することによって決定される。

【0150】

GPCR7のDOMAIN結果を、Reverse Position Specific BLASTを用いたConversed Domain Database (CDD) から収集した。このBLASTサンプルドメインは、SmartおよびPfamコレクション中に見出した。この結果を、統計およびドメインの詳細とともに、表7Eに列挙する。この結果は、このタンパク質が示された位置で以下のタンパク質ドメイン (Interproによって規定されるような) を含むことを示す：ドメイン名7tm_1 (InterPro) 7膜貫通レセプター (ロドプシンファミリー)。7tm_1 (配列番号45) の残基61~142は、表7EにおいてGPCR7 41-180と整列されている (E=5e-18)。7tm_1の残基307~377はまた、GPCR7の残基222~291と同一性を有する (E=0.001)。このことは、GPCR7の配列が、このドメイン/これらのドメインを含むことが公知の他のタンパク質の配列と類似で、かつこれらのドメインの特性と類似の特性を有することを示す。

【0151】

【表57】

10

20

30

30

40

40

、精神分裂病、躁うつ病、せん妄、痴呆、重度の精神遅滞およびジスキネジア（ハンチントン舞踏病またはジル・ド・ラ・ツレット症候群など）を含む）および／または他の病理および障害の処置に有効性を有する。G P C R様タンパク質をコードする新規な核酸および本発明のG P C R様タンパク質、またはそのフラグメントは、診断適用（ここで核酸またはタンパク質の存在または量が評価されるべきである）においてさらに有用であり得る。これらの物質はさらに、治療または診断方法における使用のために、本発明の新規な物質に免疫特異的に結合する抗体を生成するのに有用である。これらの抗体は、以下の「抗G P C R X抗体」の節において記載されるように、疎水性親水性指標チャートからの予想を用いて、当該分野で公知の方法に従って生成され得る。

【0154】

(G P C R 8)

G P C RプローブまたはホモログについてCuraGen社の配列プロフィールを用いてTblastNによって第6染色体上で新規な核酸を同定し、そしてGenBankによって利用可能になされたGenomic Daily Filesに対して実施した。核酸をさらに、エクソンの選択を含む、GenScanTMプログラムによって予想した。これらは、さらにBLAST検索を用いる類似の手段によって改変された。次いで、この配列を明白な不一致について手動で補正し、これによって全長タンパク質をコードする配列を得た。開示された、958ヌクレオチドの新規なG P C R 8核酸（またはdj408b20とも呼ばれる）は、表8Aに示される。オープンリーディングフレームは、ヌクレオチド4～6でATG開始コドンで開始し、そしてヌクレオチド955～957でTGA

【0155】

【表58】

表 8A. GPCR8 ヌクレオチド配列 (SEQ ID NO:30)
GCAATGGAAAAATCCAATGTCAGCTCAGTGTATGGTTTATCTTGGTGGGTTTCTCTGATCGTCCCAAGCTG GAGATGGTGCTCTTTACAGTAAATTTTATTCTGTATTGAGTGGCTGTGCTGGGAAATTCAACCATAATCC TTGTGTGTATATTAGACTCTCAACTTCATACCCCAATGTACTTCTTTCTGGCAAATCTTTCCTTTCTAGA TCTCTGCTTCAGTACTAGTTGCATCCCAAAATGCTGGTAAACCTCTGGGGCCCTGACAAGACTATTAGC TGCTGTGGCTGTGTTGCCAGCTTTTCTCTTTCTTCTGTCAGGGGAATTGAGTGCATCCTTCTGGCTG TCATGGCCTATGACAGCTATGCTGCAGTCTGCAACCGTTGCGCTATCTGGTCATTATGCACCTCCAGCT GTGCTAGGACTGATGGCTGCAGCTGGGGGAGTGGACTGGTCAATGCCGTTGTCATGTCACCACTAACA ATGACCCTCTCCAGAAGTGGCCGCCGCCGAGTTAACCATTTCCTCTGTGAAAAGCCAGCACTGATCAAGA TGCTTTGTTGGATGTTGCTGCAGTGGAAATGCTGGCTTTTGCTTTTGCCGTTCTCATTGTCTACTGCC CCTCACTCTATTCTTGTCTCCTACGGCTACATTGCTGCAGCTGTGCTAAGCATCAAGTCAGCTGCCAGG CAATGGAAGGCCTTCCATACCTGTAGCTCTCACCTCACAGTGGTCTCCCTGTTTATGGGAGCATCATCT ATATGTATATGCAGCCAGGAACAGTTCTTCCCAAGACCAAGGCAAGTTTCTCACTCTTCTACAACT GGTGACTCCTATGTTGAATCTGCTCATCTATACTTTAAGGAATAAGGAGGTGAAAGGAGCACTGAAGAAG GTTTGGGGAGGCAAAATGAACGAGAAATATGATAAGTTGTGAA

開示された核酸配列は、Homo sapiens OR様（第6p21.31-21.33染色体上のクローン88J8由来のgb:GENBANK-ID:HS88J8|acc:AL035402 Human DNA配列）。これは、新規な7膜貫通ウイルスレセプター（ロドプシンファミリー）（嗅覚レセプター様）タンパク質の遺伝子、嗅覚レセプター遺伝子類似の偽遺伝子、およびGTP結合タンパク質SARA（マウス）偽遺伝子を含む。これはEST、STSおよびGSS、完全な配列-Homo sapiens、4216塩基を含む）と1148塩基のうち768塩基（66%）同一である（E値=1.3e⁻⁸¹）。

【0156】

配列番号30にコードされるG P C R 8タンパク質は、317アミノ酸残基を有し、そして表8Bにおいて1文字コードを用いて示される（配列番号31）。G P C R 8についてのこのSignalP、Sortおよび／または疎水性親水性指標プロフィールによって、G P C R 8がシグナルペプチドを有すること、そして0.6850の確実性で原形質

10

20

30

40

50

膜に局在するらしいことが予期される。この Signal P は、シグナル配列が、最初の 42 アミノ酸（アミノ酸 41 と 42 との間）にすなわち、配列 V L G - N S におけるスラッシュで切断部位とともに、コードされていることを示す。これは、このタイプの膜タンパク質の代表である。

【0157】

【表59】

表8B. コードされたGPCR8タンパク質配列(配列番号31)

```
MEKSNVSSVYGFILVGFSDREPKLEMLFTVNFILYSVAVLG/NSTIILVCILDSQLHTPMYFFLANLSF
LDLCFSTSCIPQMLVNLWGPDKTISCAGCVVQLFSFLSVRGIECILLAVMAYDSYAAVCKPLRYLVIMH
LQLCLGLMAAAWGSGLVNAVVMSPMTLSRSGRRRVNHFLCEKPALIKMACLDVRAVELAFAFAVLI
VLLPLTLILVSYGYIAAAVLSIKSAARQWKAFHTCSSHLTVVSLFYGSIIYMQPGNSSSQDQKFLT
LFYNLVTPLMLNLLIYTLRNKEVKGALKVLGRQNELEKYDKL
```

10

本発明のタンパク質の全長アミノ酸配列が、ヒト由来の 320 アミノ酸残基の新規の膜貫通レセプター（ロドプシンファミリー、OR 様、H S 6 M 1 - 1 5）タンパク質（p t n r : S P T R E M B L - A C C : Q 9 Y 3 N 9）と、305 アミノ酸残基のうち 187 アミノ酸残基が同一（61%）であり、また 305 個のアミノ酸残基のうち 239 の残基が陽性（78%）であることが見出された（E 値 = $2.4 \times e^{-101}$ ）。さらに、B L A S T 解析が、表 8 C に示されるような有意な結果を生成した。開示された G P C R 8 タン

20

【0158】

【表60】

表8C. GPCR8についてのBLASTの結果

遺伝子アクセッション/識別名	タンパク質/生物種	長さ (aa)	同一性 (%)	陽性 (%)	期待値
Gi 4826521 emb CAB42853.1 (AL035402, dJ8878.1) (AJ302594-99) (AJ302600-01)	novel 7 tm receptor protein (rhodopsin fam., OR-like) (hs6M1-15) <i>Homo sapiens</i>	320	177/305 (58%)	226/305 (74%)	5e-92
Gi 12054431 emb CAC20523.1 (AJ302603)	OR <i>Homo sapiens</i>	320	176/305 (57%)	226/305 (73%)	1e-91
Gi 12054429 emb CAC20522.1 (AJ302602)	OR <i>Homo sapiens</i>	320	177/305 (58%)	225/305 (73%)	1e-91
Gi 6679170 ref NP_032788.1	OR 15 (OR3) <i>Mus musculus</i>	312	166/307 (54%)	211/307 (68%)	2e-81
Gi 12231029 sp Q15062 O2H3 HUMAN	OR 2H3 <i>Homo sapiens</i>	316	163/306 (53%)	208/306 (67%)	5e-81

30

この情報は、GPCR8を関連するタンパク質と比較する Clustal W 解析として表 8 D（第 1 行目に GPCR8 が示されている）で与えられたマルチプルアライメントで図式的に表現される。

40

【0159】

【表61】

表8D. ClustalW proteinについての情報

- 1) GPCR8 (配列番号: 31)
- 2) gi|4826521|emb|CAB42853.1| dJ88J8.1 (novel 7 TM receptor (rhodopsin family) (OR like) protein) (hs6M1-15)) Homo sapiens (配列番号: 59)
- 3) gi|12054431|emb|CAC20523.1| olfactory receptor Homo sapiens (配列番号: 60)
- 4) gi|12054429|emb|CAC20522.1| olfactory receptor Homo sapiens (配列番号: 61)
- 5) gi|6679170|ref|NP_032788.1| olfactory receptor 15 Mus musculus (配列番号: 58)
- 6) gi|12231029|sp|L5062|O2H3_HUMAN OR 2H3 (OR-LIKE PRT FAT11) (配列番号: 56)

	10	20	30	40	50	60
GPCR8						
gi 4826521						
gi 12054431						
gi 12054429						
gi 6679170						
gi 12231029						
	70	80	90	100	110	120
GPCR8						
gi 4826521						
gi 12054431						
gi 12054429						
gi 6679170						
gi 12231029						
	130	140	150	160	170	180
GPCR8						
gi 4826521						
gi 12054431						
gi 12054429						
gi 6679170						
gi 12231029						
	190	200	210	220	230	240
GPCR8						
gi 4826521						
gi 12054431						
gi 12054429						
gi 6679170						
gi 12231029						
	250	260	270	280	290	300
GPCR8						
gi 4826521						
gi 12054431						
gi 12054429						
gi 6679170						
gi 12231029						
	310	320				
GPCR8						
gi 4826521						
gi 12054431						
gi 12054429						
gi 6679170						
gi 12231029						

10

20

30

GPCR8において同一視可能なドメインの存在が、PROSITE、Blocks、Pfam、ProDomain、Printsのようなアルゴリズムを用いたサーチによって決定され、ついでInterproウェブ (<http://www.ebi.ac.uk/interpro/>) を使用してドメインマッチを組合せてInterpro番号を決定する。

40

【0160】

GPCR8に対するDOMAINの結果を、Conserved Domain Database (CDD) からReverse Position Specific BLASTを用いて収集した。このBLASTは、SmartおよびPfamのコレクション中に見出されるドメインをサンプルリングする。この結果は、表8Eに統計値およびドメインの説明と共に列挙される。この結果は、アミノ酸残基位置41~170で、GPCR8が7tm_1 (InterPro) 7回膜貫通レセプター (ロドプシンファミリー) ドメイン (Interproによって定義される) を含むことを示している。これは、GPCR8の配列が、これまたこれらのドメインを含む既知の他のタンパク質の配列と類似の特

50

【 0 1 6 1 】

【表 6 2】

gnl|Pfam|pfam00001, 7tm_1, 7 transmembrane receptor (rhodopsin family) Length = 377 (SEQ ID NO:45).

Score = 85.9 bits (211), Expect = 3e-18

GPCR8
TM7

GPCR8 LFSFISURGIECLAVMYYSRAUCKPERLVIM-HLCLCLGEPAAAGSGLVNVVM
TM7 LTLDIMCTASILNCAEIERRTAVAMPMLNTRYSSKRVRTVME-IVVVLSTISCP

GPCR8 SPL-TM1LSRSG-----
TM7 LFGLNNIDONE-----CIIANP-----AFVYSSIVSEYV---PFIVTLLVYI

本明細書中で開示されたG P C R 8タンパク質および核酸についてのこの類似性情報が、G P C R 8の嗅覚レセプターファミリーおよびG P C Rファミリーの重要な構造的特徴および／または生理学的機能の特徴を有していることを示唆している。従って、本発明の核酸およびタンパク質は、潜在的な診断および治療の適用において有用であり、そしてリサーチツールとして有用である。これらは、特定または選択的な核酸またはタンパク質の診断マーカーおよび／または予後マーカーとして役割を含み（ここでこの核酸またはこのタンパク質の存在または量が評価され）ならびに、例えば以下のような潜在的な治療的適用を含む：（i）タンパク質治療剤（i i）低分子薬物標的（i i i）抗体標的（治療学薬物標的化／細胞傷害性抗体、診断的薬物標的化／細胞傷害性抗体）および（i v）遺伝子治療（遺伝子送達／遺伝子アブレーション（a b l a t i o n））において有用な核酸（v）インビトロおよびインビボにおける組織再生を促進する組成物（v i）生物学的防衛手段。本発明のG P C R 8をコードする新規の核酸、および本発明のG P C R 8タンパク質、またはこれらのフラグメントは、さらに診断的適用において有用であり、ここでこの核酸またはこのタンパク質の量は、見積もられ得る。

【 0 1 6 2 】

開示された G P C R 8 核酸および本発明のタンパク質は、以下に関連する潜在的な治療学的応用において有用である：新生物；腺癌；リンパ腫；前立腺癌；子宮癌；免疫応答；A I D S ；喘息；クローン病；多発性硬化症；ならびにオールブライト遺伝性骨形成異常ならびに／または他の病理および障害の処置。例えば、G P C R 様タンパク質をコードする c D N A は、遺伝子治療において有用であり、そして G P C R 様タンパク質は、これらを必要としている被験体に投与する場合に有用であり得る。非限定的な例として、本発明の組成物は、以下に苦しむ患者の処置について効力を有する：新生物；腺癌；リンパ腫；前立腺癌；子宮癌；免疫応答；A I D S ；喘息；クローン病；多発性硬化症；およびオールブライト遺伝性骨形成異常の処置。G P C R 様タンパク質、および本発明の G P C R タンパク質またはこれらのフラグメントをコードする新規の核酸が、さらに診断的応用において有用であり、ここでこの核酸またはこのタンパク質の量は、評価される。これらの物質は、治療的方法または診断的方法において、免疫特異的に新規の G P C R 8 基質に結合する抗体産生においてさらに有用である。これらの抗体は、以下の「抗 G P C R X 抗体」の節に記載のように、疎水性チャートから予測を用いて、当該分野に公知の方法に従って生成され得る。

【 0 1 6 3 】

(G P C R 9)

新規の核酸を、GPCRプローブまたはGPCRホモログに対するCuraGen Corporationの配列ファイル使用して、Genbankにより提供されたGeno

me Daily Files に対して実行した TblastN により、第 11 染色体上に同定した。この核酸は、エキソン選択を含むプログラム GenScanTM によってさらに予測された。これらは、さらに BLAST 検索を使用する相同性の手段によって改変される。この配列は、次いで手動で明らかな不一致に対して修正し、これによって全長タンパクをコードする配列を得た。この開示された 946 ヌクレオチドからなる新規の GPCR9 核酸 (6-L-19-B として称される) は、表 9A 中に示される。読み枠は、ヌクレオチド 5~7 における ATG 開始コドンで始まり、ヌクレオチド 932~934 の TAA コドンで終止する。開始コドンより上流にある推定非翻訳領域および終止コドンより下流の推定非翻訳領域は、表 9A で下線により示し、開始コドンおよび終止コドンは太字で示した。

10

【0164】

配列データベースのサーチにおいて、例えば、この核酸配列が、902 塩基中に、嗅覚レセプター A16 (GENBANK-ID: GENBANK-ID: AB030896 | acc: AB030896) についての Mus musculus 遺伝子に対して同一 ($E = 8.5 \times 10^{-47}$) な 563 塩基 (62%) を有する。配列データベースのサーチにおいて、部分的な一致 (373 塩基中 353 塩基、つまり 94% 同一) がまた、Homo sapiens GPCR EST (GENBANK-ID: GENBANK-ID: AW182678 | acc: AW182678; xj45d11. x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE: 26601813' similar to TR: Q9Z1V0 Q9Z1V0 OLFACTORY RECEPTOR C6) の 306~678 ヌクレオチドと同一な GPCR9 の 1~372 ヌクレオチドと同定された。公開された配列の領域と本発明 (遺伝子) の領域の間のこの 94% の一致 ($E = 1.1 \times 10^{-69}$) は、本発明 (遺伝子) が公開された GPCR EST (部分的 mRNA) のスプライス改変体で有り得ることを示唆している。これはまた、GPCR としての GPCR9 の同定を支持している。配列データベースのサーチにおいて、Homo sapiens GPCR EST (GENBANK-ID: AA206680 | acc: AA206680; zq51c11. r1 Stragene neuroepithelium (#937231) Homo sapiens cDNA clone IMAGE: 645140 5' similar to contains L1. b2 L1 repetitive element) の 251~348 のヌクレオチドとの GPCR9 の 893~794 ヌクレオチドの部分が一致している (100 塩基中 94 塩基、94% 同一)。公開された配列のヌクレオチドと GPCR9 配列のヌクレオチドとのこの 94% の一致は、GPCR9 が GPCR EST (mRNA の部分) のスプライス改変体であり得ることを示唆している。

20

30

【0165】

【表 63】

表9A. GPCR9ヌクレオチド配列(配列番号32)
<u>G</u>TTCATGGAAATAGGAATATTGTCACCTGCTTTATTCTCCTGGGACTTTCTCAAACCAAGAACATTGAA GTTTTTTGGTTTGTATTATTTGTATTTTGCTACATTGCTATTGGATGGAAAACCTCATCATAATGATT CTATCATGTACATTGGCTAATTGACCAACCCATGTAATTTCTCCTTAATTACCTCGCACTCTCAGATCT TTGCTACATATCCACTGTGGCCCCAAGCTAATGATTGACCTACTAACAGAAAGGAAGATCGTTTCCTAT AATAACTGCATGTACAGCTATTATTCACCTCACTTCCTTGGAGACATTGAGATCTTCATACTCAAAGCAA TGGCCTATGACCACTACATAGCCATCTGCAAGCACCTGCACTACACCATCATCACGACCAAGCAAAGCTG TAACACCATCATCATAGCTTGTGTACTGGGGGATTATACACTCTGCCAGTCAGTTTCTTCTTACCATC TTCTTACCGTTCTGTGGTCTTAATGAGATAGATCAGTACTTCTGCTATGTGTATCCTCTGCTGAAGTTGG CTCGCATTGATATATACAGAATTGGTTTCTTGGTAATTGTTAATTCAAGGCTGATTCTTCTTGTGGCTTT TGTGATTTTGATGGTGTCTTATTATTTGATATTATCCACCATCAGGTTTACTCTGCTGAGAGTCATACC AAAGCTCTTCAACCTGTAGCTCTCACATAATAGTTTGGTCTTATCTTGTGCCTGCCCTCTTCATTT ACATCAGACCAAGCATAAAGTTTCCAGAAGATAAAGTGTGTTCTCTTCTGTCGCATCATTGCTCCCAT GTTCACTCTTCTATCTACATGCTGAGAAAGGTGGAGATGAAGAAGCTGTAAGGAAATGTGGTGTCTCAT CAATTGCTCTGGCAAGGAAGTAAGTGTATGAAAG

40

配列番号 32 にコードされている GPCR9 タンパク質は、309 アミノ酸残基を有して

50

おり、そして表9B（配列番号33）中に1文字コードを用いて表現されている。GPCR9についてのSignalP、Psortおよび/または疎水性プロファイルは、GPCR9がシグナルペプチドを有しており、小胞体膜に局在する傾向（0.6850の確実性を有する）があることを予測している（これに対し、原形質膜に対しては0.6850の確実性を有して局在する）。SignalPは、表9B中のスラッシュが示すようなアミノ酸38およびアミノ酸39の間、IWM-E N配列に切断部位を予測した。

【0166】

【表64】

表9B. コードされたGPCR9タンパク質配列(配列番号33)

MENRNIVTVFILLGLSQNKNIIEVFWFVLFVFCYIAIWM/ENFIIMISIMYIWLIDQPMYFFLNYLALSOLC
YISTVAPKLMIDLTERKIVSYNNCMIQLFITFLGDIIEFYLKAMAYDHYIAICKHLHYTIITTKQSCN
TIIACCTGGFIHSASQFLITIFLPCGLNEIDQYFCYVYPLLKLARIDIYRIGFLVIVNSGLISLLAFV
ILMVSYLLSTIRVYSAESHTKALSTCSSHIIIVVLFVFPALFIYIRPAITFPEDKVFVLFCAIAPME
SLLIYMLRKVMKNNAVRKMWCHQLLLARK

10

本発明のタンパク質の全長アミノ酸配列が、Mus musculus由来の嗅覚レセプターA16の302アミノ酸（ptnr:TREMBLNEW-ACC:BAA86127）と、300アミノ酸残基のうちの140残基が同一（46%）であり、300残基中193残基（64%）が陽性であることが見出された（E値=1.0e^{-7.2}）。さらなるBLAST解析が、表9Cに列举される有意な結果を生み出した。開示されたGPCR8タンパク質（配列番号33）は、多くの嗅覚レセプタータンパク質と十分な同一性を有している。

20

【0167】

【表65】

表9C. GPCR9に対するBLAST結果

遺伝子名/タンパク質名	タンパク質/生物体	長さ (aa)	同一性 (%)	陽性 (%)	期待値
GI 11496249 ref NP_067343.1 (AB030895)	嗅覚レセプター MOR18 Mus musculus	308	140/300 (46%)	185/300 (61%)	1e-59
GI 11464995 ref NP_065261.1 (AB030896)	嗅覚レセプター A16 Mus musculus	302	137/300 (45%)	185/300 (61%)	2e-59
GI 423702 pir S29710	嗅覚レセプター OR18 -	307	142/303 (46%)	186/303 (60%)	6e-57
GI 11464993 ref NP_065260.1 (AB030894)	嗅覚レセプター MOR83 Mus musculus	308	133/297 (44%)	182/297 (60%)	7e-55
GI 10644519 gb AAG21324.1 AF271051_1 (AF271051)	嗅覚レセプター Mus musculus	264	124/262 (47%)	169/262 (64%)	6e-51

30

この情報は、関連タンパク質配列とGPCR9を比較するClustalW分析として表9D（ライン1に示されるGPCR9と共に）に示される複数の配列整列において図として示される。

40

【0168】

【表66】

表. 9D. *Clusta/W* タンパク質の対の相対的:

- 1) GPCR9 (配列番号:33)
- 2) gi|11496249|ref|NP_067343.1| odorant receptor 16 *Mus musculus* (配列番号 :78)
- 3) gi|11464995|ref|NP_065261.1| gene for odorant receptor A16 *Mus musculus* (配列番号 :79)
- 4) gi|423702|pir|S29710 olfactory receptor OR18 - rat (配列番号 :80)
- 5) gi|11464993|ref|NP_065260.1| gene for odorant receptor MOR33 *Mus musculus* (配列番号 :81)
- 6) gi|10644519|gb|AA021324.1|AF271051_1 odorant receptor *Mus musculus* (配列番号 :70)

	10	20	30	40	50	60
GPCR9						
gi 11496249	ENRNITVDELLGLSKKMEVTFVFCDAIMWQNTPLKSAQIWLADCPN					
gi 11464995	DEIPNHTDEETLGLSKKMEVTFVFCDAIMWQNTPLKSAQIWLADCPN					
gi 423702	DSPRHTDEETLGLSKKMEVTFVFCDAIMWQNTPLKSAQIWLADCPN					
gi 11464993	GCNHTVDEETLGLSKKMEVTFVFCDAIMWQNTPLKSAQIWLADCPN					
gi 10644519	MGLNCTVDEETLGLSKKMEVTFVFCDAIMWQNTPLKSAQIWLADCPN					
	70	80	90	100	110	120
GPCR9						
gi 11496249	FFINVTLSGLSLDITPAKNIIDETPRSLTAMNVIDETITLQISHTLLMAY					
gi 11464995	FFINVTLSGLSLDITPAKNIIDETPRSLTAMNVIDETITLQISHTLLMAY					
gi 423702	FFINVTLSGLSLDITPAKNIIDETPRSLTAMNVIDETITLQISHTLLMAY					
gi 11464993	FFINVTLSGLSLDITPAKNIIDETPRSLTAMNVIDETITLQISHTLLMAY					
gi 10644519	FFINVTLSGLSLDITPAKNIIDETPRSLTAMNVIDETITLQISHTLLMAY					
	130	140	150	160	170	180
GPCR9						
gi 11496249	DRYVAICRPLHYTTTHHIVAVVGGGTHSTATITVFCDECCGNVAVVQI					
gi 11464995	DRYVAICRPLHYTTTHHIVAVVGGGTHSTATITVFCDECCGNVAVVQI					
gi 423702	DRYVAICRPLHYTTTHHIVAVVGGGTHSTATITVFCDECCGNVAVVQI					
gi 11464993	DRYVAICRPLHYTTTHHIVAVVGGGTHSTATITVFCDECCGNVAVVQI					
gi 10644519	DRYVAICRPLHYTTTHHIVAVVGGGTHSTATITVFCDECCGNVAVVQI					
	190	200	210	220	230	240
GPCR9						
gi 11496249	MDPLNLRRIITRECHVNSCHTSIALVIRVSVYELISLSEVFAASHKALSTC					
gi 11464995	MDPLNLRRIITRECHVNSCHTSIALVIRVSVYELISLSEVFAASHKALSTC					
gi 423702	MDPLNLRRIITRECHVNSCHTSIALVIRVSVYELISLSEVFAASHKALSTC					
gi 11464993	MDPLNLRRIITRECHVNSCHTSIALVIRVSVYELISLSEVFAASHKALSTC					
gi 10644519	MDPLNLRRIITRECHVNSCHTSIALVIRVSVYELISLSEVFAASHKALSTC					
	250	260	270	280	290	300
GPCR9						
gi 11496249	SSHTFVVLSEVFAASHKALSTC					
gi 11464995	SSHTFVVLSEVFAASHKALSTC					
gi 423702	SSHTFVVLSEVFAASHKALSTC					
gi 11464993	SSHTFVVLSEVFAASHKALSTC					
gi 10644519	SSHTFVVLSEVFAASHKALSTC					
	310					
GPCR9						
gi 11496249	SCCHOLLARK					
gi 11464995	SCCHOLLARK					
gi 423702	SCCHOLLARK					
gi 11464993	SCCHOLLARK					
gi 10644519	SCCHOLLARK					

10

20

30

GPCR9において同定可能なドメインの存在は、PROSITE、Blocks、Pfam、ProDomain、Printsといったアルゴリズムを使用する検索、次いで、Interproウェブサイト (<http://www.ebi.ac.uk/interpro/>) を使用するドメインマッチ (またはナンバー) を交差することによってInterproナンバーを決定することによって決定された。

【0169】

GPCR9に関するDOMAIN結果は、反位特異的BLASTを有する保存ドメインデータベース (CDD) から収集された。このBLASTサンプルドメインは、SmartおよびPfamコレクションにおいて見出した。この結果は、統計およびドメインの詳細と共に表9Eに列挙される。この結果は、GPCR9が、アミノ酸残基の56位~234位において7m_1 (InterPro) 7膜貫通レセプター (ロドプシンファミリー) ドメイン (InterProによって規定される) を含む。これは、GPCR9の配列が、これ/これらのドメインを含むことが公知の他のタンパク質の配列と類似の性質およびこれらのドメインの性質と類似を有することを示す。

50

症候群および／または他の病理学および傷害))を患う患者の処置に効果を有する。G P C R様タンパク質をコードする新規の核酸、本発明のG P C R様タンパク質またはそれらのフラグメントは、診断的適用においてさらに有用であり、ここで、核酸またはタンパク質の存在または量は、評価される。

【0173】

これらの材料は、治療方法または診断方法における使用のための新規のG R C R 9基質に免疫特異的に結合する抗体の産生においてさらに有用である。これらの抗体は、以下のセクションの「抗G P C R X抗体」に記載されるように、疎水性チャートからの推定を使用して、当該分野において公知の方法に従って産生され得る。

【0174】

(G R C R 10)

G R C R 10は、以下に開示される3つの類似の核酸および3つの類似のタンパク質のファミリーを含む。開示される核酸は、G P C R、O R様タンパク質をコードする。

【0175】

(G P C R 10 a)

開示される新規の核酸は、G P C Rプローブまたは相同体に関するCuraGen Corporationの配列ファイル(GenBankによって利用可能になったGenomic Daily Fileに対して駆動する)を使用するTblastNによって染色体11にあることが同定された。この核酸は、プログラムGenScanTM(エキソンの選択を含む)によってさらに推定される。これらは、BLAST検索を使用して類似度によってさらに改変された。次いで、この配列は、明白な不一致に関して手動で回収され、それによって全長をコードする配列を得た。G P C R 10 aは、表10Aに示されるように(または6-L-19-Aと称される)948bp長の核酸である(配列番号34)。O R Fは、ヌクレオチド7~9においてATG開始コドンで始まり、ヌクレオチド934~936におけるTAAコドンで終わる。開始コドンから上流および終止コドンから下流の推定非転写領域を表10Aにおいて下線を引いて示し、そして開始コドンおよび終止コドンを、太文字で示す。

【0176】

【表68】

表. 10A. GPCR10aヌクレオチド配列 (配列番号 34)
TGAGAAATGGAAAATCAAAACAATGTGACTGAATTCATTCTTCTGGGTCTCACAGAGAACCTGGAGCTGT GGAAAATATTTTCTGCTGTGTTTCTTGTGATGTATGTAGCCACAGTGTCTGGAAAATCTACTTATTGTGGT AACTATTATCACAAGTCAGAGTCTGAGGTCACTATGTATTTTCTTACCTTCTTGTCCCTTTGGAT GTCATGTTCTCATCTGTGCTGTCGTTGCCCAAGGTGATTGTAGACACCTCTCCAAGAGCACTACCATCTCTC TCAAGGCTGCTCCTCAGCCAGCTGTTTGTGGAGCATTCTTTGGTGGTGTGGGGATCATCTCTCTCACTGT GATGGCTATGACCCGCTACGTGGCCATCTGTAAGCCCTGCACTACACGATCATCATGAGTCCACGGGTG TGCTGCTAATGGTAGGAGGGGCTTGGTGGGGGATTTATGCACGCAATGATACAACTTCTCTTCATGT ATCAATACCTTCTGTGGTCTAATATCATAGATCACTTTATATGTGATTTGTTTCAGTTGTTGACACT TGCCTGCACGGACACCCACATCCTGGGCTCTTAGTTACCTCAACAGTGGGATGATGTGTGTGGCCATC TTCTTATCTTAATTGCGTCTACACGGTCATCCTATGCTCCCTGAAGTCTTACAGCTCTAAGGGCGGC ACAAAGCCCTCTCTACCTGCAGCTCCACCTCACGGTGGTTGTATTGTTCTTTGTCCCTGTATTTCTT GTACATGAGGCTGTGGTCACTACCCCATAGACAAGGCAATGGCTGTGTGAGACTCAATCATCACCCC ATGTTAAATCCCTTGATCTATACACTGAGGAATGCAGAGGTGAAAAGTCCATGAAGAACTCTGGATGA AATGGGAGGCTTTGGCTGGGAAATACTGCAATGCTGA

配列番号34によってコードされるG P C R 10 aタンパク質が、309アミノ酸残基を有し、そしてこれを、表10Bにおいて1文字コドンを使用して表す(配列番号35)。G P C R 10 aに関するSignalP、Sortおよび／または疎水性プロファイルは、G P C R 10 aが、シグナルペプチドを有し、そして0.6000の確実度で形質膜に局在するようだと推定する。SignalPは、表10Bにおいて斜線によって示される、配列VLE-NL(アミノ酸39とアミノ酸40との間)における切断部位を推定する。

【0177】

配列データベースの検索において、例えば、核酸配列が、臭気物質レセプター A 1 6 m R N A (G E N B A N K - I D : A B 0 3 0 8 9 6) に関する 9 0 9 b p M u s m u s c u l u s 遺伝子と 9 0 8 塩基のうち 6 2 5 の同一性 (6 8 %) を有する (E = 3 . 0 e - 7 5) ことを見出す。

【 0 1 7 8 】

【 表 6 9 】

表 10B. コードである GPCR10a タンパク質配列 (配列番号 35).
MENQNVTEFILLGLTENLELWKIFSAVFLVMYVATVLE/NLLIVVTIITSQSLRSPMYFFLTFSLLDVM FSSVYAPKVIQVDTLSKSTTISLKGCLTQLVEHFPGVGIIILLTVMAYDRYVAICKPLHYTIIMSFRVCC LMVGGAUVGGFMHAMIQLLFMYPFCGPNIIDHFICDLFQLLTACTDTHILGLLVTLNSGMMCVAFI ILIASYTVILCSLKSYSKGRHKALSTCSSHLTVVVLFFVPCIFLYMRPVVTHPIDKAMAVSDSIITPML NPLIYTLRNAEVKSAMKRLNMKEALAGK

10

本発明のタンパク質の全アミノ酸は、R a t t u s s p . (p t n r : T R E M B L N E W A C C : G 2 6 4 6 1 8) 由来の 3 0 7 アミノ酸残基 O R 1 8 臭気物質レセプタータンパク質に対して 3 0 2 アミノ酸残基のうち 1 8 3 の同一性 (6 0 %) 、そして 3 0 2 残基のうち 2 3 2 の陽性 (7 6 %) を有することを見出した。

【 0 1 7 9 】

(G P C R 1 0 b)

G P C R 1 0 a (6 - L - 1 9 - A) を、配列を確認するためにエクソン連結プロセスに供した。P C R プライマーを、前向きプライマーについては利用可能な最も上流の配列で開始し、そして逆向きのプライマーについては利用可能な最も下流の配列で開始することによって設計した。各場合に、この配列を、唯一であるかまたは高度に選択的であるかのいずれかである適切な配列に遭遇するまで、各末端からコード配列に内部に向かってウォーキングして試験するかあるいは、逆向きのプライマーの場合、終止コドンに達するまでウォーキングして試験した。次いで、このような適切な配列を、広範な c D N A ライブラリーに基づいた P C R 増幅において前向きおよび逆向きのプライマーとして使用した。生じたアンプリコンを、ゲル精製し、クローニングし、そして G P C R 1 0 b (6 - L - 1 9 - A 1 とまたいわれる) を提供する高度な重複に配列決定した。

20

【 0 1 8 0 】

G P C R 1 0 b (9 4 8 b p 、配列番号 3 6) についてのヌクレオチド配列を、表 1 0 C に示す。このヌクレオチド配列は、G P C R 1 0 a とは、1 ヌクレオチド変化 ((G P C R 1 0 a に対して番号を付し) T 4 0 4 > C) によって異なる。

30

【 0 1 8 1 】

【 表 7 0 】

表 10C. GPCR10b ヌクレオチド配列 (配列番号 36)
TGAGAAATGGAATCAAAACAAATGTGACTGAATTCATTCTTCTGGGTCTCACAGAACCTGGAGCTGTGGAATATT TTCTGCTGCTTTCTGTCTATGTATGAGCCACAGTGTGGAAATCTACTTATTGTGGTAATATTATCACAAGTCAGA GTCTGAGGTACACCTATGTATTTTCTTACCTTCTTGTCCCTTTTGGATGTCATGTTCTCATCTGTGTTGCCCAAG GTGATTGTAGACACCTCTCCAGAGCACTACCATCTCTCTCAAAGGCTGCCTACCCAGCTGTTTGTGGAGCAATTTCTT TGGTGGTGTGGGGATCATCTCTCACTGTGATGGCCTATGACCGCTACGTGGCCATCTGTAGCCCTGCATACAGCA TCACCATGAGTCCACGGGTGTGCTGCCTAATGGTATGAGGGGCTTGGGTGGGGGATTTATGCACGCAATGATACAACTT CTCTCATGTATCAAAATACCCCTTCTGTGGTCTAATATCATAGATCACTTTATATGTGATTGTTTTCAGTTGTGACACT TGCCCTGCACGGACACCCACATCTGGGCTCTTAGTTACCCCTCAACAGTGGGATGATGTGTGTGGCCATCTTCTTATCT TAATTGCGTCTACACGGTCATCCTATGCTCCCTGAAGTCTTACAGCTCTAAAGGGGGGACAAAGCCCTCTTACCTGC AGCTCCACCTCACGGTGGTGTATGTTCTTTGTCCCTGTATTTCTTGTACATGAGGCTGTGGTCACTCACCCCAT AGACAAGGCAATGGCTGTGTGAGCTCAATCATCACCCCATGTTAAATCCCTGTATCTATCACTGAGGAATGCAGAGG TGAAGTGCATGAAGAACTCTGGATCAATGGGAGGCTTTGGCTGGGAAATCACTGCAATGCTGA

40

コードされた G P C R 1 0 b タンパク質を、表 1 0 D に示す。この開示されたタンパク質は、3 0 9 アミノ酸の長さであり、配列番号 3 7 によって示される。G P C R 1 0 b は、G P C R 1 0 a とは、1 アミノ酸残基 (I 1 3 3 > T) によって異なる。G P C R 1 0 a のように、G P C R 1 0 b についての P 分類 (P s o r t) プロフィールは、この配列が

50

単一のペプチドであり、0.6000の確実性でおそらく細胞膜に位置することを予知する。ペプチドについて最もありえる分裂部位は、SignalPの結果に基づいて、アミノ酸39および40の間、すなわち、アミノ酸配列VLE-NLにおけるスラッシュの位置である（表10D中にスラッシュとして示される）。

【0182】

【表71】

表10D. コードされたGPCR10bタンパク質配列(配列番号37)
MENQNNVTEFILLGLTENLELWKIFSAVFLVMYVATVLE/NLLIVVTIITSQSLRSPMYFFLTFLSLLD VFMSSVVAPKVIQDTLSKSTTISLKGCLTQLFEVHFFGGVGIIILLTMAYDRYVAICKPLHYTTITMSPR VCCLMVGGAWVGGFMHAMIQLLFMYQIPFCGPNIIDHFICDLFQLLTACTDTHILGLLVTLNSGMMCV AIFLILIASYTVILCSLSKSYSSKGRHAKALSTCSSHLTVVVLFFVPCIFLYMRPVVTHPIDKAMAVSDSI ITPMLNPLIYTLRNAEVKSAMKKLWMKWEALAGK

10

表10Eは、GPCR10bについてのBLASTP結果を示す。

【0183】

【表72】

表10E. GPCR10bについてのBLASTP結果
スコア=999(351.7%)、期待値=1.2e-100、P=1.2e-100 同-性=182/302(60%)、陽性=231/302(76%) (PIR-ID:S29710 嗅覚レセプター-OR18-ラットと) スコア=757(266.5%)、期待値=5.2e-75、P=5.2e-75 同-性=144/298(48%)、陽性=200/298(67%) (ACC:O95013 WUGSC:H_DJ0855D21.1 タンパク質-Homo sapiens (ヒト)と)、312アミノ酸 スコア=757(266.5%)、期待値=5.2e-75、P=5.2e-75 同-性=144/298(48%)、陽性=200/298(67%) (ACC:O95013 WUGSC:H_DJ0855D21.1 タンパク質-Homo sapiens (ヒト)と)、312アミノ酸 スコア=667(234.8%)、期待値=1.1e-64、P=1.1e-64 同-性=131/300(43%)、陽性=194/300(64%)、フレーム=+1 (ACC:O43749 嗅覚レセプター-Homo sapiens (ヒト)と)、312アミノ酸

20

(GPCR10c)

GPCR10a(6-L-19-A)を配列を確認するためにエクソン連結プロセスに供した場合、別のヌクレオチド配列が生じた。PCRプライマーを、前向きプライマーについては利用可能な最も上流の配列で開始し、逆向きのプライマーについては利用可能な最も下流の配列で開始することによって設計した。各場合に、この配列を、唯一であるかまたは高度に選択的であるかのいずれかである適切な配列に遭遇するまで、各末端からコード配列に内部に向かってウォーキングして試験するかあるいは、逆向きのプライマーの場合、終止コドンに達するまでウォーキングして試験した。次いで、このような適切な配列を、広範なcDNAライブラリーに基づいたPCR増幅において前向きおよび逆向きのプライマーとして使用した。

30

【0184】

次いで、これらのプライマーを、以下のヒトcDNAプールに基づいたPCR増幅において使用した：副腎、骨髄、脳-扁桃、脳-小脳、脳-海馬、脳-黒質、脳-視床、脳-全体、胎児脳、胎児腎臓、胎児肝臓、胎児肺、心臓、腎臓、リンパ腫-ラージ、乳腺、脾臓、下垂体、胎盤、前立腺、唾液腺、骨格筋、小腸、脊髄、脾臓、胃、精巣、甲状腺、気管、子宮。通常、生じるアンプリコンをゲル精製し、クローニングし、高い重複性に配列決定した。全てのクローンから生じる配列を、それら自身で、CuraGen Corporationのデータベース中の他のフラグメントと、そして公開ESTとアセンブルさせた。フラグメントおよびESTを、アセンブリの別の成分とのこれらの同一性の程度が50bpを越えて少なくとも95%である場合、アセンブリのための成分としてインキュベートした。さらに、配列の痕跡を、手動で評価し、適切な場合、補正のために編集した。生じるアンプリコンをゲル精製し、クローニングし、以下に報告する配列を提供する高

40

50

い重複性に配列決定した。これは、登録番号 6-L-19-A-d a l または G P C R 1 0 c と名付けられる。

【0185】

G P C R 1 0 c (943bp、配列番号38) についてのヌクレオチド配列を、表10Fに示す。G P C R 1 0 c ヌクレオチド配列は、G P C R 1 0 a とは、5'末端で6少ないヌクレオチドならびに2ヌクレオチドの変化((G P C R 1 0 a について番号を付し) G 466>A および C 834>T) を有することによって異なる。

【0186】

【表73】

表10F. GPCR10c ヌクレオチド配列(配列番号38)
ATGGAAATCAAAACAATGTGACTGAATTCATTCTTCTGGGTCTCACAGAGAACCTGGAGCTGTGAAAA TATTTCTGCTGTGTTCTTGTGTCATGTATGTAGCCACAGTGTGGAAATCTACTTATTGTGGTAACTAT TATCACAGTCAGAGTCTGAGGTACCTATGTATTTTCTTACCTTCTGTCCCTTTTGGATGTCATG TTCTCATCTGTGTTGCCCAAGGTGATTGTAGACACCTCTCCAAGAGCACTACCATCTCTCTCAAAG GCTGCCCTACCCAGCTGTTTGTGGAGCATTTCTTGGTGGTGTGGGGATCATCCTCCTCACTGTGATGGC CTATGACCGCTACGTGGCCATCTGTAAGCCCTGCACTACAGATCATCATGAGTCCACGGGTGTGCTGC CTAATGCTAGGAGGGGCTTGGGTGGGGGATTATGCACACATGATACAACTTCTCTTCATGTATCAAA TACCCTTCTGTGGTCCCTAATATCATAGATCACTTTATATGTGATTGTTTCAGTTGTTGACACTGCGCTG CACGGACACCCACATCCTGGGCTCTTAGTTACCTCAACAGTGGGATGATGTGTGTGGCCATCTTTCTT ATCTTAATTGCGTCTACACGGTCACTCTATGCTCCCTGAAGTCTTACAGCTCTAAAGGCGGCACAAAG CCCTCTCTACCTGCAGTCCACCTCACGGTGGTGTGATTGTTCTTTGTCCCCTGATTTTCTTGTACAT GAGGCTGTGGTCACTACCCCATAGACAGGCAATGGCTGTGTGAGACTCAATCATTACCCCATGTTA AATCCCTTGATCTATACACTGAGGAATGCAGAGGTGAAAAGTGCCATGAAGAACTCTGGATGAATGGG AGGCTTTGGCTGGGAAATAACTGCAATGCTGA

10

20

G P C R 1 0 c のコード領域は、1~928ヌクレオチドであり、コードされた G P C R 1 0 c タンパク質を与え、表10Gに示される。この開示されたタンパク質は、309アミノ酸の長さであり、配列番号83に示される。G P C R 1 0 c は、G P C R 1 0 a とは1アミノ酸残基(A154>T)によって異なる。G P C R 1 0 a のように、G P C R 5 c についてのP分類プロフィールは、この配列が単一のペプチドであり、0.6000の確実性でおそらく細胞膜に位置することを予知する。ペプチドについて最もありえる分裂部位は、SignalPの結果に基づいて、アミノ酸39および40の間、すなわち、アミノ酸配列VLE-NLにおけるスラッシュの位置である(表10G中にスラッシュとして示される)。

30

【0187】

【表74】

表10G. コードされたGPCR10cタンパク質配列(配列番号83)
MENQNVTEFILLGLTENLELWKIFSAVFLVMYVATVLE/NLLIVVTIITSQSLRSPMYFFLTFLSLLDVM FSSVAPKVIYDLSKSTTISLKGCLTQLFVEHFFGGVGIILLTMAYDRYVAICKPLHYTIIMS PRVCC LMVGGAWVGGFMHTMIQLLFMYQIPFCGPNIDHFICDLFQLLTACTDTHILGLLVTLNSGMMCVAFIL ILIASYTVILCSIKSYSSKGRHKALSTCSSHLTVVVLFFVPCIFLYMRPVVTHPIOKAMAVSDSIITPML NPLIYTLRNEVKSAMKKLWMKWEALAGK

40

G P C R 1 0 について見出される可能な S N P を、表10Hに記載した。

【0188】

【表75】

表 10H: SNPs		
塩基の位置	前の塩基	後の塩基
65	T	A(2)
120	T	Gap(2)
147	T	C(2)
234	A	G(3)
412	T	C(7)
471	G	A(2)
814	A	G(3)

10

この開示された GPCR10 タンパク質（配列番号 35）は、多くの嗅覚レセプタータンパク質に良好な同一性を有する。ClustalW 分析について使用されるこの同一性情報を、表 10 I に示す。GPCR10 a、GPCR10 b または GPCR10 c として具体的にいわない場合、GPCR10 に対する任意の言及は、全ての改変体を含むとみなす。本明細書中の任意の GPCR X 改変体配列の間で異なる残基を、「a」改変体中の残基および「a」改変体に対する残基位置を示すために書く。個々の GPCR 改変体の間で異なる、以下の全ての配列アラインメント中の GPCR 残基を、四角で強調し、そして本明細書中の全てのアラインメントの改変体の残基の上に（o）記号でマークする。例えば、表 10 J の 1 行目に示されるタンパク質は、GPCR10 a についての配列を示し、そして GPCR10 b または GPCR10 c が異なる位置を（o）記号でマークし、そして四角で強調した。全ての GPCR10 タンパク質は、嗅覚レセプター（OR）タンパク質に有意な相同性を有する：

20

【0189】

【表 76】

表 10 I. GPCR10 についての BLAST 結果					
送付インデックス/識別名	タンパク質/生体	長さ (アミノ酸)	同一性 (%)	陽性 (%)	期待値
GI 11496249 ref NP_067343.1 (AB030895)	嗅覚物質 レセプター MOR18 <i>Mus musculus</i>	308	184/306 (60%)	241/306 (78%)	6e-91
GI 423702 pir IS29710	OR OR18 - rat	307	183/302 (60%)	232/302 (76%)	2e-88
GI 11464995 ref NP_065261.1 (AB030896)	嗅覚物質 レセプター A16 <i>Mus musculus</i>	302	175/302 (57%)	232/302 (75%)	8e-86
GI 11464993 ref NP_065260.1 (AB030894)	嗅覚物質 レセプター MOR83 <i>Mus musculus</i>	308	157/297 (52%)	208/297 (69%)	3e-72
GI 10644517 gb AA021323.1 (AF271050_1 (AF271050))	嗅覚物質 レセプター <i>Rattus norvegicus</i>	264	155/260 (59%)	202/260 (77%)	2e-71

30

40

この情報を、GPCR10 を関連するタンパク質配列と比較する ClustalW 分析として、表 10 J に与えられる複数の配列アラインメントに図式的に示す（GPCR10 を 1 行目に示して）。

【0190】

【表 77】

表10J. Clustal Wタンパク質についての情報

- 1) GPCR10 (配列番号35)
- 2) gi|11496249|ref|NP_067343.1| odorant receptor 16 *Mus musculus* (配列番号78)
- 3) gi|423702|pir|S29710| olfactory receptor OR18 - rat (配列番号80)
- 4) gi|11464995|ref|NP_065261.1| gene for odorant receptor A16 *Mus musculus* (配列番号79)
- 5) gi|11464993|ref|NP_065260.1| gene for odorant receptor MOR83 *Mus musculus* (配列番号81)
- 6) gi|10644517|gb|AAG21323.1|AF271050_1 odorant receptor *Rattus norvegicus* (配列番号82)

	10	20	30	40	50	60
GPCR10	ENOM	ETETFLGLGLENLEWKE	FSAN	GLVLA	EA	ELIUVITLTCSTERSFHY
gi 11496249	ETETFLGLGLENLEWKE	FSAN	GLVLA	EA	ELIUVITLTCSTERSFHY	
gi 423702	ETETFLGLGLENLEWKE	FSAN	GLVLA	EA	ELIUVITLTCSTERSFHY	
gi 11464995	ETETFLGLGLENLEWKE	FSAN	GLVLA	EA	ELIUVITLTCSTERSFHY	
gi 11464993	ETETFLGLGLENLEWKE	FSAN	GLVLA	EA	ELIUVITLTCSTERSFHY	
gi 10644517	ETETFLGLGLENLEWKE	FSAN	GLVLA	EA	ELIUVITLTCSTERSFHY	
	70	80	90	100	110	120
GPCR10	FELSL	LSLEVM	SEV	APN	VEV	TSRST
gi 11496249	FELSL	LSLEVM	SEV	APN	VEV	TSRST
gi 423702	FELSL	LSLEVM	SEV	APN	VEV	TSRST
gi 11464995	FELSL	LSLEVM	SEV	APN	VEV	TSRST
gi 11464993	FELSL	LSLEVM	SEV	APN	VEV	TSRST
gi 10644517	FELSL	LSLEVM	SEV	APN	VEV	TSRST
	130	140	150	160	170	180
GPCR10	DRYVAIC	KPLHY	IT	TH	PH	ET
gi 11496249	DRYVAIC	KPLHY	IT	TH	PH	ET
gi 423702	DRYVAIC	KPLHY	IT	TH	PH	ET
gi 11464995	DRYVAIC	KPLHY	IT	TH	PH	ET
gi 11464993	DRYVAIC	KPLHY	IT	TH	PH	ET
gi 10644517	DRYVAIC	KPLHY	IT	TH	PH	ET
	190	200	210	220	230	240
GPCR10	FGL	EL	ACT	TH	EL	GL
gi 11496249	FGL	EL	ACT	TH	EL	GL
gi 423702	FGL	EL	ACT	TH	EL	GL
gi 11464995	FGL	EL	ACT	TH	EL	GL
gi 11464993	FGL	EL	ACT	TH	EL	GL
gi 10644517	FGL	EL	ACT	TH	EL	GL
	250	260	270	280	290	300
GPCR10	EL	TL	VL	VE	FC	EL
gi 11496249	EL	TL	VL	VE	FC	EL
gi 423702	EL	TL	VL	VE	FC	EL
gi 11464995	EL	TL	VL	VE	FC	EL
gi 11464993	EL	TL	VL	VE	FC	EL
gi 10644517	EL	TL	VL	VE	FC	EL
	310					
GPCR10	ET	SE	V	G	A	L
gi 11496249	ET	SE	V	G	A	L
gi 423702	ET	SE	V	G	A	L
gi 11464995	ET	SE	V	G	A	L
gi 11464993	ET	SE	V	G	A	L
gi 10644517	ET	SE	V	G	A	L

10

20

30

GPCR10における同定可能なドメインの存在を、PROSITE、Blocks、Pfam、ProDomain、Printsのようなアルゴリズムを使用した研究によって決定し、次いで、Interproのウェブサイト (<http://www.ebi.ac.uk/interpro/>) を使用してドメインの一致 (または数) を交差させることによってInterpro数を決定する。

