

(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2004-528007

(P2004-528007A)

(43) 公表日 平成16年9月16日(2004.9.16)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
<b>C 1 2 N 15/09</b>	C 1 2 N 15/00	Z N A A
<b>A 6 1 K 38/00</b>	A 6 1 P 1/00	2 G O 4 5
<b>A 6 1 P 1/00</b>	A 6 1 P 1/04	4 B O 2 4
<b>A 6 1 P 1/04</b>	A 6 1 P 1/06	4 B O 2 9
<b>A 6 1 P 1/06</b>	A 6 1 P 1/08	4 B O 6 3
		4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 263 頁) 最終頁に続く

(21) 出願番号	特願2002-531207 (P2002-531207)	(71) 出願人	301005050
(86) (22) 出願日	平成13年9月28日 (2001.9.28)		インサイト・ゲノミックス・インコーポレイテッド
(85) 翻訳文提出日	平成15年3月28日 (2003.3.28)		アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) 国際出願番号	PCT/US2001/030661	(74) 代理人	100089266
(87) 国際公開番号	W02002/026825		弁理士 大島 陽一
(87) 国際公開日	平成14年4月4日 (2002.4.4)	(72) 発明者	ボーグン、マライア・アール
(31) 優先権主張番号	60/236,546		アメリカ合衆国カリフォルニア州94577・サンレアンドロ・サンティアゴロード 14244
(32) 優先日	