【0191】

GPCR10についてのDOMAINの結果を、Conserved Domain Database (CDD) からReverse Position Specific BLASTによって収集した。このBLASTは、SmartおよびPfam収集に見出されるドメインをサンプリングする。この結果を、統計およびドメインの説明と共に表10Kに記載する。これらの結果は、このタンパク質が以下のタンパク質ドメイン (アミノ酸位置39~213でドメイン名7tm_1 (InterPro) 7回膜貫通レセプター (ロドプシンファミリー) (Interproによって規定されるように) を示された位

40

50

【表 7 8】

スコア = 93.6 $r_{\text{スコア}}(231)$, 期待値 = $1e-20$

```
GPCR10 VAIFGLI-----  
TM7 KIYIEPRRRKRUNTKRSSRAFRANLKAPIKGNC TH |PEDMKICTVIMKSNNGSEFPVNRRRV
```

10

20

30

40

50

物；線癌；リンパ種；前立腺癌；子宮癌を含むがこれらに限定されない）、食欲不振、過食症、喘息、パーキンソン病、急性心不全、低血圧、高血圧、尿停滞、骨粗鬆症、クローン病、多発性硬化症を被っている患者の処置について、そしてオールブライト遺伝性骨形成異常症、狭心症、心筋梗塞、潰瘍、喘息、アレルギー、良性前立腺肥大、ならびに精神病性傷害および神経学的傷害（不安、精神分裂病、躁うつ病、せん妄、痴呆、深刻な精神遅滞を含む）ならびに運動障害（ハンチントン病またはジル・ド・ラ・ツレット症候群）ならびに／または他の病理学および傷害の処置について効力を有する。これらの材料はさらに、治療または診断的方法における使用のための新規 G P C R 10 物質に免疫特異的に結合する抗体の産生において有用である。これらの抗体を、以下の「抗 G P C R X 抗体」の節で記載するように、疎水性チャートからの予知を使用して当該分野で公知の方法に従って産生し得る。

10

【0194】

本発明の G P C R X 核酸およびタンパク質のまとめを、表 11 に示す。

【0195】

【表 79】

表 11: 本発明の核酸およびタンパク質のまとめ

名前	表	クローン; ホモログの記載	核酸配列番号	タンパク配列番号
GPCR1	1A, 1B, 1D, 1E, 1G, 1H	GPCR1a: ba113a10_B, 嗅覚レセプター	1	2
		GPCR1b: ba32713_A, 嗅覚レセプター	3	4
		GPCR1c: ba113a10_C, 嗅覚レセプター	5	6
GPCR2	2A, 2B	GPCR2: 11612531_L, 嗅覚レセプター	7	8
GPCR3	3A, 3B	GPCR3: ba145L22_B, 嗅覚レセプター	9	10
GPCR4	4A, 4B, 4C, 4G, 4H	GPCR4a1: dj408b20_C, 嗅覚レセプター	11	12
		GPCR4a2: dj408b20_C_dal, 嗅覚レセプター	13	17
		GPCR4a3: CG55358_03, 嗅覚レセプター	16	
GPCR5	5A, 5B, 5C, 5D, 5G, 5H	GPCR5a1: 115-a-12-A, 嗅覚レセプター	18	19
		GPCR5a2: 115-a-12-B, 嗅覚レセプター	20	21
		GPCR5a3: 115-a-12-A_dal, 嗅覚レセプター	22	23
GPCR6	6A, 6B	GPCR6: 6-L-19-C, 嗅覚レセプター	24	25
GPCR7	7A, 7B	GPCR7: dj313i6_D, 嗅覚レセプター	28	29
GPCR8	8A, 8B	GPCR8: dj408b20_A, 嗅覚レセプター	30	31
GPCR9	9A, 9B	GPCR9: 6-L-19-B, 嗅覚レセプター	32	33
GPCR10	10A, 10B, 10C, 10D, 10F, 10G	GPCR10a: 6-L-19-A, 嗅覚レセプター	34	35
		GPCR10b: 6-L-19-A1, 嗅覚レセプター	36	37
		GPCR10c: 6-L-19-A_dal, 嗅覚レセプター	38	83

20

30

(G P C R X 核酸およびポリペプチド)

本発明の 1 つの局面は、G P C R X ポリペプチドまたはその生物学的に活性なタンパク質をコードする、単離された核酸分子に関する。また、G P C R X をコードする核酸（例えば、G P C R X m R N A）を同定するためのハイブリダイゼーションプローブとしての使用について十分な核酸フラグメント、および G P C R X 核酸分子の増幅および／または変異のための P C R プライマーとしての使用のためのフラグメントもまた、本発明に含まれる。本明細書中で使用される場合、用語「核酸分子」は、D N A 分子（例えば、c D N A またはゲノム D N A）、R N A 分子（例えば、m R N A）、ヌクレオチドアナログを使用して作成される D N A または R N A のアナログ、ならびにそれらの誘導體、フラグメントおよびホモログを含むことを意図する。核酸分子は一本鎖または二本鎖であり得るが、好ましくは二本鎖 D N A である。

40

【0196】

G P C R X 核酸は、成熟 G P C R X ポリペプチドをコードし得る。本明細書中で使用される場合、本発明において開示されるポリペプチドの「成熟」形態またはタンパク質は、天然に存在するポリペプチドまたは前駆体形態またはプロプロテイン (p r o p r o t e i n) の生成をいう。天然に存在するポリペプチドまたは前駆体形態またはプロプロテインとしては、非限定的な例の目的のみとして、対応する遺伝子によりコードされる全長遺伝

50

子産物が挙げられる。あるいは、これらは、本明細書中に開示されるオープンリーディングフレームによりコードされるポリペプチド、前駆体またはプロプロテインとして規定され得る。生成物「成熟」形態は、また非限定的な例の目的として、1以上の天然に存在するプロセシング工程（これらがこの遺伝子産物が生じる細胞または宿主細胞内で起こる場合）の結果として生じる。ポリペプチドまたはタンパク質の「成熟」形態に至るこのようなプロセシング工程の例としては、オープンリーディングフレームの開始コドンによりコードされるN末端メチオニン残基の切断、またはシグナルペプチドまたはリーダー配列のタンパク質分解切断が挙げられる。従って、残基1～N（残基1がN末端メチオニンである場合）を有する前駆体ポリペプチドまたはタンパク質から生じる成熟形態は、N末端メチオニンの除去後に残基2～Nを有する。あるいは、残基1～N（ここで、N末端シグナル配列（残基1～M）が切断される）を有する前駆体ポリペプチドまたはタンパク質から生じる成熟形態は、残りの残基M+1～残基Nを有する。さらに、本明細書中で使用される場合、ポリペプチドまたはタンパク質の「成熟」形態は、タンパク質分解事象ではなく、翻訳後修飾の工程から生じ得る。このようなさらなるプロセスとしては、非限定的な例として、グリコシル化、ミリスチル化またはリン酸化が挙げられる。一般に、成熟ポリペプチドまたはタンパク質は、これらのプロセスの1つのみまたはこれらの任意の組み合わせの操作から生じ得る。

10

【0197】

用語「プローブ」とは、本明細書中で使用される場合、種々の長さの核酸配列をいい、好ましくは、特定の用途に依存して、少なくとも約10ヌクレオチド（nt）、100nt、または例えば、約6,000ntほどの大きさの間である。プローブは、同一、類似または相補的な核酸配列の検出において使用される。より長いプローブは、通常、天然供給源または組換え供給源から入手され、非常に特異的であり、そしてより長さの短いオリゴマープローブよりもはるかに遅くハイブリダイズする。プローブは、一本鎖または二本鎖であり得、そしてPCR、膜ベースのハイブリダイゼーション技術、またはELISA様の技術において特異性を有するように設計される。

20

【0198】

用語「単離された」核酸分子は、本明細書中で使用される場合、この核酸の天然の供給源中に存在する他の核酸分子から分離された核酸分子である。好ましくは、「単離された」核酸は、この核酸が由来する生物のゲノムDNA中でこの核酸に天然に隣接する配列（すなわち、この核酸の5'末端および3'末端に位置する配列）を含まない。例えば、種々の実施形態において、単離されたGPCRX核酸分子は、この核酸が由来する細胞／組織（例えば、脳、心臓、肝臓、脾臓など）のゲノムDNA中でこの核酸に天然で隣接する、約5kb、4kb、3kb、2kb、1kb、0.5kb、または0.1kb未満のヌクレオチド配列を含み得る。さらに、「単離された」核酸分子、例えば、cDNA分子は、組換え技術により産生される場合、その他の細胞物質または培養培地を実質的に含まないか、または化学的に合成される場合、化学物質前駆体もしくはその他の化学物質を実質的に含まないものであり得る。

30

【0199】

本発明の核酸分子、例えば、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38のヌクレオチド配列を有する核酸分子、または上記のヌクレオチド配列の相補体は、標準的な分子生物学的技法および本明細書で提供される配列情報を用いて単離され得る。ハイブリダイゼーションプローブとして、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38の核酸の全部または一部を使用して、GPCRX分子は、標準的なハイブリダイゼーションおよびクローニング技術（例えば、Sambrookら（編）、MOLECULAR CLONING: A LABORATORY MANUAL第2版、Cold Spring Harbor Laboratory Press、Cold Spring Harbor、NY、1989；およびAusubelら、（編）、CURRENT PROTOCOLS IN MOLECULA

40

50

R B I O L O G Y、J o h n W i l e y & S o n s、N e w Y o r k、N Y、1 9 9 3に記載されるように)を用いて単離され得る。

【0200】

本発明の核酸は、標準的なPCR増幅技法に従って、テンプレートとしてcDNA、mRNAあるいはゲノムDNA、および適切なオリゴヌクレオチドプライマーを用いて増幅され得る。このように増幅された核酸は、適切なベクター中にクローン化され、そしてDNA配列分析により特徴付けられ得る。さらに、GPCRXヌクレオチド配列に対応するオリゴヌクレオチドは、標準的な合成技術、例えば、自動化DNA合成機を用いることにより調製され得る。

【0201】

本明細書で用いられる場合、用語「オリゴヌクレオチド」は、一連の連結されたヌクレオチド残基をいい、このオリゴヌクレオチドは、PCR反応で用いられるに十分な数のヌクレオチド塩基を有する。短いオリゴヌクレオチド配列は、ゲノム配列もしくはcDNA配列に基づき得るか、またはそれから設計され得、そして特定の細胞もしくは組織において、同一、類似もしくは相補的DNAまたはRNAを増幅し、確認し、もしくはその存在を示すために用いられる。オリゴヌクレオチドは、約10nt、50nt、または100ntの長さ、好ましくは約15nt~30ntの長さを有する核酸配列の部分を含む。本発明の1つの実施形態では、100ntより少ない長さの核酸分子を含むオリゴヌクレオチドは、さらに、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38の少なくとも6個連続するヌクレオチド、またはその相補体を含む。オリゴヌクレオチドは、化学的に合成され得、そしてプローブとして用いられ得る。

【0202】

別の実施形態では、本発明の単離された核酸分子は、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38に示されるヌクレオチド配列またはこのヌクレオチド配列の一部(例えば、プローブまたはプライマーとして使用され得るフラグメント、あるいはGPCRXポリペプチドの生物学的に活性な部分をコードするフラグメント)の相補体である核酸分子を含む。配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38に示されるヌクレオチド配列に相補的である核酸分子は、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38に示されるヌクレオチド配列に十分相補的である核酸分子であり、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38に示されるヌクレオチド配列に対しミスマッチがほとんどまたは全くなく水素結合し得、それによって安定な二本鎖を形成する核酸分子である。

【0203】

本明細書で用いる場合、用語「相補的」は、核酸分子のヌクレオチド単位間のWatson-CrickまたはHoogsteen塩基対形成をいい、そして用語「結合」は、2つのポリペプチドまたは化合物または関連ポリペプチドまたは化合物またはその組合せ間の物理的または化学的相互作用を意味する。結合は、イオン性、非イオン性、ファンデルワールス性、疎水性相互作用などを含む。物理的相互作用は直接的であるかまたは間接的であり得る。間接的相互作用は、別のポリペプチドまたは化合物を介するか、またはその効果に起因し得る。直接的結合は、別のポリペプチドもしくは化合物を介して、またはその効果に起因して生じない、その他の実質的な化学的中間体がない相互作用をいう。

【0204】

本明細書中に提供されるフラグメントは、少なくとも6個の(連続する)核酸配列または少なくとも4個の(連続する)アミノ酸配列(それぞれ、核酸の場合には、特異的なハイブリダイゼーションを可能にし、またはアミノ酸の場合には、エピトープの特異的な認識を可能にするに十分な長さ)として規定され、そして全長配列未満のせいぜいいくらかの

10

20

30

40

50

部分である。フラグメントは、選択された核酸配列またはアミノ酸配列の任意の連続する部分に由来し得る。誘導体は、直接的にかまたは改変もしくは部分的置換によってか、のいずれかでネイティブな化合物から形成される核酸配列またはアミノ酸配列である。アナログは、ネイティブな化合物に類似する構造を有する（しかし、同一ではない）が、特定の成分または側鎖に関してそれとは異なる核酸配列またはアミノ酸配列である。アナログは、合成物であり得るか、または進化的に異なる起源に由来し得、そして野生型と比較して、類似の代謝活性または反対の代謝活性を有し得る。ホモログは、異なる種由来の特定の遺伝子の核酸配列またはアミノ酸配列である。

【0205】

誘導体およびアナログは、以下に記載のように、誘導体またはアナログが、修飾された核酸またはアミノ酸を含む場合、全長、または全長以外のものであり得る。本発明の核酸またはタンパク質の誘導体またはアナログは、種々の実施形態において、本発明の核酸もしくはタンパク質に、同一の大きさの核酸またはアミノ酸配列にわたって、または整列した配列に比較した場合（この整列は、当該分野で公知であるコンピューター相同性プログラムによって実施される）、少なくとも約30%、50%、70%、80%、または95%の同一性（好ましい同一性は、80~95%）で実質的に相同であるか、あるいはそのコードする核酸がストリンジェントの、中程度にストリンジェントの、または低いストリンジェントの条件下で上記のタンパク質をコードする配列の相補体にハイブリダイズし得る、領域を含む分子を包含するが、これらに限定されない。例えば、Ausubelら、CURRENT PROTOCOLS IN MOLECULAR BIOLOGY、John Wiley & Sons、New York、NY、1993、および以下を参照のこと。

【0206】

「相同な核酸配列」もしくは「相同なアミノ酸配列」、またはその改変体とは、上記で考察したように、ヌクレオチドレベルまたはアミノ酸レベルにおける相同性によって特徴付けられる配列をいう。相同なヌクレオチド配列は、GPCRXポリペプチドのアイソフォームをコードする配列をコードする。アイソフォームは、例えば、RNAの選択的スプライシングの結果として、同一の生物の異なる組織において発現され得る。あるいは、アイソフォームは、異なる遺伝子によってコードされ得る。本発明において、相同なヌクレオチド配列は、ヒト以外の種（脊椎動物が挙げられるが、これに限定されず、従って、例えば、カエル、マウス、ラット、ウサギ、イヌ、ネコ、ウシ、ウマおよび他の生物が挙げられ得る）のGPCRXポリペプチドをコードするヌクレオチド配列を含む。相同なヌクレオチド配列はまた、天然に生じる対立遺伝子改変体および本明細書に記載されるヌクレオチド配列の変異体が挙げられるが、これらに限定されない。しかし、相同的ヌクレオチド配列は、ヒトGPCRXタンパク質をコードする正確なヌクレオチド配列を含まない。相同核酸配列は、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83の保存的アミノ酸置換（以下を参照のこと）、およびGPCRX生物学的活性を保有するポリペプチドをコードする核酸配列を含む。GPCRXタンパク質の種々の生物学的活性は以下に記載されている。

【0207】

GPCRXポリペプチドは、GPCRX核酸のオープンリーディングフレームによりコードされる。ORFは、ポリペプチドに潜在的に翻訳され得るヌクレオチド配列に対応する。ORFを含む核酸のストレッチは、終止コドンによって途切れない。全タンパク質についてのコード配列を表すORFは、ATG「開始」コドンで始まり、そして3つの「終止」コドン（すなわち、TAA、TAG、またはTGA）のうちの1つで終結する。本発明の目的で、ORFは、開始コドン、終止コドンまたはその両方を有するかまたは有さない、コード配列の任意の部分であり得る。bonafide細胞タンパク質のコードのための良好な候補としてみなされるORFについて、最小サイズの要求がしばしば、例えば、50アミノ酸以上のタンパク質をコードするDNAのストレッチを設定する。

【0208】

ヒト GPCR X 遺伝子のクローニングから決定されたヌクレオチド配列は、他の細胞型（例えば、他の組織由来）における GPCR X ホモログ、ならびに他の脊椎動物由来の GPCR X ホモログを同定および／またはクローニングする際の使用のために設計されるプローブおよびプライマーの作製を可能にする。このプローブ／プライマーは、代表的には、実質的に精製されたオリゴヌクレオチドを含む。このオリゴヌクレオチドは、代表的には、配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 の少なくとも約 12 個、約 25 個、約 50 個、約 100 個、約 150 個、約 200 個、約 250 個、約 300 個、約 350 個または約 400 個の連続するセンス鎖ヌクレオチド配列；または配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 のアンチセンス鎖ヌクレオチド配列；あるいは配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 の天然に存在する変異体の少なくとも約 12 個、約 25 個、約 50 個、約 100 個、約 150 個、約 200 個、約 250 個、約 300 個、約 350 個または約 400 個の連続するセンス鎖ヌクレオチド配列に、ストリンジェントな条件下でハイブリダイズするヌクレオチド配列の領域を含む。

【0209】

ヒト GPCR X ヌクレオチド配列に基づくプローブは、同じまたは相同なタンパク質をコードする転写物またはゲノム配列を検出するために使用され得る。種々の実施形態において、このプローブはさらに、プローブに付着した標識基を含み、例えば、この標識基は放射性同位体、蛍光化合物、酵素、または酵素補因子であり得る。このようなプローブは、GPCR X タンパク質を誤って発現する細胞または組織を同定するための診断試験キットの一部として、例えば、被験体由来の細胞のサンプル中の GPCR X をコードする核酸のレベルを測定すること（例えば、GPCR X mRNA レベルを検出すること、またはゲノム GPCR X 遺伝子に変異しているかまたは欠失しているか否かを決定すること）によって使用され得る。

【0210】

「GPCR X ポリペプチドの生物学的に活性な部分を有するポリペプチド」は、本発明のポリペプチドの活性に類似するが必ずしも同一ではない活性を示すポリペプチドをいい、用量依存性の有無に関わらず、特定の生物学的アッセイにおいて測定されるような成熟形態を含む。「GPCR X の生物学的に活性な部分」をコードする核酸フラグメントは、GPCR X の生物学的活性（GPCR X タンパク質の生物学的活性は、以下に記載される）を有するポリペプチドをコードする、配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 の一部を単離し、GPCR X タンパク質のコードされた部分を発現させ（例えば、インビトロでの組換え発現によって）、そして GPCR X のコードされた部分の活性を評価することによって調製され得る。

【0211】

（GPCR X 核酸およびポリペプチドの改変体）

本発明はさらに、遺伝コードの縮重に起因して、配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 に示されるヌクレオチド配列とは異なる核酸分子を包含し、従って、配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 に示されるヌクレオチド配列によってコードされるタンパク質と同じ GPCR X タンパク質をコードする。別の実施形態において、本発明の単離された核酸分子は、配列番号 2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および 83 に示されるアミノ酸配列を有するタンパク質をコードするヌクレオチド配列を有する。

【0212】

配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30

、32、34、36、および38に示されるGPCRXヌクレオチド配列に加えて、このGPCRXポリペプチドのアミノ酸配列における変化を導くDNA配列多型は、集団（例えば、ヒト集団）内に存在し得ることが、当業者によって理解される。GPCRX遺伝子中のこのような遺伝的多型は、天然の対立遺伝子のバリエーションに起因して、集団内の個体間に存在し得る。本明細書中で使用される場合、用語「遺伝子」および「組換え遺伝子」は、GPCRXタンパク質、好ましくは脊椎動物のGPCRXタンパク質をコードするオープンリーディングフレーム（ORF）を含む核酸分子をいう。このような天然の対立遺伝子のバリエーションは、代表的には、GPCRX遺伝子のヌクレオチド配列において1～5%の変動性を生じ得る。天然の対立遺伝子バリエーションの結果であり、そしてこのGPCRXポリペプチドの機能的活性を変化させない、GPCRXポリペプチド内の任意および全てのこのようなヌクレオチドのバリエーションならびに得られるアミノ酸多型は、本発明の範囲内であると意図される。 10

【0213】

さらに、他の種由来のGPCRXタンパク質をコードし、従って、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38のヒト配列とは異なるヌクレオチド配列を有する核酸分子は、本発明の範囲内にあることが意図される。本発明のGPCRXのcDNAの天然の対立遺伝子改変体およびホモログに対応する核酸分子は、本明細書中に開示されるヒトGPCRX核酸に対するその相同性に基づいて、ストリンジェントハイブリダイゼーション条件下で、標準的なハイブリダイゼーション技術に従うハイブリダイゼーションプローブとしてヒトcDNAまたはその一部を用いて単離され得る。 20

【0214】

従って、別の実施形態では、本発明の単離された核酸分子は、少なくとも6ヌクレオチド長であり、そしてストリンジェント条件下で配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38のヌクレオチド配列を含む核酸分子にハイブリダイズする。別の実施形態では、この核酸は、少なくとも10、25、50、100、250、500、750、1000、1500、または2000以上のヌクレオチド長である。なお別の実施形態では、本発明の単離された核酸分子は、コード領域にハイブリダイズする。本明細書中で用いられる場合、用語「ストリンジェント条件下でハイブリダイズする」は、ハイブリダイゼーションおよび洗浄に関して、互いに少なくとも60%相同なヌクレオチド配列が代表的には互いにハイブリダイズしたままである条件を記載することを意図する。 30

【0215】

ホモログ（すなわち、ヒト以外の種由来のGPCRXタンパク質をコードする核酸）または他の関連配列（例えば、パラログ（paralogs））は、特定のヒト配列の全てまたは一部をプローブとし、核酸ハイブリダイゼーションおよびクローニングに関して当該分野で周知の方法を使用して、低い、中程度の、または高いストリンジェンシーのハイブリダイゼーションによって入手され得る。

【0216】

本明細書で用いられる場合、語句「ストリンジェントハイブリダイゼーション条件」は、その条件下で、プローブ、プライマーまたはオリゴヌクレオチドが、その標的配列にハイブリダイズするが、その他の配列にはハイブリダイズしない条件をいう。ストリンジェントな条件は配列依存性であり、そして異なる状況で異なる。より長い配列は、より短い配列より高い温度で特異的にハイブリダイズする。一般に、ストリンジェントな条件は、規定されたイオン強度およびpHで、特定の配列の熱融解点（ T_m ）より約5℃低いように選択される。この T_m は、標的配列に相補的なプローブの50%が、平衡状態で標的配列にハイブリダイズする（規定されたイオン強度、pHおよび核酸濃度下）温度である。標的配列は一般に過剰で存在するので、 T_m では、50%のプローブが平衡状態で占有されている。代表的には、ストリンジェントな条件は、pH7.0～8.3で、塩濃度が約1.0Mナトリウムイオンより少なく、代表的には約0.01～1.0Mナトリウムイオン 40 50

(またはその他の塩)、そして温度が短いプローブ、プライマーまたはオリゴヌクレオチド (例えば、10 nt ~ 50 nt) について少なくとも約 30 °C、そしてより長いプローブ、プライマーおよびオリゴヌクレオチドについて少なくとも約 60 °C であるような条件である。ストリンジェントな条件はまた、ホルムアミドのような、脱安定化剤の添加で達成され得る。

【0217】

ストリンジェント条件は、当業者に公知であり、そして Ausubel ら (編), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6 に見出され得る。好ましくは、この条件は、互いに少なくとも約 65 %、約 70 %、約 75 %、約 85 %、約 90 %、約 95 %、約 98 % または約 99 % 相同な配列が、代表的には互いにハイブリダイズしたままであるような条件である。ストリンジェントハイブリダイゼーション条件の非限定的な例は、6 × SSC、50 mM Tris-HCl (pH 7.5)、1 mM EDTA、0.02 % PVP、0.02 % Ficoll、0.02 % BSA、および 500 mg/ml 変性サケ精子 DNA を含む高塩緩衝液中での 65 °C でのハイブリダイゼーションであり、続いて、0.2 × SSC、0.01 % BSA 中での 50 °C での 1 回以上の洗浄を行う。ストリンジェント条件下で配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 の配列にハイブリダイズする、本発明の単離された核酸分子は、天然に存在する核酸分子に対応する。本明細書中で使用される場合、「天然に存在する」核酸分子とは、天然に存在する (例えば、天然のタンパク質をコードする) ヌクレオチド配列を有する、RNA 分子または DNA 分子をいう。

【0218】

第 2 の実施形態では、配列番号 1、4、6、またはそれらのフラグメント、アナログもしくは誘導体のヌクレオチド配列を含む核酸分子に、中程度のストリンジェンシー条件下でハイブリダイズし得る核酸配列が提供される。中程度のストリンジェンシーハイブリダイゼーション条件の非限定的な例は、55 °C での 6 × SSC、5 × デンハルト溶液、0.5 % SDS および 100 mg/ml 変性サケ精子 DNA 中でのハイブリダイゼーション、続いて 1 × SSC、0.1 % SDS 中での 37 °C での 1 回以上の洗浄である。用いられ得る中程度のストリンジェンシーの他の条件は、当該分野で周知である。例えば、Ausubel ら (編), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY および Krieger, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY を参照のこと。

【0219】

第 3 の実施形態では、配列番号 1、4、6、またはそれらのフラグメント、アナログもしくは誘導体のヌクレオチド配列を含む核酸分子に、低ストリンジェンシー条件下でハイブリダイズし得る核酸が提供される。低ストリンジェンシーハイブリダイゼーション条件の非限定的な例は、35 % ホルムアミド、5 × SSC、50 mM Tris-HCl (pH 7.5)、5 mM EDTA、0.02 % PVP、0.02 % Ficoll、0.2 % BSA、100 mg/ml の変性サケ精子 DNA、10 % (重量/容量) デキストラン硫酸中での 40 °C でのハイブリダイゼーション、続いて 2 × SSC、25 mM Tris-HCl (pH 7.4)、5 mM EDTA および 0.1 % SDS 中での 50 °C での 1 回以上の洗浄である。用いられ得る低ストリンジェンシーの他の条件は、当該分野で周知である (例えば、種交差ハイブリダイゼーションについて用いられるように)。例えば、Ausubel ら (編), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY ならびに Krieger, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press,

NY; Shil oおよびWeinberg, 1981, Proc Natl Acad Sci USA 78: 6789-6792を参照のこと。

【0220】

(保存的変異)

集団中に存在し得る、GPCRX配列の天然に存在する対立遺伝子改変体に加えて、当業者は、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38のヌクレオチド配列への変異によって変化が導入され得、それによって、このGPCRXタンパク質の機能的能力を変更することなく、コードされるGPCRXタンパク質のアミノ酸配列に変化がもたらされることをさらに理解する。例えば、「非必須」アミノ酸残基にてアミノ酸置換をもたらすヌクレオチド置換は、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83の配列において行われ得る。「非必須」アミノ酸残基は、生物学的活性を有意に変更することなく、このGPCRXタンパク質のの野生型配列から、変更され得る残基であり、一方、「必須」アミノ酸残基は、このような生物学的活性に必要とされる。例えば、本発明のGPCRXタンパク質の間で保存されているアミノ酸残基は、特に変更を受け入れ難いと予測される。保存的置換がなされ得るアミノ酸は、当該分野で周知である。

【0221】

本発明の別の局面は、活性に必須ではないアミノ酸残基における変化を含む、GPCRXタンパク質をコードする核酸分子に関する。このようなGPCRXタンパク質は、アミノ酸配列が、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83とは異なるが、生物学的活性をなお保持する。1つの実施形態では、単離された核酸分子は、タンパク質をコードするヌクレオチド配列を含み、ここで、このタンパク質は、それぞれ、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のアミノ酸配列に少なくとも約45%の相同性であるアミノ酸配列を含む。好ましくは、この核酸分子によってコードされるタンパク質は、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に対して少なくとも約60%相同であり；より好ましくは配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に対して少なくとも約70%相同であり；なおより好ましくは、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に対してして少なくとも約80%相同であり；なおより好ましくは、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に対して少なくとも約90%相同であり；そして最も好ましくは、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に対して少なくとも約95%相同である。

【0222】

配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のタンパク質に相同なGPCRXタンパク質をコードする単離された核酸分子は、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38のヌクレオチド配列に1以上のヌクレオチドの置換、付加または欠失を導入することにより作製され得、その結果、1以上のアミノ酸の置換、付加または欠失が、コードされるタンパク質に導入される。

【0223】

変異は、標準技術（例えば、部位特異的変異誘発およびPCR媒介変異誘発）によって配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に導入され得る。好ましくは、保存的アミノ酸置換が、1以上の推定非必須アミノ酸残基にて作製される。「保存的アミノ酸置換」は、アミノ酸残基が類似の側鎖を有するアミノ酸残基で置換される、アミノ酸置換である。類似の側鎖を有す

るアミノ酸残基のファミリーは、当該分野で定義されている。これらのファミリーとしては、以下が挙げられる：塩基性側鎖を有するアミノ酸（例えば、リジン、アルギニン、ヒスチジン）、酸性側鎖を有するアミノ酸（例えば、アルパラギン酸、グルタミン酸）、非荷電極性側鎖を有するアミノ酸（例えば、グリシン、アスパラギン、グルタミン、セリン、トレオニン、チロシン、システイン）、非極性側鎖を有するアミノ酸（例えば、アラニン、バリン、ロイシン、イソロイシン、プロリン、フェニルアラニン、メチオニン、トリプトファン）、 β 分枝側鎖を有するアミノ酸（例えば、トレオニン、バリン、イソロイシン）および芳香族側鎖を有するアミノ酸（例えば、チロシン、フェニルアラニン、トリプトファン、ヒスチジン）。従って、GPCRXタンパク質中の推定された非必須アミノ酸残基は、同じ側鎖のファミリー由来の別のアミノ酸残基で置換される。あるいは、別の実施形態では、変異は、GPCRXコード配列の全てまたは一部に沿って（例えば、飽和変異誘発（saturation mutagenesis）によって）ランダムに導入され得、そして得られる変異体は、GPCRXの生物学的活性についてスクリーニングされて、活性を維持する変異体を同定し得る。配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38の変異誘発に続いて、コードされるタンパク質は、当該分野で公知の任意の組換え技術によって発現され得、そしてこのタンパク質の活性が決定され得る。

10

【0224】

アミノ酸ファミリーの関連性はまた、側鎖に基づいて決定され得る。置換されたアミノ酸は、完全に保存された「強い」残基または完全に保存された「弱い」残基であり得る。保存されたアミノ酸残基の「強い」群は、以下の群のいずれかであり得る：STA、NEQK、NHQK、NDEQ、QHRK、MILV、MILF、HY、FYW（ここで、単一の文字のアミノ酸コードは、互いに置換され得るこれらのアミノ酸残基によりグループ化される）。同様に、保存された残基の「弱い」群は、以下のいずれかで有り得る：CSA、ATV、SAG、STNK、STPA、SGND、SNDEQK、NDEQHK、NEQHRK、VLIM、HFY（ここで、各群内の文字は、一文字アミノ酸コードを表する）。

20

【0225】

1つの実施形態では、変異GPCRXタンパク質は、以下についてアッセイされ得る：（i）他のGPCRXタンパク質、他の細胞表面タンパク質、または生物学的に活性なそれらの部分と、タンパク質：タンパク質相互作用を形成する能力、（ii）変異GPCRXタンパク質と、GPCRXのリガンドとの間の複合体形成；（iii）変異GPCRXタンパク質が細胞内標的タンパク質またはその生物学的に活性な部分に結合する能力；（例えば、アビジンタンパク質）。

30

【0226】

なお別の実施形態において、変異体GPCRXタンパク質は、特定の生物学的機能を調節する能力（例えば、インスリン放出の能力）についてアッセイされ得る。

【0227】

（アンチセンス核酸）

本発明の別の局面は、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38またはそのフラグメント、アナログもしくは誘導体のヌクレオチド配列を含む核酸分子にハイブリダイズし得るかまたは相補的である、単離されたアンチセンス核酸分子に関する。「アンチセンス」核酸は、タンパク質をコードする「センス」核酸に相補的である（例えば、二本鎖cDNA分子のコード鎖に相補的であるかまたはmRNA配列に相補的である）ヌクレオチド配列を含む。特定の局面では、少なくとも約10、約25、約50、約100、約250もしくは約500ヌクレオチドまたは全体のGPCRXコード鎖、またはそれらの一部のみに相補的な配列を含む、アンチセンス核酸分子が提供される。配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のGPCRXタンパク質のフラグメント、ホモログ、誘導体およびアナログをコードする核酸分子、

40

50

または配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38 の GPCR X 核酸配列に相補的なアンチセンス核酸がさらに提供される。

【0228】

1つの実施形態では、アンチセンス核酸分子は、GPCR Xタンパク質をコードするヌクレオチド配列のコード鎖の「コード領域」に対してアンチセンスである。用語「コード領域」とは、アミノ酸残基に翻訳されるコドンを含む、ヌクレオチド配列の領域をいう。別の実施形態では、このアンチセンス核酸分子は、GPCR Xタンパク質をコードするヌクレオチド配列のコード鎖の「非コード領域」に対してアンチセンスである。用語「非コード領域」とは、コード領域に隣接する、アミノ酸に翻訳されない、5'配列および3'配列をいう（すなわち、5'非翻訳領域および3'非翻訳領域ともいわれる）。 10

【0229】

本明細書中に開示される GPCR Xタンパク質をコードするコード鎖配列を考慮すれば、本発明のアンチセンス核酸は、WatsonおよびCrickまたはHoogsteenの塩基対合の規則に従って設計され得る。アンチセンス核酸分子は、GPCR X mRNAのコード領域全体に相補的であり得るが、より好ましくは、GPCR X mRNAのコード領域または非コード領域の一部にのみアンチセンスであるオリゴヌクレオチドである。例えば、アンチセンスオリゴヌクレオチドは、GPCR X mRNAの翻訳開始部位を取り囲む領域に相補的であり得る。アンチセンスオリゴヌクレオチドは、例えば、長さが約5、約10、約15、約20、約25、約30、約35、約40、約45または約50 20ヌクレオチドであり得る。本発明のアンチセンス核酸は、当該分野で公知の手順を用いて、化学的合成または酵素連結反応を用いて構築され得る。例えば、アンチセンス核酸（例えば、アンチセンスオリゴヌクレオチド）は、天然に存在するヌクレオチド、またはこの分子の生物学的安定性を増大させるように、もしくはアンチセンス核酸とセンス核酸との間で形成される二重鎖の物理的安定性を増大させるように設計された種々に改変されたヌクレオチドを用いて化学的に合成され得る（例えば、ホスホロチオエート誘導体およびアクリジン置換ヌクレオチドが用いられ得る）。

【0230】

アンチセンス核酸を作製するために用いられ得る改変されたヌクレオチドの例としては以下が挙げられる：5-フルオロウラシル、5-ブロモウラシル、5-クロロウラシル、5-ヨードウラシル、ヒポキサンチン、キサンチン、4-アセチルシトシン、5-（カルボキシヒドロキシメチル）ウラシル、5-カルボキシメチルアミノメチル2-チオウリジン、5-カルボキシメチルアミノメチルウラシル、ジヒドロウラシル、β-D-ガラクトシルクエオシン（galactosylqueosine）、イノシン、N6-イソペンテニルアデニン、1-メチルグアニン、1-メチルイノシン、2,2-ジメチルグアニン、2-メチルアデニン、2-メチルグアニン、3-メチルシトシン、5-メチルシトシン、N6-アデニン、7-メチルグアニン、5-メチルアミノメチルウラシル、5-メトキシアミノメチル-2-チオウラシル、β-D-マンノシルクエオシン（mannosylqueosine）、5'-メトキシカルボキシメチルウラシル、5-メトキシウラシル、2-メチルチオ-N6-イソペンテニルアデニン、ウラシル-5-オキシ酢酸（v）、ワイプトキソシン、シュードウラシル、クエオシン（queosine）、2-チオシトシン、5-メチル-2-チオウラシル、2-チオウラシル、4-チオウラシル、5-メチルウラシル、ウラシル-5-オキシ酢酸メチルエステル、ウラシル-5-オキシ酢酸（v）、5-メチル-2-チオウラシル、3-（3-アミノ-3-N-2-カルボキシプロピル）ウラシル、（acp3）w、および2,6-ジアミノプリン。あるいは、このアンチセンス核酸は、核酸がアンチセンス方向でサブクロニングされた発現ベクターを用いて生物学的に生成され得る（すなわち、挿入された核酸から転写されたRNAは、以下の小節にさらに記載される、目的の標的核酸に対してアンチセンス方向である）。 40

【0231】

本発明のアンチセンス核酸分子は、代表的には被験体に投与されるか、またはインサイチ 50

ユで生成され、その結果それらは、G P C R Xタンパク質をコードする細胞性m R N Aおよび／またはゲノムD N Aとハイブリダイズするか、またはそれに結合し、それによってこのタンパク質の発現を阻害する（例えば、転写および／または翻訳を阻害することによって）。ハイブリダイゼーションは、安定な二重鎖を形成する従来のヌクレオチド相補性によってか、または例えばD N A二重鎖に結合するアンチセンス核酸分子の場合には、二重らせんのメジャーグループ（m a j o r g r o o v e）における特異的相互作用を介してであり得る。本発明のアンチセンス核酸分子の投与経路の例としては、組織部位での直接的注射が挙げられる。あるいは、アンチセンス核酸分子は、選択された細胞を標的化するように改変され得、次いで全身に投与される。例えば、全身投与のために、アンチセンス分子は、それらが選択された細胞表面上に発現されたレセプターまたは抗原に特異的に結合するように改変され得る（例えば、そのアンチセンス核酸分子を細胞表面レセプターまたは抗原に結合するペプチドまたは抗体に連結することによって）。このアンチセンス核酸分子はまた、本明細書中に記載されるベクターを用いて細胞に送達され得る。十分な核酸分子を達成するために、アンチセンス核酸分子が強力なp o l I Iプロモーターまたはp o l I I Iプロモーターの制御下に置かれているベクター構築物が、好ましい。

【0232】

なお別の実施形態において、本発明のアンチセンス核酸分子は、 α -アノマー核酸分子である。 α -アノマー核酸分子は、相補的R N Aと特異的な二本鎖ハイブリッドを形成し、ここで、通常の α -ユニットとは対照的に、鎖は、互いに平行に走行する（例えば、G a u l t i e rら、1987, N u c l . A c i d s R e s . 15:6625-6641を参照のこと）。このアンチセンス核酸分子はまた、2'- α -メチルリボヌクレオチド（例えば、I n o u eら、1987, N u c l . A c i d s R e s . 15:6131-6148を参照のこと）またはキメラR N A-D N Aアナログ（例えば、I n o u eら、1987, F E B S L e t t . 215:327-330を参照のこと）を含み得る。

【0233】

（リボザイムおよびP N A部分）

核酸の改変としては、非制限的な例として、改変塩基、および糖リン酸バックボーンが改変または誘導体化された核酸が挙げられる。これらの改変は、少なくとも一部は、改変された核酸の化学的安定性を増強するために実施され、その結果、それらは、例えば、被験体での治療的適用におけるアンチセンス結合核酸として使用され得る。

【0234】

1つの実施形態において、本発明のアンチセンス核酸はリボザイムである。リボザイムは、一本鎖核酸（例えば、m R N A）を切断し得るリボヌクレアーゼ活性を有する触媒性R N A分子であり、これらは、その一本鎖核酸に対して相補領域を有する。従って、リボザイム（例えば、H a s e l h o f fおよびG e r l a c h 1988, N a t u r e 334:585-591に記載されるハンマーヘッド型リボザイム）を使用して、G P C R X m R N A転写物を触媒的に切断し、それによってG P C R X m R N Aの翻訳を阻害し得る。G P C R Xをコードする核酸に特異性を有するリボザイムは、本明細書中に開示されるG P C R X c D N Aのヌクレオチド配列（すなわち、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38）に基づいて設計され得る。例えば、活性部位のヌクレオチド配列が、G P C R Xをコードするm R N A内で切断されるヌクレオチド配列に相補的である、テトラヒメナL-19 I V S R N Aの誘導体が構築され得る。例えば、C e c hらの米国特許第4,987,071号およびC e c hらの米国特許第5,116,742号を参照のこと。G P C R X m R N Aをまた使用して、R N A分子のプールから、特異的リボヌクレアーゼ活性を有する触媒性R N Aを選択し得る。例えば、B a r t e lら（1993）S c i e n c e 261:1411-1418を参照のこと。

【0235】

あるいは、G P C R X遺伝子発現は、G P C R X核酸の調節領域（例えば、G P C R Xの

プロモーターおよび／またはエンハンサー)に相補的なヌクレオチド配列を標的化し、標的細胞中でG P C R X遺伝子の転写を妨害する三重らせん構造を形成することによって阻害され得る。例えば、H e l l e n e, 1991. A n t i c a n c e r D r u g D e s . 6 : 5 6 9 ~ 8 4 ; H e l l e n e ら, 1992. A n n . N . Y . A c a d . S c i . 6 6 0 : 2 7 - 3 6 ; M a h e r, 1992. B i o a s s a y s 1 4 : 8 0 7 - 1 5 を参照のこと。

【0236】

種々の実施形態において、G P C R X核酸は、塩基部分、糖部分またはリン酸骨格で改変され、例えば、その分子の安定性、ハイブリダイゼーションまたは可溶性を改善し得る。例えば、核酸のデオキシリボースリン酸骨格を改変して、ペプチド核酸を生成し得る（例えば、H y r u p ら, 1996. B i o o r g M e d C h e m 4 : 5 - 2 3 を参照のこと）。本明細書中で使用される場合、用語「ペプチド核酸」または「P N A」は、デオキシリボースリン酸骨格が偽ペプチド骨格によって置換され、そして4つの天然の核塩基 (n u c l e o b a s e) のみが保持されている核酸模倣物（例えば、DNA模倣物）をいう。P N A の中性の骨格は、低いイオン強度の条件下でDNAおよびRNAに対する特異的ハイブリダイゼーションを可能にすることが示されている。P N A オリゴマーの合成は、H y r u p ら, 1996. 前出; P e r r y - O ' K e e f e ら, 1996. P r o c . N a t l . A c a d . S c i . U S A 9 3 : 1 4 6 7 0 - 1 4 6 7 5 において記載されるような標準的固相ペプチド合成プロトコルを用いて行われ得る。

10

【0237】

G P C R X の P N A は、治療的適用および診断的適用において使用され得る。例えば、P N A は、例えば転写または翻訳の停止を誘導することまたは複製を阻害することによる、遺伝子発現の配列特異的調節のためのアンチセンスまたは抗遺伝子剤として使用され得る。G P C R X の P N A はまた、例えば P N A 指向性 P C R クランピングによる遺伝子における一塩基対変異の分析において；他の酵素（例えば、S 1 ヌクレアーゼ）と組み合わせて使用される場合の人工制限酵素として（H y r u p B . ら, 1996. 前出を参照のこと）；またはDNA配列およびハイブリダイゼーションのプローブもしくはプライマーとして（H y r u p ら, 1996, 前出; P e r r y - O ' K e e f e ら, 1996, 前出を参照のこと）、使用され得る。

20

【0238】

別の実施形態において、G P C R X の P N A は、例えば、それらの安定性または細胞性取り込みを増強するために、P N A に脂溶性基または他のヘルパー基を結合することによって、P N A - D N A キメラの形成によって、またはリポソームもしくは当該分野において公知の薬物送達他の技術の使用によって改変され得る。例えば、P N A およびDNAの有利な特性を組合せ得る、G P C R X の P N A - D N A キメラが生成され得る。P N A 部分が高い結合親和性および特異性を提供する一方で、そのようなキメラは、DNA認識酵素（例えば、R N a s e H およびDNAポリメラーゼ）がDNA部分と相互作用するのを可能にする。P N A - D N A キメラは、塩基のスタッキング、核塩基間の結合数および方向を考慮して選択される適切な長さのリンカーを使用して連結され得る（H y r u p ら, 1996, 前出を参照のこと）。P N A - D N A キメラの合成は、H y r u p ら, 1996, 前出およびF i n n ら, 1996, N u c l A c i d s R e s 2 4 : 3 3 5 7 - 3 3 6 3 において記載されるように行われ得る。例えば、DNA鎖は、標準的なホスホラミダイトカップリング化学を用いて固体支持体上で合成され得、そして改変されたヌクレオシドアナログ（例えば、5' - (4 - メトキシトリチル) アミノ - 5' - デオキシ - チミジンホスホルアミダイト）が、P N A とDNAの5'末端との間に使用され得る（M a g ら, 1989, N u c l A c i d R e s 1 7 : 5 9 7 3 ~ 5 9 8 8 を参照のこと）。次いで、P N A モノマーが段階様式でカップリングされ、5' P N A セグメントおよび3' DNA セグメントを有するキメラ分子を生成する（F i n n ら, 1996, 前出を参照のこと）。あるいは、5' DNA セグメントおよび3' P N A セグメントを用いて、キメラ分子が合成され得る。例えば、P e t e r s e n ら, 1975,) B i o o r

30

40

50

g. Med. Chem. Lett. 5:1119~11124を参照のこと。

【0239】

他の実施形態において、オリゴヌクレオチドは、以下のような他の付属の基を含み得る：ペプチド（例えば、インビボで宿主細胞レセプターを標的化するため）、または細胞膜を横切る輸送を容易にする因子（例えば、Lettingerら、1989、Proc. Natl. Acad. Sci. U. S. A. 86:6553~6556；Lemaîtreら、1987、Proc. Natl. Acad. Sci. 84:648~652；PCT公開番号WO88/09810を参照のこと）、または血液脳関門（例えば、PCT公開番号WO89/10134を参照のこと）。さらに、オリゴヌクレオチドは、ハイブリダイゼーション誘発切断剤（例えば、Krolら、1988、BioTechniques 6:958~976を参照のこと）、またはインターカレーター剤（例えば、Zon, 1988、Pharm. Res. 5:539~549を参照のこと）で改変され得る。この目的のために、オリゴヌクレオチドは、別の分子（例えば、ペプチド、ハイブリダイゼーション誘発架橋剤、輸送剤、ハイブリダイゼーション誘発切断剤など）に結合され得る。

【0240】

（GPCRXポリペプチド）

本発明に従うポリペプチドは、配列が配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に提供されるGPCRXタンパク質のアミノ酸配列を含むポリペプチドを含む。本発明はまた、なおそのGPCRX活性および生理学的機能を維持するタンパク質、またはその機能的フラグメントをコードするが、その任意の残基が配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に示される対応する残基から変化し得る、変異体または改変体タンパク質を含む。

【0241】

一般に、GPCRX様機能を保持するGPCRX改変体は、配列中の特定位置の残基が他のアミノ酸により置換され、そしてさらに、親タンパク質の2つの残基間にさらなる残基を挿入する可能性、および親配列から1つ以上の残基を欠失する可能性を含む。任意のアミノ酸置換、挿入または欠失が、本発明に包含される。好適な状況では、この置換は、上記で規定されるような保存的置換である。

【0242】

本発明の1つの局面は、単離されたGPCRXタンパク質、およびその生物学的に活性な部分、またはそれらの誘導体、フラグメント、アナログもしくはホモログに関する。抗GPCRX抗体を惹起するための免疫原としての使用のために適切なポリペプチドフラグメントもまた提供される。1つの実施形態において、ネイティブなGPCRXタンパク質は、標準的なタンパク質精製技術を用いる適切な精製スキームによって、細胞または組織供給源から単離され得る。別の実施形態において、GPCRXタンパク質は、組換えDNA技術によって産生される。組換え発現に代わるものとして、GPCRXタンパク質またはポリペプチドは、標準的なペプチド合成技術を用いて化学合成され得る。

【0243】

「単離された」または「精製された」ポリペプチドまたはタンパク質あるいはその生物学的に活性な部分は、GPCRXタンパク質の由来する細胞または組織供給源由来の細胞性物質または他の夾雑タンパク質を実質的に含まないか、あるいは化学合成される場合に化学前駆体または他の化学物質を実質的に含まない。用語「細胞性物質を実質的に含まない」は、GPCRXタンパク質の調製物を含み、この調製物において、GPCRXタンパク質が単離または組換え産生される細胞の細胞性成分から、GPCRXタンパク質は分離されている。1つの実施形態において、用語「細胞性物質を実質的に含まない」は、非GPCRXタンパク質（本明細書中において「夾雑タンパク質」とも呼ばれる）を約30%未満（乾燥重量にて）、より好ましくは非GPCRXタンパク質を約20%未満、なおより好ましくは非GPCRXタンパク質を約10%未満、そして最も好ましくは非GPCRX

タンパク質を約5%未満有する、GPCRXタンパク質の調製物を含む。GPCRXタンパク質またはその生物学的に活性な部分が組換え産生される場合、好ましくは、調製物はまた培養培地を実質的に含まない。すなわち、培養培地は、そのタンパク質調製物の容量の約20%未満、より好ましくは約10%未満、そして最も好ましくは約5%未満を示す。

【0244】

用語「化学前駆体または他の化学物質を実質的に含まない」は、タンパク質が、そのタンパク質の合成に参与する化学前駆体または他の化学物質から分離されているGPCRXタンパク質の調製物を含む。1つの実施形態において、用語「化学前駆体または他の化学物質を実質的に含まない」は、化学前駆体または非GPCRXの化学物質を約30%未満（乾燥重量にて）、より好ましくは化学前駆体または非GPCRX化学物質を約20%未満、なおより好ましくは化学前駆体または非GPCRX化学物質を約10%未満、そして最も好ましくは化学前駆体または非GPCRX化学物質を約5%未満有する、GPCRXタンパク質の調製物を含む。 10

【0245】

GPCRXタンパク質の生物学的に活性な部分は、全長GPCRXタンパク質より少ないアミノ酸を含み、そしてGPCRXタンパク質の少なくとも1つの活性を示す、GPCRXタンパク質のアミノ酸配列（例えば、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に示されるアミノ酸配列）に十分に相同なアミノ酸配列、またはGPCRXタンパク質のアミノ酸配列に由来するアミノ酸配列を含むペプチドを含む。代表的には、生物学的に活性な部分は、GPCRXタンパク質の少なくとも1つの活性を有するドメインまたはモチーフを含む。GPCRXタンパク質の生物学的に活性な部分は、例えば、長さが10、25、50、100またはそれより多いアミノ酸残基であるポリペプチドであり得る。 20

【0246】

さらに、タンパク質の他の領域が欠失している他の生物学的に活性な部分は、組換え技術によって調製され得、そしてネイティブなGPCRXタンパク質の機能的活性のうちの1つ以上について評価され得る。

【0247】

1つの実施形態において、GPCRXタンパク質は、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に示されるアミノ酸配列を有する。他の実施形態において、GPCRXタンパク質は、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83に実質的に相同であり、そして以下に詳細に議論されるように、天然の対立遺伝子改変体または変異誘発に起因してアミノ酸配列が異なるが、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のタンパク質の機能的活性を保持する。従って、別の実施形態において、GPCRXタンパク質は、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のアミノ酸配列に少なくとも約45%相同なアミノ酸配列を含み、そして、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のGPCRXタンパク質の機能的活性を保持するタンパク質である。 30 40

【0248】

（2つ以上の配列間の相同性の決定）

2つのアミノ酸配列または2つの核酸の相同性の百分率を決定するために、配列は、至適な比較の目的のために整列される（例えば、ギャップは、第2のアミノ酸または核酸配列との最適な整列のために、第1のアミノ酸または核酸配列に導入され得る）。次いで、対応するアミノ酸位置またはヌクレオチド位置でのアミノ酸残基またはヌクレオチドが比較される。第1の配列中の位置が、第2の配列中の対応する位置と同じアミノ酸残基またはヌクレオチドで占められる場合、分子はその位置で相同である（すなわち、本明細書中で 50

使用される場合、アミノ酸または核酸の「相同性」は、アミノ酸または核酸の「同一性」と等価である）。

【0249】

核酸配列の相同性は、2つの配列間の同一性の程度として決定され得る。相同性は、当該分野において公知のコンピュータプログラム（例えば、GCGプログラムパッケージにおいて提供されるGAPソフトウェア）を用いて決定され得る。NeedlemanおよびWunsch, 1970 J Mol Biol 48:443~453を参照のこと。核酸配列比較のための以下の設定（GAP作製ペナルティー、5.0、およびGAP伸長ペナルティー、0.3）を用いてGCG GAPソフトウェアを使用すると、上記で言及される類似の核酸配列のコード領域は、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38に示されるDNA配列のCDS（コード）部分と、好ましくは少なくとも70%、75%、80%、85%、90%、95%、98%または99%の同一性の程度を示す。

【0250】

用語「配列同一性」は、2つのポリヌクレオチド配列またはポリペプチド配列が、特定の比較領域にわたって、残基毎を基準として同一である程度をいう。用語「配列同一性のパーセンテージ」は、以下により算出される：この比較領域にわたって最適に整列された2つの配列を比較すること、両方の配列において同一の核酸塩基（例えば、核酸の場合にはA、T、C、G、U、またはI）が存在する位置の数を決定し、一致した位置の数を導くこと、この一致した位置の数を、比較領域の内の位置の総数（すなわち、ウィンドウサイズ）で除算すること、およびその結果を100で乗算して、配列同一性のパーセンテージを導くこと。用語「実質的な同一性」は、本明細書中で使用される場合、ポリヌクレオチド配列の特徴を示し、ここでこのポリヌクレオチドは、比較領域にわたり参照配列と比較して、少なくとも80%の配列同一性、好ましくは少なくとも85%の配列同一性、そして頻繁には90~95%の配列同一性、より通常には少なくとも99%の配列同一性を有する配列を含む。

【0251】

（キメラタンパク質および融合タンパク質）

本発明はまた、GPCRXキメラタンパク質または融合タンパク質を提供する。本明細書中で使用される場合、GPCRX「キメラタンパク質」またはGPCRX「融合タンパク質」は、非GPCRXポリペプチドに作動可能に連結された、GPCRXポリペプチドを含む。「GPCRXポリペプチド」は、GPCRXタンパク質（配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83）に対応するアミノ酸配列を有するポリペプチドをいうが、「非GPCRXポリペプチド」は、GPCRXタンパク質に対して実質的に相同ではないタンパク質（例えば、GPCRXタンパク質とは異なり、かつ同一でかまたは異なる生物体に由来するタンパク質）に対応するアミノ酸配列を有するポリペプチドをいう。GPCRX融合タンパク質において、このGPCRXポリペプチドは、GPCRXタンパク質のすべてまたは一部分に対応し得る。1つの実施形態では、GPCRX融合タンパク質は、GPCRXタンパク質の少なくとも1つの生物学的に活性な部分を含む。別の実施形態では、GPCRX融合タンパク質は、GPCRXタンパク質の少なくとも2つの生物学的に活性な部分を含む。なお別の実施形態においてGPCRX融合タンパク質は、少なくとも3つの生物学的に活性な部分を含む。融合タンパク質において、用語「作動可能に連結された」は、GPCRXポリペプチドおよび非GPCRXポリペプチドが、インフレームで互いに融合されていることを示すことが意図される。非GPCRXポリペプチドは、GPCRXポリペプチドのN末端またはC末端に融合され得る。

【0252】

1つの実施形態においては、この融合タンパク質は、GST-GPCRX融合タンパク質であり、ここではGPCRX配列が、GST（グルタチオンS-トランスフェラーゼ）配列のC末端に融合される。このような融合タンパク質は、組換えGPCRXポリペプチド

の精製を容易にし得る。

【0253】

別の実施形態において、この融合タンパク質は、N末端に異種シグナル配列を含むGPCRXタンパク質である。特定の宿主細胞（例えば、哺乳動物宿主細胞）において、GPCRXの発現および／または分泌は、異種シグナル配列の使用を介して増加され得る。

【0254】

なお別の実施形態において、この融合タンパク質は、GPCRX-免疫グロブリン融合タンパク質であり、ここで、GPCRX配列が、免疫グロブリンタンパク質ファミリーのメンバーに由来する配列に融合される。本発明のこのGPCRX-免疫グロブリン融合タンパク質は、薬学的組成物中に取り込まれ、そして被験体に投与されて、細胞の表面上でGPCRXリガンドとGPCRXタンパク質との間の相互作用を阻害し、それによってインビボのGPCRX媒介信号伝達を抑制し得る。このGPCRX-免疫グロブリン融合タンパク質を用い、GPCRX同族リガンドの生体利用性に影響を与え得る。GPCRXリガンド／GPCRX相互作用の阻害は、増殖障害および分化障害の両方の処置、および細胞生存を調節（例えば、促進または阻害）することに、治療的に有用であり得る。さらに、本発明のこのGPCRX-免疫グロブリン融合タンパク質は、被験体中で抗GPCRX抗体を産生するための免疫原として用いられ得、GPCRXリガンドを精製し、そしてGPCRXリガンドとのGPCRXの相互作用を阻害する分子を同定するスクリーニングアッセイで用いられ得る。

10

【0255】

本発明のGPCRXキメラまたは融合タンパク質は、標準的な組換えDNA技法により産生され得る。例えば、異なるポリペプチド配列をコードするDNAフラグメントを、従来の技法に従って、例えば、連結のための平滑末端または粘着（*stagger*）末端、適切な末端を提供するための制限酵素消化、必要に応じて粘着（*cohesive*）末端のフィルイン、所望しない連結を避けるためのアルカリホスファターゼ処理、および酵素的連結を採用することにより、インフレイムで一緒に連結する。別の実施形態では、融合遺伝子を、自動化DNA合成機を含む従来技法により合成し得る。あるいは、遺伝子フラグメントのPCR増幅を、キメラ遺伝子配列を生成するために次いでアニールおよび再増幅され得る、2つの連続遺伝子フラグメント間の相補的オーバハングを生じるアンカープライマーを用いて実施し得る（例えば、Ausbelら（編）CURRENT PROTOCOLS IN MOLECULAR BIOLOGY、John Wiley & Sons、1992を参照のこと）。さらに、融合成分（例えば、GSTポリペプチド）をすでにコードした多くの発現ベクターが市販されている。GPCRXをコードする核酸は、この融合成分がGPCRXタンパク質にインフレイム連結されるように、このような発現ベクター中にクローン化され得る。

20

30

【0256】

（GPCRXアゴニストおよびアンタゴニスト）

本発明はまた、GPCRXアゴニスト（模倣物）として、またはGPCRXアンタゴニストとして機能するGPCRXタンパク質の改変体に関する。GPCRXタンパク質の改変体（例えば、GPCRXタンパク質の離散した点変異または短縮型）が、変異誘発により生成され得る。GPCRXタンパク質のアゴニストは、天然に存在する形態のGPCRXタンパク質と実質的に同じ生物学的活性、またはその生物学的活性のサブセットを保持し得る。GPCRXタンパク質のアンタゴニストは、天然に存在する形態のGPCRXタンパク質の1つ以上の活性を、例えば、GPCRXタンパク質を含む細胞シグナル伝達カスケードの下流または上流メンバーに競合的に結合することにより阻害し得る。従って、特異的生物学的効果が、限られた機能の改変体を用いた処理により惹起され得る。1つの実施形態では、このタンパク質の天然に存在する形態の生物学活性のサブセットを有する改変体を用いた被験体の処置は、GPCRXタンパク質の天然に存在する形態を用いた処置に対して被験体におけるより少ない副作用を有する。

40

【0257】

50

GPCR Xアゴニスト（模倣物）として、またはGPCR Xアンタゴニストのいずれかとして機能するGPCR Xタンパク質の改変体は、GPCR Xタンパク質アゴニストまたはアンタゴニスト活性のためのGPCR Xタンパク質の変異体（例えば短縮型変異体）のコンビナトリアルライブラリーをスクリーニングすることにより同定され得る。1つの実施形態では、GPCR X改変体の多彩なライブラリーは、核酸レベルのコンビナトリアル変異誘発により生成され、そして多彩な遺伝子ライブラリーによりコードされる。GPCR X改変体の多彩なライブラリーは、例えば、合成オリゴヌクレオチドの混合物を、潜在的なGPCR X配列の縮重セットが、個々のポリペプチドとして、あるいは、その中にGPCR X配列のセットを含む（例えば、ファージディスプレイのための）より大きな融合タンパク質のセットとして発現可能であるように、遺伝子配列に酵素的に連結することにより産生され得る。縮重オリゴヌクレオチド配列から潜在的なGPCR X改変体のライブラリーを産生するために用いられ得る種々の方法がある。縮重遺伝子配列の化学的合成は、自動化DNA合成機中で実施され得、次いでこの合成遺伝子は、適切な発現ベクター中に連結される。遺伝子の縮重セットの使用は、1つの混合物において、潜在的なGPCR X配列の所望のセットをコードする配列のすべての供給を可能にする。縮重オリゴヌクレオチドを合成する方法は当該分野で周知である。例えば、Narang, 1983. Tetrahedron 39:3; Itakuraら, 1984. Annu. Rev. Biochem. 53:323; Itakuraら, 1984. Science 198:1056; Ikeら, 1983. Nucl. Acids Res. 11:477を参照のこと。

10

20

【0258】

（ポリペプチドライブラリー）

さらに、GPCR Xタンパク質コード配列のフラグメントのライブラリーを用いて、GPCR Xタンパク質の改変体のスクリーニングおよび引き続く選択のためのGPCR Xフラグメントの多彩な集団を生成し得る。1つの実施形態では、コード配列フラグメントのライブラリーは、GPCR Xコード配列の二本鎖PCRフラグメントを、1分子あたり約1つのみのニックが生じる条件下、ヌクレアーゼで処理すること、二本鎖DNAを変性させること、異なるニック産物からのセンス／アンチセンス対を含み得る二本鎖DNAを形成するためにこのDNAを再生すること、S1ヌクレアーゼを用いた処理により再形成された二本鎖から一本鎖部分を除去すること、および得られるフラグメントライブラリーを発現ベクター中に連結することにより生成し得る。この方法により、GPCR Xタンパク質の種々のサイズのN末端および内部フラグメントをコードする発現ライブラリーが派生し得る。

30

【0259】

点変異または短縮により作成されたコンビナトリアルライブラリーの遺伝子産物をスクリーニングするため、選択された性質を有する遺伝子産物のcDNAライブラリーをスクリーニングするための種々の技術が当該分野で公知である。このような技法を、GPCR Xタンパク質のコンビナトリアル変異誘発により生成された遺伝子ライブラリーの迅速スクリーニングに適用可能である。大きな遺伝子ライブラリーをスクリーニングするための、高スループット分析に適した最も広く用いられる技法は、代表的には、遺伝子ライブラリーを複製可能な発現ベクター中にクローニングすること、得られるベクターのライブラリーで適切な細胞を形質転換すること、およびこのコンビナトリアル遺伝子を、所望の活性の検出が、その産物が検出された遺伝子をコードするベクターの単離を容易にする条件下で発現することを含む。ライブラリー中の機能的変異体の頻度を増大する新規技法であるリクルーシブエンセブル変異誘発（REM）を、スクリーニングアッセイと組み合わせて用い、GPCR X改変体を同定し得る。例えば、ArklinおよびYourvan, 1992. Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgraveら, 1993. Protein Engineering 6:327-331を参照のこと。

40

【0260】

50

(抗GPCRX抗体)

本発明は、本発明の任意のGPCRXポリペプチドに免疫特異的に結合する、Fabまたは(Fab)₂のような抗体または抗体フラグメントを包含する。

【0261】

単離されたGPCRXタンパク質、またはその部分もしくはフラグメントは、ポリクローナル抗体およびモノクローナル抗体調製のための標準的な技術を用いて、GPCRXポリペプチドに結合する抗体を産生するための免疫原として用いられ得る。完全長のGPCRXタンパク質が用いられ得るが、あるいは、本発明は、免疫原としての使用のためのGPCRXタンパク質の抗原性ペプチドフラグメントを提供する。GPCRXの抗原性ペプチドは、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83で示されるアミノ酸配列の少なくとも4アミノ酸残基を含み、そしてこのペプチドに対して惹起された抗体がGPCRXと特異的な免疫複合体を形成するように、GPCRXのエピトープを含む。好ましくは、この抗原性ペプチドは、少なくとも6、8、10、15、20、または30個のアミノ酸残基を含む。より長い抗原性ペプチドが、用途に依存して、そして当業者に周知の方法に従い、より短い抗原性ペプチドよりもしばしば好適である。

【0262】

本発明の特定の実施形態では、抗原性ペプチドに含まれる少なくとも1つのエピトープは、タンパク質の表面上に位置するGPCRXの領域、例えば、親水性領域である。抗体産生を標的にする手段として、親水性および疎水性の領域を示すヒドロパシープロットを、例えば、フーリエ変換を伴うかまたは伴わない、Kyte DoolittleまたはHopp Woods法を含む、当該分野で周知の任意の方法により生成し得る。例えば、それらの全体が本明細書に参考として援用される、例えば、HoppおよびWoods、1981、Proc. Nat. Acad. Sci. USA 78:3824-3828; KyteおよびDoolittle 1982、J. Mol. Biol. 157:105-142を参照のこと。この各々が、それらの全体において、本明細書中で参考として援用される。

【0263】

本明細書で開示されるように、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のGPCRXタンパク質配列、またはその誘導体、フラグメント、アナログもしくはホモログを、これらのタンパク質成分に免疫特異的に結合する抗体の産生における免疫原として利用し得る。本明細書で用いる用語「抗体」は、免疫グロブリン分子、および免疫グロブリン分子の免疫学的に活性な部分、すなわち、GPCRXのような抗原に特異的に結合(免疫反応)する抗原結合部位を含む分子をいう。このような抗体は、制限されずに、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、単鎖抗体、FabおよびF(ab')₂フラグメント、およびFab発現ライブラリーを含む。特定の実施形態では、ヒトGPCRXタンパク質に対する抗体が開示される。当該分野で公知の種々の手順が、配列番号2、4、6、8、10、12、17、19、21、23、25、29、31、33、35、37、および83のGPCRXタンパク質配列、またはその誘導体、フラグメント、アナログもしくはホモログに対するポリクローナルまたはモノクローナル抗体の産生のために用いられ得る。これらのタンパク質のいくつかを以下で論議する。

【0264】

ポリクローナル抗体の産生のために、種々の適切な宿主動物(例えば、ウサギ、ヤギ、マウスまたはその他の哺乳動物)が、ネイティブタンパク質、またはその合成改変体、または前述の誘導体での注射により免疫され得る。適切な免疫原性調製物は、例えば、組換えにより発現されたGPCRXタンパク質または化学的に合成されたGPCRXポリペプチドを含有し得る。この調製物は、アジュバントをさらに含み得る。免疫学的応答を増大するために用いられる種々のアジュバントは、制限されずに、フロイントの完全および不完全アジュバント、ミネラルゲル(例えば、水酸化アルミニウム)、界面活性物質(例えば

、リゾレシチン、プルロニックポリオール、ポリアニオン、ペプチド、オイルエマルジョン、ジニトロフェノールなど）、Bacille Calmette-GuérinおよびCorynebacterium parvumのようなヒトアジュバント、または類似の免疫刺激剤を含む。所望であれば、GPCRXに対して惹起された抗体分子は、哺乳動物（例えば、血液）から単離され、そしてさらに、IgG画分を得るために、プロテインAクロマトグラフィーのような、周知の技術により精製され得る。

【0265】

本明細書で用いられる用語「モノクローナル抗体」または「モノクローナル抗体組成物」は、GPCRXの特定のエピトープと免疫反応し得る、1種類の抗原結合部位のみを含む抗体分子の集団をいう。従って、代表的には、モノクローナル抗体組成物は、それと免疫反応する特定のGPCRXタンパク質に対し、単一の結合親和性を示す。特定のGPCRXタンパク質、またはその誘導体、フラグメント、アナログもしくはホモログに対して惹起されたモノクローナル抗体の調製には、継続的な細胞株培養による抗体分子の産生を提供する任意の技術が利用され得る。このような技術は、制限されずに、ハイブリドーマ技術（例えば、Kohler & Milstein, 1975 Nature 256:495-497を参照のこと）；トリオーマ（trioma）技術；ヒトB細胞ハイブリドーマ技術（例えば、Kozborら、1983 Immunol. Today 4:72）およびヒトモノクローナル抗体を産生するためのEBVハイブリドーマ技術（例えば、Coleら、1985 MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc. 77-96頁を参照のこと）を含む。ヒトモノクローナル抗体を、本発明の実施において利用し得、そしてヒトハイブリドーマを用いることによるか（例えば、Coteら、1983. Proc Natl Acad Sci USA 80:2026-2030を参照のこと）、またはインビトロでヒトB細胞をエプスタイン-バーウイルスで形質転換することにより（例えば、Coteら、1985 MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc. 77-96頁を参照のこと）産生し得る。上記の引用の各々は、それらの全体が参考として本明細書に援用される。

【0266】

本発明によれば、技術は、GPCRXタンパク質に特異的な単鎖抗体の産生のために適合され得る（例えば、米国特許第4,946,778号を参照のこと）。さらに、方法は、Fab発現ライブラリーの構築のために適合されて（例えば、Huseら、1989 Science 246:1275-1281を参照のこと）、GPCRXタンパク質、またはその誘導体、フラグメント、アナログもしくはホモログに対する所望の特異性を有するモノクローナルFabフラグメントの迅速かつ有効な同定を可能にし得る。非ヒト抗体は、当該分野で周知の技術により「ヒト化」され得る。例えば、米国特許第5,225,539号を参照のこと。GPCRXタンパク質に対するイディオタイプを含む抗体フラグメントを、当該分野で公知の技術により産生し得、これには、制限されないで：(i) 抗体分子のペプシン消化により産生されるF(ab')₂フラグメント；(ii) F(ab')₂フラグメントのジスルフィド架橋を還元することにより産生されるFabフラグメント；(iii) 抗体分子のパパインおよび還元剤での処理により産生されるFabフラグメント；ならびに(iv) Fvフラグメントが含まれる。

【0267】

さらに、ヒトおよび非ヒト部分の両方を含むキメラ抗体およびヒト化モノクローナル抗体のような、組換え抗GPCRX抗体（これらは、標準的な組換えDNA技術を用いて作製され得る）は、本発明の範囲内である。このようなキメラ抗体およびヒト化モノクローナル抗体は、当該分野で公知の組換えDNA技術により産生され得る。この技術は例えば、以下に記載の方法を用いる：国際出願番号PCT/US86/02269；欧州特許出願番号184,187号；欧州特許出願番号171,496；欧州特許出願番号173,494；PCT国際公開番号WO 86/01533；米国特許第4,816,567号；米国特許第5,225,539号；欧州特許出願番号125,023号；Betterら

(1988) Science 240:1041~1043; Liuら (1987) Proc. Natl. Acad. Sci. USA 84:3439~3443; Liuら (1987) J. Immunol. 139:3521~3526; Sunら (1987) Proc. Natl. Acad. Sci. USA 84:214~218; Nishimuraら (1987) Cancer Res 47:999~1005; Woodら (1985) Nature 314:446~449; Shawら (1988) J. Natl. Cancer Inst. 80:1553~1559; Morrison (1985) Science 229:1202~1207; Oiら (1986) BioTechniques 4:214; Jonesら (1986) Nature 321:552~525; Verhoevenら (1988) Science 239:1534; ならびに Beidlerら (1988) J. Immunol. 141:4053~4060。上記引用の各々は、それらの全体において参考として本明細書中に援用される。 10

【0268】

1つの実施形態において、所望の特異性を保有する抗体のスクリーニングのための方法は、当該分野内で公知の酵素結合イムノソルベント検定法 (ELISA) および他の免疫学的に媒介された技術を含むが、これに限定されない。特定の実施形態において、GPCRXタンパク質の特定のドメインに対して特異的である抗体の選抜は、このようなドメインを所有しているGPCRXタンパク質のフラグメントに結合するハイブリドーマの生成により容易にされる。従って、GPCRXタンパク質、またはそれらの誘導体、フラグメント、アナログまたはホモログ内の所望のドメインについて特異的である抗体がまた、本明細書において提供される。 20

【0269】

抗GPCRX抗体は、GPCRXタンパク質の局在化および／または定量化に関する当該分野で公知の方法において用いられ得る (例えば、適切な生理学的なサンプル中のGPCRXタンパク質のレベルを測定する使用のために、診断的方法における使用のために、タンパク質の画像化における使用のために、など)。所定の実施形態において、GPCRXタンパク質、またはそれらの誘導体、フラグメント、アナログまたはホモログについての抗体 (抗体由来の結合ドメインを含む) は、薬理学的に活性な化合物 (本明細書中、以降において「治療剤」) として利用される。

【0270】

抗GPCRX抗体 (例えば、モノクローナル抗体) は、標準的技術 (例えばアフィニティークロマトグラフィまたは免疫沈降) によって、GPCRXポリペプチドを単離するために用いられ得る。抗GPCRX抗体は、細胞からの天然のGPCRXポリペプチド、そして宿主細胞において発現される組換え的に産生されたGPCRXポリペプチドの精製を容易にし得る。さらに、抗GPCRX抗体が、 (例えば、細胞の溶解産物または細胞上清における) GPCRXタンパク質を検出するために用いられて、GPCRXタンパク質の発現のアバндانسおよびパターンを評価し得る。抗GPCRX抗体は、例えば、所定の治療レジメンの有効性を決定するために、臨床試験の手順の一部として、組織におけるタンパク質レベルを診断的にモニターするために用いられ得る。検出は、抗体を検出可能物質に結合する (すなわち、物理的に連結する) ことにより容易にされ得る。検出可能物質の例としては、種々の酵素、補欠分子族、蛍光物質、発光物質、生物発光物質および放射性物質が挙げられる。適切な酵素の例としては、西洋ワサビペルオキシダーゼ、アルカリホスファターゼ、 β -ガラクトシダーゼまたはアセチルコリンエステラーゼが挙げられる; 適切な補欠分子族複合体の例としては、ストレプトアビジン/ビオチンおよびアビジン/ビオチンが挙げられる; 適切な蛍光物質の例としては、ウンベリフェロン、フルオレセイン、フルオレセインイソチオシアネート、ローダミン、ジクロロトリアジニルアミン (dichlorotriazinylamine) フルオレセイン、ダンシルクロリドまたはフィコエリトリンが挙げられる; 発光物質の例としては、ルミノールが挙げられる; 生物発光物質の例としては、ルシフェラーゼ、ルシフェリンおよびエクオリンが挙げられる; そして、適切な放射性物質の例としては ^{125}I 、 ^{131}I 、 ^{35}S または ^3H が挙げ 30 40 50

られる。

【0271】

(GPCRX組換え発現ベクターおよび宿主細胞)

本発明の別の局面は、GPCRXタンパク質、またはその誘導体、フラグメント、アナログもしくはホモログをコードする核酸を含む、ベクター、好ましくは、発現ベクターに関する。本明細書において使用する場合、用語「ベクター」は、連結された別の核酸を輸送し得る核酸分子をいう。1つの型のベクターは、「プラスミド」である。これはさらなるDNAセグメントが連結され得る環状の二本鎖DNAループをいう。他の型のベクターは、ウイルス性ベクターであり、ここでは、さらなるDNAセグメントが、ウイルスゲノムに連結され得る。特定のベクターは、ベクターが導入される宿主細胞中で自律複製し得る（例えば、細菌性の複製起点を有する細菌ベクター、およびエピソーム性哺乳動物ベクター）。他のベクター（例えば、非エピソーム性哺乳動物ベクター）は、宿主細胞への導入の際、宿主細胞のゲノムに組み込まれ、そして、これにより宿主ゲノムとともに複製される。さらに、特定のベクターは、それらが作動可能に連結される遺伝子の発現を指示し得る。このようなベクターは、本明細書において「発現ベクター」と呼ばれる。一般に、組換えDNA技術において有用な発現ベクターは、しばしばプラスミドの形態である。本明細書において、「プラスミド」および「ベクター」は、プラスミドが最も一般的に用いられる形態のベクターであるので、交換可能に使用され得る。しかし、本発明は、等価の機能を果たす、ウイルス性ベクター（例えば、複製欠損レトロウイルス、アデノウイルス、およびアデノ随伴ウイルス）のような、このような他の形態の発現ベクターを含むことを意図する。

【0272】

本発明の組換え発現ベクターは、宿主細胞における核酸の発現に適切な形態の本発明の核酸を含む。これは、組換え発現ベクターが、発現されるべき核酸配列に作動可能に連結されている、発現に用いられる宿主細胞に基づいて選択された1つ以上の調節配列を含むことを意味する。組換え発現ベクターにおいて、「作動可能に連結する」とは、目的のヌクレオチド配列が、（例えば、インビトロの転写／翻訳系において、またはベクターが宿主細胞に導入される場合は、宿主細胞において）ヌクレオチド配列の発現を可能にする様式で調節配列に連結されることを意味することを意図する。

【0273】

用語「調節配列」は、プロモーター、エンハンサーおよび他の発現制御エレメント（例えば、ポリアデニル化シグナル）を含むことを意図する。このような調節配列は、例えば、Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990)に記載されている。調節配列は、多くの型の宿主細胞においてヌクレオチド配列の構成的発現を指示する配列、および特定の宿主細胞のみにおいてヌクレオチド配列の発現を指示する配列（例えば、組織特異的調節配列）を含む。発現ベクターの設計が形質転換される宿主細胞の選択、所望のタンパク質の発現のレベルなどのような因子に依存し得ることは当業者には明白である。本発明の発現ベクターは、宿主細胞に導入され得、それにより、本明細書に記載される核酸によってコードされるタンパク質またはペプチド（融合タンパク質またはペプチドを含む）（例えば、GPCRXタンパク質、GPCRXタンパク質の変異形態、融合タンパク質など）を産生し得る。

【0274】

本発明の組換え発現ベクターは、原核生物細胞または真核生物細胞における、GPCRXタンパク質発現のために設計され得る。例えば、GPCRXタンパク質は、細菌細胞（例えば、Escherichia coli）、昆虫細胞（バキュロウイルス発現ベクターを用いる）、酵母細胞または哺乳動物細胞において発現され得る。適切な宿主細胞は、Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990)にさらに考察されている。あるいは、組換え型発現

ベクターは、例えば、T7プロモーター調節配列およびT7ポリメラーゼを用いて、インビトロで転写および翻訳され得る。

【0275】

原核生物におけるタンパク質の発現は、最も頻繁には、融合タンパク質または非融合タンパク質のいずれかの発現を指示する、構成的プロモーターまたは誘導性プロモーターを含むベクターを有する *Escherichia coli* において実施される。融合ベクターは、そこにコードされるタンパク質に、通常、組換えタンパク質のアミノ末端に、多数のアミノ酸を付加する。このような融合ベクターは、代表的には、以下の3つの目的のために役立つ：(i) 組換えタンパク質の発現を増加させるため；(ii) 組換えタンパク質の可溶性を増加させるため；および(iii) アフィニティー精製においてリガンドとして作用することによって組換えタンパク質の精製の際に補助するため。しばしば、融合発現ベクターにおいて、タンパク質分解の切断部位が、融合部分と組換えタンパク質との結合部に導入されて、融合タンパク質の精製の後に融合部分から組換えタンパク質を分離することを可能とする。このような酵素、およびその同族(cognate)の認識配列は、第Xa因子、トロンピン、およびエンテロキナーゼを含む。代表的な融合発現ベクターとしては、それぞれ、グルタチオンS-トランスフェラーゼ(GST)、マルトースE結合タンパク質、またはプロテインAを標的の組換えタンパク質に融合する、pGEX (Pharmacia Biotech Inc. ; SmithおよびJohnson (1988) Gene 67:31-40)、pMAL (New England Biolabs, Beverly, Mass.) およびpRIT5 (Pharmacia, Piscataway, N. J.) が挙げられる。

【0276】

適切な誘導性非融合 *E. coli* 発現ベクターの例には、pTrc (Amrannら (1988) Gene 69:301-315) およびpET 11d (Studierら、GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89) が挙げられる。

【0277】

E. coli における組換えタンパク質発現を最大化するための1つの戦略は、組換えタンパク質をタンパク質分解的に切断する能力を欠損した宿主細菌中でタンパク質を発現させることである。例えば、Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128を参照のこと。別の戦略は、各アミノ酸についての個々のコドンが、*E. coli* において優先的に利用されるコドンであるように発現ベクターに挿入される核酸の核酸配列を変更することである(例えば、Wadaら (1992) Nucleic Acids Res. 20:2111-2118を参照のこと)。本発明の核酸配列のこのような変更は、標準的なDNA合成技術によって実行され得る。

【0278】

別の実施形態において、GPCRX発現ベクターは、酵母発現ベクターである。酵母 *Saccharomyces cerevisiae* における発現のためのベクターの例には、pYepSec1 (Baldariら、(1987) EMBO J. 6:229-234)、pMFa (KurjanおよびHerskowitz、(1982) Cell 30:933-943)、pJRY88 (Schultzら、(1987) Gene 54:113-123)、pYES2 (Invitrogen Corporation, San Diego, Calif.)、およびpicZ (Invitrogen Corp., San Diego, Calif.) が挙げられる。

【0279】

あるいは、GPCRXは、バキュロウイルス発現ベクターを使用して、昆虫細胞中で発現され得る。培養昆虫細胞(例えば、SF9細胞)中でのタンパク質の発現のために利用可

能なバキュロウイルスベクターには、pAcシリーズ (Smithら (1983) Mol. Cell. Biol. 3: 2156-2165) および pVLシリーズ (Lucklow および Summers (1989) Virology 170: 31-39) が挙げられる。

【0280】

なお別の実施形態において、本発明の核酸は、哺乳動物発現ベクターを使用して、哺乳動物細胞中で発現される。哺乳動物発現ベクターの例には、pCDM8 (Seed (1987) Nature 329: 840) および pMT2PC (Kaufmanら (1987) EMBO J. 6: 187-195) が挙げられる。哺乳動物細胞中で使用される場合、発現ベクターの制御機能は、しばしば、ウイルスの調節エレメントによって提供される。例えば、一般に使用されるプロモーターは、ポリオーマ、アデノウイルス2、サイトメガロウイルス、およびシミアンウイルス40に由来する。原核生物細胞および真核生物細胞の両方のために適切な他の発現系については、例えば、Sambrookら、MOLECULAR CLONING: A LABORATORY MANUAL, 第2版、Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y., 1989の第16章および第17章を参照のこと。

【0281】

別の実施形態において、組換え哺乳動物発現ベクターは、特定の細胞型 (例えば、組織特異的調節エレメントを使用して核酸を発現する) 中で優先的に核酸の発現を指示し得る。組織特異的調節エレメントは、当該分野で公知である。適切な組織特異的なプロモーターの非限定的な例には、アルブミンプロモーター (肝臓特異的; Pinkertら (1987) Genes Dev. 1: 268-277)、リンパ特異的プロモーター (Calame および Eaton (1988) Adv. Immunol. 43: 235-275)、T細胞レセプターの特定のプロモーターにおいて (Winoto および Baltimore (1989) EMBO J. 8: 729-733) および免疫グロブリンの特定のプロモーターにおいて (Banerjiら (1983) Cell 33: 729-740; Queen および Baltimore (1983) Cell 33: 741-748)、ニューロン特異的プロモーター (例えば、神経フィラメントプロモーター; Byrne および Ruddie (1989) Proc. Natl. Acad. Sci. USA 86: 5473-5477)、膵臓特異的プロモーター (Edlundら (1985) Science 230: 912-916)、ならびに乳腺特異的プロモーター (例えば、乳清プロモーター; 米国特許第4, 873, 316号および欧州出願公開第264, 166号) が挙げられる。発達的に調節されるプロモーターもまた含まれる (例えば、マウス hox プロモーター (Kessel および Gruss (1990) Science 249: 374-379) および a-フェトプロテインプロモーター (Campes および Tilghman (1989) Genes Dev. 3: 537-546))。

【0282】

本発明はさらに、アンチセンス方向で発現ベクターにクローニングされた本発明のDNA分子を含む組換え発現ベクターを提供する。すなわち、そのDNA分子は、GPCRX mRNA に対してアンチセンスであるRNA分子の発現 (DNA分子の転写によって) を可能にする様式で調節配列に作動可能に連結される。種々の細胞型においてアンチセンスRNA分子の持続的な発現を指示する、アンチセンス方向でクローニングされた核酸に作動可能に連結された調節配列 (例えば、ウイルスプロモーターおよび/もしくはエンハンサー) が選択され得るか、または、アンチセンスRNAの構成的発現、組織特異的発現、または細胞特異的発現を指示する調節配列が、選択され得る。アンチセンス発現ベクターは、組換えプラスミド、ファージミド、または弱毒化されたウイルスの形態であり得、ここではアンチセンス核酸は、高効率調節領域の制御下で産生され、その活性は、ベクターが導入される細胞型によって決定され得る。アンチセンス遺伝子を使用する遺伝子発現の調節の考察については、例えば、Weintraubら、「Antisense RNA

as a molecular tool for genetic analysis」、Reviews-Trends in Genetics、第1巻(1)1986を参照のこと。

【0283】

本発明の別の局面は、本発明の組換え発現ベクターが導入された宿主細胞に関する。用語「宿主細胞」および「組換え宿主細胞」は、本明細書中で、交換可能に使用される。このような用語は、特定の被験細胞をいうのみでなく、そのような細胞の子孫または潜在的な子孫をもいうことが理解される。変異または環境の影響のいずれかに起因して、特定の改変は次の世代において生じ得るので、このような子孫は、実際には、親の細胞と同一でないかもしれないが、なお、本明細書中で使用される場合のこの用語の範囲内に含まれる。 10

【0284】

宿主細胞は、任意の原核生物細胞または真核生物細胞であり得る。例えば、GPCRXタンパク質は、細菌細胞（例えば、E. coli）、昆虫細胞、酵母または哺乳動物細胞（例えば、チャイニーズハムスター卵巣細胞（CHO）またはCOS細胞）において発現され得る。他の適切な宿主細胞は、当業者に公知である。

【0285】

ベクターDNAは、従来の形質転換またはトランスフェクション技術を介して原核生物細胞または真核生物細胞に導入され得る。本明細書中で使用される場合、用語「形質転換」および「トランスフェクション」とは、外来性の核酸（例えば、DNA）を宿主細胞に導入するために当該分野で容認されている種々の技術をいうことを意図し、これらには、リン酸カルシウムまたは塩化カルシウム共沈、DEAEデキストラン媒介トランスフェクション、リポフェクション、またはエレクトロポレーションが含まれる。宿主細胞を形質転換またはトランスフェクションするための適切な方法は、Sambrookら（MOLECULAR CLONING: A LABORATORY MANUAL、第2版、Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y., 1989）および他の実験室マニュアルに見出され得る。 20

【0286】

安定な哺乳動物細胞のトランスフェクションのために、使用される発現ベクターおよびトランスフェクション技術に依存して、細胞のほんの一部のみが外来DNAをそのゲノムに組み込み得ることが知られている。これらの要素を同定および選択するために、選択マーカー（例えば、抗生物質に対する耐性）をコードする遺伝子が、一般的には目的の遺伝子とともに宿主細胞に導入される。種々の選択マーカーには、薬物（例えば、G418、ハイグロマイシン、およびメトトレキサート）に対する耐性を付与するものが含まれる。選択マーカーをコードする核酸は、GPCRXをコードするベクターと同じベクター上で宿主細胞に導入され得るか、あるいは別々のベクター上で導入され得る。導入された核酸で安定にトランスフェクトされた細胞は、薬物選択によって同定され得る（例えば、選択マーカー遺伝子を取り込んだ細胞は生存するが、他の細胞は死滅する）。 30

【0287】

本発明の宿主細胞（例えば、培養中の原核生物宿主細胞および真核生物宿主細胞）は、GPCRXタンパク質を産生（すなわち、発現）するために使用され得る。従って、本発明はさらに、本発明の宿主細胞を使用して、GPCRXタンパク質を産生するための方法を提供する。1つの実施形態において、本発明の方法は、GPCRXタンパク質が産生されるような適切な培地中で、本発明の宿主細胞（ここに、GPCRXタンパク質をコードする組換え発現ベクターが導入されている）を培養する工程を包含する。別の実施形態において、本発明はさらに、培地または宿主細胞からGPCRXタンパク質を単離する工程を包含する。 40

【0288】

（トランスジェニックGPCRX動物）

本発明の宿主細胞はまた、非ヒトトランスジェニック動物を作製するために使用され得る 50

。例えば、1つの実施形態において、本発明の宿主細胞は、GPCRXタンパク質コード配列が導入された、受精卵母細胞または胚性幹細胞である。次いで、このような宿主細胞は、外因性のGPCRX配列がゲノムに導入された非ヒトトランスジェニック動物、または内因性のGPCRX配列が変更された相同組換え動物を作製するために使用され得る。このような動物は、GPCRXタンパク質の機能および／または活性を研究するため、ならびにGPCRXタンパク質活性のモジュレーターを同定および／または評価するために有用である。本明細書中で使用される場合、「トランスジェニック動物」とは、非ヒト動物であり、好ましくは哺乳動物であり、より好ましくは、ラットまたはマウスのような齧歯類であり、ここでは、その動物の1つ以上の細胞が、導入遺伝子を含む。トランスジェニック動物の他の例としては、非ヒト霊長類、ヒツジ、イヌ、ウシ、ヤギ、ニワトリ、両生類などが挙げられる。導入遺伝子は、細胞（この細胞からトランスジェニック動物を発生させた）のゲノムに組み込まれかつ成熟動物のゲノムに残存し、それによって、トランスジェニック動物の1つ以上の細胞型または組織においてコードされた遺伝子産物の発現を指示する、外因性のDNAである。本明細書中で使用される場合、「相同（homologous）組換え動物」とは、非ヒト動物であり、好ましくは哺乳動物であり、より好ましくはマウスであり、ここでは内因性のGPCRX遺伝子は、内因性の遺伝子と、その動物の発生の前にその動物の細胞（例えば、その動物の胚細胞）に導入された外因性のDNA分子との間の相同組換えによって変更されている。

【0289】

本発明のトランスジェニック動物は、GPCRXをコードする核酸を、例えば、マイクロインジェクション、レトロウイルス感染によって、受精卵母細胞の雄性前核に導入すること、およびその卵母細胞を偽妊娠雌性養育動物（foster animal）中で発達させることを可能にすることによって作製され得る。配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38のヒトのGPCRX cDNA配列は、非ヒト動物のゲノムに導入遺伝子として導入され得る。あるいは、ヒトのGPCRX遺伝子の非ヒトホモログ（例えば、マウスのGPCRX遺伝子）は、ヒトのGPCRX cDNAに対するハイブリダイゼーションに基づいて単離され得（上記にさらに記載された）、そして導入遺伝子として使用され得る。イントロン配列およびポリアデニル化シグナルもまた導入遺伝子中に含まれ得、その導入遺伝子の発現の効率を増大させ得る。組織特異的調節配列（単数または複数）は、特定の細胞に対して、GPCRXタンパク質の発現を指示するように、GPCRX導入遺伝子に作動可能に連結され得る。胚の操作およびマイクロインジェクションを介するトランスジェニック動物（特に、マウスのような動物）を作製するための方法は、当該分野で慣習的になっており、そして例えば、米国特許第4,736,866号；同第4,870,009号；および同第4,873,191号；ならびにHogan 1986、MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. に記載されている。同様の方法が、他のトランスジェニック動物の作製のために使用される。トランスジェニック創始（founder）動物は、そのゲノムにおけるGPCRX導入遺伝子の存在および／またはその動物の組織もしくは細胞中のGPCRX mRNAの発現に基づいて同定され得る。次いで、トランスジェニック創始動物は、導入遺伝子を有するさらなる動物を育種するために使用され得る。さらに、GPCRXタンパク質をコードする導入遺伝子を有するトランスジェニック動物は、さらに、他の導入遺伝子を有する他のトランスジェニック動物に交雑され得る。

【0290】

相同組換え動物を作製するために、GPCRX遺伝子の少なくとも一部を含むベクターを調製する。ここで、この遺伝子は、欠失、付加、または置換が導入され、それによってそのGPCRX遺伝子が増えられている（例えば、機能的に破壊されている）。GPCRX遺伝子はヒト遺伝子（例えば、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38のcDNA）であり得る

が、より好ましくは、ヒトのGPCRX遺伝子の非ヒトホモログである。例えば、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38のヒトのGPCRX遺伝子のマウスホモログは、マウスゲノムにおける内因性のGPCRX遺伝子を変更するのに適切な相同組換えベクターを構築するために使用され得る。1つの実施形態において、そのベクターは、相同組換えに際して、その内因性のGPCRX遺伝子が、機能的に破壊されるように設計される（すなわち、機能的タンパク質をもはやコードしない；「ノックアウト」ベクターともいわれる）。

【0291】

あるいは、このベクターは、相同組換えに際して、内因性GPCRX遺伝子が発現されるか、あるいはさもなければ、変更されるが、なお機能的タンパク質をコードするように、設計され得る（例えば、上流の調節領域が変更され、それによって内因性GPCRXタンパク質の発現を変更し得る）。相同組換えベクターにおいて、GPCRX遺伝子の変更された部分は、GPCRX遺伝子のさらなる核酸によって、その5'末端および3'末端で隣接されることにより、相同組換えが、ベクターによって保有される外因性GPCRX遺伝子と胚幹細胞中の内因性GPCRX遺伝子との間で起こることを可能にする。さらなる隣接するGPCRX核酸は、内因性遺伝子との首尾よい相同組換えのために十分な長さである。典型的に、数キロベースの隣接DNA（5'末端および3'末端の両方における）が、ベクターに含まれる。例えば、相同組換えベクターの説明についてThomasら（1987）Cell 51:503を参照のこと。次いで、このベクターは、（例えば、エレクトロポレーションによって）胚性幹細胞株に導入され、そして導入されたGPCRX遺伝子が内因性GPCRX遺伝子と相同組換えした細胞が、選択される（例えば、Lira（1992）Cell 69:915を参照のこと）。

【0292】

次いで、選択された細胞が、動物（例えば、マウス）の未分化胚芽細胞へ注入され、凝集キメラを形成する。例えば、Bradley（1987）のTeratocarcinomas and Embryonic Stem Cells: A Practical Approach、Robertson編、IRL、Oxford、113頁～152頁を参照のこと。次いで、キメラ胚が、適切な偽妊娠雌性養育動物に移植され得、そしてその胚は分娩に到る。その生殖細胞中に相同組換えされたDNAを保有する子孫が、動物を育種するために使用され得、ここでこの動物の全ての細胞は、導入遺伝子の生殖系列伝達によって、相同組換えされたDNAを含む。相同組換えベクターおよび相同組換え動物を構築するための方法が、さらに以下に記載される；Bradley（1991）Curr. Opin. Biotechnol. 2:823-829；PCT国際公開番号：WO90/11354；WO91/01140；WO92/0968；およびWO/93/04169。

【0293】

別の実施形態において、導入遺伝子の調節された発現を可能にする選択された系を含む非ヒトトランスジェニック動物が産生され得る。このような系の1つの例は、バクテリオファージP1のcre/loxPリコンビナーゼ系である。cre/loxPリコンビナーゼ系の説明については、例えば、Laksora（1992）Proc. Natl. Acad. Sci. USA 89:6232-6236を参照のこと。リコンビナーゼ系の別の例は、Saccharomyces cerevisiaeのFLPリコンビナーゼ系である（O'Gormanら（1991）Science 251:1351-1355を参照のこと）。cre/loxPリコンビナーゼ系が導入遺伝子の発現を調節するために使用される場合、Creリコンビナーゼおよび選択されたタンパク質の両方をコードする導入遺伝子を含む動物が必要となる。このような動物は、例えば、2つのトランスジェニック動物（一方は選択されたタンパク質をコードした導入遺伝子を含み、他方はリコンビナーゼをコードした導入遺伝子を含む）を交配することによる「二重の」トランスジェニック動物の構築によって提供され得る。

【0294】

本明細書中に記載される非ヒトトランスジェニック動物のクローンがまた、W i l m u t t ら (1 9 9 7) N a t u r e 3 8 5 : 8 1 0 - 8 1 3 に記載される方法に従って、産生され得る。簡単には、トランスジェニック動物からの細胞 (例えば、体細胞) が単離され得、そして増殖サイクル (g r o w t h c y c l e) から出て、G₀ 期に入るように誘導され得る。次いで、静止性細胞が、例えば、電気パルスの使用により、その静止性細胞が単離されたのと同種の動物から摘出された卵母細胞へ融合され得る。次いで、この再構築された卵母細胞は培養され、これにより、これは桑実胚または未分化胚芽細胞に発達し、次いで、偽妊娠雌性養育動物に移される。この雌性養育動物から産まれた子孫は、その細胞 (例えば、その体細胞) が単離された動物のクローンである。

【0295】

10

(薬学的組成物)

本発明の G P C R X 核酸分子、G P C R X タンパク質、および抗 G P C R X 抗体 (これはまた、本明細書中で、「活性化化合物」とも呼ばれる)、ならびにそれらの誘導體、フラグメント、アナログ、およびホモログが、投与に適した薬学的組成物に組み込まれ得る。このような組成物は、典型的に、核酸分子、タンパク質または抗体、および薬学的に受容可能なキャリアを含む。本明細書中で使用される場合、「薬学的に受容可能なキャリア」は、薬学的な投与に適合可能な、あらゆる溶媒、分散媒、コーティング、抗菌剤および抗真菌剤、等張剤および吸収遅延剤 (d e l a y i n g a g e n t) などを含むことが意図される。適切なキャリアは、R e m i n g t o n ' s P h a r m a c e u t i c a l S c i e n c e s (当該分野の標準的参考文献であって、本明細書中に参考として援用される) の最新版に記載される。このようなキャリアまたは希釈剤の好ましい例としては、水、生理食塩水、フィンガー (f i n g e r) 溶液、デキストロース溶液、および 5 % ヒト血清アルブミンが挙げられるが、これらに限定されない。リポソームおよび非水性ビヒクル (例えば、不揮発性油) もまた、使用され得る。薬学的に活性な物質に対するこのような媒体および薬剤の使用は、当該分野で周知である。任意の従来の媒体または薬剤が活性化化合物と不適合である場合を除いて、組成物におけるその使用が意図される。補助的な活性化化合物もまた、組成物へ組み込まれ得る。

20

【0296】

本発明の薬学的組成物は、その意図される投与経路と適合可能であるように処方される。投与経路の例には、非経口 (例えば、静脈内、皮内、皮下) 投与、経口 (例えば、吸入) 投与、経皮 (すなわち、局所的) 投与、経粘膜 (t r a n s m u c o s a l) 投与、および直腸投与が挙げられる。非経口適用、皮内適用または皮下適用のために使用される溶液または懸濁液は、以下の成分を含み得る：注射用水、生理食塩水溶液、不揮発性油、ポリエチレングリコール、グリセリン、プロピレングリコールまたは他の合成溶媒のような、滅菌希釈剤；ベンジルアルコールまたはメチルパラベンのような、抗菌剤；アスコルビン酸または重亜硫酸ナトリウムのような、抗酸化剤；エチレンジアミン四酢酸 (E D T A) のような、キレート剤；酢酸塩、クエン酸塩またはリン酸塩のような、緩衝液、および塩化ナトリウムまたはデキストロースのような、張度の調整のための薬剤。p H は、塩酸または水酸化ナトリウムのような酸または塩基で調整され得る。非経口調製物は、ガラスまたはプラスチック製のアンプル、使い捨てシリンジ、または多用量のバイアル中に入れられ得る。

30

40

【0297】

注入用途に適した薬学的組成物は、滅菌水溶液 (ここでは、水溶解性) または分散媒 (d i s p e r s i o n)、および滅菌注入可能溶液または分散媒の即座の調製のための滅菌粉末を含む。静脈内投与において、適切なキャリアには、生理食塩水、静菌水、C r e m o p h o r E L^{T M} (B A S F、P a r s i p p a n y、N. J.) またはリン酸緩衝化生理食塩水 (P B S) が挙げられる。全ての場合において、組成物は、無菌でなければならず、そして容易な注入性 (s y r i n g e a b i l i t y) が存在する程度に流動性であるべきである。これは、製造および保存の条件下で安定でなければならず、そして、細菌および真菌のような微生物の汚染行為に対して保護されなければならない。このキャ

50

リアは、例えば、以下を含む溶媒または分散媒であり得る：水、エタノール、ポリオール（例えば、グリセロール、プロピレングリコール、および液体ポリエチレングリコールなど）ならびにそれらの適切な混合物。適切な流動性が、例えば、レシチンのようなコーティングの使用によって、分散媒の場合には、要求される粒子サイズを維持することによって、および界面活性剤を使用することによって維持され得る。微生物の作用の防止は、種々の抗菌剤および抗真菌剤（例えば、パラベン、クロロブタノール、フェノール、アスコルビン酸、チメロサルなど）によって達成され得る。多くの場合、組成物中に等張剤（例えば、砂糖、ポリアルコール（例えば、マンニトール（*m a n i t o l*）、ソルビトール）、塩化ナトリウム）を含むことが好ましい。注入可能組成物の吸収期間の延長は、組成物に吸収を遅延させる薬剤（例えば、モノステアリン酸アルミニウムおよびゼラチン）を含ませることによってもたらされ得る。

10

【0298】

滅菌注射可能溶液は、必要量のこの活性化合物（例えば、GPCRXタンパク質または抗GPCRX抗体）を、適切な溶媒中で、上記で列挙される成分の1つまたは組み合わせと共に組み込み、必要な場合、続いて濾過滅菌することによって調製され得る。一般的に、分散媒は、活性化合物を、基本（*b a s i c*）の分散媒と上記で列挙される成分から必要とされる他の成分とを含む滅菌ビヒクル中へ組み込むことによって調製される。滅菌注射可能液剤の調製のための滅菌粉末の場合において、調製方法は、真空乾燥および凍結乾燥であり、これにより、予め滅菌濾過されたその溶液から活性成分および任意のさらなる所望の成分の粉末を得る。

20

【0299】

経口組成物は、一般的に、不活性希釈剤または食用キャリアを含む。これらは、ゼラチンカプセルに封入され得るか、または錠剤へと圧縮され得る。経口治療投与の目的のために、この活性化合物は、賦形剤とともに組み込まれ得、そして錠剤、トローチ剤、またはカプセル剤の形態で使用され得る。経口組成物はまた、うがい薬としての使用のために流体キャリアを使用して調製され得、ここで、この流体キャリア中のこの化合物は、経口的に適用され、そして音を立てられ（*s w i s h*）、そして吐き出されるか、または飲み込まれる。薬学的に適合性の結合剤、および／またはアジュバント物質が、この組成物の一部として含まれ得る。錠剤、丸剤、カプセル剤、トローチ剤などが、任意の以下の成分または同様の性質を有する化合物を含み得る：結合剤（例えば、微結晶セルロース、ガムトラガントまたはゼラチン）；賦形剤（例えば、デンプンまたはラクトース）、崩壊剤（例えば、アルギン酸、プリモゲル（*P r i m o g e l*）、またはコーンスターチ）；滑沢剤（例えば、ステアリン酸マグネシウムまたはステロテス（*S t e r o t e s*））；潤滑剤（*g l i d a n t*）（例えば、コロイド状二酸化ケイ素）；甘味剤（例えば、スクロースまたはサッカリン）；あるいは香味剤（例えば、ペパーミント、サリチル酸メチル、またはオレンジフレーバー）。

30

【0300】

吸入による投与について、この化合物は、適切な噴霧剤（例えば、二酸化炭素のような気体）を含む圧縮容器またはディスペンサー、あるいは噴霧器から、エアロゾル噴霧の形態で送達される。

40

【0301】

全身的投与はまた、経粘膜手段または経皮手段により得る。経粘膜投与または経皮投与について、浸透される障壁に適切な浸透剤が、処方物において使用される。このような浸透剤は、一般的に、当該分野で公知であり、そして、例えば、経粘膜投与については、界面活性剤、胆汁酸塩、およびフシジン酸誘導体が挙げられる。経粘膜投与は、鼻噴霧または坐剤の使用によって達成され得る。経皮投与については、この活性化合物は、当該分野で一般的に公知である軟膏剤（*o i n t m e n t*）、軟膏（*s a l v e*）、ゲル、またはクリーム剤へ処方される。

【0302】

この化合物はまた、直腸送達のための坐剤（例えば、ココアバターおよび他のグリセリド

50

のような従来の坐剤基剤と共に) または貯留 (r e n t e n t i o n) 浣腸の形態で調製され得る。

【0303】

1つの実施形態において、この活性化合物は、身体からの迅速な排出に対してこの化合物を保護するキャリアを用いて調製され(例えば、制御放出処方物)、これには、移植片およびマイクロカプセル化された送達系が挙げられる。酢酸エチレンビニル、ポリ無水物、ポリグリコール酸、コラーゲン、ポリオルトエステル、およびポリ乳酸のような、生分解性、生体適合性ポリマーが使用され得る。このような処方物の調製のための方法は、当業者には明らかである。これらの物質はまた、Alza Corporation and GPCR X a Pharmaceuticals, Inc. から商業的に入手され得る。リボソーム懸濁剤(ウイルス抗原に対するモノクローナル抗体を感染させた細胞へ標的化されるリボソームを含む)がまた、薬学的に受容可能なキャリアとして使用され得る。これらは、例えば米国特許第4, 522, 811号に記載されるような、当業者に公知の方法に従って調製され得る。

10

【0304】

投与の容易さおよび投薬量の均一性のために、投薬単位形態で、経口組成物または非経口組成物を処方することが、特に有益である。本明細書で使用される投薬単位形態は、処置される被験体に対する単位投薬量として適切な物理的に個々の単位をいい;各単位は、必要とされる薬学的キャリアに関連して所望の治療的効果を生じるように計算された、所定量の活性化合物を含む。本発明の投薬単位形態についての詳細は、この活性化合物の固有の特徴、および達成される特定の治療効果、ならびに個体の処置のためのこのような活性化合物を調合する分野に固有の制限によって決定されるか、あるいはこれらに直接依存する。

20

【0305】

本発明の核酸分子は、ベクターに挿入され得、そして遺伝子治療ベクターとして使用され得る。遺伝子治療ベクターは、被験体へ、例えば静脈内注射、局所投与(米国特許第5, 328, 470号を参照のこと)によって、または定位注射(例えば、Chenら(1994) Proc. Natl. Acad. Sci. USA 91:3054-3057を参照のこと)によって送達され得る。遺伝子治療ベクターの薬学的調製物には、受容可能な希釈剤中の遺伝子治療ベクターが挙げられ得るか、または遺伝子送達ビヒクルが組み込まれる除放性マトリックスが挙げられ得る。あるいは、完全な遺伝子送達ベクターが組換え細胞からインタクトで産生され得(例えば、レトロウイルスベクター)、薬学的調製物は、遺伝子送達系を産生する1以上の細胞を含み得る。

30

【0306】

薬学的組成物は、投与のための使用説明書と共に、容器、包装、またはディスペンサーに含まれ得る。

【0307】

(スクリーニングおよび検出の方法)

本発明の単離された核酸分子は、以下にさらに説明されるように、GPCRXタンパク質(例えば、遺伝子治療適用における宿主細胞中の組換え発現ベクターを介する)を発現するため、GPCRX mRNA(例えば、生物学的サンプルにおいて)またはGPCRX遺伝子における遺伝子損傷を検出するため、ならびにGPCRX活性を調節するために使用され得る。さらに、GPCRXタンパク質は、GPCRXタンパク質の活性または発現を調節する薬物または化合物をスクリーニングするために、ならびにGPCRXタンパク質の不十分もしくは過剰な産生によって、あるいはGPCRX野生型タンパク質と比較して減少した活性もしくは異常な活性を有するGPCRXタンパク質形態の産生によって特徴付けられる障害を処置するために使用され得る。さらに、本発明の抗GPCRX抗体が、GPCRXタンパク質を検出および単離するため、ならびにGPCRX活性を調節するために使用され得る。

40

【0308】

50

本発明は、さらに、本明細書中に記載のスクリーニングアッセイによって同定される新規の薬剤、および上記で本明細書中で記載されるような処置のためのそれらの使用に関する。

【0309】

(スクリーニングアッセイ)

本発明は、調節因子、すなわち、GPCRXタンパク質に結合するか、あるいは例えば、GPCRXタンパク質の発現またはGPCRXタンパク質の活性に対して刺激効果または阻害効果を有する、候補または試験化合物または薬剤（例えば、ペプチド、ペプチド模倣物、低分子または他の薬物）を同定するための方法（本明細書中において「スクリーニングアッセイ」とも称される）を提供する。本発明はまた、本明細書中に記載のスクリーニングアッセイにおいて同定された化合物を含む。

【0310】

1実施形態において、本発明は、GPCRXタンパク質またはポリペプチド、あるいはその生物学的に活性な部分の膜結合形態に結合するか、またはその膜結合形態の活性を調節する、候補化合物もしくは試験化合物をスクリーニングするためのアッセイを提供する。本発明の試験化合物は、当該分野において公知のコンビナトリアルライブラリーにおける多数のアプローチのいずれかを使用して得られ得、これには、以下が挙げられる：生物学的ライブラリー；空間的に位置付け可能な並行固相もしくは溶液相ライブラリー；逆重畳を要する合成ライブラリー法；「1ビーズ1化合物」ライブラリー法；およびアフィニティークロマトグラフィー選択を使用する合成ライブラリー法。生物学的ライブラリーアプローチはペプチドライブラリーに限定されるが、他の4つのアプローチは、ペプチド、非ペプチドオリゴマーもしくは化合物の低分子ライブラリーに適用可能である（例えば、Lam (1997) *Anticancer Drug Design* 12:145を参照のこと）。

【0311】

本明細書中で使用される場合、「低分子」とは、約5kD未満の分子量、最も好ましくは約4kD未満の分子量を有する組成物をいうように意味される。低分子は、例えば、核酸、ペプチド、ポリペプチド、ペプチド模倣体、炭水化物、脂質または他の有機分子もしくは無機分子であり得る。化学的および／または生物学的混合物（例えば、真菌、細菌または藻類抽出物）のライブラリーは、当該分野で公知であり、そして本発明のアッセイのいずれかを用いてスクリーニングされ得る。

【0312】

分子ライブラリーの合成のための方法の例は、当該分野において、例えば以下に見出され得る：DeWittら (1993) *Proc. Natl. Acad. Sci. USA* 90:6909；Erbら (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91:11422；Zuckermannら (1994) *J. Med. Chem.* 37:2678；Chora (1993) *Science* 261:1303；Carrellら (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059；Carrellら (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061；およびGallopら (1994) *J. Med. Chem.* 37:1233。

【0313】

化合物のライブラリーは、溶液中で（例えば、Houghten (1992) *Biotechniques* 13:412~421）、あるいはビーズ上（Lam (1991) *Nature* 354:82~84）、チップ上（Fodor (1993) *Nature* 364:555~556）、細菌（Ladner 米国特許第5,223,409号）、孢子（Ladner 米国特許第5,233,409号）、プラスミド（Cullら (1992) *Proc. Natl. Acad. Sci. USA* 89:1865~1869）またはファージ上（ScottおよびSmith (1990) *Science* 249:386~390；Devlin (1990) *Science* 249:404~406；Cwirllaら (1990) *Proc. Natl. Acad. Sci. USA* 87:6

378～6382; Felici (1991) J. Mol. Biol. 222:301～310; Ladner、米国特許第5,233,409号)において提示され得る。

【0314】

1つの実施形態において、アッセイは細胞ベースのアッセイであり、ここで、膜結合形態のGPCRXタンパク質、またはその生物学的に活性な部分を細胞表面上に発現する細胞が、試験化合物と接触され、そしてこの試験化合物が、GPCRXタンパク質に結合する能力が、決定される。例えば、細胞は、哺乳動物起源または酵母細胞であり得る。この試験化合物がGPCRXタンパク質に結合する能力の決定は、例えば、その試験化合物を放射性同位体または酵素標識とカップリングさせて、その結果、この試験化合物のGPCRXタンパク質またはその生物学的に活性な部分に対する結合が、複合体におけるその標識化合物を検出することによって決定され得ることによって達成され得る。例えば、試験化合物は、 ^{125}I 、 ^{35}S 、 ^{14}C 、または ^3H で直接的または間接的のいずれかで標識され得、そしてその放射性同位体が、放射線放射の直接の計数により、またはシンチレーション計数により、検出され得る。あるいは、試験化合物は、例えば、西洋ワサビペルオキシダーゼ、アルカリホスファターゼ、またはルシフェラーゼで酵素的に標識され得、そしてこの酵素的標識が、適切な基質の生成物への転換を決定することにより、検出され得る。1つの実施形態において、このアッセイは、膜結合形態のGPCRXタンパク質、またはその生物学的に活性な部分をその細胞表面上に発現する細胞を、GPCRXと結合する既知の化合物と接触させて、アッセイ混合物を形成する工程、このアッセイ混合物に試験化合物を接触させる工程、ならびにこの試験化合物がGPCRXタンパク質と相互作用する能力を決定する工程を包含し、ここで、この試験化合物がGPCRXタンパク質と相互作用する能力を決定する工程によって、この試験化合物が、既知の化合物と比較して、GPCRXまたはその生物学的に活性な部分と優先的に結合する能力を決定する工程を包含する。

【0315】

別の実施形態において、アッセイは、細胞ベースのアッセイであり、これは、膜結合形態のGPCRXタンパク質またはその生物学的に活性な部分を細胞表面上で発現する細胞を、試験化合物と接触させる工程、ならびにこの試験化合物が、GPCRXタンパク質またはその生物学的に活性な部分の活性を調節（例えば、刺激または阻害）する能力を決定する工程を包含する。この試験化合物が、GPCRXまたはその生物学的に活性な部分の活性を調節する能力の決定は、例えば、GPCRXタンパク質が、GPCRX標的分子に結合またはこれら標的分子と相互作用する能力を決定することによって、達成され得る。本明細書中において使用する場合には、「標的分子」とは、GPCRXタンパク質が天然に結合または相互作用する分子であり、例えば、GPCRX相互作用タンパク質を発現する細胞表面上の分子、第二の細胞表面上の分子、細胞外環境中の分子、細胞膜の内面と会合する分子、または細胞質分子である。GPCRX標的分子は、非GPCRX分子あるいは本発明のGPCRXタンパク質またはポリペプチドであり得る。1つの実施形態において、GPCRX標的分子は、シグナル伝達経路の構成要素であり、これは、細胞膜を通過して細胞内への細胞外シグナル（例えば、化合物が膜結合GPCRX分子に結合することにより発生するシグナル）の伝達を促進する。その標的は、例えば、触媒活性を有する第二の細胞内タンパク質、または下流シグナル伝達分子のGPCRXとの会合を容易にするタンパク質であり得る。

【0316】

GPCRXタンパク質がGPCRX標的分子に結合するかまたはその標的分子と相互作用する能力の決定は、直接的結合を決定するための上記方法の1つにより、達成され得る。1つの実施形態において、GPCRXタンパク質がGPCRX標的分子に結合するかまたはその標的分子と相互作用する能力の決定は、その標的分子の活性を決定することにより、達成され得る。例えば、この標的分子の活性は、その標的の細胞性セカンドメッセンジャー（すなわち、細胞内 Ca^{2+} 、ジアシルグリセロール、IP3など）の誘導を検出すること、適切な基質への標的の触媒活性／酵素活性を検出すること、レポーター遺伝子（

検出可能なマーカー（例えば、ルシフェラーゼ）をコードする核酸に作動可能に連結された GPCR X 応答性調節エレメントを含む）の誘導を検出すること、または細胞応答（例えば、細胞生存度、細胞分化、または細胞増殖）を検出することにより、決定され得る。

【0317】

なお別の実施形態において、本発明のアッセイは、無細胞（cell-free）アッセイであり、GPCR X タンパク質またはその生物学的に活性な部分を試験化合物に接触させる工程、ならびにその試験化合物が GPCR X タンパク質またはその生物学的に活性な部分と結合する能力を決定する工程を包含する。この試験化合物の、GPCR X タンパク質への結合は、上記のように、直接的または間接的にのいずれかで決定され得る。1つのこのような実施形態において、このアッセイは、GPCR X タンパク質またはその生物学的に活性な部分を、GPCR X を結合する既知の化合物に接触させて、アッセイ混合物を形成する工程、このアッセイ混合物を試験化合物に接触させる工程、ならびにその試験化合物が GPCR X タンパク質と相互作用する能力を決定する工程を包含し、ここで、この試験化合物が GPCR X タンパク質と相互作用する能力を決定する工程は、この試験化合物が、既知の化合物と比較して、GPCR X、またはその生物学的に活性な部分と優先的に結合する能力を決定する工程を包含する。

10

【0318】

なお別の実施形態において、アッセイは、無細胞アッセイであり、GPCR X タンパク質またはその生物学的に活性な部分を試験化合物と接触させる工程、ならびにその試験化合物が GPCR X タンパク質またはその生物学的に活性な部分の活性を調節（例えば、刺激または阻害）する能力を決定する工程を包含する。この試験化合物が GPCR X の活性を調節する能力の決定は、例えば、GPCR X タンパク質が、GPCR X 標的分子に結合する能力を、直接的結合の決定のための上記方法の1つによって決定することにより、達成され得る。代替の実施形態において、この試験化合物が GPCR X タンパク質の活性を調節する能力の決定は、GPCR X タンパク質が、GPCR X 標的分子をさらに調節する能力を決定することにより、達成され得る。例えば、この標的分子の適切な基質に対する触媒活性／酵素活性は、上記のように決定され得る。

20

【0319】

さらに別の実施形態において、無細胞アッセイは、GPCR X タンパク質またはその生物学的に活性な部分を、GPCR X タンパク質を結合する既知の化合物に接触させてアッセイ混合物を形成する工程、このアッセイ混合物を試験化合物と接触させる工程、ならびにこの試験化合物が GPCR X タンパク質と相互作用する能力を決定する工程を包含し、ここで、この試験化合物が GPCR X タンパク質と相互作用する能力の決定は、GPCR X タンパク質が、GPCR X 標的分子と優先的に結合するか、またはその標的分子の活性を優先的に調節する能力を決定する工程を包含する。

30

【0320】

本発明の無細胞アッセイは、GPCR X タンパク質の、可溶性の形態または膜結合形態の両方の使用に受け入れられる。膜結合形態の GPCR X タンパク質を含む無細胞アッセイの場合には、GPCR X タンパク質の膜結合形態が溶液中に維持されるように、可溶化剤を利用することが所望され得る。このような可溶化剤の例には、非イオン性界面活性剤が挙げられ、例えば、n-オクタシルグルコシド、n-ドデシルグルコシド、n-ドデシルマルトシド、オクタノイル-N-メチルグルカミド、デカノイル-N-メチルグルカミド、Triton（登録商標）X-100、Triton（登録商標）X-114、Thesit（登録商標）、イソトリデシルポリ（エチレングリコールエーテル）_n（Isotridecyl poly（ethylene glycol ether）_n、N-ドデシル-N，N-ジメチル-3-アンモニオ-1-プロパンスルホネート、3-（3-コラミドプロピル）ジメチルアンミニオール-1-プロパンスルホネート（3-（3-cholamidopropyl）dimethylammonio-1-propane sulfonate）（CHAPS）、または3-（3-コラミドプロピル）ジメチルアンミニオール-2-ヒドロキシ-1-プロパンスルホネート（3-（3-cholamido

40

50

propyl) dimethylamminiol-2-hydroxy-1-propane sulfonate) (CHAPS O) である。

【0321】

本発明の上記アッセイ方法の1つより多い実施形態において、GPCRXタンパク質またはその標的分子のいずれかを固定して、そのタンパク質の一方または両方の非複合形態からの複合形態の分離を促進し、そしてそのアッセイの自動化に適合させることが、所望され得る。試験化合物の、GPCRXタンパク質への結合、または候補化合物の存在下および非存在下での、GPCRXタンパク質の標的分子との相互作用は、これらの反応物を収容するために適切な任意の容器内で、達成され得る。このような容器の例には、マイクロタイタープレート、試験管、および微小遠心管が挙げられる。1実施形態において、そのタンパク質の一方または両方がマトリックスに結合することを可能にするドメインを付加する融合タンパク質が、提供され得る。例えば、GST-GPCRX融合タンパク質またはGST標的融合タンパク質は、グルタチオンセファロースビーズ (Sigma Chemical, St. Louis, MO) またはグルタチオン誘導体化マイクロタイタープレート上に吸着され得、次いでこれらは、試験化合物と合わせられるか、あるいは試験化合物および未吸着標的タンパク質またはGPCRXタンパク質のいずれかと合わせられ、そしてこの混合物が、複合体形成を誘導する条件下 (例えば、塩およびpHに関して生理学的条件) でインキュベートされる。インキュベーション後、ビーズまたはマイクロタイタープレートウェルを洗浄して、結合していないあらゆる成分を除去し、ビーズの場合にはマトリックスを固定し、例えば上記のように、複合体を直接的または間接的のいずれかで決定する。あるいは、複合体がマトリックスから解離され得、そしてGPCRXタンパク質の結合レベルまたは活性レベルを、標準的な技術を使用して決定し得る。

10

20

【0322】

タンパク質をマトリックスに固定するための他の技術もまた、本発明のスクリーニングアッセイにおいて使用され得る。例えば、GPCRXタンパク質、またはその標的分子のいずれかが、ビオチンとストレプトアビジンとの結合を利用して、固定され得る。ビオチン化GPCRXタンパク質、または標的分子は、当該分野において周知の技術を使用して、ビオチン-NHS (N-ヒドロキシスクシンイミド) から調製され得 (例えば、ビオチン化キット、Pierce Chemicals, Rockford, Ill.)、そしてストレプトアビジンでコーティングした96ウェルのプレート (Pierce Chemical) のウェルに固定され得る。あるいは、GPCRXタンパク質、または標的分子と反応性であるがGPCRXタンパク質のその標的分子への結合を妨害しない抗体が、そのプレートのウェルに誘導体化され得、そして結合していない標的またはGPCRXタンパク質が、抗体の結合によってウェル内に捕捉され得る。このような複合体を検出するための方法には、GST固定複合体に関しての上記で述べた方法に加えて、GPCRXタンパク質または標的分子と反応性の抗体を使用する、複合体の免疫検出、ならびにGPCRXタンパク質、または標的分子に関する酵素活性の検出に依存する酵素結合アッセイが挙げられる。

30

【0323】

別の実施形態において、GPCRXタンパク質発現のモジュレーターは、細胞を候補化合物と接触させ、そして細胞中のGPCRX mRNAまたはタンパク質の発現を決定する方法において同定される。候補化合物の存在下でのGPCRX mRNAまたはタンパク質の発現レベルは、候補化合物の非存在下でのGPCRX mRNAまたはタンパク質の発現レベルと比較される。次いで、候補化合物は、この比較に基づいて、GPCRX mRNAまたはタンパク質発現のモジュレーターとして同定され得る。例えば、GPCRX mRNAまたはタンパク質の発現が候補化合物の非存在下より、その存在下における方が大きい (すなわち、統計的に有意に大きい) 場合、この候補化合物は、GPCRX mRNAまたはタンパク質の発現の刺激物質として同定される。あるいは、GPCRX mRNAまたはタンパク質の発現が候補化合物の非存在下よりその存在下の方が少ない (統計的に有意に少ない) 場合、この候補化合物は、GPCRX mRNAまたはタンパク質

40

50

の発現のインヒビターとして同定される。細胞中の G P C R X m R N A またはタンパク質の発現レベルは、G P C R X m R N A またはタンパク質を検出するために本明細書中に記載の方法によって決定され得る。

【0324】

本発明のなお別の局面において、G P C R X タンパク質は、ツーハイブリッドアッセイまたはスリーハイブリッドアッセイにおいて「餌 (b a i t) タンパク質」として使用されて (例えば、米国特許第 5, 283, 317 号; Z e r v o s ら (1993) C e l l 72:223-232; M a d u r a ら、1993 J. B i o l. C h e m. 268:12046-12054; B a r t e l ら、1993 B i o t e c h n i q u e s 14:920-924; I w a b u c h i ら、1993 O n c o g e n e 8:1693-1696; および B r e n t W O 94/10300 を参照のこと)、G P C R X (「G P C R X 結合タンパク質」または「G P C R X - b p」) に結合するか、またはこれと相互作用し、そして G P C R X 活性を調節する他のタンパク質を同定し得る。このような G P C R X 結合タンパク質はまた、例えば、G P C R X 経路の上流または下流エレメントとして G P C R X タンパク質によるシグナル伝達に参与する可能性がある。

【0325】

ツーハイブリッドシステムは、分離可能な D N A 結合ドメインおよび活性化ドメインからなる、大部分の転写因子のモジュール的性質に基づく。簡潔には、このアッセイは、2つの異なる D N A 構築物を利用する。一方の構築物においては、G P C R X をコードする遺伝子が既知の転写因子 (例えば、G A L - 4) の D N A 結合ドメインをコードする遺伝子に融合される。他方の構築物においては、D N A 配列のライブラリー由来の、未同定タンパク質 (「餌食 (p r e y)」または「サンプル」) をコードする D N A 配列が、既知の転写因子の活性化ドメインをコードする遺伝子に融合される。「餌」および「餌食」タンパク質がインビボで相互作用して、G P C R X 依存性複合体を形成し得る場合、この転写因子の D N A 結合ドメインおよび活性化ドメインは、非常に近くにある。近位にあることにより、転写因子に応答性の転写調節部位に作動可能に連結されたレポーター遺伝子 (例えば、L a c Z) の転写が可能となる。レポーター遺伝子の発現が検出され得、そして機能的転写因子を含む細胞コロニーが、単離され得、そして G P C R X と相互作用するタンパク質をコードするクローニングされた遺伝子を得るために使用され得る。

【0326】

本発明はさらに、上記のスクリーニングアッセイにより同定される新規な薬剤および本明細書中に記載されるような処置のためのこの薬剤の使用に関する。

【0327】

(検出アッセイ)

本明細書中で同定される c D N A 配列の一部またはフラグメント (および対応する完全な遺伝子配列) は、ポリヌクレオチド試薬として多くの方法で使用され得る。例えば、これらの配列を使用して、(i) 染色体上にそれぞれの遺伝子をマッピングし得; 従って遺伝病と関連する遺伝子領域を位置決定し得る; (i i) 微量の生物学的サンプルから個体を同定し得る (組織型決定); および (i i i) 生物学的サンプルの法医学的識別を助け得るが、これに限定されない。これらの適用は、以下の節において記載される。

【0328】

(染色体マッピング)

一旦遺伝子の配列 (または配列の一部) が単離されると、この配列を用いて染色体上に遺伝子の位置をマッピングし得る。このプロセスは、染色体マッピングとよばれる。従って、本明細書中に記載の G P C R X 配列の一部またはフラグメント、配列番号 1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および 38、またはそのフラグメントもしくは誘導体を用いて、それぞれ、G P C R X 遺伝子の位置を染色体上にマッピングし得る。G P C R X 配列を染色体にマッピングすることは、これらの配列と、疾患と関連する遺伝子とを相関付ける際の重要な第一歩である。

【0329】

簡潔には、G P C R X 遺伝子は、G P C R X 配列から P C R プライマー（好ましくは、15～25bpの長さ）を調製することにより染色体にマッピングし得る。G P C R X 配列のコンピューター分析を用いて、ゲノムDNAにおける1つを超えるエキソンにまたがらないプライマーを迅速に選択し得、従って、増幅プロセスを複雑にし得る。次いで、これらのプライマーは、個々のヒト染色体を含む体細胞ハイブリッドのP C R スクリーニングのために使用され得る。G P C R X 配列に対応するヒト遺伝子を含むハイブリッドのみが増幅されたフラグメントを生じる。

【0330】

体細胞ハイブリッドは、異なる哺乳動物由来の体細胞（例えば、ヒトおよびマウス細胞）を融合することにより調製される。ヒトおよびマウスのハイブリッドが増殖および分裂するにつれ、それらは、無作為な順序で徐々にヒト染色体を失うが、マウス染色体を維持する。マウス細胞は増殖できないが、ヒト細胞は増殖できる培地を使用することにより、それらは特定の酵素を失うので、必要な酵素をコードする遺伝子を含む1つのヒト染色体が維持される。種々の培地を使用することによって、ハイブリッド細胞株のパネルが確立され得る。パネルにおける各細胞株は、単一のヒト染色体または少数のヒト染色体いずれかおよびマウス染色体の完全なセットを含み、これにより、個々の遺伝子を特定のヒト染色体にマッピングすることが容易に可能になる（D' E u s t a c h i o ら（1983）S c i e n c e 220:919-924を参照のこと）。ヒト染色体のフラグメントのみを含む体細胞ハイブリッドはまた、転座および欠失を伴うヒト染色体を使用することにより生成され得る。

【0331】

体細胞ハイブリッドのP C R マッピングは、特定の染色体に特定の配列を割り当てるためには迅速な手順である。1つのサーマルサイクラーを用いて一日に3つ以上の配列が割り当てられ得る。G P C R X 配列を使用して、オリゴヌクレオチドプライマーを設計すると、下位位置決定（s u b l o c a l i z a t i o n）は、特定の染色体由来のフラグメントのパネルを用いて達成され得る。

【0332】

中期染色体スプレッドに対するDNA配列の蛍光インサイチュハイブリダイゼーション（F I S H）をさらに使用して、一工程で正確な染色体位置を提供し得る。染色体スプレッドは、コルセミド（染色体紡錘体を破壊する）のような化学物質により分裂が中期においてブロックされた細胞を使用して行われ得る。この染色体は、トリプシンで短時間処理され得、次いで、ギムザ染色され得る。薄いバンドおよび濃いバンドのパターンが各染色体で発生し、その結果、この染色体は、個々に同定され得る。F I S H 技術は、500または600塩基ほどの長さのDNA配列と共に使用され得る。しかし、1,000塩基より大きなクローンは、簡単な検出に十分なシグナル強度で、独特の染色体位置に結合する可能性がより高い。好ましくは、1,000塩基、そしてより好ましくは、2,000塩基が妥当な時間量で良好な結果を得るために十分である。この技術の総説については、V e r m a ら、H U M A N C H R O M O S O M E S : A M A N U A L O F B A S I C T E C H N I Q U E S （P e r g a m o n P r e s s , N e w Y o r k 1988）を参照のこと。

【0333】

染色体マッピングの試薬は、単一の染色体またはその染色体上の単一部位を印付けするために個々に使用され得るか、または試薬のパネルは複数部位および／または複数染色体を印付けするために使用され得る。遺伝子の非コード領域に対応する試薬は、実際に、マッピング目的のために好ましい。コード配列は、遺伝子ファミリー内に保存されており、従って、染色体マッピングの間に交叉ハイブリダイゼーションの機会が増大する可能性が高い。

【0334】

一旦配列が正確な染色体位置にマッピングされると、その配列の染色体上の物理的位置は

、遺伝子マップデータと相関し得る。このようなデータは、例えば、McKusick, MENDELIAN INHERITANCE IN MAN (Johns Hopkins University Welch Medical Libraryを通じてオンラインで入手可能)において見いだされる。次いで、同じ染色体領域にマッピングされる遺伝子と疾患との間の関係は、例えば、Egelandら(1987) Nature, 325:783-787に記載される連鎖分析(物理的に隣接する遺伝子の同時遺伝)によって同定され得る。

【0335】

さらに、GPCRX遺伝子と関連する疾患に罹患した個体と、この疾患に罹患していない個体との間のDNA配列における差異が決定され得る。罹患した個体のいくらかまたは全てにおいて変異が認められるが、非罹患個体のいずれにおいても認められない場合、この変異は特定の疾患の候補薬剤である可能性が高い。罹患個体と非罹患個体の比較は、一般に、染色体における構造的変化(例えば、染色体スプレッドから可視であるか、またはそのDNA配列に基づいたPCRを用いて検出可能な欠失または転座)を最初に探すことを包含する。最終的に、いくらかの個体に由来する遺伝子の完全な配列決定を行って、変異の存在を確認し、かつ多型に由来する変異を区別する。

10

【0336】

(組織型決定(tissue typing))

本発明のGPCRX配列はまた、わずかな生物学的サンプルから個体を識別するために使用され得る。この技術において、個体のゲノムDNAは、1つ以上の制限酵素で消化され、そして同定のために独特のバンドを生成するためにサザンブロット上でプローブされる。本発明の配列は、RFLP(米国特許第5,272,057号に記載の「制限フラグメント長多型」)のためのさらなるDNAマーカーとして有用である。

20

【0337】

さらに、本発明の配列を用いて、個体のゲノムの選択された部分について実際の塩基ごとにDNA配列を決定する代替的技術を提供し得る。従って、本明細書中に記載のGPCRX配列を用いて、配列の5'末端および3'末端から2つのPCRプライマーを調製し得る。次いで、これらのプライマーを使用して、個体のDNAを増幅し得、引き続いて、配列決定し得る。

【0338】

このように調製された個体由来の対応するDNA配列のパネルは、各個体が、対立遺伝子差異に起因するこのようなDNA配列の独特のセットを有するので、唯一の個体識別を提供し得る。本発明の配列は、個体由来および組織由来の配列のこのような識別を得るために使用され得る。本発明のGPCRX配列は、ヒトゲノムの部分を独特に表す。対立遺伝子変異は、これらの配列のコード領域においてある程度生じ、そして非コード領域においてより大きな程度に生じる。個々のヒト間での対立遺伝子変異は、各500塩基につき約1回の頻度で生じると見積もられる。対立遺伝子変異の多さは、制限フラグメント長多型(RFLP)を含む単一ヌクレオチド多型(SNP)に起因する。

30

【0339】

本明細書中で記載の配列の各々は、ある程度、標準物質(これに対して個体からのDNAが識別の目的で比較され得る)として使用され得る。より多くの多型が非コード領域で生じるので、個体を区別するために、それほど多くの配列が必要であるわけではない。非コード配列は、おそらく10~1,000プライマーのパネルを用いてポジティブな個体識別を不自由なく提供し得る。これらのプライマーは、各々が100塩基の増幅された非コード配列を生じる。推定コード配列(例えば、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38における配列)が使用される場合、ポジティブな個体識別に関するプライマーのより適切な数は、500~2,000である。

40

【0340】

(予測医療)

50

本発明はまた、診断アッセイ、予後アッセイ、薬物ゲノム (p h a r m a c o g e n o m i c s) およびモニタリング臨床試験が、予後 (予測) の目的に使用され、これによって個体を予防的に処置する、予測医療の分野に関する。従って、本発明の1つの局面は、G P C R Xタンパク質および／または核酸の発現、ならびにG P C R Xの活性を、生物学的サンプル (例えば、血液、血清、細胞、組織) の関連で決定するための診断アッセイに関し、これによって、異常なG P C R Xの発現または活性に関連して、個体が疾患または障害に罹患するかどうか、あるいは障害を発症するリスクがあるかどうかを決定する。本発明はまた、個体が、G P C R Xのタンパク質、核酸の発現または活性と関連した障害を発症するリスクがあるかどうかを決定するための予後的 (または予測的) アッセイを提供する。例えば、G P C R Xの遺伝子における変異が、生物学的サンプルにおいてアッセイされ得る。このようなアッセイは、予後的または予測的な目的に使用され得、これによってG P C R Xのタンパク質、核酸の発現または生物学的活性によって特徴付けられるかまたは関連した障害の発病の前に個体を予防的に処置する。

10

【0341】

本発明の別の局面は、個体におけるG P C R Xのタンパク質、核酸の発現あるいは活性を決定するための方法を提供し、これによって、その個体についての適切な治療的または予防的試薬 (本明細書において「薬物ゲノム」とよばれる) を選択する。薬物ゲノムは、個体の遺伝型 (例えば、特定の試薬に対して応答する個体の能力を決定するために試験された個体の遺伝型) に基づいて個体の治療的または予防的処置のための試薬 (例えば、薬物) の選択を可能にする。

20

【0342】

本発明のなお別の局面は、臨床試験におけるG P C R Xの発現または活性に対する試薬 (例えば、薬物、化合物) の影響をモニタリングすることに関する。

【0343】

これらおよび他の試薬は、以下の節でさらに詳細に記載される。

【0344】

(診断アッセイ)

生物学的サンプルにおけるG P C R Xの存在または非存在を検出するための例示的な方法は、試験被験体から生物学的サンプルを得る工程、およびその生物学的サンプルをG P C R Xのタンパク質またはG P C R Xタンパク質をコードする核酸 (例えば、m R N A、ゲノムD N A) を検出し得る化合物もしくは薬剤とを接触させ、その結果、G P C R Xの存在が、その生物学的サンプルにおいて検出される、工程を包含する。G P C R Xのm R N AもしくはゲノムD N Aを検出するための薬剤は、G P C R X m R N AもしくはゲノムD N Aにハイブリダイズし得る、標識された核酸プローブである。この核酸プローブは、例えば、全長のG P C R Xの核酸 (例えば、配列番号1、3、5、7、9、11、13、16、18、20、22、24、28、30、32、34、36、および38の核酸またはそれらの一部) (例えば、少なくとも、15、30、50、100、250もしくは500ヌクレオチド長のオリゴヌクレオチドであり、ストリンジェントな条件下でG P C R Xのm R N AまたはゲノムD N Aと特異的にハイブリダイズするに十分である核酸) であり得る。本発明の診断アッセイにおける使用のための他の適切なプローブは本明細書において記載されている。

30

40

【0345】

G P C R Xのタンパク質を検出するための薬剤は、G P C R Xのタンパク質に結合し得る抗体であり、好ましくは、検出可能な標識を有する抗体である。抗体は、ポリクローナルであり得るか、またはより好ましくはモノクローナル抗体であり得る。インタクトな抗体またはそのフラグメント (例えば、F a bまたはF (a b ')₂) が使用され得る。用語「標識 (された)」とは、プローブまたは抗体に関して、検出可能な物質をそのプローブもしくは抗体にカップリングさせる (すなわち、物理的に連結する) ことによってそのプローブまたは抗体を直接標識すること、ならびに、直接標識された別の試薬との反応性によってそのプローブもしくは抗体の間接的な標識をすることを包含することが意図される

50

。間接的な標識の例としては、蛍光標識された二次抗体を用いた一次抗体の検出、および蛍光標識されたストレプトアビジンを用いて検出され得るようにDNAプローブのビオチンを用いた末端標識が挙げられる。用語「生物学的サンプル」とは、被験体から単離された、組織、細胞および生物学的流体ならびに被験体に存在する組織、細胞および流体を含むことが意図される。すなわち、本発明の検出方法を用いて、GPCRXのmRNA、タンパク質またはゲノムDNAを、生物学的サンプル中で、インビトロおよびインビボで検出し得る。例えば、GPCRXのmRNAの検出のためのインビトロ技術としては、ノーザンハイブリダイゼーションおよびインサイチュハイブリダイゼーションが挙げられる。GPCRXのタンパク質の検出のためのインビトロ技術としては、酵素連結免疫吸着アッセイ(ELISA)、ウェスタンブロット、免疫沈降および免疫蛍光が挙げられる。GPCRXのゲノムDNAを検出するためのインビトロ技術としては、サザンハイブリダイゼーションが挙げられる。さらに、GPCRXのタンパク質の検出のためのインビボ技術としては、標識された抗GPCRX抗体を被験体に導入することが挙げられる。例えば、その抗体は、放射性マーカーを用いて標識され得る。被験体における放射性マーカーの存在および位置は、標準的な画像化技術によって検出され得る。

10

【0346】

1つの実施形態において、この生物学的サンプルは、その試験被験体からのタンパク質分子を含む。あるいは、その生物学的サンプルは、その試験被験体からのmRNA分子またはその試験被験体からのゲノムDNA分子を含み得る。好ましい生物学的サンプルは、被験体から従来の手段によって単離された末梢血白血球サンプルである。

20

【0347】

別の実施形態において、本発明の方法はさらに、コントロール被験体からコントロール生物学的サンプルを得る工程、そのコントロールサンプルを、GPCRXのタンパク質、mRNAもしくはゲノムDNAを検出し得る化合物もしくは薬剤と接触させ、その結果、GPCRXのタンパク質、mRNAもしくはゲノムDNAの存在がその生物学的サンプルにおいて検出される、工程、およびそのコントロールサンプルにおけるGPCRXのタンパク質、mRNAもしくはゲノムDNAの存在と、その試験サンプルにおけるGPCRXのタンパク質、mRNAもしくはゲノムDNAの存在とを比較する工程を包含する。

【0348】

本発明はまた、生物学的サンプルにおけるGPCRXの存在を検出するためのキットを包含する。例えば、このキットは、以下を備え得る：生物学的サンプルにおいてGPCRXのタンパク質またはmRNAを検出し得る、標識された化合物もしくは薬剤；そのサンプルにおいてGPCRXの量を決定するための手段；およびそのサンプルにおいて、標準と、GPCRXの量とを比較するための手段。この化合物または薬剤は、適切な容器内に包装され得る。このキットは、さらに、GPCRXのタンパク質または核酸を検出するためにキットを用いるための指示書を備え得る。

30

【0349】

(予後アッセイ)

本明細書において記載された診断方法をさらに利用して、GPCRXの異常発現または異常活性に関連した疾患もしくは障害を有するか、またはその発症の危険性を有する被験体を同定し得る。例えば、本明細書に記載されるアッセイ(例えば、上述の診断アッセイまたは下記のアッセイ)を利用して、GPCRXのタンパク質、核酸の発現または活性に関連する障害を有するかまたはその発症の危険性を有する被験体を同定し得る。あるいは、この予後アッセイを利用して、疾患または障害を有するかまたはその発症の危険性を有する被験体を同定し得る。従って、本発明は、GPCRXの異常発現または異常活性に関連する疾患もしくは障害を同定するための方法を提供する。ここで、試験サンプルは、被験体から得られ、そしてGPCRXのタンパク質または核酸(例えば、mRNA、ゲノムDNA)が検出され、ここで、GPCRXのタンパク質または核酸の存在は、GPCRXの異常発現または異常活性に関連する疾患または障害を有するかまたはその発症の危険性を有する被験体についての診断指標である。本明細書において使用される場合「試験サンプ

40

50

ル」とは、目的の被験体から得られた生物学的サンプルをいう。例えば、試験サンプルは、生物学的流体（例えば、血清）、細胞サンプル、または組織であり得る。

【0350】

さらに、本明細書に記載される予後アッセイを使用して、被験体に薬剤（例えば、アゴニスト、アンタゴニスト、ペプチド模倣物、タンパク質、ペプチド、核酸、低分子、または他の薬物候補）を投与してGPCRXの異常発現または異常活性に関連する疾患または障害を処置し得るか否かを、決定し得る。例えば、このような方法を使用して、被験体が障害のための薬剤で有効に処置され得るか否かを決定し得る。従って、本発明は、GPCRXの異常発現または異常活性に関連する障害についての薬剤を用いて、被験体が有効に処置され得るか否かを決定するための方法を提供する。ここで、試験サンプルが得られ、そしてGPCRXのタンパク質または核酸が検出される（例えば、ここで、GPCRXのタンパク質または核酸の存在は、この薬剤が投与されてGPCRXの異常発現または異常活性に関連する障害が処置され得る被験体についての、診断指標である）。 10

【0351】

本発明の方法はまた、GPCRX遺伝子における遺伝的損傷を検出し、それによって、その損傷遺伝子を有する被験体が異常な細胞増殖および／または分化によって特徴付けられる障害についての危険性を有するか否かを決定するためにも使用され得る。種々の実施形態において、この方法は、その被験体からの細胞のサンプルにおいて、GPCRXタンパク質をコードする遺伝子の統合性に影響を与える少なくとも1つの変更によって特徴付けられる遺伝的損傷の存在または非存在、あるいはGPCRX遺伝子の誤発現を検出する工程を包含する。例えば、そのような遺伝的損傷は、以下の少なくとも1つの存在を確認することによって検出され得る：（i）GPCRX遺伝子からの1つ以上のヌクレオチドの欠失；（ii）GPCRX遺伝子への1つ以上のヌクレオチドの付加；（iii）GPCRX遺伝子の1つ以上のヌクレオチドの置換、（iv）GPCRX遺伝子の染色体再配置；（v）GPCRX遺伝子のメッセンジャーRNA転写物のレベルにおける変更、（vi）GPCRX遺伝子の異常改変（例えば、ゲノムDNAのメチル化パターンの異常改変）、（vii）GPCRX遺伝子のメッセンジャーRNA転写物の非野生型スプライシングパターンの存在、（viii）GPCRXタンパク質の非野生型レベル、（ix）GPCRX遺伝子の対立遺伝子の欠失、ならびに（x）GPCRXタンパク質の不適切な翻訳後修飾。本明細書において記載されるように、当該分野において、GPCRX遺伝子における損傷を検出するために使用され得る、多数の公知のアッセイ技術が存在する。好ましい生物学的サンプルは、従来手段によって被験体から単離された末梢血白血球サンプルである。しかし、有核細胞を含む任意の生物学的サンプルが使用され得、これには、例えば、頬粘膜細胞が挙げられる。 20

【0352】

特定の実施形態において、損傷の検出は、ポリメラーゼ連鎖反応（PCR）（例えば、米国特許第4,683,195号および同第4,683,202号を参照のこと）（例えば、アンカーPCRまたはRACE PCR）、あるいは、連結連鎖反応（LCR）（例えば、Landegranら（1988）Science 241:1077-1080；およびNakazawaら（1994）Proc. Natl. Acad. Sci. USA 91:360-364を参照のこと）におけるプローブ／プライマーの使用を包含する。 40
後者は、GPCRX遺伝子における点変異を検出するために特に有用であり得る（Abravayaら、1995 Nucleic Acids Res 23:675-682を参照のこと）。この方法は、患者から細胞のサンプルを収集する工程、核酸（例えば、ゲノム、mRNAまたはその両方）をそのサンプルの細胞から単離する工程、GPCRXの遺伝子に特異的にハイブリダイズする1つ以上のプライマーとその核酸サンプルとを、GPCRX遺伝子（存在する場合）のハイブリダイゼーションおよび増幅が生じるような条件下で接触させる工程、ならびに増幅産物の存在もしくは非存在を検出する工程、またはその増幅産物のサイズを検出する工程およびその長さをコントロールサンプルと比較する工程を包含し得る。PCRおよび／またはLCRは、本明細書中に記載される変異を検出 50

するために使用される技術のいずれかとともに予備的増幅工程として使用されることが望ましくあり得ることが予想される。

【0353】

代替的な増幅方法としては、以下が挙げられる：当業者に周知な技術を用いた、その増幅された分子の検出の前の、自己維持配列複製（Guatelliら、1990、Proc. Natl. Acad. Sci. USA 87:1874-1878を参照のこと）、転写増幅系（Kwohら、1989、Proc. Natl. Acad. Sci. USA 86:1173-1177を参照のこと）、Qレプリカーゼ（Lizardiら、1988、BioTechnology 6:1197を参照のこと）、または他の任意の核酸増幅方法。これらの検出スキームは、核酸分子が非常に極少数で存在する場合に、そのような核酸分子の検出のために特に有用である。

10

【0354】

代替の実施形態において、サンプル細胞からのGPCRX遺伝子における変異は、制限酵素切断パターンにおける変更によって同定され得る。例えば、サンプルおよびコントロールのDNAが単離され、増幅され（必要に応じて）、1つ以上の制限エンドヌクレアーゼを用いて消化され、そしてフラグメント長の大きさがゲル電気泳動によって決定され、そして比較される。サンプルDNAとコントロールDNAとの間のフラグメント長の大きさにおける差異は、そのサンプルDNAにおける変異を示す。さらに、配列特異的なリボザイムの使用（例えば、米国特許第5,493,531号を参照のこと）を使用して、リボザイム切断部位の発生または喪失によって特異的な変異の存在についてスコア付けし得る。

20

【0355】

他の実施形態において、GPCRXにおける遺伝子変異は、サンプル核酸およびコントロール核酸（例えば、DNAまたはRNA）を、数百または数千のオリゴヌクレオチドプローブを含む高密度アレイに対してハイブリダイズさせることによって同定され得る（Croninら（1996）Human Mutation 7:244-255；Kozalら（1996）Nat. Med. 2:753-759を参照のこと）。例えば、GPCRXにおける遺伝子変異は、Croninら（前出）に記載されるように光生成DNAプローブを含む二次元アレイにおいて同定され得る。手短には、プローブの第一ハイブリダイゼーションアレイを用いて、サンプルおよびコントロールにおける長いストレッチのDNAにわたって走査し、連続的に重複するプローブの線形アレイを作成することによって、その配列間の塩基変化を同定し得る。この工程は、点変異の同定を可能にする。この工程に第二のハイブリダイゼーションアレイが続き、これは、検出される全ての改変体または変異体に相補的な、より小さな特化されたプローブアレイを用いることによる特定の変異の特徴付けを可能にする。各変異アレイは、一方が野生型遺伝子に対して相補的であり、そして他方が変異遺伝子に対して相補である並行プローブセットから構成される。

30

【0356】

なお別の実施形態において、当該分野で公知の種々の配列決定反応のいずれかを使用して、GPCRX遺伝子を直接配列決定し得、そしてサンプルGPCRX配列と対応する野生型（コントロール）配列とを比較することによって、変異を検出し得る。配列決定反応の例としては、MaxamおよびGilbert（1977）Proc. Natl. Acad. Sci. USA 74:560またはSanger（1977）Proc. Natl. Acad. Sci. USA 74:5463によって開発された技術に基づくものが挙げられる。診断アッセイを実施する場合、種々の自動化配列決定手順のいずれかを利用し得ることもまた意図される（例えば、Naevveら（1995）Biotechniques 19:448を参照のこと）。これらには、質量分析法による配列決定法（例えば、PCT国際公開番号WO 94/16101；Cohenら（1996）Adv. Chromatography 36:127-162；およびGriffinら（1993）Appl. Biochem. Biotechnol. 38:147-159を参照のこと）が含まれる。

40

50

【0357】

GPCR X 遺伝子における変異を検出するための他の方法としては、切断剤からの保護を使用して、RNA/RNAもしくはRNA/DNAのヘテロ二重鎖に基づくミスマッチ塩基を検出する方法が挙げられる (Myersら (1985) Science 230:1242を参照のこと)。一般に、「ミスマッチ切断」の当該分野の技術は、野生型のGPCR X 配列を含む (標識された) RNAまたはDNAを、組織サンプルから得られた潜在的な変異体RNAまたはDNAとハイブリダイズさせることによって形成されるヘテロ二重鎖を提供する工程によって始まる。この二本鎖の二重鎖を、二重鎖の一本鎖領域 (例えば、そのコントロールとサンプルの鎖との間の塩基対ミスマッチに起因して存在するもの) を切断する薬剤を用いて処理する。例えば、RNA/DNA二重鎖を、RNAseを用いて処理し得、そしてDNA/DNAハイブリッドを、そのミスマッチ領域を酵素的に消化することに対して、S₁ヌクレアーゼを用いて処理し得る。他の実施形態において、DNA/DNAまたはRNA/DNAのいずれかの二重鎖を、ミスマッチ領域を消化するために、ヒドロキシルアミンまたは四酸化オスミウム、およびピペリジンを用いて処理し得る。そのミスマッチ領域の消化後、次いで、得られた物質を変性ポリアクリルアミドゲル上で、大きさにより分離して、変異の部位を決定する。例えば、Cottonら (1988) Proc Natl Acad Sci USA 85:4397; Saleebaら (1992) Methods Enzymol 217:286-295を参照のこと。1つの実施形態において、コントロールのDNAまたはRNAは、検出のために標識され得る。

10

20

【0358】

なお別の実施形態において、ミスマッチ切断反応は、二本鎖DNAにおけるミスマッチ塩基対を認識する1つ以上のタンパク質 (いわゆる「DNAミスマッチ修復」酵素) を、細胞のサンプルから得られたGPCR X cDNAにおける点変異を検出およびマッピングするために規定された系において使用する。例えば、E. coliのmutY酵素は、G/AミスマッチでAを切断し、そしてHeLa細胞からのチミジンDNAグリコシラーゼは、G/TミスマッチでTを切断する (Hsuら (1994) Carcinogenesis 15:1657~1662)。例示的な実施形態に従って、GPCR X 配列 (例えば、野生型GPCR X 配列) に基づくプローブは、試験細胞由来のcDNAまたは他のDNA産物にハイブリダイズされる。二重鎖は、DNAミスマッチ修復酵素を用いて処理され、そしてその切断産物 (もしあれば) は、電気泳動プロトコルなどから検出され得る。例えば、米国特許第5,459,039号を参照のこと。

30

【0359】

他の実施形態において、電気泳動の移動度における変化は、GPCR X 遺伝子における変異を同定するために使用される。例えば、一本鎖コンホメーション多型 (SSCP) は、変異体と野生型核酸との間の電気泳動の移動度における差異を検出するために使用され得る (Oritaら (1989) Proc Natl Acad Sci USA:86:2766、またCotton (1993) Mutat Res 285:125~144; Hayashi (1992) Genet Anal Tech Appl 9:73~79を参照のこと)。サンプルおよびコントロールGPCR X 核酸の一本鎖DNAフラグメントは、変性され、そして再生される。一本鎖核酸の二次構造は、配列に従って変化し、電気泳動の移動度において得られる変化は、1つの塩基変化の検出さえも可能にする。DNAフラグメントは、標識され得るか、または標識されたプローブを用いて検出され得る。アッセイの感度は、DNAよりもむしろ、二次構造が配列中の変化に対してより感受的であるRNAを使用することによって増強され得る。1つの実施形態において、本発明の方法は、ヘテロ二重鎖分析を利用して、電気泳動の移動度における変化に基づいて二本鎖のヘテロ二重鎖分子を分離する。例えば、Keenら (1991) Trends Genet 7:5を参照のこと。

40

【0360】

なお別の実施形態において、一定勾配の変性剤を含有するポリアクリルアミドゲルにおけ

50

る変異体または野生型フラグメントの移動は、変性勾配ゲル電気泳動 (D G G E) を使用してアッセイされる。例えば、Myersら (1985) Nature 313:495 を参照のこと。D G G E が分析の方法として使用される場合、DNA は、例えば、PCR により約 40 bp の高融点 GC リッチ DNA の GC クランプを付加することによって、完全には変性されないことを確実に改変される。さらなる実施形態において、温度勾配は、コントロールおよびサンプル DNA の移動度における差異を同定するために、変性剤勾配の代わりに使用される。例えば、Rosenbaum および Reissner (1987) Biophys Chem 265:12753 を参照のこと。

【0361】

点変異を検出するための他の技術の例としては、以下が挙げられるが、これらに限定されない：選択的オリゴヌクレオチドハイブリダイゼーション、選択的増幅、または選択的プライマー伸長。例えば、オリゴヌクレオチドプライマーは、既知の変異が中心的に配置されるように調製され得、次いで、完全なマッチが見出される場合にのみハイブリダイゼーションを許容する条件下で標的 DNA にハイブリダイズされる。例えば、Saikiら (1986) Nature 324:163) ; Saikiら (1989) Proc Natl Acad Sci USA 86:6230 を参照のこと。このような対立遺伝子特異的オリゴヌクレオチドは、このオリゴヌクレオチドがハイブリダイズ膜に付着され、そして標識された標的 DNA とハイブリダイズされる場合に、PCR 増幅された標的 DNA または多くの異なる変異にハイブリダイズされる。

10

【0362】

あるいは、選択的 PCR 増幅に依存する対立遺伝子特異的増幅技術は、本発明と合わせて使用され得る。特異的増幅についてのプライマーとして使用されるオリゴヌクレオチドは、分子の中心において (その結果、増幅は、差次的ハイブリダイゼーションに依存する) (Gibbsら (1989) Nucleic Acids Res 17:2437~2448) か、あるいは適切な条件下でミスマッチが妨げられ得るかまたはポリメラーゼ伸長を減少し得る、1つのプライマーの 3' の最末端で、目的の変異を保有し得る (Prossner (1993) Tibtech 11:238)。さらに、切断に基づく検出を行うために、変異領域に新規な制限部位を導入することは、切断に基づく検出を行うために望ましくあり得る。例えば、Gaspariniら (1992) Mol Cell Probes 6:1 を参照のこと。特定の実施形態において、増幅はまた、増幅用 Taq リガーゼを使用して実施され得ることが予測される。例えば、Barany (1991) Proc Natl Acad Sci USA 88:189 を参照のこと。このような場合において、連結は、5' 配列の 3' 末端に完全なマッチが存在する場合にのみ生じ、増幅の存在または非存在を探索することによって、特定の部位における既知の変異の存在を検出することを可能にする。

20

30

【0363】

本明細書中に記載される方法は、例えば、本明細書中に記載れる少なくとも1つのプローブ核酸または抗体試薬を含む、予めパッケージングされた診断キットを利用することによって実施され得、これは、例えば、GPCRX 遺伝子を含む疾患または疾病の症状または家族病歴を示す患者を診断するための臨床的設定において簡便に使用され得る。

40

【0364】

さらに、GPCRX が発現される任意の細胞型または組織 (好ましくは、末梢血白血球) は、本明細書中に記載される予後アッセイにおいて利用され得る。しかし、有核細胞を含む任意の生物学的サンプル (例えば、頬粘膜細胞を含む) が、使用され得る。

【0365】

(薬理ゲノム学 (Pharmacogenomics))
GPCRX 活性 (例えば、GPCRX 遺伝子発現) に対する刺激性または阻害性の影響を有する因子、すなわちモジュレーターは、本明細書中に記載されるスクリーニングアッセイによって同定されるように、障害 (例えば、癌、異常 GPCR X 活性に関連する免疫障害) を (予防的または治療的に) 処置するために個体に投与され得る。このような処置と

50

合わせて、個体の薬理ゲノム学（すなわち、個体の遺伝子型と外来化合物または薬物に対するその個体の応答との間の関係についての研究）が、考慮され得る。治療剤の代謝における差異は、薬理的に活性な薬物の用量と血中濃度との間の関係を変更することによって、重篤な毒性または治療の失敗を導き得る。従って、個体の薬理ゲノム学は、個体の遺伝子型の考慮に基づく予防的または治療的処置のために有効な薬剤（例えば、薬物）の選択を許容する。このような薬理ゲノム学は、さらに、適切な投薬量および治療剤レジメンを決定するために使用され得る。従って、GPCRXタンパク質の活性、GPCRX核酸の発現、あるいは個体におけるGPCRX遺伝子の変異含量が決定されて、それによって個体の治療的または予防的処置のために適切な薬剤を選択し得る。

【0366】

薬理ゲノム学は、罹患した人における変更された薬物の性質および異常な作用に起因して、薬物に応答する臨床的に有意な遺伝性変更を扱う。例えば、Eichelbaum、1996、Clin. Exp. Pharmacol. Physiol, 23:983~985 および Linder、1997、Clin. Chem, 43:254~266を参照のこと。一般に、2つの型の薬理ゲノム学状態が、区別され得る。薬物が身体に作用する方法を変更する1つの因子として伝達される遺伝的状态（変更された薬物作用）、または身体が薬物に作用する方法を変更する1つの因子として伝達される遺伝的状态（変更された薬物代謝）。これらの薬理ゲノム学状態は、稀な欠損としてか、または多型としてのいずれかで生じ得る。例えば、グルコース-6-リン酸デヒドロゲナーゼ（G6PD）欠損は、一般的な遺伝性酵素病であり、この主な臨床的合併症は、酸化剤薬物（抗マラリア剤、スルホンアミド、鎮痛薬、ニトロフラン）の摂取およびソラマメの摂食後の溶血である。

【0367】

例示的な実施形態として、薬物代謝酵素の活性は、薬物作用の強度および持続期間の両方の主要な決定因子である。薬物代謝酵素（例えば、N-アセチルトランスフェラーゼ2（NAT2）およびシトクロムP450酵素CYP2D6およびCYP2C19）の遺伝的多型の発見は、幾人かの患者が予期される薬物効果を得ないか、または標準的かつ安全な用量の薬物を摂取した後に過大な薬物応答および深刻な毒性を示すことに関しての説明を提供した。これらの多型は、集団において2つの表現型（高い代謝能を持つ人（extensive metabolizer）（EM）および低い代謝能を持つ人（poor metabolizer）（PM））で表現される。PMの有病率は、異なる集団の間で異なる。例えば、CYP2D6をコードする遺伝子は高度に多型であり、そしていくらかの変異がPMにおいて同定されており、この全ては機能的CYP2D6の非存在に至る。CYP2D6およびCYP2C19の低い代謝能を持つ人は、彼らが標準的な用量を受けるときに、かなり頻繁に過大な薬物応答および副作用を経験する。代謝産物が活性な治療的部分である場合、そのCYP2D6形成代謝産物であるモルヒネによって媒介されるコデインの鎮痛効果について実証されるように、PMは治療的応答を示さない。他の極端なものは、標準的な用量に応答しない、いわゆる超迅速な代謝能を持つ人である。最近、超迅速な代謝の基準となる分子は、CYP2D6遺伝子増幅に起因していることが同定されている。

【0368】

従って、GPCRXのタンパク質の活性、GPCRXの核酸の発現、あるいは個体におけるGPCRXの遺伝子の変異内容を決定して、それによって、その個体の治療的または予防的処置のために適切な薬剤を選択し得る。さらに、薬理ゲノム学の研究を使用して、個体の薬物応答性の表現型の同定に対して薬物代謝酵素をコードする多型対立遺伝子の遺伝子型を適用し得る。この知見は、用量または薬物選択に適用される場合、有害な反応または治療の失敗を回避し得、従って、被験体をGPCRXの調節因子（例えば、本明細書中に記載される例示的なスクリーニングアッセイの1つによって同定される調節因子）を用いて処置する場合に治療的または予防的効率を増強し得る。

【0369】

（臨床試験中の効果のモニタリング）

G P C R Xの発現または活性（例えば、異常な細胞増殖および／または分化を調節する能力）に対する薬剤（例えば、薬物、化合物）の影響をモニタリングすることは、基本的な薬物スクリーニングおよび臨床試験に適用され得る。例えば、本明細書中に記載されるようなスクリーニングアッセイによって決定される、G P C R Xの遺伝子発現、タンパク質レベルを増加するため、またはG P C R X活性をアップレギュレートする薬剤の効力は、減少したG P C R Xの遺伝子発現、タンパク質レベル、またはダウンレギュレートしたG P C R Xの活性を示す被験体の臨床試験においてモニターされ得る。あるいは、スクリーニングアッセイによって決定される、G P C R Xの遺伝子発現、タンパク質レベルを減少、またはG P C R Xの活性をダウンレギュレートする薬剤の効力は、増加したG P C R Xの遺伝子発現、タンパク質レベル、またはアップレギュレートしたG P C R Xの活性を示す被験体の臨床試験においてモニターされ得る。このような臨床試験において、G P C R Xの発現または活性、および好ましくは、例えば、細胞増殖または免疫障害に關与するような他の遺伝子が、「リードアウト（読み出し）（read out）」、すなわち、特定の細胞の免疫応答性のマーカーとして使用され得る。

10

【0370】

例えば、G P C R Xを含む遺伝子（これは、G P C R X活性（例えば、本明細書中に記載されるようなスクリーニングアッセイにおいて同定される）を調節する薬剤（例えば、化合物、薬物または低分子）を用いる処置によって、細胞内で調節される）が同定され得るが、これらに限定されない。従って、細胞性増殖障害に対する薬剤の効果を研究するために、例えば、臨床試験において、細胞が単離され得、そしてR N Aが調製され得、そしてG P C R Xおよびこの障害に關与する他の遺伝子の発現のレベルについて分析され得る。遺伝子発現のレベル（すなわち、遺伝子発現パターン）は、本明細書中に記載されるように、ノーザンブロット分析もしくはR T - P C Rによるか、あるいは産生されるタンパク質の量を測定することによるか、本明細書中に記載されるような方法の1つによるか、あるいはG P C R Xまたは他の遺伝子の活性のレベルを測定することによって、定量され得る。この様式で、この遺伝子発現パターンは、この薬剤に対する細胞の生理学的応答の指標であるマーカーとして作用し得る。従って、この応答状態は、この薬剤を用いる個体の処置の前、および処置の間の種々の時点で、決定され得る。

20

【0371】

1つの実施形態において、本発明は、薬剤（例えば、アゴニスト、アンタゴニスト、タンパク質、ペプチド、核酸、ペプチド模倣物、低分子、または本明細書中に記載されるスクリーニングアッセイによって同定される他の薬物候補物）を用いる、被験体の処置の効力をモニタリングするための方法を提供し、これは、以下の工程を包含する：（i）薬剤の投与の前に、被験体から投与前サンプルを得る工程；（i i）この投与前サンプルにおいて、G P C R Xのタンパク質、m R N A、またはゲノムD N Aの発現のレベルを検出する工程；（i i i）この被験体から1つ以上の投与後サンプルを得る工程；（i v）この投与後サンプルにおいて、G P C R Xのタンパク質、m R N A、またはゲノムD N Aの発現または活性のレベルを検出する工程；（v）この投与前サンプルにおけるG P C R Xのタンパク質、m R N A、またはゲノムD N Aの発現または活性のレベルを、この投与後サンプルにおけるG P C R Xのタンパク質、m R N A、またはゲノムD N Aの発現または活性のレベルと比較する工程；ならびに（v i）従って、この被験体に対する薬剤の投与を変更する工程。例えば、この薬剤の増加した投与は、検出されるよりも高いレベルにG P C R Xの発現または活性を増加することが（すなわち、この薬剤の効力を増加すること）望ましくあり得る。あるいは、この薬剤の減少した投与は、検出されるよりも低いレベルにG P C R Xの発現または活性を減少することが（すなわち、この薬剤の効力を減少すること）望ましくあり得る。

30

40

【0372】

（処置方法）

本発明は、異常なG P C R Xの発現または活性に關連する障害の危険性のある（または感受性）か、またはこの障害を有する被験体を処置する予防的および治療的の両方の方法を

50

提供する。このように関連する疾患または障害としては、例えば、アテローム性動脈硬化症、高血圧症、先天性心欠損症、大動脈狭窄症、心房中隔欠損症（A S D）、房室（A - V）管欠損症、動脈管、肺動脈弁狭窄症、大動脈弁下部狭窄症、心室中隔欠損症（V S D）、弁疾患、結節硬化症、強皮症、肥満症、移植、副腎脳白質ジストロフィー、先天性副腎過形成、前立腺癌、新形成；腺癌、リンパ腫、子宮癌、受胎能、血友病、凝固能亢進、特発性血小板減少性紫斑病、免疫不全症、対宿主性移植片病、A I D S、気管支炎ぜん息、クローン病；多発性硬化症、オールブライト遺伝性骨形成異常症の処置、ならびに他の疾患、障害および状態など。

【0373】

これらの処置方法は、以下により詳細に記載される。

10

【0374】

（疾患および障害）

（その疾患または障害に罹患していない被験体と比較して）増加したレベルまたは生物学的活性によって特徴付けられる疾患および障害は、活性を拮抗する（すなわち、低減または阻害する）治療剤を用いて処置され得る。活性を拮抗する治療剤は、治療的または予防的な様式で、投与され得る。利用され得る治療剤としては、以下が挙げられるが、これらに限定されない：（i）上記ペプチド、またはそのアナログ、誘導体、フラグメントもしくはホモログ；（i i）上記ペプチドに対する抗体；（i i i）上記ペプチドをコードする核酸；（i v）相同組換えによって上記ペプチドの内因性機能を「ノックアウトする」ために利用される、アンチセンス核酸および「機能不全性」である（すなわち、上記ペプチドに対するコード配列のコード配列内の異種挿入に起因する）核酸の投与、（例えば、C a p e c c h i、1989、S c i e n c e 244：1288～1292を参照のこと）；または（v）上記ペプチドとその結合パートナーとの間の相互作用を変化させる、調節因子（すなわち、インヒビター、アゴニストおよびアンタゴニスト（本発明のさらなるペプチド模倣物または本発明のペプチドに対して特異的な抗体を含む））。

20

【0375】

（その疾患または障害に罹患していない被験体と比較して）減少したレベルまたは生物学的活性によって特徴付けられる疾患および障害は、活性を増加させる（すなわち、活性に対するアゴニストである）治療剤を用いて処置され得る。活性をアップレギュレートする治療剤は、治療的または予防的な様式で、投与され得る。利用され得る治療剤としては、以下が挙げられるが、これらに限定されない：上記ペプチド、またはそのアナログ、誘導体、フラグメントもしくはホモログ；あるいはバイオアベイラビリティを増加させるアゴニスト。

30

【0376】

増加したレベルまたは減少したレベルは、ペプチドおよび／またはR N Aを定量することによって、容易に検出され得る。この定量は、患者の組織サンプルを（例えば、生検組織から）入手し、そしてそのサンプルを、その発現したペプチド（または上記ペプチドのm R N A）のR N Aレベルまたはペプチドレベル、構造および／または活性をインビトロでアッセイすることによる。当該分野において周知の方法としては、以下が挙げられるが、これらに限定されない：イムノアッセイ（例えば、ウェスタンブロット分析、ドデシル硫酸ナトリウム（S D S）ポリアクリルアミドゲル電気泳動が後に続く免疫沈降、免疫細胞化学などによる）および／またはm R N Aの発現を検出するためのハイブリダイゼーションアッセイ（例えば、ノーザンアッセイ、ドットブロット、インサイチュハイブリダイゼーションなど）。

40

【0377】

（予防的方法）

1つの局面において、本発明は、被験体において異常なG P C R Xの発現または活性と関連する疾患または状態を、G P C R Xの発現または少なくとも1つのG P C R X活性を調節する薬剤をこの被験体に投与することによって予防するための方法を提供する。異常なG P C R Xの発現または活性によって引き起こされるかまたはこれらに起因する、疾患に

50

かかる危険がある被験体は、例えば、本明細書中に記載の診断アッセイまたは予後アッセイのいずれか、またはそれらの組み合わせによって、同定され得る。予防薬剤の投与は、疾患または障害が予防されるか、あるいはその進行を遅らせられるように、このGPCRX異常の特徴である症状の発現の前に行い得る。このGPCRX異常の型に依存して、例えば、GPCRXアゴニスト薬剤またはGPCRXアンタゴニスト薬剤が、その被験体を処置するために使用され得る。その適切な薬剤は、本明細書中に記載のスクリーニングアッセイに基づいて決定され得る。本発明の予防方法は、以下の小節において、さらに議論される。

【0378】

(治療方法)

本発明の別の局面は、治療目的のためにGPCRXの発現または活性を調節する方法に関する。本発明の調節方法は、細胞を、その細胞に関するGPCRXタンパク質活性の活性のうちの1つ以上を調節する薬剤と接触させる工程を包含する。GPCRXタンパク質活性を調節する薬剤は、核酸またはタンパク質、GPCRXタンパク質の天然に存在する同族リガンド、ペプチド、GPCRXペプチド模倣物、または他の低分子のような、本明細書中に記載されるような薬剤であり得る。1つの実施形態において、この薬剤は、GPCRXタンパク質活性のうちの1つ以上を刺激する。このような刺激薬剤の例としては、活性なGPCRXタンパク質、およびその細胞に導入されたGPCRXをコードする核酸分子が挙げられる。別の実施形態において、この薬剤は、GPCRXタンパク質活性のうちの1つ以上を阻害する。このような阻害薬剤の例としては、アンチセンスGPCRX核酸分子、および抗GPCRX抗体が挙げられる。これらの調節方法は、インビトロで（例えば、その薬剤とともにその細胞を培養することによって）、あるいはインビボで（例えば、被験体にその薬剤を投与することによって）実施され得る。このように、本発明は、GPCRXのタンパク質または核酸分子の、異常な発現または異常な活性によって特徴付けられる、疾患または障害に罹患した個体を処置する方法を提供する。1つの実施形態において、この方法は、GPCRXの発現または活性を調節する（例えば、アップレギュレートまたはダウンレギュレートする）薬剤（例えば、本明細書中に記載のスクリーニングアッセイによって同定される薬剤）あるいはそのような薬剤の組み合わせを投与する工程を包含する。別の実施形態において、この方法は、GPCRXのタンパク質または核酸分子を、低減したかまたは異常な、GPCRXの発現または活性を補償するための治療として、投与する工程を包含する。

【0379】

GPCRX活性の刺激は、GPCRXが異常にダウンレギュレートされている状況、および／またはGPCRX活性の増加が有益な効果を有するようである状況において、望ましい。このような状況の1つの例は、被験体が、異常な細胞増殖および／または細胞分化によって特徴付けられる障害（例えば、癌または免疫関連）を有する場合である。このような状況の別の例は、被験体が妊娠性疾患（例えば、プレ克蘭プシア（preclampsia））を有する場合である。

【0380】

(治療剤の生物学的効果の決定)

本発明の種々の実施形態において、適切なインビトロまたはインビボアッセイを行って、特定の治療剤の効果およびその投与が罹患組織の処置を示すか否かを決定する。

【0381】

種々の特定の実施形態において、インビトロアッセイが患者の障害に関与する代表的な細胞型で行われ、所定の治療剤がこの細胞型に対して所望の効果を発揮するか否かを決定し得る。治療において使用する化合物は、ヒト被験体において試験する前に適切な動物モデル系において試験され得る。これらの動物モデル系としては、以下が挙げられるが、これらに限定されない：ラット、マウス、ニワトリ、ウシ、サル、ウサギなど。同様に、インビボ試験については、当該分野で公知の任意の動物モデル系が、ヒト被験体に対する投与の前に使用され得る。

10

20

30

40

50

【0382】

(本発明の組成物の予防的用途および治療的用途)

本発明のGPCRX核酸およびタンパク質は、以下を含むが、これらに限定されない種々の障害に関連する強力な予防的適用および治療的適用に有用である：代謝障害、糖尿病、肥満症、感染症、食欲不振、癌関連癌、神経変性疾患、アルツハイマー病、パーキンソン病、免疫障害、造血障害、および種々の異常脂肪性浮腫 (dyslipidemias)、肥満に関連する代謝障害、X代謝症候群 (metabolic syndrome X)、および慢性疾患と関連する消耗疾患、および種々の癌。

【0383】

例のように、本発明のGPCRXタンパク質をコードするcDNAは、遺伝子治療に有用であり、そしてこのタンパク質は、遺伝子治療を必要とする被験体に投与される場合、有用であり得る。非制限的例示として、本発明の組成物は、上記の障害に罹患した患者の処置についての効力を有する。

10

【0384】

本発明のGPCRXタンパク質をコードする新規の核酸、およびGPCRXタンパク質の両方、またはそれらのフラグメントは、診断適用に有用で有り得る。ここで、核酸またはタンパク質の存在または量が評価される。さらなる用途は、抗菌分子（すなわち、いくつかのペプチドは、抗菌特性を有することが見出されている）としての用途である。これらの材料は、治療的または診断的な方法における用途に関する本発明の新規の物質に免疫特異的に結合する抗体の生成においてさらに有用である。

20

【0385】

(等価物)

特定の実施形態は、詳細に本明細書中に開示されているが、これは、例証の目的のための例として開示され、そして前記の添付される特許請求の範囲に関して、制限することを意図しない。特に、種々の置換、変更、および改変が、特許請求の範囲により規定されるような発明の精神および範囲から逸脱することなく、本発明に対してなされ得ることが、本発明により意図される。核酸出発物質、目的のクローン、ライブラリーの型の選択は、本明細書中に記載される実施形態の知見を用いると当業者に慣用的な内容であると考えられる。他の局面、利点、および改変は、添付の特許請求の範囲の範囲内であると考えられる。

30

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number
WO 01/64879 A2(51) International Patent Classification: **C12N 15/12**,
C07K 14/705, 16/28, C12N 1/21, 5/10, G01N 33/68,
C12Q 1/68, A61K 38/17, 31/70, 39/395

(21) International Application Number: PCT/US01/06474

(22) International Filing Date: 28 February 2001 (28.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/185,674	28 February 2000 (28.02.2000)	US
60/185,535	28 February 2000 (28.02.2000)	US
60/186,585	3 March 2000 (03.03.2000)	US
60/186,604	3 March 2000 (03.03.2000)	US
60/186,606	3 March 2000 (03.03.2000)	US
60/186,715	3 March 2000 (03.03.2000)	US
60/186,716	3 March 2000 (03.03.2000)	US
60/186,717	3 March 2000 (03.03.2000)	US
60/186,719	3 March 2000 (03.03.2000)	US
60/186,827	3 March 2000 (03.03.2000)	US
60/218,323	14 July 2000 (14.07.2000)	US
60/218,433	14 July 2000 (14.07.2000)	US
60/220,517	24 July 2000 (24.07.2000)	US
60/223,897	9 August 2000 (09.08.2000)	US
60/260,020	5 January 2001 (05.01.2001)	US
60/264,849	26 January 2001 (26.01.2001)	US
09/795,271	27 February 2001 (27.02.2001)	US

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

US	60/185,674 (CIP)
Filed on	28 February 2000 (28.02.2000)
US	60/185,535 (CIP)
Filed on	28 February 2000 (28.02.2000)
US	60/186,585 (CIP)
Filed on	3 March 2000 (03.03.2000)
US	60/186,604 (CIP)
Filed on	3 March 2000 (03.03.2000)
US	60/186,606 (CIP)
Filed on	3 March 2000 (03.03.2000)
US	60/186,715 (CIP)
Filed on	3 March 2000 (03.03.2000)

US	60/186,716 (CIP)
Filed on	3 March 2000 (03.03.2000)
US	60/186,717 (CIP)
Filed on	3 March 2000 (03.03.2000)
US	60/186,719 (CIP)
Filed on	3 March 2000 (03.03.2000)
US	60/186,827 (CIP)
Filed on	3 March 2000 (03.03.2000)
US	60/218,323 (CIP)
Filed on	14 July 2000 (14.07.2000)
US	60/218,433 (CIP)
Filed on	14 July 2000 (14.07.2000)
US	60/220,517 (CIP)
Filed on	24 July 2000 (24.07.2000)
US	60/223,897 (CIP)
Filed on	9 August 2000 (09.08.2000)
US	60/260,020 (CIP)
Filed on	5 January 2001 (05.01.2001)
US	60/264,849 (CIP)
Filed on	26 January 2001 (26.01.2001)
US	Not furnished (CIP)
Filed on	27 February 2001 (27.02.2001)

(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

(73) Inventors/Applicants (for US only): PADIGARU, Murallikrishna [IN/US], 1579 Rhineland Avenue, Apartment 6E, Bronx, NY 10561 (US); MAJUMDER, Kamod [IN/US], 140 Silver Hill Lane, Stamford, CT 06905 (US); BURGESS, Catherine, F. [US/US], 90 Carriage Hill Drive, Watersfield, CT 06109 (US); VERNET, Corine, A., M. [FR/US], 4830 N.W. 43rd Street P#253, Gainesville, FL 32660 (US); FERNANDES, Elma [IN/US], 77 Florence Road, #2, Branford, CT 06405 (US); SHIMKETS, Richard, A. [US/US], 191 Leste Street, West Haven, CT 06516 (US); TCHERNEV, Velizar, T. [BG/US], 1216 S.W. 2nd Avenue, Apt. 156, Gainesville, FL 32601 (US); MISHRA, Vishnu [IN/US], 309 East Main Street, #306, Branford, CT 06405 (US); CASMAN, Stacie [US/US], 155 Knollwood Drive, Wallingford, CT 06492 (US); SPYTER, Kimberly, A. [US/US], 28 Court Street #1, New Haven, CT 06511 (US); ZERHUSEN, Bryan [US/US], 35 Jefferson Road, Apartment #210, Branford, CT 06405 (US).

[Continued on next page]

WO 01/64879 A2

(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 immunospecifically 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 nucleic acids and proteins.

WO 01/64879 A2



(74) Agent: FLRIFI, Ivan, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Poppo, P.C., One Financial Center, Boston, MA 02111 (US).

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, GW, ML, MR, NE, SN, TD, TG).

(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, EC, EG, ES, FI, GB, GD, GE, GR, GM, HR, HU, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TH, TM, TR, TT, TZ, UA, UG, US, UZ, VM, YU, ZA, ZW.

Published:

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

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

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

WO 01/64879

PCT/US01/06474

NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME**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 GPCR_X, or GPCR1, GPCR2, GPCR3, GPCR4, GPCR5, GPCR6, GPCR7, GPCR8, GPCR9, and GPCR10 nucleic acids and polypeptides. These nucleic acids and polypeptides, as well as derivatives, homologs, analogs and fragments thereof, will hereinafter be collectively designated as "GPCR_X" nucleic acid or polypeptide sequences.

In one aspect, the invention provides an isolated GPCR_X nucleic acid molecule encoding a GPCR_X polypeptide that includes a nucleic acid sequence that has identity to the nucleic acids disclosed in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38. In some embodiments, the GPCR_X 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 GPCR_X nucleic acid sequence. The invention also includes an isolated nucleic acid that encodes a GPCR_X 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:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83. 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38.

Also included in the invention is an oligonucleotide, *e.g.*, an oligonucleotide which includes at least 6 contiguous nucleotides of a GPCR_X nucleic acid (*e.g.*, SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38) or a complement of said oligonucleotide.

Also included in the invention are substantially purified GPCR_X polypeptides (SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83). In certain embodiments,

WO 01/64879

PCT/US01/06474

the GPCR_X polypeptides include an amino acid sequence that is substantially identical to the amino acid sequence of a human GPCR_X polypeptide.

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

5 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 GPCR_X nucleic acid, a GPCR_X polypeptide, or an antibody specific for a GPCR_X polypeptide. In a further aspect, the invention includes, in one or more containers, a therapeutically- or prophylactically-effective amount of this pharmaceutical composition.

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

15 In another aspect, the invention includes a method of detecting the presence of a GPCR_X 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 GPCR_X polypeptide within the sample.

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

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

25 In a further aspect, the invention provides a method for modulating the activity of a GPCR_X polypeptide by contacting a cell sample that includes the GPCR_X polypeptide with a compound that binds to the GPCR_X 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.

30 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

WO 01/64879

PCT/US01/06474

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 GPCR_X nucleic acid, a GPCR_X polypeptide, or a GPCR_X-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 Osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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 GPCR_X may be useful in gene therapy, and GPCR_X 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 Osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, allergies, benign

WO 01/64879

PCT/US01/06474

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.

5 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,
10 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 GPCR_X polypeptide and determining if the test compound binds to said GPCR_X polypeptide. Binding of the test compound to the GPCR_X polypeptide indicates the test compound is a modulator of activity, or of latency or predisposition to the aforementioned
15 disorders or syndromes.

Also within the scope of the invention is a method for screening for a modulator of activity, or of latency or predisposition to an 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
20 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 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
25 GPCR_X nucleic acid. Expression or activity of GPCR_X polypeptide is then measured in the test animal, as is expression or activity of the protein in a control animal which recombinantly-expresses GPCR_X polypeptide and is not at increased risk for the disorder or syndrome. Next, the expression of GPCR_X polypeptide in both the test animal and the control animal is compared. A change in the activity of GPCR_X polypeptide in the test animal relative to the
30 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 GPCR_X polypeptide, a GPCR_X nucleic acid, or both, in a subject (*e.g.*, a human subject). The method includes

WO 01/64879

PCT/US01/06474

measuring the amount of the GPCR α polypeptide in a test sample from the subject and comparing the amount of the polypeptide in the test sample to the amount of the GPCR α polypeptide present in a control sample. An alteration in the level of the GPCR α 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 GPCR α polypeptide, a GPCR α nucleic acid, or a GPCR α -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 identify 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.

WO 01/64879

PCT/US01/06474

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, GPCR9, and GPCR10. The nucleic acids, and their encoded polypeptides, are collectively designated herein as "GPCRX".

The novel GPCR_X nucleic acids of the invention include the nucleic acids whose sequences are provided in Tables 1A, 1D, 1G, 2A, 3A, 4A, 4C, 4G, 5A, 5C, 5G, 6A, 7A, 8A, 9A, 10A, 10C and 10F, inclusive ("Tables 1A - 10F"), or a fragment, derivative, analog or homolog thereof. The novel GPCR_X proteins of the invention include the protein fragments whose sequences are provided in Tables 1B, 1E, 1H, 2B, 3B, 4B, 4H, 5B, 5D, 5H, 6B, 7B, 8B, 9B, 10B, 10D, and 10G, inclusive ("Tables 1B - 10G"). The individual GPCR_X 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.

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

WO 01/64879

PCT/US01/06474

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).

GPCR1

GPCR1 includes a family of three novel G-protein coupled receptor ("GPCR") proteins disclosed below. The disclosed proteins have been named GPCR1a, GPCR1b, and GPCR1c and are related to olfactory receptors.

GPCR1a

A disclosed GPCR1a nucleic acid of 1019 nucleotides is shown in Table 1A. The disclosed GPCR1a open reading frame ("ORF") begins at the ATG initiation codon at nucleotides 27-29, shown in bold in Table 1A. The encoded polypeptide is alternatively referred to herein as GPCR1a or as ba113a10_B. The disclosed GPCR1a ORF terminates at a TAG codon at nucleotides 984-986. As shown in Table 1A, putative untranslated regions 5' to the start codon and 3' to the stop codon are underlined, and the start and stop codons are in bold letters.

Table 1A. GPCR1a nucleotide sequence (SEQ ID NO:1).

<u>CGTTCACTTCACAGAGGAGATACACT</u>ATG<u>GTAA</u>GTGCAATCAGACAGCCTCTGTGACCGAGTTTATTC TCGTGGGCGCTCTGTGCGCAACCAAGCTGGAGAAACCTTCTTTGTGCTCACTCTGTGATGTACCTGG TGAATCTTACCTGGGCAATGGGGTCTCATGCTGATGACCTGTGTCACCTCCACCTGACATGCCATGT ACTTCTTCTCTGGGAACTCTCTCTCTGGACATCTGCTATAGAACATCTGCATGCTCCCTCATGCTTG ACAGCTTCTTGACCCCGAGSAAAGCCATCTCTCTCTGACCTGTGACAGATGTCTCTCTCTCTG CCATGAGGCGACAGAGTGTCTCTCTGACATGATGGGCTTGAATCGCTACGTGGGCTCTGCAACC CCTTTAGGACCCCTGTGCTCATGAGCAAGGCTCCTACATGACCATTAAGGCTGGGGCTCTGGGTAG CTGGAAGCACTGCTTCATGCTGACAGATCCTTGCAATGAGGCTGCCCTCTCTGAGACACATCA TCAAGCACTTCACTGTGATGATTCTGGCTGCTTGAAGTTGGCTGTGCTGATATCTCTGTCAATCTCA TCAGATGGGATGACCAATGTGATCTTCTGCGGCTCCGCTCTGTTCACTCTCTCTCTCTCTCTCT TCATCATTTGCCACATCTGTAGGATCCCTCACTGAGGGGAGGAAAGGCTTCTCCACCTGCTCTG CCGACCTACATTCCTGCTTCTCTCTATGGGACCTCTCTCTGATATGAGGAGGAGGCTCTCTCTCTCT ACCGCTGGGGGAGACAGGAGAGCTTGGAGACAAATCTATTTCTCTCTCTCTCTCTCTCTCTCTCT CCACTGCTGAGGCGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT TTTACGAATGCTTTGCTTGTGATGTTTGGGGGACAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
--

A disclosed encoded GPCR1a protein has 319 amino acid residues, referred to as the GPCR1a protein. The GPCR1a protein was analyzed for signal peptide prediction and cellular localization. SignalP results predict that GPCR1a is cleaved between position 44 and 55 of

WO 01/64879

PCT/US01/06474

SEQ ID NO:2, i.e., at the slash in the amino acid sequence GNG-VL-. Psort and Hydropathy profiles also predict that GPCR1 contains a signal peptide and is likely to be localized at the plasma membrane (certainty of 0.6000). The disclosed GPCR1 polypeptide sequence is presented in Table 1B using the one-letter amino acid code.

5

Table 1B. Encoded GPCR1a protein sequence (SEQ ID NO:2).

NVSANQVAVSTEFILGLSAHPKLSKTFEVLILMLVTLI GNG/VLZLMTVSHSLKMPHYFFLGMLSL FLDCYTTSSVPLILDSFLTERKTISTACAVQNFSLFAMGATECVLLSHMAFDRYVAICNPLRYFVVM SKAAYMFKAGSAVAGSTASMTSLAMRLPFCCNTINHTCELLAVLKACADISVNVISHSVTNV IETGVWLFISFSYVFATILRIIPSAKGRKAEFTCSARLT/VVVFYGTLLFMVGGPKSKDELGADKQ DLAKLLISLTFYGVTPKLNPIIYSLRKRQVKAAYSDLIQKSCFA
--

GPCR1a was initially identified on chromosome 9 with a TblastN analysis of a proprietary sequence file for a G-protein coupled receptor probe or homolog, which was run against the Genomic Daily Files made available by GenBank. A proprietary software program (GenScan™) was used to further predict the nucleic acid sequence and the selection of exons. The resulting sequences were further modified by means of similarities using BLAST searches. The sequences were then manually corrected for apparent inconsistencies, thereby obtaining the sequences encoding the full-length protein.

A region of the GPCR1a nucleic acid sequence has 873 of 1016 bases (85%) identical to a sequence coding for a partial *Mus musculus* olfactory receptor mRNA (1731 bp), with an E-value of $3.6e^{-160}$ (GENBANK-ID: AJ133427). In all BLAST alignments herein, the "E-value" or "Expect" value is a numeric indication of the probability that the aligned sequences could have achieved their similarity to the BLAST query sequence by chance alone, within the database that was searched. For example, the probability that the subject ("Sbjct") retrieved from the GPCR1a BLAST analysis, e.g., the *Mus musculus* olfactory receptor, matched the Query GPCR1a sequence purely by chance is 3.6×10^{-160} .

A BLASTX search was performed against public protein databases. The full amino acid sequence of the protein of the invention was found to have 274 of 316 amino acid residues (86%) identical to, and 295 of 316 residues (93%) positive with, the 319 amino acid olfactory receptor protein from *Mus musculus* (p1nr:TRFMBLNEW-ACC:CAB55392, E= 1.8 e-142). The disclosed GPCR1a protein (SEQ ID NO:2) has good identity with a number of olfactory receptor proteins. For example, GPCR1a has 256/316 (81%) amino acids identical with the 319 amino acid *Mus musculus* olfactory receptor 37a protein, and 376/316 (87%) amino acids identical, to (Expect = e-128, gi|11276075|ref|NP_062346.1|) olfactory receptor

WO 01/64879

PCT/US01/06474

37a from *Mus musculus*. The disclosed protein is also similar to the olfactory proteins disclosed in Table 1C.

Table 1C. BLAST results for GPCR1a

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi111361981 ref NP_062349.1	Olfactory Receptor (OR) 37c <i>Mus musculus</i>	319	255/318 (80%)	275/318 (86%)	e-127
gi111276077 ref NP_062347.1	OR 37b <i>Mus musculus</i>	318	249/317 (78%)	276/317 (86%)	e-127
gi111276079 ref NP_062348.1	OR 37c <i>Mus musculus</i>	318	251/314 (79%)	274/314 (86%)	e-127
gi110092669 ref NP_063950.1	OR Family 2, Subfamily 5, member 2 <i>Rattus norvegicus</i>	309	238/306 (77%)	261/306 (84%)	e-118
gi137696241 gi AAC64588.1 (AAC61565)	OR <i>Rattus norvegicus</i>	227	206/227 (90%)	217/227 (94%)	e-102

GPCR1b

A disclosed GPCR1b (also referred to as ba32713_A) nucleic acid of 1015 nucleotides is shown in Table 1D. An open reading frame was identified beginning with an ATG initiation codon at nucleotides 17-19 and ending with a TAG codon at nucleotides 971-973. A putative untranslated region upstream from the initiation codon and downstream from the termination codon is underlined in Table 1D, and the start and stop codons are in bold letters.

Table 1D. GPCR1b Nucleic acid sequence (SEQ ID NO:3).

<u>ACAGAGGAGTACACTATGGTAAGTGC</u> CAATCAGACAGCCTCTGTGACCGAGTTTATCTCTCGGGCTC TCTGCCACCCAAAGCTGGAGAAACGTTCTTTGGCTCATCCTGCTGATGTAAGTGGTATCTACTAG GCAATGGGGTCTCATCTGATGACTGTTCCAACTCCACCTGCACATGCCCATGTACTCTCTCTGGG GAACCTCTCTCTCTTGGACATCTGTATACACATCTCAGTCCCGCTGATCTTGACAGCTTCTTGACC CCGAGGAACCATCTCTCTTCTGAGCCCTGTCAGTGCAGATGTCTCTCTCTCTTTCCTGATGGGACACAG AGTGTCTCTCTGAGCATGATGGCTTTGATCGCTACGTGGCTCATCTGCAACCCCTTAGGTACCTCTGT OCTCATGAGCAAGGCTGCTACATGCCCTTAGCTGCCGCTCTGGGTAGCTGAGGACATGCTCTCTCTG GTGACAGATCTCTTGAATGAGGCTGCCCTTCTGGGAGACACATCATCAACCACTTCACTGTGAGA TTCTGGCTGTCCTGAAGTTGGCTGTGCTGATATCTCTGCAATGTGATCAGTATGGGAGTGAACCAATGT GATCTCTCTGGGGGTCCCGGTTCTGTTTATCTCTTCTCTCTATCTCTCATCATGCGACCACTCTGAGG ATCCCTCAGCTGAGGGAGGAAAGCCCTTCTGCCCTGCTCTGCCACCTTCACAGTGTGGTGTATCT TCTATGAGACATCTCTCTCATGTATGGGAGGCCAAGTCTAAGGACCCCTGAGGGAGCAGACAAACAGA CCTTGACAGAAACATCTCTCTCTTCTATGGGTGTGTGACCCCATGCTCAAGCCCATCTCATCTACAGC CTGAGGACAGAGATGTAAGGCTGCTGTGAGGAGCTTGATATTTGAGAAATGCTTTCCTAGTGATGTT TGGGGAACAGATGCTCTCATAGCTCTTTGGCTCT

In a search of sequence databases, it was found, for example, that the nucleic acid sequence has 869 of 1015 bases (85%) identical to a 1731 bp *Mus musculus* OR 37d pseudogene (GENBANK-ID: MMU133427|acc:AJ133427, E = 5.5 e-161). It was also found

WO 01/64879

PCT/US01/06474

that the nucleic acid has 505 of 791 bases identical (63%) to a partial human mRNA for olfactory receptor protein (GENBANK-ID:HSOLFMI[acc:Y14442, E= 9.5 e-49).

The encoded protein having 318 amino acid residues is presented using the one-letter code in Table 1E. The full amino acid sequence of the protein of the invention was found to have 274 of 315 amino acid residues (86%) identical to, and 296 of 315 residues (93%) positive with, a 319 amino acid residue olfactory receptor protein from *Mus musculus* (ptnr: TREMBLNEW-ACC:CAB55592, E= 6.3 e-144).

The disclosed GPCR1b protein differs from the disclosed GPCR1a protein at only two positions. At positions 145 and 146, GPCR1a has HK, while GPCR1b has a deletion (Δ) and an I.

Table 1E. Encoded GPCR1b protein sequence (SEQ ID NO:4).

```

MVSANQTASVTEFILLGLSAHPKLEKTFVLLILLWLVILLNG/VLTLLMTVNSHLHMDWYFFLGNLSPL
DICVTTSSVPLILDSFLTRKTISSACAVQMSLSFANGATKCVLLSMMAFDHYVATCNPLRYPPVMSKA
AYMPFAAGSWAGSTASMVQTSIAMRLPFQGDNIINHFTCEILAVLKACADISVNVISMGVTVNIFLGV
PVLEFSFHYVFIATILRIPSAEGKKAKSTCSARLTVVVIFYGTLEFNVQKPKSKDPLGADKQDLADKL
LSLEYGVVTWLNPIVYSLRNKDVKAARDLIFQKCF

```

A PSORT analysis predicts that the ba32713_A protein (GPCR1b) is localized in the plasma membrane with a certainty of 0.6000, or with lower certainty in the mitochondrial inner membrane, the mitochondrial intermembrane space or the Golgi body. It is also predicted that the protein has a signal peptide with the most likely cleavage site between residues 44 and 45: GNG-VL, indicated by a slash in Table 1E.

A BLASTX search was performed against public protein databases. The full amino acid sequence of the protein of the invention was found to have 274 of 315 amino acid residues (86%) identical to, and 296 of 315 residues (93%) positive with, the 319 amino acid olfactory receptor protein from *Mus musculus* (ptnr:TREMBLNEW-ACC:CAB55592, E= 6.3 e-144). The disclosed GPCR1b protein (SEQ ID NO:4) has good identity with a number of olfactory receptor proteins, as shown in Table 1F.

Table 1F. BLAST results for GPCR1b

Gene Index/ Identifier	Protein/ Organism	Length [aa]	Identity [%]	Positives [#]	Expect
gi 11276076 ref NP_062346.1	OR 37a <i>Mus musculus</i>	319	256/315 (80%)	277/315 (87%)	e-129
gi 11464981 ref NE_062346.1	(OR) 37e <i>Mus musculus</i>	319	255/317 (80%)	276/317 (86%)	e-129

WO 01/64879

PCT/US01/06474

gi111276077 ref NP_062347.1	OR 37b <i>Mus musculus</i>	318	248/316 (78%)	277/316 (87%)	a-128
gi111276079 ref NP_062348.1	OR 37c <i>Mus musculus</i>	318	251/313 (80%)	275/313 (87%)	a-128
gi13052668 ref NP_063950.1	OR Family 2, Subfamily 3, member 2 <i>Homo sapiens</i>	309	238/305 (78%)	262/305 (85%)	a-119

GPCR1c

A disclosed GPCR1c (also referred to as *ba113a10_C*) nucleic acid of 1003 nucleotides is shown in Table 1G. An open reading frame was identified beginning with an ATG initiation codon at nucleotides 26-28 and ending with a TGA codon at nucleotides 974-976. Putative untranslated regions 5' to the start codon and 3' to the stop codon are underlined in Table 1G and the start and stop codons are in bold letters.

In a search of sequence databases, it was found, for example, that the nucleic acid sequence has 719 of 899 bases (79%) identical to a *Mus musculus* GPCR mRNA (GENBANK-ID: AJ133428, E = 1.9e-120).

Table 1G. GPCR1c Nucleic acid sequence (SEQ ID NO:5).

TTTCTCAAGTGCATGGAACACCATGCTCAGTCC AATCGAGCCCTCCCTGTGCTGGGGTTCCTTCTCTGGGGCTCTCTGCCCATCCAAAGCTGGAGAGAGAT TUTTGCTGCTATCCTGCTGATGTACCTGGTGAATCTACTGGGCAATGGGCTCCTCATCCTGGTGACCAT CTTTGACTCCCGGCTTGAACACACCATGTACTTCTTCTGGGCAATGGGCTCCTCATCCTGGTGACCAT ACAACCTCCTCATCCTTGAACAGCTTCCCTGACCCAGGAAACCATCTCCTTCTCAGCCCTGTGAGTAC AGATTTCTTCTCCCTTGGCATGGGAGTCCACAGAGTGTCTTCTCTGAGCATGATGGGCTTGAATGCTA CGTGCCCATCTGCAACCCCTTTGGTACCTGAGATCATGAACAAAGCTACTTATGTGCCCATGGCTGCT GGCTCCTGGTACCTGGAAGCTTCACTGCCATGGTGAGCAACCCCTTGCATGGAGGCTGCCCTTCTGAG GAGACAAATCATCAATCACTTCACTGTGAGATTTCTGCTTCTGAAATTTGGCTGTGCTGATATCTTC TCTCAATGTGATCAGTATGGGAGTGGCCATGTGATCTTCTGGGGGCTCCTCTTCTGATATCTTTC TCTATGTCTTCTATCTTGGGAGTGGCCATGTGATCTTCTGGGGGCTCCTCTTCTGATATCTTTC CTTCTCTTCTGAGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCTGAT CTTATATGGGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGG GCTATGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCT TCTATTTTCAAGAACTCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCTGATATGGCTTCT
--

The disclosed GPCR1c protein having 316 amino acid residues is presented using the one-letter code in Table 1H. An analysis using the PSORT program predicts that the *ba113a10_C* protein localizes in the plasma membrane with a certainty=0.6400; it may also be localized in the Golgi body with a moderate certainty. It is also predicted that protein has a signal peptide whose most likely cleavage site is between residues 44 and 45: GNG-VI, indicated by a slash in Table 1H.

WO 01/64879

PCT/US01/06474

Table IH. Encoded GPCR1c protein sequence (SEQ ID NO:6).

MVSEHQZSPVLSGLLLGLCAHKLKENTFVLLILNLYLTLGLNG/VLLVTLDSRIOTPHYFFECN
 LAFIDICVTFSSRI TASLTERKT:SFSAQAVQNELSLAMGATDCVLLSMNAFDRYVALCNELWYPSVMMK
 ATYYVPWAGSSVAGSLTAMVQTPCALRLTFCGGNIINHFTCEILAVLKLKADISVNNISYGVANVIFUG
 VPVLFISSESYVILLATILRLPSAEGKKASTCSAHLTVVIVFYGTILFMYGKPKSKDPLGADKQDLADK
 LLSLFYGVYTPMLNP:IIYSLRNKEVKKAA/PNLYEQKRLQ

A BLASTX search was performed against public protein databases. The full amino acid sequence of the protein of the invention was found to have 270 of 317 amino acid residues (85%) identical to, and 287 of 317 residues (90%) positive with, a 319 amino acid residue OR from *Mus musculus* (ptnr:TREMBLNEW-ACC:CAB55596, E= 5.2 e-138). The disclosed GPCR1b protein (SEQ ID NO:6) has good identity with a number of olfactory receptor proteins, as shown in Table II.

Table II. BLAST results for GPCR1c

Gene Index/ Identiflex	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi114449811ref NP_062349.1	OR 37a <i>Mus musculus</i>	319	242/317 (76%)	257/317 (80%)	e-120
gi11276075ref NP_062346.1	OR 37a <i>Mus musculus</i>	319	241/319 (75%)	256/319 (80%)	e-119
gi11276077ref NP_062347.1	OR 37b <i>Mus musculus</i>	316	236/316 (78%)	255/316 (80%)	e-116
gi11276079ref NP_062348.1	OR 37c <i>Mus musculus</i>	318	235/308 (76%)	250/308 (80%)	e-110
gi11092569ref NP_063950.1	OR Family 2, Subfamily 5, member 2 <i>Homo sapiens</i>	309	220/309 (71%)	242/309 (78%)	e-108

The amino acids differences between the three GPCR1 proteins are shown in Table IJ. Deletions are marked by a delta (Δ). The differences between the three proteins appear to be localized to a few distinct regions. Thus, these proteins may have similar functions, such as serving as olfactory or chemokine receptors (see below).

WO 01/64879

PCT/US01/06474

Table 1J. Differences for GPCR1 Proteins															
Position	4	8	9	11	12	14	49	51	52	53	55	57	58		
GPCR1a	A	A	S	T	S	I	H	V	S	N	H	H	M		
GPCR1b	A	A	S	T	S	I	H	V	S	N	H	H	M		
GPCR1c	S	S	P	L	G	L	V	T	L	D	R	D	T		
Position	79	80	81	82	84	85	86	106	132	135	138	141	143		
GPCR1a	V	P	L	I	D	S	E	F	R	V	S	A	M		
GPCR1b	V	P	L	I	D	S	E	F	R	V	S	A	M		
GPCR1c	S	A	D	A	T	A	S	L	W	E	N	T	V		
Position	145	146	156	157	158	163	166	204	251	252	304	311	313		
GPCR1a	H	R	T	A	S	S	M	T	V	I	D	D	I		
GPCR1b	A	T	T	A	S	S	M	T	V	I	D	D	I		
GPCR1c	A	M	L	T	A	P	L	A	I	V	E	N	V		
Position	317	319	320												
GPCR1a	C	A	A												
GPCR1b	C	A	A												
GPCR1c	R	L	Q												

A ClustalW analysis comparing disclosed proteins of the invention with related OR protein sequences is given in Table 1K, with GPCR1a shown on line 1, and GPCR1c on line 2.

- 5 In the ClustalW alignment of the GPCR1a protein, as well as all other ClustalW analyses herein, the black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties), whereas non-highlighted amino acid residues are less conserved and can potentially be mutated to a much broader extent without altering protein structure or function. Unless specifically addressed as
- 10 GPCR1a GPCR1b, or GPCR1c, any reference to GPCR1 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant and the residue position with respect to the "a" variant. GPCR residues in all following sequence alignments that differ between the individual GPCR variants are highlighted with a box and marked with the (o) symbol above the variant residue
- 15 in all alignments herein. For example, the protein shown in line 1 of Table 1K depicts the sequence for GPCR1a, and the positions where GPCR1b differs are marked with a (o) symbol and are highlighted with a box. All GPCR1 proteins have significant homology to olfactory receptor (OR) proteins: 37a, 37b, 37c, and 37d from *Mus musculus* and to a human OR protein member 2 from family 2, subfamily S (see also Tables 1C, 1F, and 1I).

29

WO 01/64879

PCT/US01/06474

Table 1K. ClustalW Analysis of GPCR1

5

1) Novel GPCR1a (SEQ ID NO:2)
 2) Novel GPCR1c (SEQ ID NO:6)
 3) gi11276075|refNP_062346.1 olfactory receptor 37c *Mus musculus* (SEQ ID NO:39)
 4) gi11276077|refNP_062347.1 olfactory receptor 37b *Mus musculus* (SEQ ID NO:40)
 5) gi11464981|refNP_062349.1 olfactory receptor 37c *Mus musculus* (SEQ ID NO:41)
 6) gi11276079|refNP_062348.1 olfactory receptor 37c *Mus musculus* (SEQ ID NO:42)
 7) gi10092669|refNP_063950.1 OR, fam. 2, subfam. S, mem. 2 *H. sapiens* (SEQ ID NO:43)

10

The GPCR1 proteins also have regions of identity with a 321 amino acid human transmembrane receptor (gi|6691937|emb|CAB65797.1|bA150A6.2, novel 7 transmembrane

WO 01/64879

PCT/US01/06474

receptor (rhodopsin family), olfactory receptor-like protein (hs6M1-21)), as shown in line 3 of in Table 1L (SEQ ID NO: 44).

Table 1L. Alignment of GPCR1a & c with SEQ ID NO:44					
	10	20	30	40	50
GPCR1a					
GPCR1c					
gi16691937					
	70	80	90	100	110
GPCR1a					
GPCR1c					
gi16691937					
	130	140	150	160	170
GPCR1a					
GPCR1c					
gi16691937					
	190	200	210	220	230
GPCR1a					
GPCR1c					
gi16691937					
	250	260	270	280	290
GPCR1a					
GPCR1c					
gi16691937					
	310	320			
GPCR1a					
GPCR1c					
gi16691937					

- 5 The presence of identifiable domains in GPCR1, as well as all other GPCR proteins, was determined by searches using software algorithms such as PROSITE, DOMAIN, Blocks, Pfam, ProDomain, and 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 results, e.g., for GPCR1a as disclosed in Table 1M, were collected from the
- 10 Conserved Domain Database (CDD) with Reverse Position Specific BLAST analyses. This BLAST analysis software samples domains found in the Smart and Pfam collections. For Table 1M 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
- 15 following groups of amino acids: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW.

WO 01/64879

PCT/US01/06474

Table 1M lists the domain description from DOMAIN analysis results against GPCR1a. The region from amino acid residue 53 through 239 (SEQ ID NO:2) most probably (E = 2e⁻¹⁹) contains a "seven transmembrane receptor (rhodopsin family) fragment" domain, aligned here with residues 12-180 of the 7tm_1 entry (TM7, SEQ ID NO:45, see Table 1N for the complete sequence) of the Pfam database. This indicates that the GPCR1 sequence has properties similar to those of other proteins known to contain this domain as well as to the 377 amino acid 7tm domain itself. GPCR1a also has identity to another region of the TM7 protein. The region from amino acid residue 226 through 298 (of SEQ ID NO:2) aligns with amino acid residues 310-377 of TM7 (E = 3e⁻⁴). GPCR1b and GPCR1c also align to this domain: residues 53-238 of GPCR1b align with residues 12-180 of TM7 (E = 6e⁻¹⁹) and residues 225-297 of GPCR1b align with residues 310-377 of TM7 (E = 3e⁻⁴); residues 55-235 of GPCR1c align with residues 14-180 of TM7 (E = 2e⁻¹²) and residues 222-294 of GPCR1c align with residues 310-377 of TM7 (E = 2e⁻⁴).

Table 1M. Domain Analysis of GPCR1a	
Subject: 7 transmembrane receptor (rhodopsin family) fragment	
Gn1Pfam pfam00001: Length = 377	
Score = 89.7 bits (221), Expect = 2e-19	
GPCR1a	
TM7	GNLVNKAQSSLEAQTITTEGKCEVVAQVAATVPEETSEGTSMWSTKIAQAP
GPCR1a	
TM7	
GPCR1a	
TM7	
GPCR1a	
TM7	
GPCR1a	
TM7	

15

The representative member of the 7 transmembrane receptor family is the D2 dopamine receptor from *Bos taurus* (SWISSPROT: locus D2DR_BOVIN, accession P20288; gene index 118205). The D2 receptor is an integral membrane protein and belongs to Family 1 of G-protein coupled receptors. The activity of the D2 receptor is mediated by G proteins which inhibit adenylyl cyclase. Chio *et al.*, Nature 343:255-269 (1990). Residues 51-427 of this 444 amino acid protein are considered to be the representative TM7 domain, shown in Table 1N.

20

WO 01/64879

PCT/US01/06474

Table 1N Amino Acid sequence for TM7 (SEQ ID NO:45)

CNVLVCNAVSREKALQTTITNYLVSAYADLLVATLMPWVYVLEVVGEWKFSEIHCOIF VTLDVNMKCTASTILNLCALSIDRYTAVAMEMLYNTRYSSKRRVTNVIATNWVLEITISCPN LFGLNATDQNECLIANPAEVVYSSINSEYVPELVILLVYIKIYINLRBRKRKVNTRKSSR AFRANLKAPLKGCTHPEDAKLCTVIMKSNSSFPVMRRHVEAARRAQELEMENLSSTSEF ERTRYSPFPSSHQCLLPDPSSHGLHSTFDSFAPPEKNHAKTVNPKIAKIFELQSMFNG KTRCSLCTMSRRKLSQKEKKACQMLATVLGVFTICMLPFFITHILNIHCDCNIPPVLYS AFTWLGIVMSAVNFIY
--

The 7 transmembrane receptor family includes a number of different proteins, including, for example, serotonin receptors, dopamine receptors, histamine receptors, adrenergic 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); Expression information for GPCR_X RNA was derived using tissue sources including, but not limited to, proprietary database sources, public EST sources, literature sources, and/or RACE sources, as described in the Examples.

The nucleic acids and proteins of GPCR₁ are useful in potential therapeutic applications implicated in various GPCR- or olfactory receptor (OR)-related pathologies and/or disorders. For example, a cDNA encoding the G-protein coupled receptor-like protein may be useful in gene therapy, and the G-protein coupled receptor-like protein may be useful when administered to a subject in need thereof. The novel nucleic acid encoding GPCR₁ protein, 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. The GPCR_X nucleic acids and proteins are useful in potential diagnostic and therapeutic applications implicated in various diseases and disorders described below and/or other pathologies. For example, the compositions of the present invention will have efficacy for treatment of patients suffering from: 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,

WO 01/64879

PCT/US01/06474

scleroderma, obesity, transplantation, adrenoleukodystrophy, congenital adrenal hyperplasia, fertility, hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies, graft versus host disease, bronchial asthma, and other diseases, disorders and conditions of the like. By way of nonlimiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from neoplasm, adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune response, AIDS, asthma, Crohn's disease, multiple sclerosis, and Albright Hereditary Osteodystrophy. Additional GPCR-related diseases and disorders are mentioned throughout the Specification.

Further, the protein similarity information, expression pattern, and map location for GPCR1 suggests that GPCR1 may have important structural and/or physiological functions characteristic of 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.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel GPCR1 substances 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-GPCR1 Antibodies" section below.

GPCR2

An additional GPCR-like protein of the invention, referred to herein as GPCR2, is an Olfactory Receptor ("OR")-like protein. The novel nucleic acid of 1254 nucleotides (U1612531_1, SEQ ID NO:7) encoding a novel G-protein coupled receptor-like protein is shown in Table 2A. SeqCalling fragments for GPCR2 came from placenta, indicating that it may be expressed in tissues important for female health.

[illegible]

10

Table 2B. Encoded GPCR2 protein sequence (SEQ ID NO:8).

```

MGMRLGLQNSTNFTVLCGLITPAFGGLLEAFVSTFVVAITA/NIYMIILLINKDSRLHNTMYFLLSLS
INDLTITCITVPMQDQLSRLDSTISPMGCAQFLVLTLLGGEFFLLGLMAYDVA/VQCNPLR/ELLM
NRKRLVGLVGLVQGLSDGLMTPLRTFMSPFDRSEIHHFCEI(PAVK)SLCTDCTSLYETLMAYACVL
ITLPSVLSIVSYTHILLVHRNMGAGGRKATCACTSSHIMVVFYFGAAYTNVQZHSYHTPEKDKVSA
EYTLTPDI/NPLNLTYSNLKNDLDAIRLKLQ/GMCTSCDRLMTVY

```

12

WO 01/64879

PCT/US01/06474

Table 2C. GPCR2 BLAST results				
Sequences producing High-scoring Segment Pairs:	Reading Frame	High Score	Smallest sum Prob.	P(N)
Gen:15PT2980-ACC:CA5849 OLFATORY RECEPTOR - HOMO SAPIENS...	437	1.3e+02	1	
Gen:SWISSPROT-ACC:P27225 OLFATORY RECEPTOR 15 (OR15) - N...	712	2.7e+70	1	
Gen:SPTR2980-ACC:CA5849 OLFATORY RECEPTOR 4 - GALINUS GALLUS...	709	7.3e+70	1	
Gen:TRMBL1200-ACC:CA585593 OLFATORY RECEPTOR - Mus musculus...	706	1.2e+69	1	
Gen:TRMBL1200-ACC:CA585594 OLFATORY RECEPTOR - Mus musculus...	704	1.9e+69	1	
Gen:SPTR2980-ACC:Q43384 OLF RECEPTOR - RATTUS NORVEGICUS...	702	3.1e+69	1	
Gen:TRMBL1200-ACC:CA585592 OLFATORY RECEPTOR - Mus musculus...	698	8.3e+69	1	
Gen:TRMBL1200-ACC:CA585596 OLFATORY RECEPTOR - Mus musculus...	697	1.1e+68	1	
Gen:SWISSPROT-ACC:Q05155 OLFATORY RECEPTOR-LIKE PROTEIN...	695	1.7e+68	1	
Gen:SWISSPROT-ACC:Q13807 OLFATORY RECEPTOR-LIKE PROTEIN...	694	2.2e+68	1	
Gen:SWISSPROT-ACC:Q13807 OLFATORY RECEPTOR-LIKE PROTEIN...	689	7.5e+68	1	
Gen:TRMBL1200-ACC:CA585596 OLFATORY RECEPTOR-LIKE PROTEIN...	685	2.0e+67	1	

Other BLAST results including the sequences used for ClustalW analysis is presented in Table 2D.

5

Table 2D. BLAST results for GPCR2					
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (#)	Expect
gi1396382 gb AA013314.1 (AF107427)	OR 23 <i>Mus musculus</i>	223	156/223 (69%)	184/223 (81%)	2e-78
gi12921628 gb AA035611.1 (U06573)	OR <i>Homo sapiens</i>	216	151/216 (69%)	177/216 (81%)	5e-75
gi112007425 gb AA045193.1 (AF321234)	T2 OR <i>Mus musculus</i>	316	155/298 (52%)	205/258 (68%)	2e-72
gi112007424 gb AA045197.1 (AF321234)	T3 OR <i>Mus musculus</i>	315	156/304 (51%)	209/304 (68%)	2e-72
gi112007425 gb AA045193.1 (AF321234)	T4 OR <i>Mus musculus</i>	319	152/304 (50%)	207/304 (68%)	9e-72
gi112007422 gb AA045193.1 (AF321234)	T1 OR <i>Mus musculus</i>	316	156/305 (51%)	207/305 (67%)	2e-67

This information is presented graphically in the multiple sequence alignment given in Table 2E (with GPCR2 being shown on line 1) as a ClustalW analysis comparing GPCR2 with related protein sequences.

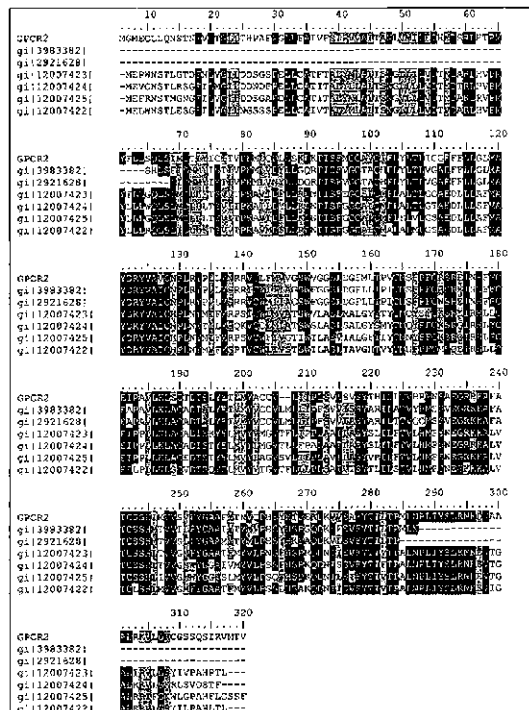
10

WO 01/64879

PCT/US01/06474

Table 2E. Information for the ClustalW proteins:

- 1) Novel GPCR2 (SEQ ID NO:8)
- 2) gB3983382|gB|AAD13319.1| olfactory receptor *R3 Mus musculus* (SEQ ID NO:46)
- 3) gB2921628|gB|AACV861.1| olfactory receptor *Human sapiens* (SEQ ID NO:47)
- 4) gB12007423|gB|AAG45196.1| T2 olfactory receptor *Mus musculus* (SEQ ID NO:48)
- 5) gB12007424|gB|AAG45197.1| T3 olfactory receptor *Mus musculus* (SEQ ID NO:49)
- 6) gB12007425|gB|AAG45198.1| T4 olfactory receptor *Mus musculus* (SEQ ID NO:50)
- 7) gB12007422|gB|AAG45195.1| T1 olfactory receptor *Mus musculus* (SEQ ID NO:51)



DOMAIN results for GPCR2 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the

WO 01/64879

PCT/US01/06474

Smart and Pfam collections. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to have two segments with homology to GPCR2. The region from amino acid residue 43 through 238 aligns with amino acids 2 through 181 of the "seven transmembrane receptor (rhodopsin family) fragment" domain (SEQ ID NO:45, E= 2e-22), and GPCR2 amino acids 226-289 aligned with residues 313-377 of the 7tm_1 entry (SEQ ID NO:45, E= .008) of the Pfam database. This indicates that the GPCR2 sequence has properties similar to those of other proteins known to contain this domain as well as to the 7tm_1 domain itself.

The disclosed GPCR2 is expressed in tissues that are important in female reproductive health and hence GPCR2 may serve as a drug target for, e.g., premature labor, endometriosis, and *in vitro* fertilization. The homology to the olfactory receptors suggests that an endogenous small molecule ligand regulates this gene and hence drugs structurally similar to the endogenous ligand could serve as agonists and antagonists to regulate the biological effects of GPCR2.

The nucleic acids and proteins of the invention 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 olfactory receptor-like protein may be useful in gene therapy, and the olfactory receptor-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 neoplasia, adenocarcinoma, lymphoma, uterus cancer, immune response, AIDS, asthma, Crohn's disease, multiple sclerosis, and Albright Hereditary Osteodystrophy. Other GPCR-related diseases and disorders are contemplated.

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-GPCR Antibodies" section below. In one embodiment, a contemplated GPCR3 epitope is from about amino acids 105 to 140. These novel proteins can be used in assay systems for functional analysis of various human disorders, which will help in understanding of pathology of the disease and development of new drug targets for various disorders.

PCT/US01/06474

An additional GPCR-like protein of the invention, referred to herein as GPCR3, is an Olfactory Receptor ("OR")-like protein. The novel nucleic acid was identified on chromosome 6 by TblastN using CuraGen Corporation's sequence file for GPCR probe or homolog, run against the Genomic Daily Files made available by GenBank. The nucleic acid was further predicted by the program GenScan™, including selection of exons. These were further modified by means of similarities using BLAST searches. The sequences were then manually corrected for apparent inconsistencies, thereby obtaining the sequences encoding the full-length protein. The novel nucleic acid of 957 nucleotides (ba145122_B, SEQ ID NO:9) encoding a novel olfactory receptor-like protein is shown in Table 3A. An open reading frame (ORF) was identified beginning with an ATG initiation codon at nucleotides 10-12 and ending with a TGA codon at nucleotides 955-957. 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.

[illegible]

The disclosed ba145122_B nucleic acid sequence has 592 of 934 bases (63%) identical to a GPCR mRNA (GENBANK-ID: HUMORLMHC| acc: L35475) from *Homo sapiens* (B= 2.4 e⁻⁵¹).

The disclosed GPCR3 polypeptide (SEQ ID NO:10) encoded by SEQ ID NO:9 is 315 amino acid residues and is presented using the one-letter code in Table 3B. The first 70 amino acids of the disclosed GPCR3 protein were analyzed for signal peptide prediction and cellular localization. SignalP results predict that GPCR3 is cleaved between position 46 and 47 of

WO 01/64879

PCT/US01/06474

SEQ ID NO:10, i.e., at the slash in the amino acid sequence NSA-LV. Psort and Hydropathy profiles also predict that GPCR3 contains a signal peptide and is likely to be localized at the plasma membrane (certainty of 0.6000).

Table 3B. Encoded GPCR3 protein sequence (SEQ ID NO:10).

NKANYSAEERFLLLESDNPSPLOPVLPAZVLGCLLTITNSA/LWLLAVRCPRLICPHYYFLGHALVIA
GFTTSVVPZLLANLRCPALNLPBSHCTAQLCAIALGSAZCVLLAWMAIDRAARVCRPIRYAGIVSPRLC
ETLASNHLGSLTNSVAQCALIABRPICAPRLGHFICELFALLKIRGGEGDTTENQMFARVILLIP
PAVILASVCAVARAVCCNRFSGGRRAVGTGGSHLCAYCLFYGSAIYTYLQPAQRYNQARGKIVSLFTYV
VTFALNPLIYTLRNKKVKGARRLRLSLGRGQAGQ

5

A BLASTX search was performed against public protein databases. The full amino acid sequence of the protein of the invention was found to have 253 of 308 amino acid residues (82%) identical to, and 264 of 308 residues (85%) positive with, the 309 amino acid residue MM17M1-6, 7 transmembrane receptor (OR-like protein) from *Mus musculus* (ptnr:SPTRMBL-ACC:Q9WV09, SEQ ID NO:52) (E = 1.5 e⁻¹²⁸). The alignment of these proteins is shown in Table 3C.

10

Table 3C. Alignment of GPCR3 with QpWV09 (SEQ ID NO:52).

	10	20	30	40	50	60
GPCR3						
Q9WV09						
	70	80	90	100	110	120
GPCR3						
Q9WV09						
	130	140	150	160	170	180
GPCR3						
Q9WV09						
	190	200	210	220	230	240
GPCR3						
Q9WV09						
	250	260	270	280	290	300
GPCR3						
Q9WV09						
	310					
GPCR3						
Q9WV09						

The disclosed GPCR3 protein (SEQ ID NO:10) also has good identity with a number of olfactory receptor proteins, as shown in Table 3D.

15

WO 01/64879

PCT/US01/06474

Table 3D. BLAST results for GPCR3					
Gene Index/ Identifier	Protein/ Organism	Length [aa]	Identity [%]	Positives [%]	Expect
gi15051404 emb CA B45612.1 (AL078656)	573K1.15 - 7TM1-6 (novel 7 transmembrane receptor) <i>Mus musculus</i>	305	229/296 (77%)	240/296 (80%)	e-108
gi112054359 emb C AC20487.1 (AJ302567) (AJ302568)	OR <i>Homo sapiens</i>	312	145/298 (48%)	191/298 (63%)	2e-69
gi112054355 emb C AC20485.1 (AJ302565) (AJ302566) (AJ3025 69) (AL30257C)	OR <i>Homo sapiens</i>	312	144/298 (48%)	191/298 (63%)	3e-69
gi112231029 sp C1 202210245 HUMAN	OR 2H3 <i>Homo sapiens</i>	316	143/294 (48%)	185/294 (62%)	3e-68
gi19799920 gb BAF 58752.1 AF211940 1 AF211940	OR <i>Homo sapiens</i>	303	142/287 (49%)	182/287 (63%)	5e-68

This information is presented graphically in the multiple sequence alignment given in Table 3E (with GPCR3 being shown on line 1) as a ClustalW analysis comparing GPCR3 with related protein sequences.

5

Table 3E. Information for the ClustalW proteins:

- 1) Novel GPCR3 (SEQ ID NO:10)
 2) gi15051404|emb|CA B45612.1| 573K1.15 (nm) 7M1-6 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor (like) protein)) *Mus musculus* (SEQ ID NO:53)
 3) gi112054359|emb|CAC20487.1| olfactory receptor *Homo sapiens* (SEQ ID NO:34)
 4) gi112054355|emb|CAC20485.1| olfactory receptor *Homo sapiens* (SEQ ID NO:55)
 5) gi112231029|sp|C1202210245| HUMAN OLFACTORY RECEPTOR 2H3 (OLFACTORY RECEPTOR-LIKE PROTEIN FAT11) (SEQ ID NO:56)
 6) gi19799920|gb|BAF58752.1|AF211940_1 olfactory receptor *Homo sapiens* (SEQ ID NO:57)

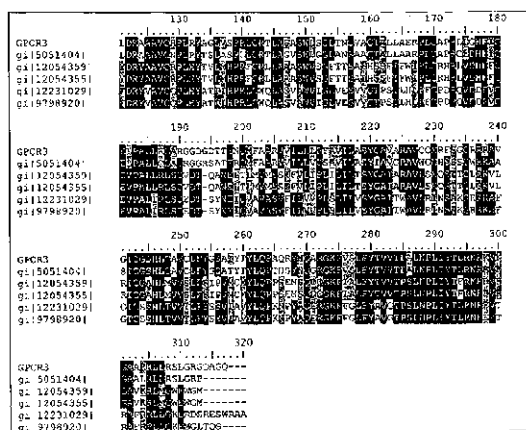
10

	10	20	30	40	50	60
GPCR3	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi15051404	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi112054359	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi112054355	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi112231029	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi19799920	---MKA	---RREAV	---LGLSP	---GR	---R	---D

	70	80	90	100	110	120
GPCR3	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi15051404	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi112054359	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi112054355	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi112231029	---MKA	---RREAV	---LGLSP	---GR	---R	---D
gi19799920	---MKA	---RREAV	---LGLSP	---GR	---R	---D

WO 01/64879

PCT/US01/06474



DOMAIN results for GPCR3 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 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to have two segments with significant homology to GPCR3. The region from amino acid residue 40 through 226 aligns with amino acids 1 through 167 of the "seven transmembrane receptor (rhodopsin family) fragment" domain (SEQ ID NO:45, E=1e-15), and GPCR3 amino acids 219-290 aligned with residues 305-377 of the 7tm_1 entry of the Pfam database (E=.004). This indicates that the GPCR3 sequence has properties similar to those of other proteins known to contain this domain as well as to the 7tm_1 domain itself.

The nucleic acids and proteins of the invention 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 olfactory receptor-like protein may be useful in gene therapy, and the olfactory receptor-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 neoplasm, adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune response, AIDS, asthma, Crohn's disease, multiple sclerosis, and Albright Hereditary Osteodystrophy. Other GPCR-related diseases and disorders are contemplated.

PCT/US01/06474

5

10

GPCR4 includes a family of three nucleic acids disclosed below. The disclosed nucleic acids encode a GPCR-like protein.

GPCR4a

15

Table 4A. GPCR4a1 Nucleotide Sequence (SEQ ID NO:11).

[illegible]

25

WO 01/64879

PCT/US01/06474

(OR2C1) gene (GENBANK-ID: AF098654) ($E = 3.8e^{-75}$). In a search of sequence databases, partial matches were also identified, e.g., the minus strand of nucleotides 719-947 had 229 of 229 bases (100%) identical to a 1320 bp *synthetic* GPCR mRNA (PATENT-ID: T72050), the sequence around marker 2B8 in HH region of chromosome 6p2.1, and the same region of GPCR4a1 had 229 of 229 bases (100%) identical to a *synthetic* GPCR mRNA (PATENT-ID: T72050), also sequence around marker 2B8 in HH region of chromosome 6p2.1 (E value in both cases is $9.6e^{-47}$).

The GPCR4a polypeptide (SEQ ID NO:12) encoded by SEQ ID NO:11 is presented using the one-letter amino acid code in Table 4B. The Psort profile for GPCR4a predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6000. The most likely cleavage site for a GPCR4a peptide is between amino acids 41 and 42, i.e., at the slash in the amino acid sequence LLG-NM, based on the SignalP result.

Table 4B. GPCR4a protein sequence (SEQ ID NO:12)

```

MENDNTSSFDGFTLVCFSDRPHLELIVEVVLIFYLITLLS/NMTIVLLSALDSRLHTPMYFFLANLSY
LQMCPTTGSIFQMCYNLWCFDKTISYVGCALQLEYVLALGSECVLLAVKAYORYAAVCKPLRYTILME
PRLOGQLASVAVLSGFGNSLIMAPQTIMLPRCGHRRVDFLCHEPALISKACVDTHMLEALASALATFI
ILAPLILLISYGVGGIVLRKAGAGRKKAFNTCSBLIVVSLFYGTIIYKYLQANTYSQDQGGKFLT
LFYTTVTPSNPLITYTIRKNDVKEAKKVLGKGSAL

```

The predicted GPCR4a1 sequence, above, was 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. 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 suitable sequences were then used as the forward and reverse primers in a PCR amplification based on a wide range of cDNA libraries. The resulting amplicon was gel purified, cloned and sequenced to high redundancy, as described in the Examples.

The cloned sequence is disclosed as an alternative embodiment of GPCR4a2 (SEQ ID NO:13), referred to herein as the GPCR4a2 and reported in Table 4C. This 945 nucleotide sequence (GPCR4a2) is alternatively referred to herein as dj408b20_C_da1. This nucleic acid is two nucleotides shorter than GPCR4a1 at the 5' UTR. However, GPCR4a2 encodes the same 312 amino acid protein (GPCR4a, SEQ ID NO:12).

PCT/US01/06474

[illegible]

5

10

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (#)	Expect
ACC:Q9Y3N9 DJ83J8.1	Novel 7 Transmembrane Receptor (Rhodopsin Family, OR-Like) Regn-15 Homo Sapiens	320	197/305 (64%)	243/305 (79%)	5.4e- 109
ACC:P23275	OLFACTORY RECEPTOR 15 (OR1) Mus musculus	312	190/308 (61%)	241/308 (78%)	5.5e- 109
ACC:Q76001 D78719.7	OR-LIKE PROTEIN (HSGN-3) Rano. 3447025	311	193/302 (62%)	244/302 (80%)	1.5e- 103

10

Sequences producing high-scoring Segment Pairs:	Reading Frame	High Score	Smallest Prob. E(N)	N
PtNC:SPKTEMBG-ACCQ9V399 D3848.1 (NOVEL 7 TRANSMEMBRAN...)	+3	1076	5.6e-130	1
PtNC:SMTS9PROT-ACC623275 0LFCRGY RECEPTOR L5 IQR3)	+3	1031	3.3e-123	1
PtNC:SPTRFNXP-ACC016081 NOC0189.7 (OLFACTORY RECEPTOR...)	+3	1027	4.7e-103	1
PtNC:SPTRFBVB-ACC056571 OLFACTORY RECEPTOR-LIKE PROT...	+3	992	6.5e-99	1
PtNC:SPTRFBVB-ACC055938 D271MEL.2 (HSM19-17 (3 TRANS...)	+3	888	1.2e-92	1
PtNC:SPTRFB4B-ACCQ9QV14 573R1.1 (SM170.2 (NOVEL 7 TR...	+3	983	3.2e-98	1
PtNC:SPTRFB4B-ACCQ9QV14 573R1.2 (SM170.3 (NOVEL 7 TR...	+3	981	6.6e-98	1

WO 01/64879

PCT/US01/06474

Possible SNPs found for GPCR4a2 are listed in Table4F.

Table 4F: SNPs		
Base Position	Base Before	Base After
44	T	C(20)
147	T	C(2)
220	T	C(2)
271	T	C(3)
432	A	G(2)
452	C	T(2)
493	A	G(2)
771	T	C(3)

GPCR4b

The target sequence identified as dj408b20_C (GPCR4a1) was again 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. The cDNA coding for the sequence was cloned by polymerase chain reaction (PCR) using the following primers:

ATACAAGTTCTTTCGAAGGCTTCATCC (SEQ ID NO:14) and

CCCTTACTATATTTCTGCACTCCCTT (SEQ ID NO:15) on pool 1 of human cDNAs

containing the following: 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.

Primers were designed based on in silico predictions for the full length or part (one or more exons) of the DNA/protein sequence of the invention or by translated homology of the predicted exons to closely related human sequences or to sequences from other species.

Usually multiple clones were sequenced to derive the sequence which was then assembled similar to the SeqCalling process. In addition, sequence traces were evaluated manually and edited for corrections if appropriate. The PCR product derived by exon linking was cloned into the pCR2.1 vector from Invitrogen. The bacterial clone

115843::DJ408B20 C.698322.D10 has an insert covering the entire open reading frame cloned into the pCR2.1 vector from Invitrogen.

WO 01/64879

PCT/US01/06474

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. Typically, the resulting amplicons were gel purified, cloned and sequenced to high redundancy as described in the examples. 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 provided a third nucleic acid encoding a GPCR4 protein. The nucleic acid is referred to as GPCR4a3 or CG55358-03. This nucleic acid is 932 nucleotides long (SEQ ID NO:16, Table 4G) and is 16 nucleotides shorter in the 5' UTR than GPCR2a1. An open reading frame of the mature protein was identified beginning with an ACA codon which codes for the amino acid threonine at nucleotides 3-5 and ending with a TAG codon at nucleotides 924-926. Putative untranslated regions, if any, are found upstream from the initiation codon and downstream from the termination codon. One silent base substitution is present: C767 is T752 in GPCR4a3. GPCR4b, the protein encoded by GPCR4a3 is identical to GPCR4a, except for the 5 amino acid deletion at the N terminus, as shown in Table 4H (SEQ ID NO:17).

The GPCR Olfactory Receptor disclosed herein 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

WO 01/64879

PCT/US01/06474

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.

In a search of sequence databases, it was found, for example, that the disclosed

GPCR4a3 nucleic acid sequence has 614 of 913 bases (67%) identical to a gb:GENBANK-

- 5 ID:AF098664[acc:AF098664.1 mRNA from Homo sapiens (Homo sapiens olfactory receptor-like protein (OR2C1) gene, complete cds, E= 2.1e-71).

Table 4G. GPCR4a3 Nucleotide Sequence (SEQ ID NO:16)

ATACAGATTCTTTCGAAAGCTTCATCCTGGTGGGCTTCTCTGATCCTCCACCTAGAGC	80
TGATCGCTCTTTGTGGTTGTCTCTCTTTTATCTGCTGACTCTTCTTGGCAACATGACCA	120
TTGTCTTGGCTTTCAGCTCTGGATTCCTGGCTGCACACCAATCTATTTCTTTTGGCAA	160
ACCTCTCATTCCTGGACATGTGTTCAGGACAGGTTCCATCCCTCAGATGCTTAGAAGC	200
TTTGGGGTCCAGATAGGACCATGAGCTATGTGGTTCTGTGCACTCAGCTCTATTTGGCC	240
TGTCCTGAGGAGGAGTGGAGTGTCTCTCTGCTGCTCATGGCATTGTACCCCTATGGCTG	280
CTCTCTGCAAGCCCTCTCTCTACGCAATCATCTGACCCGCTCTCTCTGTCCTGTCAGCTG	320
CTTACGTGACAGGCTGAGTGGCTTTGGCAATCTCTCTCATATAGGCTACCCGACATCTGA	360
TGCTACCCCTGCTG:GGGACAGAGAGTGGACCACTTTCTCTGTGAGATGGCACTCA	400
TTGATATGGCTCTCTAGACACCATGATGCTTGAAGGCACTGCTTTTGGCTTACAACTCT	440
CTATCATCTGTGGACACCTCATCTCTCATCTCTTTCTTAAGGTTACGTTGGAAGACAG	480
TGCTTGAAGTCAAGTCACTGCTGCTGGGCAAGAGAGCTTCAACACTTGGCACTTGGCATG	520
TAATTTCTTGTCTCTCTCTCTATGCTACAACTATATACACTAGAGTGGCAAGCAATA	560
CTTATTCGCAACAGGAGGCAAGTTTCTTACCTCTCTCTACCAATCTCTCTCTCTCTG	600
TTTACCCCTGATCTATACACTAGAAACAAGATCTTAAAGACCCATGAAGAACCTCC	640
TAGGAGAGGAGAGTCCAGGATATGTAAGGG	680

The SignalP, Psort and/or Hydropathy profile for the disclosed GPCR4b Olfactory

- 10 Receptor-like protein predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6000. The SignalP shows a signal sequence is coded for in the first 36 amino acids with a cleavage site at between amino acids 36 and 37, as indicated by a slash between LLG/NM in Table 4H. This is typical of this type of membrane protein.

15

Table 4H. GPCR4b Amino Acid Sequence (SEQ ID NO:17)

TSSEFGILVGFSDRPHLELVFVVLIFYLLTLLG/NMTIVLLSALDSRLHTPMYFFLAN	60
LSFLDMCFITGSLPQMLYNLVGPKTISYVGCALQYPVALGGVECVLLAVMAYDRYAA	120
VCKPLHYTIIMHPRLCGQLASVAVLSGFGNSLIMAPQTLMLPRCGHRRVDHFLCEMPALI	180
GMACVDTHMLEALAFALAJFLLAPLILLLSYGYVGGTVLRIKSAAGRKAFNTCSSHL	240
IVVSLFYGTIIYMYLQANTYSQDQGRFLTFYTVTPSVNPLIYTLRNKDVKEAMKKVL	300
GKGSAAE	307

The full amino acid sequence of the disclosed GPCR4b protein of the invention was found to have 195 of 299 amino acid residues (65%) identical to, and 239 of 299 amino acid residues (79%) similar to, the 320 amino acid residue pmr:SPREMBL-ACC:Q9Y3N9

- 20 protein from Homo sapiens (DJ88J8.1, NOVEL 7 TRANSMEMBRANE RECEPTOR

WO 01/64879

PCT/US01/06474

(RHODOPSIN FAMILY) (OLFACTORY RECEPTOR LIKE) PROTEIN) (HS6M1-15), E= 1.6e-107).

In following positions, one or more consensus positions (Cons. Pos.) of the GPCR4a3 nucleotide sequence have been identified as SNPs. "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": Cons. Pos.: 422 Depth: 18 Change: T > C, Putative Allele Freq.: 0.333; Cons. Pos.: 546 Depth: 15 Change: T > C, Putative Allele Freq.: 0.133; Cons. Pos.: 753 Depth: 8 Change: C > T, Putative Allele Freq.: 0.250.

Unless specifically addressed as GPCR4a or GPCR4b, any reference to GPCR4 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant and the residue position with respect to the "a" variant. In all following sequence alignments, the GPCR4a protein sequence was used.

The disclosed GPCR4 protein (SEQ ID NO:12) also has good identity with a number of olfactory receptor proteins. The identity information used for ClustalW analysis is presented in Table 4I.

Table 4I. BLAST results for GPCR4

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (#)	Expect
GI16679170 ref NP_032789.1	OR 15 (OR3) <i>Mus musculus</i>	312	179/308 (56%)	216/308 (69%)	1e-93
GI14828521 emb CAB42853.1 (AL035402, d38849.1) (AJ302594-99) (AJ302600-01)	novel * tm receptor protein (rhodopsin fam., OR-like) (hsGM1 15) <i>Homo sapiens</i>	320	174/305 (57%)	213/305 (69%)	2e-93
GI112054431 emb CAC29523.1 (AJ302603)	OR <i>Homo sapiens</i>	320	173/305 (56%)	213/305 (69%)	3e-93
GI112054429 emb CAC29522.1 (AJ302602)	OR <i>Homo sapiens</i>	320	173/305 (56%)	213/305 (69%)	7e-93
GI112054347 emb CAC29478.1 (AJ302598)	OR <i>Homo sapiens</i>	311	170/302 (56%)	211/302 (69%)	6e-90

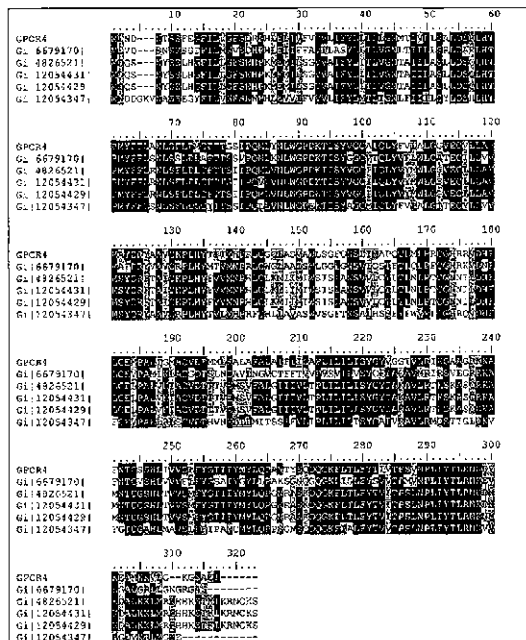
This information is presented graphically in the multiple sequence alignment given in Table 4J (with GPCR4 being shown on line 1) as a ClustalW analysis comparing GPCR4 with related OR sequences.

WO 01/64879

PCT/US01/06474

Table 4J Information for the ClustalW proteins:

- 1) GPCR4 (SEQ ID NO:12)
 2) gi|6679170|ref|NP_032788.1| olfactory receptor 15 Mus musculus (SEQ ID NO:58)
 3) gi|4826521|emb|CAB42853.1| d16818.1 (novel 7 transmembrane receptor (rhodopsin family) (OR like) protein) (het541-15)) Homo sapiens (SEQ ID NO:59)
 4) gi|12054431|emb|CAC20523.1| olfactory receptor Homo sapiens (SEQ ID NO:60)
 5) gi|12054429|emb|CAC20522.1| olfactory receptor Homo sapiens (SEQ ID NO:61)
 6) gi|12054347|emb|CAC20478.1| olfactory receptor Homo sapiens (SEQ ID NO:62)



DOMAIN results for GPCR4 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. Two regions of GPCR4 have identity to the 377 amino acid 7TM

WO 01/64879

PCT/US01/06474

domain, as described above. The 7tm_1, a seven transmembrane receptor (rhodopsin family), SEQ ID NO:45, above, was shown to have homology to GPCR4. Residues 1-120 of 7TM align with residues 41-159 of GPCR4 (E=7e-22, shown in Table 2K) and residues 224-290 of GPCR4 have identity with residues 310-377 of 7TM (E=2e-04).

5

Table 4K: Domain Alignment between GPCR4 and 7TM.

GPCR4	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000
7TM	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

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 below. For example, a cDNA encoding the olfactory receptor-like protein may be useful in gene therapy, and the olfactory receptor-like protein may be useful when administered to a subject in need thereof. The protein similarity information, expression pattern, and map location for the Olfactory Receptor-like protein and nucleic acid disclosed herein suggest that this Olfactory Receptor may have important structural and/or physiological functions characteristic of the Olfactory Receptor 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.

The nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications implicated in various diseases and disorders described below and/or other pathologies. 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, corticosteroidogenic 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

WO 01/64879

PCT/US01/06474

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; and treatment of Albright Hereditary Osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, allergies, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation, dentatorubro-pallidolusian 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 the OR-like protein may be useful in gene therapy, and the OR-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 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 Osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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. Other GPCR-4 diseases and disorders are contemplated.

The novel nucleic acid encoding OR-like protein, and the OR-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 and other diseases, disorders and conditions of the like. 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

WO 01/64879

PCT/US01/06474

known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCR Antibodies" section below. In one embodiment, a contemplated GPCR4 epitope is from about amino acids 65 to 85. In another embodiment, a GPCR4 epitope is from about amino acids 115 to 130. In additional embodiments, GPCR4 epitopes are from amino acids 155 to 175, from 215 to 240, from 250 to 275 and from amino acids 280 to 310. These novel proteins can also be used to develop assay system for functional analysis.

GPCR5

GPCR5 includes a family of three similar nucleic acids and three similar proteins disclosed below. The disclosed nucleic acids encode GPCR, OR-like proteins.

GPCR5a

The disclosed novel GPCR5a nucleic acid of 1003 nucleotides (also referred to as 115-a-12-A) is shown in Table 5A. An ORF begins with an ATG initiation codon at nucleotides 6-8 and ends with a TAA codon at nucleotides 999-1001. A putative untranslated region upstream from the initiation codon and downstream from the termination codon is underlined in Table 5A, and the start and stop codons are in bold letters.

Table 5A. GPCR5a Nucleotide Sequence (SEQ ID NO:18)

<u>AGACCATGAGTCTGATGGGACCAACAGCTAGTGCACACAGAGTTGGTCTGGGAGGGCTCCCAATCT</u>
<u>CAATGAGCGGAGAGCTGGAACTTATTTCTGTGTTTCTCTTGTCTATCTCCTGAATCTGACAGGCAATGTG</u>
<u>TTGATTTGGGCGCTGCTAAGGGCTGATCTGACATCCAGAGCCCTATGTACTCTTTCTGGTAACCTGT</u>
<u>CGTGCTAGAGATAGTGTGCTGCTTCTGTCTATCTCCAAAGATGCTTAGCAATTCTCTCCCAAGCCACACA</u>
<u>CACTATTTCTCTTGTGCACTATATCACCCAAATTCATTTCTACTCTCTTCTGGGGCTCCGATCTCTA</u>
<u>CTGTGCTGCTGATGCTGGGATGGCTACCTGGCCATCTGTCTCTCTCTGCGCTACCCCTTGTCTCATGA</u>
<u>GGAGGCTGTGTGCTTCTGTGTGGCTTGGCTCTGCTGGTGGGGAGTCTGCTGTGTGTGTGTGTGTGTGT</u>
<u>ACTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT</u>
<u>CCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT</u>
<u>TCAATGTATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT</u>
<u>GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT</u>
<u>GGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT</u>
<u>TAAATGCGGACATTTCTCTATGTGCGGCCATCCAGAGTGGTCTGTGTGTGTGTGTGTGTGTGTGTGT</u>
<u>AGCTTTGAGGAGACATGTTTAGGAGGTAGTGGCAGGCGTTTAGGGAATCTTTACTTGTATAAATGTCT</u>
<u>AGTGAGAGGACATAACTAAAT</u>

The GPCR5a protein encoded by SEQ ID NO:18 has 331 amino acid residues and is presented using the one-letter code in Table 5B. The Psort profile for GPCR5a predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6000, it may also localize to the Golgi body. The most likely cleavage site for a peptide is between amino acids 54 and 55, i.e., at the slash in the amino acid sequence VRA-DT (shown as a slash in Table 5B) based on the SignalP result.

WO 01/64879

PCT/US01/06474

Table 5B. Encoded GPCR5a protein sequence (SEQ ID NO:19)

```

MSDQGNHSSDPTSEFVLACLPMLNSARVELFSVFLVYLNLITGNVLIVGV9RA/DTRLQTEHYFF
LQNLSCLEILLTGVLIETWILSNFLSPONTISFADCIQFYFYFFLGASEFILLAVMSADRYLAIC
HPLRYPLMSGAVCIKVVNLACWVGLVPLGPTVAVALLPCKQGA/VQHFCDSPQLRLACTN
TKXLSEIDQVILASLIVVSSLLITAVSYGLEVLAVLSIPASGQKAEFTCTSHLIVVTLFYGSAT
FLVVRPSQSGSVDTMNAVITITFTVPLLNPFYALPNEQWKEALKDMERKVVAGVLGNLLEDKC
LSEKAVK

```

The disclosed nucleic acid sequence for GPCR5 has 604 of 934 bases (64%) identical to and 604 of 934 bases (64%) positive with *Rattus norvegicus* olfactory receptor protein mRNA (936 bp) (GENBANK-ID: RATOLFPROD acc:M64378) ($E=1.1e^{-45}$).

- 5 The full GPCR5 amino acid sequence has 149 of 304 amino acid residues (49 %) identical to, and 201 of 304 residues (66%) positive with, the 313 amino acid residue olfactory receptor from *Mus musculus* (ptnr: SPTREMBL-ACC: Q9Z1V0) ($E=2.6e^{-74}$).

GPCR5b

- 10 GPCR5a (115-a-12-A) was subjected to an 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. 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
- 15 suitable sequences were then employed as the forward and reverse primers in a PCR amplification based on a wide range of cDNA libraries. The resulting amplicon was gel purified, cloned and sequenced to high redundancy to provide GPCR5b, which is also referred to as 115-a-12-B.

- 20 The nucleotide sequence for GPCR5b (1004 bp, SEQ ID NO:20) is presented in Table 5C. The nucleotide sequence differs from GPCR5a by the addition of a T between A5 and A6, and by 6 nucleotide changes (numbered with respect to GPCR5a) C131>T; T186>C; G472>A; T579>A; A687>T; C799>T.

WO 01/64879

PCT/US01/06474

Table 5C. GPCR5b Nucleotide Sequence (SEQ ID NO:20)

```

AGCATATGAGTCCTGATGGCAACACAGTAGTGAATCCAGACAGCTTCCCTCCGACGGCTCCCAA
ATCTCAGACAGCGAAGAGTGGAAATATTTCTGTGTCTCTCTGTCTTCTCTCTGAAATCTGTAGG
CAATGCTTGTATGTTGGGGGTGTAAAGGCTGTACTTGGACTACAGACCCGAAATTAATCTTTCTG
GGTAACTCTCTCTGCTAGAGATATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
TCTCAAGGCAACACACTATTTCTTTTGGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
GGCTCCGAGTTCTTACTGCTTGGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
GGCTCCGCTTCTCTGATGAGTGGGCTGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
ACGACACTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
GAGACAGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
GGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CAGACTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
ATCGATTCATCTATGCTTACGTAATCAACCAACCAAGGCTTCTGAGGAGATGCTTACGAAAGT
AGTGGCAGGCTTTAGGGAATCTTTTACTTGTAAATGCTCTGAGTGAAGACAGTAAAGTAA

```

The encoded GPCR5b protein is presented in Table 5D. The disclosed protein is 331 amino acids long and is denoted by SEQ ID NO:21. GPCR5b differs from GPCR5a by 3 amino acid residues: T42>I; V151>I; P265>S. Like GPCR5a, the Psort profile for GPCR5b predicts that this sequence has a signal peptide and is likely to be localized at the plasma membranes with a certainty of 0.6000. The most likely cleavage site for a peptide is between amino acids 54 and 55, i.e., at the slash in the amino acid sequence VRA-DT (shown as a slash in Table 5D) based on the SignalP result.

Table 5D. Encoded GPCR5b protein sequence (SEQ ID NO:21)

```

MSFDGNHSSDPTFVLAGLHINSARVELFSVFLVYLLNLGNVLIVGVVRA/DTEIQTFFHYFFLGN
LSGLE/LITSVIIPKMLSNPLSRQHTISFACTITQFYFFFLGASEFLLAVMSADRYLAICHPLRYP
LMSGVAVCFVVALACWVSGLIPIVLSPTVAVAILPFCXGAVVQRFCDGSPILLRLACTINTKLEETDF
VLASLVIVSSLLITAVSYCLIVLAVLSIPGASGQKAFSTQSHLVVTI.FYGSALFLYVRSSQSCSV
DTNNAVTVITTFVTPLNFFIYALRNEQKEALXDMPEKVVAGVLGNLLDKCLSPRAVK

```

10

BLASTP (Non-Redundant Composite database) analysis of the best hits for alignments with GPCR5b are listed in Table 5E.

Table 5E. BLASTP results for GPCR5b

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (t)	Expect
SPTREMBL- ACC:Q921V6	OLFACTORY RECEPTOR C6 <i>Mus musculus</i>	313	150/304 (49%)	203/304 (66%)	9.4e-76
SWISSPROT- ACC:P23267	OLFACTORY RECEPTOR- LIKE PROTEIN P6 <i>Rattus norvegicus</i>	311	152/301 (50%)	195/301 (65%)	5.2e-75
SPTREMBL- ACC:Q9V325 DJ994K9.5	RGS6L-1P (RHOGEF 3 TM (RHODOPSIN FAMILY) (OR LIKE PROTEIN; <i>Homo sapiens</i>)	306	145/301 (48%)	197/301 (65%)	1.1e-67

WO 01/64879

PCT/US01/06474

A BLASTX was also performed to determine the proteins that have significant identity with GPCR4a. The BLASTX results are shown in Table 5F.

Table 5F. BLASTX results for GPCR5b

Sequences producing high-scoring Segment Pairs:	Reading Frame	High Score	Smallest Sum Prot P(N)	n
ptm:SPTRMBL-ACC:096190 OLFACTORY RECEPTOR CG - Mus M...	+1	764	5.6e-75	1
ptm:SWISSPROT-ACC:223267 OLFACTORY RECEPTOR-LIKE PROT...	+1	757	3.1e-74	1
ptm:SWISSPROT-ACC:223270 OLFACTORY RECEPTOR-LIKE PROT...	+1	707	2.4e-68	1
ptm:SPTRMBL-ACC:091395 OJB94F3.5 (H5H1-1) (ROR1) ? ...	+	688	6.9e-67	1
ptm:SPTRMBL-ACC:070271 OLFACTORY RECEPTOR-LIKE PROTE...	+1	689	5.3e-67	1
ptm:SPTRMBL-ACC:093007 WGGSC15_030669B10.3 PROTEIN ...	+1	680	4.5e-66	1
ptm:SPTRMBL-ACC:013036 CHICK OLFACTORY RECEPTOR 7 - ...	+1	675	1.5e-65	1
ptm:SPTRMBL-ACC:095222 OLFACTORY RECEPTOR - Homo sap...	+2	672	3.1e-65	1
ptm:SPTRMBL-ACC:070269 OLFACTORY RECEPTOR-LIKE PROTE...	+2	670	3.1e-65	1
ptm:SPTRMBL-ACC:057597 CHICK OLFACTORY RECEPTOR 7 - ...	+1	669	5.5e-65	1
ptm:SPTRMBL-ACC:070270 OLFACTORY RECEPTOR-LIKE PROTE...	+2	668	8.4e-65	1
ptm:SPTRMBL-ACC:070270 OLFACTORY RECEPTOR-LIKE PROTE...	+2	668	8.4e-65	1
ptm:SPTRMBL-ACC:070270 OLFACTORY RECEPTOR-LIKE PROTE...	+2	666	1.4e-64	1
ptm:SPTRMBL-ACC:090886 ODOANT RECEPTOR S1 - Mus mus...	+1	665	1.7e-64	1
ptm:SPTRMBL-ACC:090806 OLFACTORY RECEPTOR 4 - Gallus...	+1	665	1.7e-64	1
ptm:SWISSPROT-ACC:237071 OLFACTORY RECEPTOR-LIKE PROT...	+1	646	1.8e-62	1

GPCR5c

Another nucleotide sequence resulted when GPCR5a (115-a-12-A) was subjected to an 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. 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 suitable sequences were then employed as the forward and reverse primers in a PCR amplification based on a wide range of cDNA libraries. The resulting amplicon was gel purified, cloned and sequenced to high redundancy to provide the sequence reported below, which is designated as Accession Number 115_A_12_A_da1, or GPCR5c.

Human tissues providing SeqCalling Fragments of the clone include Pool One: 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 tissue origin of the clone is RACE(asm:126603384).

PCV/USOI/06474

5

Table 5G. GPCR5c Nucleotide Sequence (SEQ ID NO:22)

[illegible]

10

15

Table 5H. Encoded GPCR5h protein sequence (SEO ID NO:23)

MSPDGHHSDPTFLVLAGLPNLNSARVELFSVFLVYLPNLIGVFLVGVVRA/DTRLQTFMYFFLGNLSC
*TLLTSVTLRPMNLNN*SKMTLSRACTORYEYFFLGASEFLLVMSADRYLAICHELRYPLMSG

WO 01/64879

PCT/US01/06474

that differ between the individual GPCR variants are highlighted with a box and marked with the (o) symbol above the variant residue in all alignments herein. For example, the protein shown in line 1 of Table 5J depicts the sequence for GPCR5a, and the positions where GPCR5b or GPCR5c differs are marked with a (o) symbol and are highlighted with a box. All GPCR5 proteins have significant homology to olfactory receptor (OR) proteins:

Table 5I. BLAST results for GPCR5

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (#)	Expect
G1129091 sp P23267 OLFR_RAT	OR L5 (OR3) <i>Rattus norvegicus</i>	311	144/301 (47%)	189/301 (61%)	7e-57
G16754932 ref NP_035121.1 (AF102525)	OR 49, OR56 <i>Mus musculus</i>	313	145/305 (47%)	192/305 (62%)	3e-56
G11242465 ref NP_035113.1 (AF105007) (AF321234)	OR 41, OR7 <i>Mus musculus</i>	254	136/304 (44%)	181/304 (59%)	9e-51
G117363437 ref NP_039229.1	OR, family 10, subfamily C, member 1 <i>Homo sapiens</i>	306	139/301 (46%)	186/301 (61%)	2e-59
G112007431 gb AA045202.1 (AF321236) (AF321238)	m50 OR <i>Mus musculus</i>	316	129/303 (42%)	181/303 (59%)	4e-59

This information is presented graphically in the multiple sequence alignment given in Table 5J (with GPCR5 being shown on line 1) as a ClustalW analysis comparing GPCR5 with related protein sequences.

Table 5J Information for the ClustalW proteins:

- 1) GPCR5 (SEQ ID NO:19)
- 2) g1129091|sp|P23267|OLFR_RAT OLFACTORY RECEPTOR-LIKE PROTEIN F6 (SEQ ID NO:63)
- 3) g16754932|ref|NP_035121.1| olfactory receptor 49 *Mus musculus* (SEQ ID NO:64)
- 4) g11242465|ref|NP_035113.1| olfactory receptor 41 *Mus musculus* (SEQ ID NO:65)
- 5) g117363437|ref|NP_039229.1| olfactory receptor, family 10, subfamily C, member 1 *Homo sapiens* (SEQ ID NO:66)
- 6) g112007431|gb|AA045202.1| m50 olfactory receptor *Mus musculus* (SEQ ID NO:67)

	10	20	30	40	50	60
GPCR5						
G1129091						
G16754932						
G11242465						
G117363437						
G112007431						
	70	80	90	100	110	120
GPCR5						
G1129091						

WO 01/64879

PCT/US01/06474

G1167549321	
G1172421651	
G1173643371	
G1120074311	
GPCR5	
G11290911	
G1167549321	
G1172421651	
G1173643371	
G1120074311	
GPCR5	
G11290911	
G1167549321	
G1172421651	
G1173643371	
G1120074311	
GPCR5	
G11290911	
G1167549321	
G1172421651	
G1173643371	
G1120074311	
GPCR5	
G11290911	
G1167549321	
G1172421651	
G1173643371	
G1120074311	

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 5K with the statistics and domain description.

Residues 1-115 of 7tm_1 (SEQ ID NO:45) are aligned with GPCR5 43-156 (E = 1e-20), in Table 5K. Residues 314-377 of 7tm_1 also have identity with residues 231-293 (E=2e-05) of GPCR5.

Table 5K. DOMAIN results for GPCR5

gn113famipfam00701. 7tm_1, 7 transmembrane receptor (rhodopsin family) (SEQ ID NO: 45) Length = 377
Statistics for GPCR5: Score = 93.6 bits (231), Expect = 1e-20
GPCR5
7tm7
GPCR5
7tm7

10

WO 01/64879

PCT/US01/06474

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 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 infections such as 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 Osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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. 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 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 Osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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 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

WO 01/64879

PCT/US01/06474

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. This novel protein also has value in development of powerful assay system for functional analysis of various human disorders, which will help in understanding of pathology of the disease and development of new drug targets for various disorders.

GPCR6

The disclosed novel GPCR6 nucleic acid of 948 nucleotides (also referred to as 6-L-19-C) is shown in Table 6A. An open reading begins with an ATG initiation codon at nucleotides 7-9 and ends with a TAG codon at nucleotides 940-942. A putative untranslated region 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. GPCR6 Nucleotide Sequence (SEQ ID NO:24)

```

GCGGACATGGGCAAGGAAACCTGCACCCCTGGCTGAGTTCAATCTCCTGGACTATCAGATGTCCTG
AGTTGAGAGCTGCGCTCTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
GATTGCACTGATTCAGTTCAGTCTCTGGCTCCGACGACGACGACGACGACGACGACGACGACGACGAC
GTAGATTCTGCTACTGCTCAATATATGTCGCAAAATGTTGCTATATATCTTAAACAGGACGACGACGAC
TCTCTTCTCTAGGCTGATGAGGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
GGCGGATGAGGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
AAATGAGTCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
TAGCTCTAGGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
AAGCTCTGCTTGTCTCTGATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
GTACCTCTATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
AGGCGAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG
CCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
ACGCTCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
GAAATGATGGGCTCCAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT

```

The disclosed nucleic acid sequence has 617 of 915 bases (67%) identical to a *G. gallus* cor4 olfactory receptor 4 DNA (GENBANK-ID: GGCOR4GEN|acc:X94744) (E value = 8.7×10^{-65}).

The GPCR6 protein encoded by SEQ ID NO:24 has 318 amino acid residues, and is presented using the one-letter code in Table 6B (SEQ ID NO:25). The SignalP, Psort and/or Hydropathy profile for GPCR6 predict that GPCR6 has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6000. The SignalP shows a signal sequence is coded for in the first 41 amino acids, i.e., with a cleavage site at the slash in the sequence TLL-AN, between amino acids 40 and 41. This is typical of this type of membrane protein.

WO 01/64879

PCT/US01/06474

Table 6B. Encoded GPCR6 protein sequence (SEQ ID NO:25).

```

MGERCTTVAEFILGLSDWPELRVCLFLFLIYGVTL/ANLGHIALIQVSSRLHTPWFFLSHSSVD
FCYSGIIVPKMLANIPMDKHAISFLGHVQVFLFCQCVVLEVFLAMWYDRFVAICNPLCYTVTM3WNV
RVPLASCCYFGGTVCOSIIHLGLALRIFFYRSMVINRFFCLPEVLSLACSDITYNETLFLVAT2NESVT
IMIIILSYLLIITLTKMGSAREGRKAFSTCASHLTATVFMGTVLSIYCRFSQNSGDDADKATVVFYTV
VIPELRSVTYSLAPKOVKEALRKVMGSEIHS

```

The full amino acid sequence of the protein of the invention was found to have 166 of 307 amino acid residues (54%) identical to, and 217 of 307 residues (70%) positive with, the 314 amino acid residue human olfactory receptor-like protein OLF1 (p1nr:SWISSPROT-ACC:Q13606) (E value = $5.8e^{-86}$).

The GPCR6 target sequence identified previously (6_L_19_C) was 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. 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.

The cDNA coding for the sequence was cloned by polymerase chain reaction (PCR) using the following primers: CACTGIGGCTGAGTTCATTCTCCTT (SEQ ID NO:26) and TCTTCCCTAGGAGTGAATTTGGAGC (SEQ ID NO:27) on the following pool of human cDNAs: Pool 1 - 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. Primers were designed based on in silico predictions for the full length or part (one or more exons) of the DNA/protein sequence of the invention or by translated homology of the predicted exons to closely related human sequences or to sequences from other species. Usually multiple clones were sequenced to derive the sequence which was then assembled similar to the SeqCalling process. In addition, sequence traces were evaluated manually and edited for corrections if appropriate.

The PCR product derived by exon linking was cloned into the pCR2.1 vector from Invitrogen. The bacterial clone 55446::6_L_19_C.698018.M1 has an insert covering the entire open reading frame cloned into the pCR2.1 vector from Invitrogen.

Usually the resulting amplicons were gel purified, cloned and sequenced to high redundancy. The resulting sequences from all clones were assembled with themselves, with other fragments in CuraGen Corporation's database and with public ESTs. Fragments and

WO 01/64879

PCT/US01/06474

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 below, which is designated Accession Number CG50383_01 which does not differ from GPCR6 (6_L_19_C) in amino acid or nucleotide sequence.

The disclosed GPCR6-Olfactory Receptor-like protein 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, tissues that express MHC II and III, nervous tissue, 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 SeqCaling sources, Public EST sources, Literature sources, and/or RACE sources.

The disclosed GPCR6 protein (SEQ ID NO:25) has good identity with a number of olfactory receptor proteins. The identity information used for ClustalW analysis is presented in Table 6C. The GPCR6 protein has significant identity to olfactory receptor (OR) proteins:

Table 6C. BLAST results for GPCR6					
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (#)	Expect
G15729960 ref NP_006028.1	OR fam. 5, subfam. 1, mem 1 <i>Rattus norvegicus</i>	314	154/307 (50%)	200/307 (64%)	2e-72
G12495054 sp:Q95155 OLF2 CANFA	OR-like prt OLF2 <i>Canis familiaris</i>	312	152/306 (49%)	207/306 (66%)	8e-70
G111692519 gb AAG38856.1 AF282271.1 (AF282271)	OR 41, E11 <i>Mus musculus</i>	314	150/305 (49%)	199/305 (63%)	8e-69
G13746443 gi1AAC63369.1 (AF045677)	OR, OR93ch <i>Ran. trielodytes</i>	314	151/306 (49%)	199/306 (64%)	1e-68
G13746448 gi1AAC63971.1 (AF045580)	OR OR93Gib <i>Nyctobates lar</i>	313	149/305 (49%)	198/305 (64%)	3e-68

WO 01/64879

PCT/US01/06474

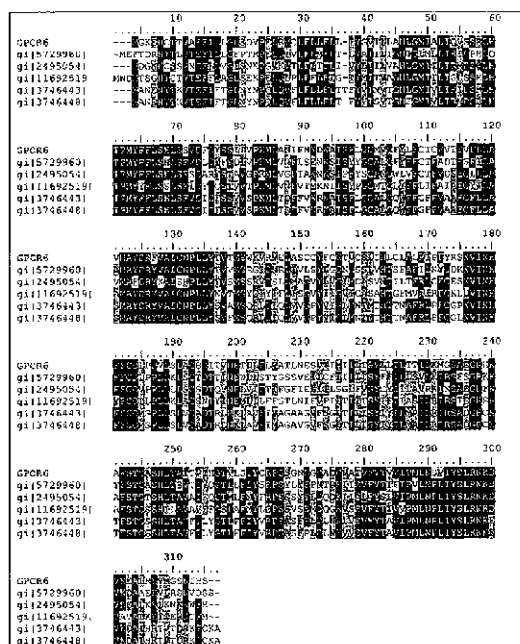
This information is presented graphically in the multiple sequence alignment given in Table 6D (with GPCR6 being shown on line 1) as a ClustalW analysis comparing GPCR6 with related protein sequences.

Table 6D Information for the ClustalW proteins:

5

10

- 1) GPCR6 (SEQ ID NO:25)
- 2) gi5729960|ref|NP_006628.1| OR, family 5, subfamily 1, member 1 Homo sapiens (SEQ ID NO:25)
- 3) gi2495054|sp|Q95155|OLF2_CANFA OR-LIKE PROTEIN OLF2 (SEQ ID NO:25)
- 4) gi11692519|gb|AAC39856.1|AF282271_1 odorant receptor K11 Mus musculus (SEQ ID NO:25)
- 5) gi3746443|gb|AAC63969.1|olfactory receptor OR93Ch [Pan troglodytes] (SEQ ID NO:25)
- 6) gi3746448|gb|AAC63971.1|olfactory receptor OR93Gib [Hylotates lar] (SEQ ID NO:25)



PC-T/USOL/06474

5 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 6E with the statistics and domain description. The results indicate that this protein contains the 7tm_1 (InterPro) 7 transmembrane receptor (rhodopsin family) (as defined by Interpro) at residues 42-203, which align with residues 2-158 of the 7TM domain. This indicates that the sequence of GPCR6 has properties similar to those of other proteins known to contain this/these domain(s) and similar to the properties of these domains.

[illegible]

49

WO 01/64879

PCT/US01/06474

encoding GPCR6, and the GPCR6 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.

The nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications implicated in various diseases and disorders described below and/or other pathologies. 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, corticosterogenic 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 neoplasia; 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 osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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 the OR-like protein may be useful in gene therapy, and the OR-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 bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to Neoplasia; 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 osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, allergies, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety,

PCV/USOI/06474

The novel nucleic acid encoding OR-like protein, and the OR-like protein of the
5 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.

GPCR7

15 A novel GPCR nucleic acid was identified by TblastN using CuraGen Corporation's sequence file for GPCR probes or homologs, and run against the Genomic Daily Files made available by GenBank. The nucleic acid was further predicted by the program GenScan™, including selection of exons. These were further modified by means of similarities using BLAST searches. The sequences were then manually corrected for apparent inconsistencies, 20 thereby obtaining the sequences encoding the full-length protein. The disclosed novel GPCR7 nucleic acid of 1013 nucleotides (also referred to as dj313i6_D) is shown in Table 7A. An open reading begins with an ATG initiation codon at nucleotides 5-7 and ends with a TAG codon at nucleotides 997-999. A putative untranslated region upstream from the initiation codon and downstream from the termination codon are underlined in Table 7A, and the start 25 and stop codons are in bold letters.

Table 7A. GPCR7 Nucleotide Sequence (SEQ ID NO:28)

[illegible]

WO 01/64879

PCT/US01/06474

The disclosed nucleic acid sequence has 615 of 939 bases (65%) identical to a *Homo sapiens* olfactory receptor-like protein (OR2C1) gene (GENBANK-ID: AF098664) (E value = $1.7e^{-87}$).

The GPCR7 protein encoded by SEQ ID NO:28 has 327 amino acid residues, and is presented using the one-letter code in Table 7B (SEQ ID NO:29). The SignalP, Psort and/or Hydropathy profile for GPCR7 predict that GPCR7 has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6000. The SignalP shows a signal sequence is coded for in the first 44 amino acids, i.e., with a cleavage site at the slash in the sequence GNG-TI, between amino acids 43 and 44. This is typical of this type of membrane protein.

Table 7B. Encoded GPCR7 protein sequence (SEQ ID NO:29).

MERANDSTFSGFLLGFSNRPQLETALFVVLLFLYFLSFLGNG/TITLISIVDPR/HTPMYFFTSNLS
SEMDLCLTCTVPQTANNFKGDKTITYGGCVTQLFIALGLGGGVECVLLSAMAURYAAVCREGLIY
MYSMLPQLCLGLVVTWLTGFGNSVTQALTHTLFLCCKNQVDHFFCEVPVMLRLSCTNTSINLALI
FAVSYFFLVFLSLLASVGHITRAVLKTKSAQGRQKAFGTGSHRLVVIIFFGTLISMYLQPESSY
SQDVNKSIALEYTLVCPLLNPLITLRNKEVKGATKNTSGECHRQMRKLTOGLOFOTTVH

The full amino acid sequence of the protein of the invention was found to have 183 of 304 amino acid residues (60%) identical to, and 229 of 304 residues (75%) positive with, the 313 amino acid residue OL1 receptor protein from *Rattus norvegicus* (plnr:SPTREMBL-ACC:Q63394) (E value = $2.6e^{-97}$). Further BLAST analysis produced the significant results listed in Table 7C. The disclosed GPCR7 protein (SEQ ID NO:29) has good identity with a number of olfactory receptor proteins.

Table 7C. BLAST results for GPCR7

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect c
g1:111177961 (trc:NP_068632.1) (L34074)	OR <i>Rattus norvegicus</i>	313	168/304 (55%)	209/304 (69%)	2e-86
g1:110944516 (emb:CA014158.1) (AL133267) c0408929.2	novel T77-OR family (MS6M1-32) <i>Homo sapiens</i>	313	165/304 (53%)	207/304 (67%)	2e-95
g1:120544111 (emb:CA020523.1) (AF102593)	OR #1, RL1 <i>Homo sapiens</i>	357	166/304 (54%)	206/304 (67%)	1e-83
g1:120543931 (emb:Ch020504.1) (AF302584 - 582)	OR <i>Homo sapiens</i>	357	165/304 (54%)	206/304 (67%)	4e-83
g1:30804671 (emb:CA011427.1) (Z99744)	OR <i>Homo sapiens</i>	310	165/304 (54%)	206/304 (67%)	5e-83

WO 01/64879

PCT/US01/06474

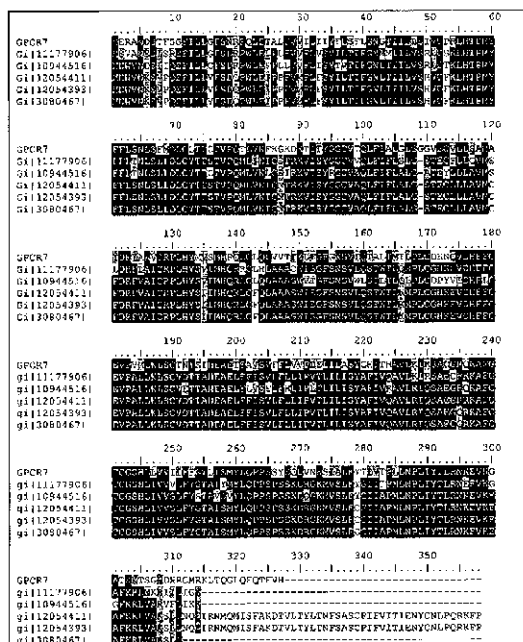
This information is presented graphically in the multiple sequence alignment given in Table 7D (with GPCR7 being shown on line 1) as a ClustalW analysis comparing GPCR7 with related protein sequences.

5

Table 7D. Information for the ClustalW proteins:

10

- 1) GPCR7 (SEQ ID NO:29)
- 2) gi11177906ref|NP_068612.1| Olfactory receptor Rattus norvegicus (SEQ ID NO:73)
- 3) gi10944516emb|CAC14158.1| dJ408B10.2 (novel) 7 TM receptor (olfactory family) (hS6M1-32) (Homo sapiens) (SEQ ID NO:74)
- 4) gi12054411emb|CAC20513.1| olfactory receptor Homo sapiens (SEQ ID NO:75)
- 5) gi12054393emb|CAC20504.1| olfactory receptor Homo sapiens (SEQ ID NO:76)
- 6) gi13080467emb|CAH11427.1| olfactory receptor Homo sapiens (SEQ ID NO:77)



PC7/US01/06474

DOMAIN results for GPCR7 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 7E with the statistics and domain description. The results indicate that this protein contains the following protein domains (as defined by Interpro) at the indicated positions: domain name 7tm_1 (InterPro) 7 transmembrane receptor (rhodopsin family). Residues 61-142 of 7tm_1 (SEQ ID NO:45) are aligned with GPCR7 41-180 (E = 5e-18) in Table 7E. Residues 307-377 of 7tm_1 also have identity with residues 222-291 (E=0.001) of GPCR7. This indicates that the sequence of GPCR7 has properties similar to those of other proteins known to contain this/these domain(s) and similar to the properties of these domains.

[illegible]

20 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

WO 01/64879

PCT/US01/06474

regeneration in vitro and in vivo (vi) biological defense weapon. The novel nucleic acid encoding GPCR7, and the GPCR7 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.

- 5 The nucleic acids and proteins of the invention are useful in potential therapeutic applications implicated in used in the treatment of infections such as 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, 10 hypertension, urinary retention, osteoporosis, Crohn's disease; multiple sclerosis; and treatment of Albright hereditary osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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 15 syndrome and/or other pathologies and disorders.

- The disclosed GPCR7 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 20 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, 25 hypertension, urinary retention, osteoporosis, Crohn's disease; multiple sclerosis; and treatment of Albright hereditary osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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 30 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

PC-T/USOL/06474

GPCRs

A novel nucleic acid was identified on chromosome 6 by TblastN using CuraGen Corporation's sequence file for GPCR probes or homologs and, run against the Genomic Daily Files made available by GenBank. The nucleic acid was further predicted by the program GenScan™, including selection of exons. These were further modified by means of similarities using BLAST searches. The sequences were then manually corrected for apparent inconsistencies, thereby obtaining the sequences encoding the full-length protein. The disclosed novel GPCR8 nucleic acid of 958 nucleotides (also referred to as dj408b20_A) is shown in Table 8A. An open reading begins with an ATG initiation codon at nucleotides 4-6 and ends with a TGA codon at nucleotides 955-957. A putative untranslated region upstream from the initiation codon and downstream from the termination codon are underlined in Table 8A, and the start and stop codons are in bold letters.

[illegible]

The disclosed nucleic acid sequence has 768 of 1148 bases (66%) identical to a *Homo sapiens* OR-like (gb:GENBANK-ID:HS88J8|acc:AL035402 Human DNA sequence from clone 88J8 on chromosome 6p21.31-21.33. Contains a gene for a novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein, pseudogene similar to olfactory receptor genes and a GTP binding protein SARA (mouse) pseudogene. Contains ESTs, an STS and GSS, complete sequence - *Homo sapiens*, 47216 bp.) (E value = 1.3e⁻⁸¹).

The GPCR8 protein encoded by SEQ ID NO:30 has 317 amino acid residues, and is presented using the one-letter code in Table 8B (SEQ ID NO:31). The SignalP, Psort and/or

WO 01/64879

PCT/US01/06474

Hydropathy profile for GPCR8 predict that GPCR8 has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6850. The SignalP shows a signal sequence is coded for in the first 42 amino acids, i.e., with a cleavage site at the slash in the sequence VLG-NS, between amino acids 41 and 42. This is typical of this type of membrane protein.

Table 8B. Encoded GPCR8 protein sequence (SEQ ID NO:31).

```
MEKENSSVYGFILVGFSDRPKLELVLTUNFILYSVAVLG/NETTILVCTLDGQHTMYFFLANLSEF
LDLCFSTSCIPQMLVNLWGPDKTISCAGCVVQLFSLVSVAGIKCTILAVKAYCSYAAVCKPLRFLYIMH
LQLCLGLIAAAWGGGLNNAVMSPLTMTLSRSCRRRVHFLCEKPLLIRMACLDRAVENLAFAPAVLI
VLPLTLFLVSYGLAABVLSIPSAARQWKAHTCSSHLTVVSLFYGSIIYHMHQPCNSGGQDQKPLT
LEFYNITPMINLLLYPLRNKEVFGALKKYLGRQNELEKYDKL
```

The full amino acid sequence of the protein of the invention was found to have 187 of 305 amino acid residues (61%) identical to, and 239 of 305 residues (78%) positive with, the 320 amino acid residue novel transmembrane receptor (rhodopsin family, OR-like, HS6M1-15) protein from *Homo sapiens* (pnr:SP TREMBL-ACC:Q9Y3N9) (E value = 2.4×10^{-10}). Further BLAST analysis produced the significant results listed in Table 8C. The disclosed GPCR8 protein (SEQ ID NO:31) has good identity with a number of olfactory receptor proteins.

Table 8C. BLAST results for GPCR8

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
G14926521 emb CAB47853.1 (AJ035402, d78879.1) (AJ302594-99) (AJ302600-01)	novel ? tm receptor protein (rhodopsin fam., OR-like) (hs6M1-15) <i>Homo sapiens</i>	320	177/305 (58%)	226/305 (74%)	5e-52
G1112054431 emb CAC20523.1 (AJ302603)	OR <i>Homo sapiens</i>	320	176/305 (55%)	226/305 (73%)	3e-51
G1112054429 emb CAC20522.1 (AJ302602)	OR <i>Homo sapiens</i>	320	177/305 (55%)	225/305 (73%)	1e-51
G116679170 ref NP_052789.1	OR 15 (OR15) <i>Mus musculus</i>	312	166/305 (54%)	211/307 (68%)	2e-51
G112231029 sp Q15062 Q2M3 HUMAN	OR 2B3 <i>Homo sapiens</i>	316	163/306 (53%)	201/308 (65%)	5e-51

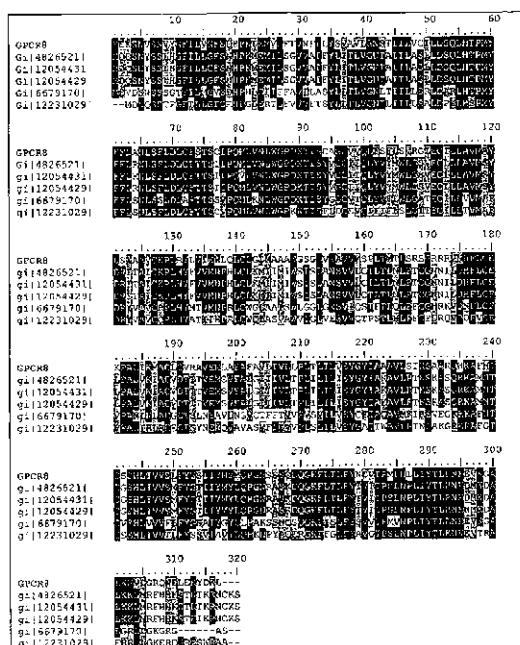
This information is presented graphically in the multiple sequence alignment given in Table 8D (with GPCR8 being shown on line 1) as a ClustalW analysis comparing GPCR8 with related protein sequences.

WO 01/64879

PCT/US01/06474

Table 8D. Information for the ClustalW proteins:

- 1) GPCR3 (SEQ ID NO:31)
- 2) gi4826521|emb|CAC42853.1|dJ8838.1 (novel 7 TM receptor (rhodopsin family) (OR like) protein) (hu5M1.15) Homo sapiens (SEQ ID NO:59)
- 3) gi12054429|emb|CAC20523.1| olfactory receptor Homo sapiens (SEQ ID NO:60)
- 4) gi12054429|emb|CAC20523.1| olfactory receptor Homo sapiens (SEQ ID NO:61)
- 5) gi6679170|ref|NP_032783.1| olfactory receptor 15 Mus musculus (SEQ ID NO:58)
- 6) gi12231029|ref|Q15062|O2H3_HUMAN OR 2H3 (OR-LIKE PRT FAT11) (SEQ ID NO:56)



The presence of identifiable domains in GPCR3 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/>).

PC-T/USOL/06474

5

[illegible]

15

25

WO 01/64879

PCT/US01/06474

treatment of Albright hereditary osteodystrophy and/or other pathologies and disorders. 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 neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer; immune response; AIDS; asthma; Crohn's disease; multiple sclerosis; and treatment of Albright hereditary osteodystrophy. 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 immuno-specifically to the novel GPCR substances 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.

15 GPCR9

A novel nucleic acid was identified on chromosome 11 by TblastN using CuraGen Corporation's sequence file for GPCR probe or homolog, run against the Genomic Daily Files made available by GenBank. The nucleic acid was further predicted by the program GenScan™, including selection of exons. These were further modified by means of similarities using BLAST searches. The sequences were then manually corrected for apparent inconsistencies, thereby obtaining the sequences encoding the full-length protein. The disclosed novel GPCR9 nucleic acid of 946 nucleotides (also referred to as 6-L-19-B) is shown in Table 9A. An open reading begins with an ATG initiation codon at nucleotides 5-7 and ends with a TAA codon at nucleotides 932-934. A putative untranslated region 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.

In a search of sequence databases, it was found, for example, that the nucleic acid sequence has 563 of 902 bases (62%) identical to (E= 8.5e-47) a *Mus musculus* gene for odorant receptor A16 (GENBANK-ID: GENBANK-ID: AB030896|acc:AB030896). In a search of sequence databases, partial matches (353 of 373 bases, 94% identical) were also identified with the nucleotides 1- 372 of GPCR9 identical with nucleotides 306-678 of a *Homo sapiens* GPCR EST (GENBANK-ID: GENBANK-ID: AW182678|acc:AW182678; xj45d11.x1 Soares_NFL_T_GBC_S1 *Homo sapiens* cDNA clone IMAGE:2660181 3' similar

WO 01/64879

PCT/US01/06474

to TR:Q9Z1V0 Q9Z1V0 OLFACTORY RECEPTOR C6). This 94% match ($E=1.1e-69$) between regions of the public sequence and regions of the present invention (gene) suggests that the present invention (gene) could be a splice variant of the public GPCR EST (partial mRNA). This also supports identification of GPCR9 as a GPCR. In a search of sequence databases, partial matches (94 of 100 bases, 94% identical) of nucleotides 893-794 of GPCR9 with nucleotides 251-348 of a *Homo sapiens* GPCR EST (GENBANK-ID: AA206680|acc:AA206680:7q51c11.r1 Stratagene neuroepithelium (#937231) *Homo sapiens* cDNA clone IMAGE:645140 5' similar to contains L1.b2 L1 repetitive element). This 94% match between nucleotides of the public sequence and nucleotides of the GPCR9 sequence suggests that GPCR9 may be a splice variant of the public GPCR EST (partial mRNA).

Table 9A. GPCR9 Nucleotide Sequence (SEQ ID NO:32)

```

GTTTCATGCAAAATAGCAATATTGTCACCTGCTTTATTTCTCTGGGACTTTCTCAAAACAGCAACATGAA
GTTTTTTGGTTTGTATTATTATTATTTGCTACATGCTACTTGGATGGAAGACTTCTATUAC'AAAGATTI
CTATCATGTATCATTTGGCTATTGACCAACCCATGATTTCTTCTTAATTACCTCGCACTCTCAGATCT
TCTTACATATCCACTGTGGGCCCCCAAGCTAATGATTGACCTACTAACAAGAGGAGATGTTTCTAT
AATRACTGATGATACAGCTATTATGACCTACCTCTGAGAGCATGAGAGATCTTCACTCAAAAGCA
TGGCTATGAGCTATGATGAGGATCTGCAAGCACTGGGCTACACCTCATCAGACCAAGCAAGCTG
CAACGCCATCATCAGAGCTCTTGTACTGGGGGATTTATACACTCTGCCATGAGTTCTCTTTACCATG
CTCTTACCTTCTGCTTAAATGACATAGATCAGTACTTCTGCTATGCTATCTCTCTGCTGAGTGG
CTGGCATTTGATATACAGAAATGGTTCCTGCTATTGTTAATTCAGGCTGATTTCTTGTGCTGCTT
CGTATTTGATGCTGCTTATATTATATATATATGACCAATCAGGCTTACTCTGCTGAGATCAGAC
AAGGCTCTTTGAGCTGAGCTCTGACATATATGATCTGCTGCTATGCTTGTGCTGCTGCTGCTGCT
AGCTCAGACCACTATATTTTTCAGAGATATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
GTTGAGCTCTCTTATCCATAGCTGAGAGCTGAGATGAGAGCTGCTGAGGAAATCTGCTGCTCAT
CAATTCCTCTGCTGAGAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

```

The GPCR9 protein encoded by SEQ ID NO:32 has 309 amino acid residues, and is presented using the one-letter code in Table 9B (SEQ ID NO:33). The SignalP, Psort and/or Hydropathy profile for GPCR9 predict that GPCR9 has a signal peptide and is likely to be localized at the endoplasmic reticulum membrane with a certainty of 0.6850 or to the plasma membrane with a certainty of 0.6400. The SignalP predicts a cleavage site at the sequence IWM-EN between amino acids 38 and 39 as indicated by the slash in Table 9B.

Table 9B. Encoded GPCR9 protein sequence (SEQ ID NO:33)

```

MERNKIVTIFLLGLSQKRIEVPWFVLEVFYCAIWM/ENFIMHSINYIMLEJZERYEELNYLALSULC
YISTVAPKRIATDILERRIVSYNMCIGLETTHELGDIEFIIIPANAYDHYIATGRIHLYTIIITRQSCN
TITTAGCTGGFTHSAGSPALITIEFNCSLMLIDTCTVIFLLHARIDYRIGIYVNSGITSILAPV
ELMISVYLLISTINYSASKEKRLSTNSHIVWVLFVERLEFYTIRPAITTFEDKVVLFCALIAWF
SLIYMLRKVEKNAVRNMCQILLARK

```

The full amino acid sequence of the protein of the invention was found to have 140 of 300 amino acid residues (46%) identical to, and 193 of 300 residues (64%) positive with, the

WO 01/06479

PCT/US01/06474

302 amino acid residue odorant receptor A16 protein from *Mus musculus* (ptm:TREMBLNEW-ACC:BAA86127) (E value = $1.0e^{-73}$). Further BLAST analysis produced the significant results listed in Table 9C. The disclosed GPCR8 protein (SEQ ID NO:33) has good identity with a number of olfactory receptor proteins.

5

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (#)	Expect
GI11496249 ref NP_067343.1 (AB030895)	odorant receptor MOR16 <i>Mus musculus</i>	308	140/300 (46%)	195/300 (61%)	1e-59
GI11464995 ref NP_065261.1 (AB030896)	odorant receptor A16 <i>Mus musculus</i>	302	137/300 (45%)	185/300 (61%)	2e-59
GI1423702 ptm S29710	olfactory receptor OR18 - rat	307	142/303 (46%)	136/303 (45%)	6e-57
GI11464993 ref NP_065260.1 (AB030894)	odorant receptor MOR83 <i>Mus musculus</i>	308	133/297 (44%)	182/297 (60%)	7e-55
GI10644519 gb AA021324.1 (AF271051)	odorant receptor <i>Mus musculus</i>	264	124/262 (47%)	169/262 (64%)	6e-51

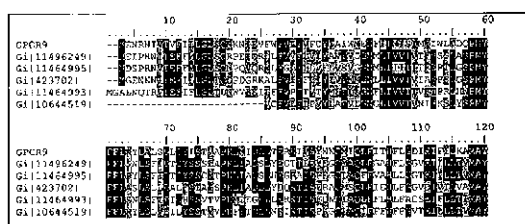
This information is presented graphically in the multiple sequence alignment given in Table 9D (with GPCR9 being shown on line 1) as a ClustalW analysis comparing GPCR9 with related protein sequences.

10

Table 9D. Information for the ClustalW proteins:

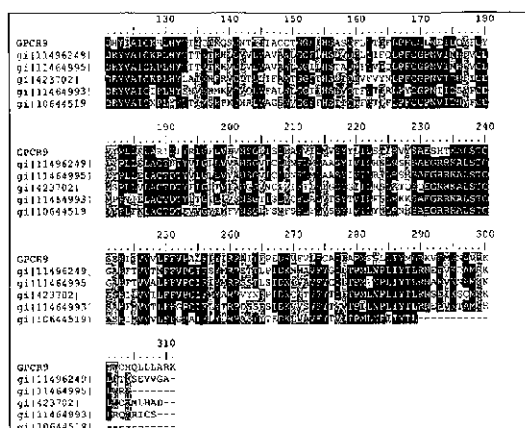
- 1) GPCR9 (SEQ ID NO:33)
- 2) gi11496249|ref|NP_067343.1| odorant receptor 16 *Mus musculus* (SEQ ID NO:78)
- 3) gi11464995|ref|NP_065261.1| gene for odorant receptor A16 *Mus musculus* (SEQ ID NO:79)
- 4) gi1423702|ptm|S29710 olfactory receptor OR18 - rat (SEQ ID NO:80)
- 5) gi11464993|ref|NP_065260.1| gene for odorant receptor MOR83 *Mus musculus* (SEQ ID NO:81)
- 6) gi10644519|gb|AA021324.1| odorant receptor *Mus musculus* (SEQ ID NO:70)

15



WO 01/64879

PCT/US01/06474



The presence of identifiable domains in GPCR9 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 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 9E with the statistics and domain description. The results indicate that GPCR9 contains the 7tm_1 (InterPro) 7 transmembrane receptor (rhodopsin family) domain (as defined by Interpro) at amino acid positions residues 56-234. This indicates that the sequence of GPCR9 has properties similar to those of other proteins known to contain this/these domain(s) and similar to the properties of these domains.

PCT/US01/06474

[illegible]

The nucleic acids and proteins of the invention are useful in potential therapeutic applications implicated in used in the treatment of infections such as 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 osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, allergies, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe

WO 01/64879

PCT/US01/06474

mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders.

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 bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to neoplasms; 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 osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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 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 GPCR9 substances 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-GPCR9 Antibodies" section below.

25 GPCR10

GPCR10 includes a family of three similar nucleic acids and three similar proteins disclosed below. The disclosed nucleic acids encode GPCR, OR-like proteins.

GPCR10a

The disclosed novel nucleic acid was identified on chromosome 11 by TblastN using CuraGen Corporation's sequence file for GPCR probe or homolog, run against the Genomic Daily Files made available by GenBank. The nucleic acid was further predicted by the program GenScan™, including selection of exons. These were further modified by means of similarities using BLAST searches. The sequences were then manually corrected for apparent

PCT/US01/06474

5

[illegible]

19

19

Table 10B. Encoded GPCR10a protein sequence (SEQ ID NO:35).

MENQNVTEFILLGHTERLEKWKISAVFLNFWYAVILE/SELIVVTITSGSIRSPMYFFELFSLLDVH
 ESYVARKVIVPDMKISLITSGELGKIVTEVFPGWGLILTPGADVAIVAKCKPLRYITLMSFWKVC
 INWGSAGWGHMAGKILQLFQYVTFEPCFNTIDFELCFQMLTACTATGALGLGLTGLTGLTGLTGLTGLT
 ITLITATVIVLGSLSKSGKGRHIALDTGSSHLTVVVFLEVFQFLVYRPPVTHIPIDSNMAYSDSIITPEL
 NPLVLTNLEKSGKSKKKNKGNKALGKSG

26

PCT/US01/06474

GFCR10b

15 The nucleotide sequence for GPCR10b (948 bp, SEQ ID NO:36) is presented in Table 10C. The nucleotide sequence differs from GPCR10a by one nucleotide change (numbered with respect to GPCR10a) T404 >C.

[illegible]

67

WO 01/64879

PCT/US01/06474

Table 10D. Encoded GPCR10b protein sequence (SEQ ID NO:37)

MRHQNVPFETLLGCTENLETWRTFSAVFLNHYVATULE/NLLTVVPFLLTSQSLASRMVFFLFLSLLD
 VHFSSVAPKIVVDTLSKSTTISLKGCLTQLEVEHFECGVGIILLTVHAGIRYVAICKPLHYTTTKSPR
 VCCLNVCCAWVCGEMHMIQLENNYQIPFCGZHIIDRFICDCEQLLTACTCTHLLGLLVTLNLSQEMCV
 AIFLLILIASVTVITCSLKSYSRGRHKALSTCSHLTVVVVDFVPCIFDVTSPVVTHEIDKAMAVSDSI
 ITPMLNPLIYTLNABEVKSAVRKEMKMEFLACK

Table 10E presents the BLASTP results for GPCR10b.

Table 10E. BLASTP Results for GPCR10b

Score = 999 (351.7 bits), Expect = 1.2e-100, P = 1.2e-100 Identities = 182/302 (60%), Positives = 231/302 (76%) with PIR-IDS29710 olfactory receptor OR18 - rat
 Score = 737 (266.5 bits), Expect = 5.2e-75, P = 5.2e-75 Identities = 144/298 (48%), Positives = 200/298 (67%) with ACC:O95013 WUGSCH_DJ0855D21.1 PROTEIN - Homo sapiens (Human), 312 aa.
 Score = 737 (266.5 bits), Expect = 5.2e-75, P = 5.2e-75 Identities = 144/298 (48%), Positives = 200/298 (67%) with ACC:O95013 WUGSCH_DJ0855D21.1 PROTEIN - Homo sapiens (Human), 312 aa.
 Score = 667 (234.8 bits), Expect = 1.1e-64, P = 1.1e-64 Identities = 131/300 (43%), Positives = 194/300 (64%), Frame = +1 with ACC:O43749 OLFACTORY RECEPTOR - Homo sapiens (Human), 312 aa.

5 **GPCR10c**

Another nucleotide sequence resulted when GPCR10a (6-L-19-A) was subjected to an 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. In each case, the sequence was examined, walking
 10 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 suitable sequences were then employed as the forward and reverse primers in a PCR amplification based on a wide range of cDNA libraries.

These primers were then employed in PCR amplification based on the following pool
 15 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,
 20 cloned and sequenced to high redundancy. 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.

WO 01/64879

PCT/US01/06474

In addition, sequence traces were evaluated manually and edited for corrections if appropriate. The resulting amplicon was gel purified, cloned and sequenced to high redundancy to provide the sequence reported below, which is designated as Accession Number 6-L-19-A-daf, or GPCR10c.

- 5 The nucleotide sequence for GPCR10c (943 bp, SEQ ID NO:38) is presented in Table 10F. The GPCR10c nucleotide sequence differs from GPCR10a by having six fewer nucleotides at the 5' end and two nucleotide changes: (numbered with respect to GPCR10a) , G466>A and C834>T.

Table 10F. GPCR10c Nucleotide Sequence (SEQ ID NO:38)

```

ATGGAAATGCAAAACATGTGACTGAATTCATTCTTGGCTCAGACAGACCTGAGCTGTGGAAA
TATTTCTGCTGTGTTTCTTGTGATGTATCTAGCCACAGTGGTGAAGAAATGACTTATTCGGTATGAT
TATCAGATCAGAGCTCTGGGTCACTATGATTTTCTCTGCTCTGGTCCCTTTGGAGTCTATG
TTCTCATCTGTCTTGGCCCAAGGTGATTTAGACACCTCTCCAGAGCACTTCCCTCTCTCAAG
GCTGGCTCACCCAGCTGTTTGTGGAGCATTTCTTTGGTGTGGGGATCATCTCCCTCACTGTGATGGC
CTATGACCTGACCTGGCCATCTGTAGCCCTTGCACTACAGCATCATCATGAGTCAAGGATGTGCTGC
CTAATGGTAGGAGGGGCTTGGTGGGGGGATTTATCCACACATCATACAACTCTCTCTCATGTATCAAA
TAGCCCTTCTGCTGCTAATATCATAGATCATCTTATATGTGATTTGTTTCACTGTGTGACATCTGCTG
CAGGACACCATCTCTGCGCTCTTATGTTACCTTACAGGTGGGATGATGTGTTGGCACTTCTCTT
ATCTTAAATGGGCTTACAGGTCTATGCTCTCTGAGATCTTACAGCTCTAAGGCGGCGACAAAG
CCTCTCTACCTGACGCTCCGAGCTGACGCTGCTTGTATTTCTGCTTCTGCTGCTGCTATCTCTGTAGAT
GAGGCTGTGCTCACTCACCCATAGACAGGCAATGGCTGTCTCAGACTCAATCATACCCATGTTA
AATCCCTTGATCTATACACTGAGGATGCGAGCTGAAAAGTCCATGAGAACTCTGGATGAATGGG
ASGCTTTGGCTGGAAATACCTGCAATGCTGA

```

10

The coding region of GPCR10c is from nucleotide 1 to 928, giving the encoded GPCR10c protein, as presented in Table 10G. The disclosed protein is 309 amino acids long and is denoted by SEQ ID NO: 83. GPCR10c differs from GPCR10a by one amino acid residue: A154>T. Like GPCR10a, the Psort profile for GPCR10c predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6000. The most likely cleavage site for a peptide is between amino acids 39 and 40, i.e., at the slash in the amino acid sequence VLE-NL (shown as a slash in Table10G) based on the SignalP result.

15

Table 10G. Encoded GPCR10c protein sequence (SEQ ID NO:83)

```

MENQRYTEFILLGLTENLELXIFSAVELVRYVATVLE/NLLIVVTIITSQIRGMYFFLTPLSLLOVM
FSSVVAQKVIYDILEKSTTSGKGLTCLFVHFEGGVGIIETVMAYRYVATCKPLRYTILMSFRVCG
LWVGGAAYGSPMTWIGLFTQIPFDEPILDRPHGULGLLACTDRIIGCLVILNSGQMCYATPL
LLASVTIICSLKSYSGRFAISVSSHCTVVLFFVPCILFTMRVFTPLIRARAVSUSITRM
NPILYTCNAEYKSRMKKIKMKWERLACK

```

20

Possible SNPs found for GPCR10 are listed in Table10H.

WO 01/64879

PCT/US01/06474

Table 10H: SNPs

Base Position	Base Before	Base After
65	T	A(2)
120	T	Gap(2)
147	T	C(2)
234	A	G(3)
412	T	C(7)
471	G	A(2)
814	A	G(3)

The disclosed GPCR10 protein (SEQ ID NO:35) has good identity with a number of olfactory receptor proteins. The identity information used for ClustalW analysis is presented in Table 10I. Unless specifically addressed as GPCR10a, GPCR10b, or GPCR10c, any reference to GPCR10 is assumed to encompass all variants. Residue differences between a GPCRX variant sequences herein are written to show the residue in the "a" variant and the residue position with respect to the "a" variant. GPCR residues in all following sequence alignments that differ between the individual GPCR variants are highlighted with a box and marked with the (o) symbol above the variant residue in all alignments herein. For example, the protein shown in line 1 of Table 10J depicts the sequence for GPCR10a, and the positions where GPCR10b or GPCR10c differs are marked with a (o) symbol and are highlighted with a box. All GPCR10 proteins have significant homology to olfactory receptor (OR) proteins:

Table 10I. BLAST results for GPCR10

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
G:11496249:ref NP_063443.1 (AR330895)	Odorant receptor MOR19 <i>Mus musculus</i>	308	164/308 (60%)	241/308 (78%)	6e-91
G:14237021:ref NP_052971.1	OR OR18 - rat	307	163/302 (60%)	232/302 (77%)	2e-88
G:11466995:ref NP_052611.1 AB030896	Odorant receptor A16 <i>Mus musculus</i>	302	175/302 (57%)	232/302 (75%)	8e-86
G:11466993:ref NP_052606.1 AB030894	Odorant receptor MOR83 <i>Mus musculus</i>	308	157/297 (52%)	208/297 (69%)	3e-72
G:10644517:ref AA021323.1 AF271050.1 (MT21050)	Odorant receptor <i>Rattus norvegicus</i>	264	135/260 (51%)	202/260 (77%)	7e-71

5

10

- [illegible]

WO 01/64879

PCT/US01/06474

The presence of identifiable domains in GPCR10 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/>).

- 5 DOMAIN results for GPCR10 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 10K with the statistics and domain description. The results indicate that this protein contains the following protein domains (as defined by Interpro) at the indicated positions: domain name 7tm_1 (InterPro) 7
- 10 transmembrane receptor (rhodopsin family) at amino acid positions 39 to 213. This indicates that the sequence of GPCR10 has properties similar to those of other proteins known to contain this domain and similar to the properties of this domain.

Table 10K. DOMAIN results for GPCR10

gn11Pfam:pfam0001, 7tm_1, 7 transmembrane receptor (rhodopsin family) Length = 370 (SEQ ID NO:45). Score = 93.5 bits (231), Expect = 1e-20
GPCR10 7M7
GPCR10 7M7
GPCR10 7M7
GPCR10 7M7

- 15 The similarity information for the GPCR10 protein and nucleic acid disclosed herein suggest that GPCR10 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
- 20 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

WO 01/64879

PCT/US01/06474

regeneration in vitro and in vivo (vi) biological defense weapon. The novel nucleic acid encoding GPCR10, and the GPCR10 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.

5 The nucleic acids and proteins of the invention are useful in potential therapeutic applications implicated in used in the treatment of infections such as 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, 10 hypertension, urinary retention, osteoporosis, Crohn's disease; multiple sclerosis; and treatment of Albright hereditary osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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 15 syndrome and/or other pathologies and disorders. 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 20 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; 25 multiple sclerosis; and treatment of Albright hereditary osteodystrophy, angina pectoris, myocardial infarction, ulcers, asthma, 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. These materials are further 30 useful in the generation of antibodies that bind immuno-specifically to the novel GPCR10 substances 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-GPCR Antibodies" section below.

WO 01/64879

PCT/US01/06474

A summary of the GPCR_X nucleic acids and proteins of the invention is provided in Table 11.

TABLE 11: Summary Of Nucleic Acids And Proteins Of The Invention

Name	Tables	Clone; Description of Homolog	Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO
GPCR1	1A, 1B, 1D, 1E, 1G, 1H	GPCR1a: ba113a10_B, olfactory receptor	1	2
		GPCR1b: ba32713_A, olfactory receptor	3	4
		GPCR1c: ba113a10_C, olfactory receptor	5	6
GPCR2	2A, 2B	GPCR2: 11612531_L, olfactory receptor	7	8
GPCR3	3A, 3B	GPCR3: ba145122_B, olfactory receptor	9	10
GPCR4	4A, 4B, 4C, 4G, 4H	GPCR4a1: dj408620_C, olfactory receptor	11	12
		GPCR4a2: dj408620_C_da1, olfactory receptor	13	17
		GPCR4a3: CG55358_03, olfactory receptor	16	
GPCR5	5A, 5B, 5C, 5D, 5G, 5H	GPCR5a1: 115-a-12-A, olfactory receptor	18	19
		GPCR5a2: 115-a-12-B, olfactory receptor	20	21
		GPCR5a3: 115-a-12-A_da1, olfactory receptor	22	23
GPCR6	6A, 6B	GPCR6: 6-L-19-C, olfactory receptor	24	25
GPCR7	7A, 7B	GPCR7: dj31316_D, olfactory receptor	28	29
GPCR8	8A, 8B	GPCR8: dj408620_A, olfactory receptor	30	31
GPCR9	9A, 9B	GPCR9: 6-L-19-B, olfactory receptor	32	33
GPCR10	10A, 10B, 10C, 10D, 10F, 10G	GPCR10a: 6-L-19-A, olfactory receptor	34	35
		GPCR10b: 6-L-19-A1, olfactory receptor	36	37
		GPCR10c: 6-L-19-A_da1, olfactory receptor	38	38

5 GPCR_X Nucleic Acids and Polypeptides

One aspect of the invention pertains to isolated nucleic acid molecules that encode GPCR_X polypeptides or biologically active portions thereof. Also included in the invention are nucleic acid fragments sufficient for use as hybridization probes to identify GPCR_X-encoding nucleic acids (e.g., GPCR_X mRNAs) and fragments for use as PCR primers for the amplification and/or mutation of GPCR_X 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 GPCR_X nucleic acid can encode a mature GPCR_X polypeptide. As used herein, a "mature" form of a polypeptide or protein disclosed in the present invention is the product of a

WO 01/64879

PCT/US01/06474

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

5 "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

10 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

15 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.

20 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

25 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.

30 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 GPCR α 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

WO 01/64879

PCT/US01/06474

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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 as a hybridization probe, GPCR_X 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 GPCR_X 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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

WO 01/64879

PCT/US01/06474

NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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 GPCR α polypeptide). A nucleic acid molecule that is complementary to the nucleotide sequence shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, is one that is sufficiently complementary to the nucleotide sequence shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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

WO 01/64879

PCT/US01/06474

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 GPCR_X 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 GPCR_X 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 GPCR_X 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, as well as a polypeptide possessing GPCR_X biological activity. Various biological activities of the GPCR_X proteins are described below.

An GPCR_X polypeptide is encoded by the open reading frame ("ORF") of an GPCR_X 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 GPCR_X genes allows for the generation of probes and primers designed for use in identifying and/or cloning

WO 01/64879

PCT/US01/06474

GPCRX homologues in other cell types, e.g. from other tissues, as well as GPCR_X 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38; or an anti-sense strand nucleotide sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38; or of a naturally occurring mutant of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38.

Probes based on the human GPCR_X 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 GPCR_X protein, such as by measuring a level of an GPCR_X-encoding nucleic acid in a sample of cells from a subject e.g., detecting GPCR_X mRNA levels or determining whether a genomic GPCR_X gene has been mutated or deleted.

"A polypeptide having a biologically-active portion of an GPCR_X 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 GPCR_X" can be prepared by isolating a portion SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 that encodes a polypeptide having an GPCR_X biological activity (the biological activities of the GPCR_X proteins are described below), expressing the encoded portion of GPCR_X protein (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of GPCR_X.

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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 due to degeneracy of the genetic code and thus encode the same GPCR_X proteins as that encoded by the nucleotide sequences shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino

WO 01/64879

PCT/US01/06474

acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83.

In addition to the human GPCR_X nucleotide sequences shown in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 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 GPCR_X polypeptides may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the GPCR_X 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 GPCR_X protein, preferably a vertebrate GPCR_X protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the GPCR_X genes. Any and all such nucleotide variations and resulting amino acid polymorphisms in the GPCR_X polypeptides, which are the result of natural allelic variation and that do not alter the functional activity of the GPCR_X polypeptides, are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding GPCR_X 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the GPCR_X cDNAs of the invention can be isolated based on their homology to the human GPCR_X 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38. 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 GPCR_X proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or

WO 01/64879

PCT/US01/06474

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 (T_m) for the specific sequence at a defined ionic strength and pH. The T_m 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 T_m , 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 or fragments, analogs or derivatives thereof, under

WO 01/64879

PCT/US01/06474

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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 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 GPCR_X 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 thereby leading to changes in the amino acid sequences of the encoded GPCR_X proteins, without altering the functional ability of said GPCR_X 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequences of the GPCR_X 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 GPCR_X proteins of the invention

WO 01/64879

PCT/US01/06474

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 GPCR_X proteins that contain changes in amino acid residues that are not essential for activity. Such GPCR_X proteins differ in amino acid sequence from SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83. 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83; more preferably at least about 70% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83; still more preferably at least about 80% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83; even more preferably at least about 90% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83; and most preferably at least about 95% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83.

An isolated nucleic acid molecule encoding an GPCR_X protein homologous to the protein of SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83 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

WO 01/64879

PCT/US01/06474

aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted non-essential amino acid residue in the GPCR_X 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 GPCR_X coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for GPCR_X biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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, QIRK, 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, NEQHKK, VLIM, HPFY, wherein the letters within each group represent the single letter amino acid code.

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

In yet another embodiment, a mutant GPCR_X 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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 GPCR_X coding strand, or to only a portion thereof. Nucleic acid molecules

WO 01/64879

PCT/US01/06474

encoding fragments, homologs, derivatives and analogs of an GPCR α protein of SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, or antisense nucleic acids complementary to an GPCR α nucleic acid sequence of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, are additionally provided.

5 In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an GPCR α 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
10 encoding the GPCR α 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 GPCR α protein disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and
15 Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of GPCR α mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of GPCR α mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of GPCR α mRNA. An antisense oligonucleotide can be, for example,
20 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
25 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-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-
30 2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylguanosine, 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-methoxymethyl-2-thiouracil, beta-D-mannosylguanosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil,

WO 01/64879

PCT/US01/06474

2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methyl ester, 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 GPCR 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

WO 01/64879

PCT/US01/06474

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 GPCR_X mRNA transcripts to thereby inhibit translation of GPCR_X mRNA. A ribozyme having specificity for an GPCR_X-encoding nucleic acid can be designed based upon the nucleotide sequence of an GPCR_X cDNA disclosed herein (i.e., SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38). 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 GPCR_X-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.* GPCR_X 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, GPCR_X gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the GPCR_X nucleic acid (e.g., the GPCR_X promoter and/or enhancers) to form triple helical structures that prevent transcription of the GPCR_X 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 GPCR_X 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

WO 01/64879

PCT/US01/06474

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 GPCR $\times$ 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 GPCR $\times$ 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.*, S₁ 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 GPCR $\times$ 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 GPCR $\times$ 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

WO 01/64879

PCT/US01/06474

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 GPCR_X polypeptides whose sequences are provided in SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83. 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83 while still encoding a protein that maintains its GPCR_X activities and physiological functions, or a functional fragment thereof.

In general, an GPCR_X variant that preserves GPCR_X-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 GPCR_X 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-GPCR_X antibodies. In one embodiment, native GPCR_X proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, GPCR_X proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, an GPCR_X 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 GPCR_X 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 GPCR_X proteins in which the protein is

WO 01/64879

PCT/US01/06474

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 GPCR_X proteins having less than about 30% (by dry weight) of non-GPCR_X proteins (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-GPCR_X proteins, still more preferably less than about 10% of non-GPCR_X proteins, and most preferably less than about 5% of non-GPCR_X proteins. When the GPCR_X 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 GPCR_X protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of GPCR_X 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 GPCR_X proteins having less than about 30% (by dry weight) of chemical precursors or non-GPCR_X chemicals, more preferably less than about 20% chemical precursors or non-GPCR_X chemicals, still more preferably less than about 10% chemical precursors or non-GPCR_X chemicals, and most preferably less than about 5% chemical precursors or non-GPCR_X chemicals.

Biologically-active portions of GPCR_X proteins include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequences of the GPCR_X proteins (e.g., the amino acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83) that include fewer amino acids than the full-length GPCR_X proteins, and exhibit at least one activity of an GPCR_X protein. Typically, biologically-active portions comprise a domain or motif with at least one activity of the GPCR_X protein. A biologically-active portion of an GPCR_X 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 GPCR_X protein.

In an embodiment, the GPCR_X protein has an amino acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83. In other embodiments, the GPCR_X protein is substantially homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, and retains the functional activity of the protein of SEQ ID

WO 01/64879

PCT/US01/06474

NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail, below.

Accordingly, in another embodiment, the GPCR_X 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, and retains the functional activity of the GPCR_X proteins of SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83.

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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38.

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.,

WO 01/64879

PCT/US01/06474

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 GPCR_X chimeric or fusion proteins. As used herein, an GPCR_X "chimeric protein" or "fusion protein" comprises an GPCR_X polypeptide operatively-linked to a non-GPCR_X polypeptide. An "GPCR_X polypeptide" refers to a polypeptide having an amino acid sequence corresponding to an GPCR_X protein (SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83), whereas a "non-GPCR_X polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the GPCR_X protein, e.g., a protein that is different from the GPCR_X protein and that is derived from the same or a different organism. Within an GPCR_X fusion protein the GPCR_X polypeptide can correspond to all or a portion of an GPCR_X protein. In one embodiment, an GPCR_X fusion protein comprises at least one biologically-active portion of an GPCR_X protein. In another embodiment, an GPCR_X fusion protein comprises at least two biologically-active portions of an GPCR_X protein. In yet another embodiment, an GPCR_X fusion protein comprises at least three biologically-active portions of an GPCR_X protein. Within the fusion protein, the term "operatively-linked" is intended to indicate that the GPCR_X polypeptide and the non-GPCR_X polypeptide are fused in-frame with one another. The non-GPCR_X polypeptide can be fused to the N-terminus or C-terminus of the GPCR_X polypeptide.

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

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

WO 01/64879

PCT/US01/06474

In yet another embodiment, the fusion protein is an GPCR_X-immunoglobulin fusion protein in which the GPCR_X sequences are fused to sequences derived from a member of the immunoglobulin protein family. The GPCR_X-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between an GPCR_X ligand and an GPCR_X protein on the surface of a cell, to thereby suppress GPCR_X-mediated signal transduction *in vivo*. The GPCR_X-immunoglobulin fusion proteins can be used to affect the bioavailability of an GPCR_X cognate ligand. Inhibition of the GPCR_X ligand/GPCR_X 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 GPCR_X-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-GPCR_X antibodies in a subject, to purify GPCR_X ligands, and in screening assays to identify molecules that inhibit the interaction of GPCR_X with an GPCR_X ligand.

An GPCR_X 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 GPCR_X-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the GPCR_X protein.

GPCR_X Agonists and Antagonists

The invention also pertains to variants of the GPCR_X proteins that function as either GPCR_X agonists (*i.e.*, mimetics) or as GPCR_X antagonists. Variants of the GPCR_X protein can be generated by mutagenesis (e.g., discrete point mutation or truncation of the GPCR_X protein). An agonist of the GPCR_X protein can retain substantially the same, or a subset of,

WO 01/64879

PCT/US01/06474

the biological activities of the naturally occurring form of the GPCR α protein. An antagonist of the GPCR α protein can inhibit one or more of the activities of the naturally occurring form of the GPCR α protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the GPCR α 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 GPCR α proteins.

10 Variants of the GPCR α proteins that function as either GPCR α agonists (*i.e.*, mimetics) or as GPCR α antagonists can be identified by screening combinatorial libraries of mutants (*e.g.*, truncation mutants) of the GPCR α proteins for GPCR α protein agonist or antagonist activity. In one embodiment, a variegated library of GPCR α variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of GPCR α variants can be produced by, for example,

15 enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential GPCR α sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of GPCR α sequences therein. There are a variety of methods which can be used to produce libraries of potential GPCR α 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 GPCR α sequences. Methods for synthesizing degenerate oligonucleotides are well-known within the art. See, *e.g.*, Narang, 1983, *Tetrahedron* 39: 3;

25 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 GPCR α protein coding sequences can be used to generate a variegated population of GPCR α fragments for screening and subsequent selection of variants of an GPCR α protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an GPCR α 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

WO 01/64879

PCT/US01/06474

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 GPCR χ 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 GPCR χ 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 GPCR χ 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-GPCR χ Antibodies

The invention encompasses antibodies and antibody fragments, such as F_{ab} or $(F_{ab})_2$, that bind immunospecifically to any of the GPCR χ polypeptides of said invention.

An isolated GPCR χ protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind to GPCR χ polypeptides using standard techniques for polyclonal and monoclonal antibody preparation. The full-length GPCR χ proteins can be used or, alternatively, the invention provides antigenic peptide fragments of GPCR χ proteins for use as immunogens. The antigenic GPCR χ peptides comprises at least 4 amino acid residues of the amino acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83 and encompasses an epitope of GPCR χ such that an antibody raised against the peptide forms a specific immune complex with GPCR χ .

Preferably, the antigenic peptide comprises at least 6, 8, 10, 15, 20, or 30 amino acid residues. Longer antigenic peptides are sometimes preferable over shorter antigenic peptides, depending on use and according to methods well known to someone skilled in the art.

WO 01/64879

PCT/US01/06474

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of GPCR_X that is located on the surface of the protein (e.g., a hydrophilic region). 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 incorporated herein by reference in their entirety).

As disclosed herein, GPCR_X protein sequences of SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, or derivatives, fragments, analogs or homologs thereof, may be utilized as immunogens in the generation of antibodies that immunospecifically bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically-active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen, such as GPCR_X. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} and F_{(ab')2} fragments, and an F_{ab} expression library. In a specific embodiment, antibodies to human GPCR_X proteins are disclosed. Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies to an GPCR_X protein sequence of SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, or a derivative, fragment, analog or homolog thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenetic preparation can contain, for example, recombinantly-expressed GPCR_X protein or a chemically-synthesized GPCR_X polypeptide. 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., lysollecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum*, or similar immunostimulatory agents. If desired, the antibody molecules directed against GPCR_X can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

WO 01/64879

PCT/US01/06474

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of GPCR_X. A monoclonal antibody composition thus typically displays a single binding affinity for a particular GPCR_X protein with which it immunoreacts. For preparation of monoclonal antibodies directed
 5 towards a particular GPCR_X protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see, e.g., Kohler & Milstein, 1975, *Nature* 256: 495-497); the trioma technique; the
 10 human B-cell hybridoma technique (see, e.g., Kozbor, et al., 1983, *Immunol. Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see, e.g., 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 invention and may be produced by using human hybridomas (see, e.g., Cole, et al., 1983, *Proc Natl Acad Sci USA*
 15 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see, e.g., Cole, et al., 1985, In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Each of the above citations is incorporated herein by reference in their entirety.

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an GPCR_X protein (see, e.g., U.S. Patent No. 4,946,778).
 20 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 an GPCR_X protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See, e.g., U.S. Patent No. 5,225,539. Antibody fragments
 25 that contain the idiotype to an GPCR_X protein may be produced by techniques known in the art including, but not limited to: (i) an F_{ab} fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an F_{ab} 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.

30 Additionally, recombinant anti-GPCR_X antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in International Application

WO 01/64879

PCT/US01/06474

No. PCT/US86/02269; European Patent Application No. 184,187; European Patent Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Patent No. 4,816,567; U.S. Pat. No. 5,225,539; European Patent Application No. 125,023; Better, *et al.*, 1988. *Science* 240: 1041-1043; Liu, *et al.*, 1987. *Proc. Natl. Acad. Sci. USA* 84: 3439-3443; Liu, *et al.*, 1987. *J. Immunol.* 139: 3521-3526; Sun, *et al.*, 1987. *Proc. Natl. Acad. Sci. USA* 84: 214-218; Nishimura, *et al.*, 1987. *Cancer Res.* 47: 999-1005; Wood, *et al.*, 1985. *Nature* 314: 446-449; Shaw, *et al.*, 1988. *J. Natl. Cancer Inst.* 80: 1553-1559; Morrison (1985) *Science* 229:1202-1207; Oi, *et al.* (1986) *BioTechniques* 4:214; Jones, *et al.*, 1986. *Nature* 321: 552-525; Verhoecyan, *et al.*, 1988. *Science* 239: 1534; and Beidler, *et al.*, 1988. *J. Immunol.* 141: 4053-4060. Each of the above citations are incorporated herein by reference in their entirety.

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 GPCR_X protein is facilitated by generation of hybridomas that bind to the fragment of an GPCR_X protein possessing such a domain. Thus, antibodies that are specific for a desired domain within an GPCR_X protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

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

WO 01/64879

PCT/US01/06474

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 ^3H .

GPCRX Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an GPCR α 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

WO 01/64879

PCT/US01/06474

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 include 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., GPCR proteins, mutant forms of GPCR proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of GPCR proteins in prokaryotic or eukaryotic cells. For example, GPCR 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

WO 01/64879

PCT/US01/06474

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 GPCR_X expression vector is a yeast expression vector.

Examples of vectors for expression in yeast *Saccharomyces cerevisiae* 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, GPCR_X 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

WO 01/64879

PCT/US01/06474

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 (Carnes 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 GPCR α 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

WO 01/64879

PCT/US01/06474

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.

5 A host cell can be any prokaryotic or eukaryotic cell. For example, GPCR_X 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.

10 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
15 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
20 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
25 GPCR_X 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) GPCR_X protein. Accordingly, the invention further provides
30 methods for producing GPCR_X 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 GPCR_X protein has been introduced) in a suitable medium such that GPCR_X protein is produced. In another embodiment, the method further comprises isolating GPCR_X protein from the medium or the host cell.

WO 01/64879

PCT/US01/06474

Transgenic GPCR_X 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 GPCR_X protein-coding sequences have been introduced.

5 Such host cells can then be used to create non-human transgenic animals in which exogenous GPCR_X sequences have been introduced into their genome or homologous recombinant animals in which endogenous GPCR_X sequences have been altered. Such animals are useful for studying the function and/or activity of GPCR_X protein and for identifying and/or evaluating modulators of GPCR_X protein activity. As used herein, a "transgenic animal" is a
10 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
15 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 GPCR_X 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
20 of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing GPCR_X-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 GPCR_X cDNA sequences of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28,
25 30, 32, 34, 36, and 38 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a non-human homologue of the human GPCR_X gene, such as a mouse GPCR_X gene, can be isolated based on hybridization to the human GPCR_X 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
30 tissue-specific regulatory sequence(s) can be operably-linked to the GPCR_X transgene to direct expression of GPCR_X 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

WO 01/64879

PCT/US01/06474

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 GPCR_X transgene in its genome and/or expression of GPCR_X 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 GPCR_X 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 GPCR_X gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the GPCR_X gene. The GPCR_X gene can be a human gene (*e.g.*, the cDNA of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38), but more preferably, is a non-human homologue of a human GPCR_X gene. For example, a mouse homologue of human GPCR_X gene of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 can be used to construct a homologous recombination vector suitable for altering an endogenous GPCR_X gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous GPCR_X 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 GPCR_X 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 GPCR_X protein). In the homologous recombination vector, the altered portion of the GPCR_X gene is flanked at its 5'- and 3'-termini by additional nucleic acid of the GPCR_X gene to allow for homologous recombination to occur between the exogenous GPCR_X gene carried by the vector and an endogenous GPCR_X gene in an embryonic stem cell. The additional flanking GPCR_X 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 then introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced GPCR_X gene has homologously-recombined with the endogenous GPCR_X 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

WO 01/64879

PCT/US01/06474

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 G₀ 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 blastocyst 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 GPCR_X nucleic acid molecules, GPCR_X proteins, and anti-GPCR_X 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,

WO 01/64879

PCT/US01/06474

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, buffer 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,

WO 01/64879

PCT/US01/06474

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

5 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 mannitol, 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.

10 Sterile injectable solutions can be prepared by incorporating the active compound (e.g., an GPCR_X protein or anti-GPCR_X 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.

20 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.

WO 01/64879

PCT/US01/06474

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
5 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
10 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
15 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
20 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
25 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
30 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.

WO 01/64879

PCT/US01/06474

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).

5 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.

10 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 GPCR_X protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect GPCR_X mRNA (e.g., in a biological sample) or a genetic lesion in an GPCR_X gene, 15 and to modulate GPCR_X activity, as described further, below. In addition, the GPCR_X proteins can be used to screen drugs or compounds that modulate the GPCR_X protein activity or expression as well as to treat disorders characterized by insufficient or excessive production of GPCR_X protein or production of GPCR_X protein forms that have decreased or aberrant activity compared to GPCR_X wild-type protein (e.g., diabetes (regulates insulin release); 20 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-GPCR_X antibodies of the invention can be used to 25 detect and isolate GPCR_X proteins and modulate GPCR_X 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*.

30

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,

WO 01/64879

PCT/US01/06474

peptidomimetics, small molecules or other drugs) that bind to GPCR proteins or have a stimulatory or inhibitory effect on, e.g., GPCR protein expression or GPCR protein activity. The invention also includes compounds identified in the screening assays described herein.

5 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 GPCR 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
10 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.

15 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
20 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; Carrell, *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).

WO 01/64879

PCT/US01/06474

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of GPCR_X 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 GPCR_X protein determined. The cell, for example, can of mammalian origin or a yeast cell.

- 5 Determining the ability of the test compound to bind to the GPCR_X 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 GPCR_X 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 ¹²⁵I, ³⁵S, ¹⁴C, or ³H, 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 GPCR_X protein, or a biologically-active portion thereof, on the cell surface with a known compound which binds GPCR_X 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 GPCR_X protein, wherein determining the ability of the test compound to interact with an GPCR_X protein comprises determining the ability of the test compound to preferentially bind to GPCR_X protein or a biologically-active portion thereof as compared to the known compound.

- 20 In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of GPCR_X 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 GPCR_X protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of GPCR_X or a biologically-active portion thereof can be accomplished, for example, by determining the ability of the GPCR_X protein to bind to or interact with an GPCR_X target molecule. As used herein, a "target molecule" is a molecule with which an GPCR_X protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses an GPCR_X 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 GPCR_X target molecule can be a non-GPCR_X molecule or an GPCR_X protein or polypeptide of the invention. In one embodiment, an GPCR_X target molecule is a component of a signal transduction pathway that facilitates

WO 01/64879

PCT/US01/06474

transduction of an extracellular signal (e.g. a signal generated by binding of a compound to a membrane-bound GPCR $\times$ molecule) through the cell membrane and into the cell. The target, for example, can be a second intracellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with GPCR $\times$.

5 Determining the ability of the GPCR $\times$ protein to bind to or interact with an GPCR $\times$ target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the GPCR $\times$ protein to bind to or interact with an GPCR $\times$ 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
 10 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 on an appropriate substrate, detecting the induction of a reporter gene (comprising an GPCR $\times$ -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,
 15 or cell proliferation.

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

In still another embodiment, an assay is a cell-free assay comprising contacting GPCR $\times$ 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
 30 GPCR $\times$ protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of GPCR $\times$ can be accomplished, for example, by determining the ability of the GPCR $\times$ protein to bind to an GPCR $\times$ 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 GPCR $\times$ protein can

WO 01/64879

PCT/US01/06474

be accomplished by determining the ability of the GPCR α protein further modulate an GPCR α 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 GPCR α protein or biologically-active portion thereof with a known compound which binds GPCR α 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 GPCR α protein, wherein determining the ability of the test compound to interact with an GPCR α protein comprises determining the ability of the GPCR α protein to preferentially bind to or modulate the activity of an GPCR α target molecule.

The cell-free assays of the invention are amenable to use of both the soluble form or the membrane-bound form of GPCR α protein. In the case of cell-free assays comprising the membrane-bound form of GPCR α protein, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of GPCR α 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[®], isotridecylpoly(ethylene glycol ether)_m, N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl) dimethylammonio-1-propane sulfonate (CHAPS), or 3-(3-cholanidopropyl)dimethylammonio-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 GPCR α 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 GPCR α protein, or interaction of GPCR α 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-GPCR α 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 GPCR α 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

WO 01/64879

PCT/US01/06474

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 GPCR_X protein binding or activity determined using standard techniques.

5 Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the GPCR_X protein or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated GPCR_X protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well-known within the art (*e.g.*, biotinylation kit, 10 Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with GPCR_X protein or target molecules, but which do not interfere with binding of the GPCR_X protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or GPCR_X protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in 15 addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the GPCR_X protein or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the GPCR_X protein or target molecule.

In another embodiment, modulators of GPCR_X protein expression are identified in a 20 method wherein a cell is contacted with a candidate compound and the expression of GPCR_X mRNA or protein in the cell is determined. The level of expression of GPCR_X mRNA or protein in the presence of the candidate compound is compared to the level of expression of GPCR_X mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of GPCR_X mRNA or protein expression 25 based upon this comparison. For example, when expression of GPCR_X 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 GPCR_X mRNA or protein expression. Alternatively, when expression of GPCR_X mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, 30 the candidate compound is identified as an inhibitor of GPCR_X mRNA or protein expression. The level of GPCR_X mRNA or protein expression in the cells can be determined by methods described herein for detecting GPCR_X mRNA or protein.

In yet another aspect of the invention, the GPCR_X 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;

WO 01/64879

PCT/US01/06474

Zervos, *et al.*, 1993. *Cell* 72: 223-232; Madura, *et al.*, 1993. *J. Biol. Chem.* 268: 12046-12054; Barte, *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 GPCR_X ("GPCR_X-binding proteins" or "GPCR_X-bp") and modulate GPCR_X activity. Such GPCR_X-binding proteins are also likely to be involved in the propagation of signals by the GPCR_X proteins as, for example, upstream or downstream elements of the GPCR_X 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 GPCR_X 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 GPCR_X-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 GPCR_X.

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

WO 01/64879

PCT/US01/06474

chromosome mapping. Accordingly, portions or fragments of the GPCR_X sequences, SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, or fragments or derivatives thereof, can be used to map the location of the GPCR_X genes, respectively, on a chromosome. The mapping of the GPCR_X sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, GPCR_X genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the GPCR_X sequences. Computer analysis of the GPCR_X 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 GPCR_X 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 GPCR_X 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.

WO 01/64879

PCT/US01/06474

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 GPCR_X 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 GPCR_X 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

WO 01/64879

PCT/US01/06474

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 GPCR_X 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 GPCR_X 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38 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 GPCR_X protein and/or nucleic acid expression as well as GPCR_X 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 GPCR_X

WO 01/64879

PCT/US01/06474

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 GPCR_X protein, nucleic acid expression or activity. For example, mutations in an GPCR_X 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 GPCR_X protein, nucleic acid expression, or biological activity.

Another aspect of the invention provides methods for determining GPCR_X 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 GPCR_X 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 GPCR_X 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 GPCR_X protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes GPCR_X protein such that the presence of GPCR_X is detected in the biological sample. An agent for detecting GPCR_X mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to GPCR_X mRNA or genomic DNA.

The nucleic acid probe can be, for example, a full-length GPCR_X nucleic acid, such as the nucleic acid of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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

WO 01/64879

PCT/US01/06474

GPCRX mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting GPCR_X protein is an antibody capable of binding to GPCR_X 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 GPCR_X mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of GPCR_X mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of GPCR_X protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, and immunofluorescence. *In vitro* techniques for detection of GPCR_X genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of GPCR_X protein include introducing into a subject a labeled anti-GPCR_X 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 GPCR_X protein, mRNA, or genomic DNA, such that the presence of GPCR_X protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of GPCR_X protein, mRNA or genomic DNA in the control sample with the presence of GPCR_X protein, mRNA or genomic DNA in the test sample.

WO 01/64879

PCT/US01/06474

The invention also encompasses kits for detecting the presence of GPCR_X in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting GPCR_X protein or mRNA in a biological sample; means for determining the amount of GPCR_X in the sample; and means for comparing the amount of GPCR_X 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 GPCR_X 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 GPCR_X 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 GPCR_X protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk of developing a disease or disorder. Thus, the invention provides a method for identifying a disease or disorder associated with aberrant GPCR_X expression or activity in which a test sample is obtained from a subject and GPCR_X protein or nucleic acid (e.g., mRNA, genomic DNA) is detected, wherein the presence of GPCR_X protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant GPCR_X 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 GPCR_X 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 GPCR_X expression or activity in which a test sample is obtained and GPCR_X protein or nucleic acid is detected (e.g., wherein the presence of GPCR_X protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant GPCR_X expression or activity).

WO 01/64879

PCT/US01/06474

The methods of the invention can also be used to detect genetic lesions in an GPCR_X 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 GPCR_X-protein, or the misexpression of the GPCR_X 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 GPCR_X gene; (ii) an addition of one or more nucleotides to an GPCR_X gene; (iii) a substitution of one or more nucleotides of an GPCR_X gene, (iv) a chromosomal rearrangement of an GPCR_X gene; (v) an alteration in the level of a messenger RNA transcript of an GPCR_X gene, (vi) aberrant modification of an GPCR_X 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 GPCR_X gene, (viii) a non-wild-type level of an GPCR_X protein, (ix) allelic loss of an GPCR_X gene, and (x) inappropriate post-translational modification of an GPCR_X 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 GPCR_X 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., Landegren, *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 GPCR_X-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 GPCR_X gene under conditions such that hybridization and amplification of the GPCR_X 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.

WO 01/64879

PCT/US01/06474

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, Kwok, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 1173-1177); Q β 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 GPCR χ 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 GPCR χ 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 GPCR χ 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 GPCR χ gene and detect mutations by comparing the sequence of the sample GPCR χ 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

WO 01/64879

PCT/US01/06474

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).

5 Other methods for detecting mutations in the GPCR_X 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 GPCR_X sequence with
10 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 S₁ nuclease to enzymatically digesting the mismatched regions. In other embodiments, either
15 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
20 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 GPCR_X cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli*
25 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 GPCR_X sequence, *e.g.,* a wild-type GPCR_X 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
30 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 GPCR_X 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,

WO 01/64879

PCT/US01/06474

1993. *Mutat. Res.* 285: 125-144; Hayashi, 1992. *Genet. Anal. Tech. Appl.* 9: 73-79.

Single-stranded DNA fragments of sample and control GPCR 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

WO 01/64879

PCT/US01/06474

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 GPCR_X gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which GPCR_X 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 GPCR_X activity (e.g., GPCR_X 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 GPCR_X protein, expression of GPCR_X nucleic acid, or mutation content of GPCR_X genes in an

WO 01/64879

PCT/US01/06474

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
 5 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
 10 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
 15 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
 20 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
 25 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 GPCR_X protein, expression of GPCR_X nucleic acid, or mutation content of GPCR_X genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition,
 30 pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness

WO 01/64879

PCT/US01/06474

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 GPCR $\times$ modulator, such as a modulator identified by one of the exemplary screening assays described herein.

5

Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of GPCR $\times$ (*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 GPCR $\times$ gene expression, protein levels, or upregulate GPCR $\times$ activity, can be monitored in clinical trials of subjects exhibiting decreased GPCR $\times$ gene expression, protein levels, or downregulated GPCR $\times$ activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease GPCR $\times$ gene expression, protein levels, or downregulate GPCR $\times$ activity, can be monitored in clinical trials of subjects exhibiting increased GPCR $\times$ gene expression, protein levels, or upregulated GPCR $\times$ activity. In such clinical trials, the expression or activity of GPCR $\times$ 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 GPCR $\times$, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) that modulates GPCR $\times$ 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 GPCR $\times$ 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 GPCR $\times$ 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

WO 01/64879

PCT/US01/06474

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 GPCR α 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 GPCR α protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the GPCR α protein, mRNA, or genomic DNA in the pre-administration sample with the GPCR α 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 GPCR α 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 GPCR α 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 GPCR α 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, neoplasia, 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 Osteodystrophy, 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,

WO 01/64879

PCT/US01/06474

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

5 "knockout" endogenous 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.

10 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,

15 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

20 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).

25 Prophylactic Methods

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant GPCR_X expression or activity, by administering to the subject an agent that modulates GPCR_X expression or at least one GPCR_X activity. Subjects at risk for a disease that is caused or contributed to by aberrant GPCR_X expression or activity

30 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 GPCR_X aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending upon the type of GPCR_X aberrancy, for example, an GPCR_X agonist or GPCR_X antagonist agent can be used for

WO 01/64879

PCT/US01/06474

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.

5 Therapeutic Methods

Another aspect of the invention pertains to methods of modulating GPCR_X 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 GPCR_X protein activity associated with the cell. An agent that modulates GPCR_X protein activity can be an
 10 agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of an GPCR_X protein, a peptide, an GPCR_X peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more GPCR_X protein activity. Examples of such stimulatory agents include active GPCR_X protein and a nucleic acid molecule encoding GPCR_X that has been introduced into the cell. In another embodiment, the agent inhibits one
 15 or more GPCR_X protein activity. Examples of such inhibitory agents include antisense GPCR_X nucleic acid molecules and anti-GPCR_X 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
 20 activity of an GPCR_X 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) GPCR_X expression or activity. In another embodiment, the method involves administering an GPCR_X protein or nucleic acid molecule as therapy to compensate for reduced or aberrant
 25 GPCR_X expression or activity.

Stimulation of GPCR_X activity is desirable in situations in which GPCR_X is abnormally downregulated and/or in which increased GPCR_X 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
 30 associated disorders). Another example of such a situation is where the subject has a gestational disease (e.g., preeclampsia).

WO 01/64879

PCT/US01/06474

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.

- 5 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 GPCR_X 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.

- 20 As an example, a cDNA encoding the GPCR_X 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 GPCR_X protein, and the GPCR_X 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.

WO 01/64879

PCT/US01/06474

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.

WO 01/64879

PCT/US01/06474

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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83;
 - (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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83; and
 - (d) a variant of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83.
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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38.
4. The polypeptide of claim 1, wherein the amino acid sequence of said variant comprises a conservative amino acid substitution.

WO 01/64879

PCT/US01/06474

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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83;
 - (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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83;
 - (d) a variant of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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.

WO 01/64879

PCT/US01/06474

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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38.
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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38;
 - (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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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, 16, 18, 20, 22, 24, 28, 30, 32, 34, 36, and 38, 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.

WO 01/64879

PCT/US01/06474

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.

WO 01/64879

PCT/US01/06474

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 GPCR_X-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 GPCR_X-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 GPCR_X-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired

WO 01/64879

PCT/US01/06474

the nucleic acid of claim 5 in an amount sufficient to treat or prevent said GPCR_X-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 GPCR_X-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 GPCR_X-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.

WO 01/64879

PCT/US01/06474

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.

WO 01/64879

PCT/US01/06474

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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, wherein said method comprises:

WO 01/64879

PCT/US01/06474

- 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83;
 - 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83, 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, 17, 19, 21, 23, 25, 29, 31, 33, 35, 37, and 83;
 - 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.

【国際公開パンフレット（コレクトバージョン）】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number
WO 01/64879 A3

- (51) International Patent Classification: C12N 15/12, C07K 14/705, 16/28, C12N 1/21, S10, G01N 33/68, C12Q 1/68, A61K 38/17, 31/70, 39/395
- (21) International Application Number: PCT/US01/06474
- (22) International Filing Date: 26 February 2001 (26.02.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- | | | | | | |
|------------|-------------------------------|----|------------------|---------------------------|------------------|
| 60/185,674 | 28 February 2000 (28.02.2000) | US | 60/186,716 (CIP) | US | 60/186,716 (CIP) |
| 60/185,535 | 28 February 2000 (28.02.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/186,585 | 3 March 2000 (03.03.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/186,604 | 3 March 2000 (03.03.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/186,606 | 3 March 2000 (03.03.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/186,715 | 3 March 2000 (03.03.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/186,716 | 3 March 2000 (03.03.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/186,717 | 3 March 2000 (03.03.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/186,719 | 3 March 2000 (03.03.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/186,827 | 3 March 2000 (03.03.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/218,323 | 14 July 2000 (14.07.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/218,435 | 14 July 2000 (14.07.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/220,517 | 24 July 2000 (24.07.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/223,897 | 9 August 2000 (09.08.2000) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/260,030 | 5 January 2001 (05.01.2001) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 60/264,849 | 26 January 2001 (26.01.2001) | US | Filed on | 3 March 2000 (03.03.2000) | US |
| 09/795,271 | 27 February 2001 (27.02.2001) | US | Filed on | 3 March 2000 (03.03.2000) | US |
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:
- | | |
|----------|-------------------------------|
| US | 60/185,674 (CIP) |
| Filed on | 28 February 2000 (28.02.2000) |
| US | 60/185,535 (CIP) |
| Filed on | 28 February 2000 (28.02.2000) |
| US | 60/186,585 (CIP) |
| Filed on | 3 March 2000 (03.03.2000) |
| US | 60/186,604 (CIP) |
| Filed on | 3 March 2000 (03.03.2000) |
| US | 60/186,606 (CIP) |
| Filed on | 3 March 2000 (03.03.2000) |
| US | 60/186,715 (CIP) |
| Filed on | 3 March 2000 (03.03.2000) |
| US | 60/186,716 (CIP) |
| Filed on | 3 March 2000 (03.03.2000) |
| US | 60/186,717 (CIP) |
| Filed on | 3 March 2000 (03.03.2000) |
- (71) Applicant (for all designated States except US): CITRAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th floor, New Haven, CT 06511 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): PADIGARU, Muradhdhara [IN/US]; 1579 Rhineland Avenue, Apartment 64E, Bronx, NY 10561 (US); MAJUMDER, Komud [IN/US]; 140 Silver Hill Lane, Stamford, CT 06905 (US); BURGESS, Catherine, E. [US/US]; 90 Carriage Hill Drive, Waterfield, CT 06109 (US); VERNET, Carine, A., M. [FR/US]; 4830 N.W. 43rd Street P9253, Gainesville, FL 32660 (US); FERNANDES, Elina [IN/US]; 77 Florence Road, #2, Branford, CT 06405 (US); SHIMKETS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US); TCHERNEV, Velizar, E. [BG/US]; 1216 S.W. 2nd Avenue, Apt. 156, Gainesville, FL 32601 (US); MISHRA, Vishnu [IN/US]; 309 East Main Street, #306, Branford, CT 06405 (US); CASMAN, Stacie [US/US]; 155 Knollwood Drive, Wallingford, CT 06492 (US); SPYTEK, Kimberly, A. [US/US]; 28 Court Street #1, New Haven, CT 06511 (US); ZERHUSEN, Bryan [US/US]; 45 Jefferson Road, Apartment #2-10, Branford, CT 06405 (US).

[Continued on next page]

(54) Title: G-PROTEINS COUPLED RECEPTOR RELATED POLYPEPTIDES

(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 immunospecifically 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 use of these novel human nucleic acids and proteins.



WO 01/64879 A3

WO 01/64879 A3



(74) Agent: ELRIFI, Ivor, R.; Miniz, Levin, Cohn, Ferris, Glosky and Popco, P.C., One Financial Center, Boston, MA 02111 (U.S.).

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, LJ, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(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.

Published:

with international search report

(88) Date of publication of the international search report: 10 May 2002

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

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.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		Int. Appl. No. PCT/US 01/06474						
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/705 C07K16/28 C12N1/21 C12M5/10 G01N33/68 C12Q1/68 A61K38/17 A61K31/70 A61K39/395								
According to International Patent Classification (IPC), to the extent that it is not covered by the prior art.								
B. FIELDS SEARCHED Abstracts searched: IPC 7 C12N C07K G01N								
Documents searched: Other than minimum of documents to the extent that such documents are included in the fields searched.								
Electronic data base consulted during the international search (name of data base and, where practical, search terms used): EPO-Internal, SEQUENCE SEARCH, EMBL, BIOSIS, NPI Data, PAJ								
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category</th> <th>Document or other information, where appropriate, of the relevant document</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>EP 0 867 508 A (SMITHKLINE BEECHAM CORP) 30 September 1998 (1998-09-30) the whole document --- -/-</td> <td>1-24, 26-52</td> </tr> </tbody> </table>			Category	Document or other information, where appropriate, of the relevant document	Relevant to claim No.	A	EP 0 867 508 A (SMITHKLINE BEECHAM CORP) 30 September 1998 (1998-09-30) the whole document --- -/-	1-24, 26-52
Category	Document or other information, where appropriate, of the relevant document	Relevant to claim No.						
A	EP 0 867 508 A (SMITHKLINE BEECHAM CORP) 30 September 1998 (1998-09-30) the whole document --- -/-	1-24, 26-52						
☒ Further documents are listed in the continuation of part C. ☒ Document or other information are listed in part C.								
Subject categories of cited documents: "A" document relating to the general state of the art and is not considered to be of particular interest. "E" document relating to the state of the art and is not considered to be of particular interest. "I" document which may raise doubts on priority, but which is not considered to be of particular interest. "O" document relating to an invention which is not considered to be of particular interest. "P" document published prior to the international filing date but later than the priority date.								
Date of the actual date of the international search: 30 October 2001								
Date of mailing of the international search report: 31. Oct. 2002								
Name and mailing address of the ISA: European Patent Office (EPO) 5101, rue de la Woluwe 1200, Brussels, Belgium Tel: +31 (0) 22 343 2446, Fax: +31 (0) 22 343 2446 Telex: +31 (0) 22 343 2446								
Authorized officer: LE CORNEC N.D.R.								

Form PCT/ISA/210 (Rev. 01/97)

INTERNATIONAL SEARCH REPORT

International Application No.
PL/US 01/06474

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication where appropriate of the relevant categories	Relevant to
X	DATABASE ENBL [Online] MMU133427, accession number AJ133427, 9 September 1999 (1999-09-09) J. STROTMAN ET AL: "Olfactory receptor; or37d gene; pseudogene" XP002181511 cited in the application * 84.8% identity in 1021 nt overlap with seq ID no.1 (5-1013:41-1056) * * 85% identity in 1015 nt overlap with seq ID no.3 (3-1011:47-1061) * abstract -& J. STROTMAN ET AL: "Small subfamily of olfactory receptor genes: Structural features, expression pattern and genomic organization" GENE, vol. 236, 1999, pages 281-291, XP004178064 ELSEVIER BIOMEDICAL PRESS, AMSTERDAM, NL ISSN: 0378-1119	1,5-7,9, 10,12,14
X	DATABASE ENBL [Online] MMU133428, accession number AJ133428, 9 September 1999 (1999-09-09) J. STROTMAN: "Olfactory receptor or37e gene" XP002181512 * 79.7% identity in 902 nt overlap with seq ID no.5 (108-1002:152-1047) * abstract -& J. STROTMAN ET AL: "Small subfamily of olfactory receptor genes: structural features, expression pattern and genomic organization" GENE, vol. 236, 1999, pages 281-291, XP004178064	1,5-7,9, 10,12,14
X	DATABASE ENBL [Online] AL135841, 21 December 1999 (1999-12-21) N. CORBY: "human DNA sequence from clone RP11-327L3 on chromosome9p13.1-13-3" XP002181513 * 84.9% identity in 997 nt overlap with seq ID no.1 (995-3:81731-82724) * abstract	1,5-7,9, 10,12, 14,19
A	ROUQUIER S ET AL: "Distribution of olfactory receptor genes in the human genome" NATURE GENETICS, NEW YORK, NY, US, vol. 18, no. 3, March 1998 (1998-03), pages 243-250, XP002111208 ISSN: 1061-4036	

Form PCT/ISA(2) distribution on or before 15th July 1999

INTERNATIONAL SEARCH REPORT

International Application No.
PL/US 01/06474

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication where appropriate of the relevant passages	Relevant to claim No.
A	DATABASE EMBL [Online] AL138834, accession number AL138834, 4 February 2000 (2000-02-04) D. LLOYD: "Human DNA sequence from clone RP11-113A10 on chromosome 9" XP002181514 * 99.9% identity in 1019 nt overlap with sequence ID no.1 (1019-1:27692-28709) * * 99.9% identity in 1303 bp overlap with seq ID no.5 (1093-1:16125-17126) * * 90.34% identity in 807 nt overlap with seq ID no.3 (1014-217:16121-16919)* abstract	5-11,14
E	WO 01 27158 A (FUCHS TANJA ; GLUSMAN GUSTAVO (IL); LANCET DORON (IL); YEDA RES & D) 19 April 2001 (2001-04-19) sequences ID no.733, 734, 735, 1665, 1666, 1667 and 1670 claims	1-24, 26-52

File No. 1754279 (continuation of 5000000) (2nd 1992)

INTERNATIONAL SEARCH REPORT		National Application No. PCT/US 01/06474
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)		
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1.	☒ Claims Nos.	because they relate to subject matter not required to be searched by this Authority namely: Although claims 26-37 and 48-49 are directed to a method of treatment of the human/animal body (rule 39.1 IV PCT), the search has been carried out and based on the alleged effects of the compound/composition.
2.	☒ Claims Nos. 25	because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out. Specifically: see FURTHER INFORMATION sheet PCT/ISA/210
3.	☐ Claims Nos.	because they are dependent claims and are not drafted in accordance with the caption and introductory sentences of claim 5, 4(a).
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)		
This International Searching Authority found multiple inventions in this International Application, as follows: see additional sheet		
1.	☐	As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.	☐	As all searchable claims could be searched without effort requiring an additional fee, this Authority did not make payment of any additional fee.
3.	☐	As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims in which fees were paid, specifically claims Nos.:
4.	☒	No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claim Nos. 1-52 all partially
Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees.		

International Application No. PCT/US 01/06474

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box 1.2

Claims Nos.: 25

Claim 25 refers to a method for modulating the activity of the polypeptide GPCR using a compound that binds to said polypeptide without giving a true technical characterization. Moreover, no such compounds are disclosed in the application. In consequence, the scope of said claim is ambiguous and vague and its subject-matter is not sufficiently disclosed and supported (Art. 5/6 PCT).

no meaningful search can be carried out for such purely speculative claim whose wording is, in fact, a mere recitation of the result to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

International Application No. PCT/US 01/06474

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-52 all partially

Olfactory receptors GPCR1a, GPCR1b and GPCR1c represented respectively by the amino acid sequences ID no.2, 4 and 6 and their respective corresponding DNA sequences ID no.1, 3 and 5. Antibodies. Therapeutic and diagnostic uses.

2. Claims: 1-52 all partially

Olfactory receptor GPCR2 represented respectively by the amino acid sequences ID no.8 and its respective corresponding DNA sequence ID no.7. Antibodies. Therapeutic and diagnostic uses.

3. Claims: 1-52 all partially

Olfactory receptor GPCR3 represented respectively by the amino acid sequence ID no.10 and its respective corresponding DNA sequence ID no.9. Antibodies. Therapeutic and diagnostic uses.

4. Claims: 1-52 all partially

Olfactory receptors GPCR4a and GPCR4b represented respectively by the amino acid sequences ID no.12 and no.17 and their corresponding DNA sequences ID no.11, 13 and 16. Antibodies. Therapeutic and diagnostic uses.

5. Claims: 1-52 all partially

Olfactory receptors GPCR5a, GPCR5b and GPCR5c represented respectively by the amino acid sequences ID no.19, 21 and 23 and their respective corresponding DNA sequences ID no.18, 20 and 22. Antibodies. Therapeutic and diagnostic uses.

6. Claims: 1-52 all partially

Olfactory receptor GPCR6 represented by the amino acid sequence ID no.25 and its corresponding DNA sequence ID no. 24. Antibodies. Therapeutic and diagnostic uses.

7. Claims: 1-52 all partially

Olfactory receptor GPCR7 represented by the amino acid sequence ID no.29 and its corresponding DNA sequence ID no. 28. Antibodies. Therapeutic and diagnostic uses.

International Application No. PCT/US 01/06474

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

8. Claims: 1-52 all partially

Olfactory receptor GPCR8 represented by the amino acid sequence ID no.31 and its corresponding DNA sequence ID no. 30. Antibodies. Therapeutic and diagnostic uses.

9. Claims: 1-52 all partially

Olfactory receptor GPCR9 represented by the amino acid sequence ID no.33 and its corresponding DNA sequence ID no. 32. Antibodies. Therapeutic and diagnostic uses.

10. Claims: 1-52 all partially

Olfactory receptor GPCR10a, GPCR10b and GPCR10c represented respectively by the amino acid sequences ID no.35, 37 and 83 and their respective corresponding DNA sequence ID no.34, 36 and 38. Antibodies. Therapeutic and diagnostic uses.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 01/06474

PC/US 01/06474

Patent document cited in search report		Publication date	Patent form & numbers	Publication date
EP 0867508	A	30-09-1998	US 5874243 A	23-02-1999
			EP 0867508 A2	30-09-1998
			JP 11028094 A	02-02-1999
WO 0127158	A	19-04-2001	AU 1326201 A	13-02-2001
			AU 2963701 A	23-04-2001
			AU 5118160 A	13-02-2001
			WO 0107093 A1	01-02-2001
			WO 0107094 A1	01-02-2001
			WO 0127158 A2	19-04-2001

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
A 6 1 P 3/00	A 6 1 P 3/04	4 C 0 8 4
A 6 1 P 3/04	A 6 1 P 3/10	4 C 0 8 5
A 6 1 P 3/10	A 6 1 P 7/00	4 H 0 4 5
A 6 1 P 7/00	A 6 1 P 7/06	
A 6 1 P 7/06	A 6 1 P 9/04	
A 6 1 P 9/04	A 6 1 P 9/10	
A 6 1 P 9/10	A 6 1 P 9/12	
A 6 1 P 9/12	A 6 1 P 11/06	
A 6 1 P 11/06	A 6 1 P 13/10	
A 6 1 P 13/10	A 6 1 P 19/08	
A 6 1 P 19/08	A 6 1 P 19/10	
A 6 1 P 19/10	A 6 1 P 25/00	
A 6 1 P 25/00	A 6 1 P 25/14	
A 6 1 P 25/14	A 6 1 P 25/16	
A 6 1 P 25/16	A 6 1 P 25/22	
A 6 1 P 25/22	A 6 1 P 25/24	
A 6 1 P 25/24	A 6 1 P 25/28	
A 6 1 P 25/28	A 6 1 P 31/04	
A 6 1 P 31/04	A 6 1 P 31/10	
A 6 1 P 31/10	A 6 1 P 31/18	
A 6 1 P 31/18	A 6 1 P 33/02	
A 6 1 P 33/02	A 6 1 P 35/00	
A 6 1 P 35/00	A 6 1 P 37/00	
A 6 1 P 37/00	C 0 7 K 14/705	
C 0 7 K 14/705	C 0 7 K 16/28	
C 0 7 K 16/28	C 1 2 N 1/15	
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 Q 1/68	A
C 1 2 N 5/10	G 0 1 N 33/15	Z
C 1 2 Q 1/68	G 0 1 N 33/50	Z
G 0 1 N 33/15	G 0 1 N 33/53	D
G 0 1 N 33/50	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/53	N
G 0 1 N 33/566	G 0 1 N 33/566	
// C 1 2 P 21/08	C 1 2 N 5/00	A
	A 6 1 K 37/02	
	C 1 2 P 21/08	

- (31)優先権主張番号 60/186,604
(32)優先日 平成12年3月3日(2000.3.3)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/186,606
(32)優先日 平成12年3月3日(2000.3.3)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/186,715
(32)優先日 平成12年3月3日(2000.3.3)

- (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/186,716
 (32)優先日 平成12年3月3日(2000.3.3)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/186,717
 (32)優先日 平成12年3月3日(2000.3.3)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/186,719
 (32)優先日 平成12年3月3日(2000.3.3)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/186,827
 (32)優先日 平成12年3月3日(2000.3.3)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/218,323
 (32)優先日 平成12年7月14日(2000.7.14)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/218,435
 (32)優先日 平成12年7月14日(2000.7.14)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/220,517
 (32)優先日 平成12年7月24日(2000.7.24)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/223,897
 (32)優先日 平成12年8月9日(2000.8.9)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/260,020
 (32)優先日 平成13年1月5日(2001.1.5)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/264,849
 (32)優先日 平成13年1月26日(2001.1.26)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 09/795,271
 (32)優先日 平成13年2月27日(2001.2.27)
 (33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),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

(72)発明者 マジウムデル, クムド
 アメリカ合衆国 コネチカット 06905, スタンフォード, シルバー ヒル レーン 140

(72)発明者 バージェス, キャサリン イー.
 アメリカ合衆国 コネチカット 06109, ウェザーズフィールド, キャリッジ ヒル ドライブ 90

(72)発明者 ヴェルネ, コリーヌ アー. エム.
 アメリカ合衆国 フロリダ 32060, ゲインズビル, エヌ. ダブリュー. 43アールデ
 ィー ストリート 4830 ピーナンバー253

- (72)発明者 フェルナンデス, エルマ
アメリカ合衆国 コネチカット 06405, ブランフォード, フローレンス ロード 77,
ナンバー2
- (72)発明者 シムケッツ, リチャード エイ.
アメリカ合衆国 コネチカット 06516, ウェスト ハイブン, リート ストリート 1
91
- (72)発明者 チェルネフ, ヴェリザル テー.
アメリカ合衆国 フロリダ 32601, ゲインズビル, エス. ダブリュー. 2エヌディー
アベニュー 1216, アpartment 156
- (72)発明者 ミシュラ, ヴィシュヌ
アメリカ合衆国 コネチカット 06405, ブランフォード, イースト メイン ストリ
ート 309, ナンバー306
- (72)発明者 キャスマン, ステイシー
アメリカ合衆国 コネチカット 06492, ウォーリングフォード, ノウルウッド ドライ
ブ 155
- (72)発明者 スパイテック, キンバリー エイ.
アメリカ合衆国 コネチカット 06511, ニュー ハイブン, コート ストリート 28
ナンバー1
- (72)発明者 ツェアファーゼン, ブライアン
アメリカ合衆国 コネチカット 06405, ブランフォード, ジェファーソン ロード 4
5, アpartment ナンバー2-10

Fターム(参考) 2G045 AA34 AA35 CB01 CB02 DA13 DA36 FB02 FB03
4B024 AA01 AA11 BA41 BA63 CA04 DA01 DA02 DA05 DA11 GA01
GA11 GA18 GA19 HA08 HA11
4B063 QA01 QA18 QQ42 QR32 QR55 QS34
4B064 AG20 AG26 AG27 CA02 CA05 CA10 CA11 CA19 CA20 CC24
DA01 DA13
4B065 AA01X AA57X AA87X AA93Y AB01 AB02 CA24 CA25 CA44 CA46
4C084 AA02 BA22 BA23 NA14 ZA01 ZA06 ZA08 ZA12 ZA16 ZA18
ZA36 ZA38 ZA40 ZA42 ZA43 ZA45 ZA51 ZA59 ZA66 ZA68
ZA70 ZA81 ZA96 ZA97 ZB07 ZB26 ZB33 ZB35 ZB38 ZC21
ZC35 ZC55
4C085 AA14 CC21
4H045 AA10 AA11 AA20 AA30 BA09 CA40 DA50 DA75 DA76 EA20
EA50 FA72 FA73 FA74

专利名称(译)	G蛋白偶联受体相关多肽		
公开(公告)号	JP2004500098A	公开(公告)日	2004-01-08
申请号	JP2001564362	申请日	2001-02-28
[标]申请(专利权)人(译)	CURAGEN CORP		
申请(专利权)人(译)	Kyurajen公司		
[标]发明人	パディガルマラリダラ マジウムデルクムド バージェスキャサリンイー ヴェルネコリーヌアーエム フェルナンデスエルマ シムケッツリチャードエイ チェルネフヴェリザルター ミシュラヴィシユヌ キャスマンステイシー スパイテックキンバリーエイ ツェアフォーゼンブライアン		
发明人	パディガル, マラリダラ マジウムデル, クムド バージェス, キャサリン イー. ヴェルネ, コリーヌ アー. エム. フェルナンデス, エルマ シムケッツ, リチャード エイ. チェルネフ, ヴェリザル テー. ミシュラ, ヴィシユヌ キャスマン, ステイシー スパイテック, キンバリー エイ. ツェアフォーゼン, ブライアン		
IPC分类号	G01N33/50 A61K38/00 A61K39/00 A61K39/395 A61P1/04 A61P1/14 A61P3/00 A61P3/04 A61P3/10 A61P7/00 A61P7/06 A61P9/04 A61P9/10 A61P9/12 A61P11/06 A61P13/10 A61P19/08 A61P19/10 A61P25/00 A61P25/14 A61P25/16 A61P25/22 A61P25/24 A61P25/28 A61P31/04 A61P31/10 A61P31 /18 A61P33/02 A61P35/00 A61P37/00 C07K14/705 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12N5 /10 C12N15/09 C12N15/12 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	A61K38/00 A61K39/00 A61P1/04 A61P1/14 A61P3/00 A61P3/04 A61P11/06 A61P13/10 A61P19/08 A61P19/10 A61P25/00 A61P25/14 A61P25/16 A61P25/22 A61P25/24 A61P25/28 A61P31/04 A61P31 /10 A61P31/18 A61P33/02 A61P35/00 C07K14/705		
FI分类号	C12N15/00.ZNA.A A61K39/395.N A61P1/04 A61P1/14 A61P3/00 A61P3/04 A61P3/10 A61P7/00 A61P7/06 A61P9/04 A61P9/10 A61P9/12 A61P11/06 A61P13/10 A61P19/08 A61P19/10 A61P25/00 A61P25/14 A61P25/16 A61P25/22 A61P25/24 A61P25/28 A61P31/04 A61P31/10 A61P31/18 A61P33 /02 A61P35/00 A61P37/00 C07K14/705 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/53.N G01N33/566 C12N5/00.A A61K37/02 C12P21/08		
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/CB01 2G045/CB02 2G045/DA13 2G045/DA36 2G045/FB02 2G045 /FB03 4B024/AA01 4B024/AA11 4B024/BA41 4B024/BA63 4B024/CA04 4B024/DA01 4B024/DA02 4B024/DA05 4B024/DA11 4B024/GA01 4B024/GA11 4B024/GA18 4B024/GA19 4B024/HA08 4B024 /HA11 4B063/QA01 4B063/QA18 4B063/QQ42 4B063/QR32 4B063/QR55 4B063/QS34 4B064/AG20 4B064/AG26 4B064/AG27 4B064/CA02 4B064/CA05 4B064/CA10 4B064/CA11 4B064/CA19 4B064		

/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA87X 4B065/AA93Y 4B065/AB01 4B065/AB02 4B065/CA24 4B065/CA25 4B065/CA44 4B065/CA46 4C084/AA02 4C084/BA22 4C084/BA23 4C084/NA14 4C084/ZA01 4C084/ZA06 4C084/ZA08 4C084/ZA12 4C084/ZA16 4C084/ZA18 4C084/ZA36 4C084/ZA38 4C084/ZA40 4C084/ZA42 4C084/ZA43 4C084/ZA45 4C084/ZA51 4C084/ZA59 4C084/ZA66 4C084/ZA68 4C084/ZA70 4C084/ZA81 4C084/ZA96 4C084/ZA97 4C084/ZB07 4C084/ZB26 4C084/ZB33 4C084/ZB35 4C084/ZB38 4C084/ZC21 4C084/ZC35 4C084/ZC55 4C085/AA14 4C085/CC21 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA09 4H045/CA40 4H045/DA50 4H045/DA75 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA73 4H045/FA74

优先权	60/185674 2000-02-28 US
	60/185535 2000-02-28 US
	60/186585 2000-03-03 US
	60/186604 2000-03-03 US
	60/186606 2000-03-03 US
	60/186715 2000-03-03 US
	60/186716 2000-03-03 US
	60/186717 2000-03-03 US
	60/186719 2000-03-03 US
	60/186827 2000-03-03 US
	60/218323 2000-07-14 US
	60/218435 2000-07-14 US
	60/220517 2000-07-24 US
	60/223897 2000-08-09 US
	60/260020 2001-01-05 US
	60/264849 2001-01-26 US
	09/795271 2001-02-27 US

外部链接 [Espacenet](#)

摘要(译)

本文公开了编码G偶联的蛋白受体相关多肽的核酸序列。还公开了由这些核酸序列编码的多肽和与多肽免疫特异性结合的抗体，以及上述多肽，多核苷酸或抗体的片段的衍生物，变体，突变体。本发明进一步公开了用于诊断，治疗和预防涉及这些新型人核酸和蛋白质中的任何一种的疾病的治疗，诊断和研究方法。

表 1A. GPCR1a ヌクレオチド配列 (配列番号 1).
CCTTCAGTTGACAGAGGAGATACACTATGGTAAGTGCCAAATCAGACAGCCTCTGTGACCGAGTTTATTC TCCTGGGCTCTCTGCCCACCCAAAGCTGGAGAAAACGTTCTTTGCTCATCCTGCTGATGTACCTGG TGATCCTACTGGGCAATGGGTCCTCATCCTGATGACTGTGTCCAATCCCACCTGCACATGCCCATGT ACTTCTTCCTGGGGAACCTCTCCTTCCTGGACATCTGCTATACAACATCCTCAGTCCCCCTCATCCTTG ACAGCTTCTTGACCCCAAGAAAACCACTCCTTCTCAGCCTGTGCAGTGCAGATGTTCTCTCCTTTG CCATGGGAGCCACAGAGTGTGTTCTCCTGAGCATGATGGCCTTTGATCGCTACGTGGCCATCTGCAACC CCCTTAGGTACCCCTGTGCTCATGAGCAAGGCTGCCTACATGCCCATTAAGGCTGCCGCTCCTGGGTAG CTGGAAGCACTGCTTCCATGGTGCAGACATCCCTTGCAATGAGGCTGCCCTTCTGTGGAGACAACATCA TCAACCACTTCACCTGTGAGATTCTGGCTGTCTTGAAGTTGGCCTGTGCTGATACTCTGTCAATGTGA TCAGTATGGGAGTGACCAATGTGATCTTCCTGGGGTCCCGGTTCTGTTCACTCTTTCTCTATGTCT TCATCATTTGCCACCATCCTGAGGATCCCTCAGCTGAGGGGAGGAAAAAGGCTTC7CCACCTGCTCTG CCCACCTCACAGTCGTGGTCACTTCTATGGGACCATCCTTTCATGTATGGGAAGCCCAAGTCTAAGG ACCCGCTGGGGGACAGACAAGCAAGACCTTGAGACAAAGCTCATTTCCCTTTTCTATGGGGTGGTGACCC CCATGCTCAACCCCATCATCTACAGCCTGAGGAACAAGGATGTAAAGGCTGCTGTGAGGGACTTGATAT TTCAGAAATGCTTTGCCCTACTGATGTTTGGGGAACAGATGTCCTCATAGCTC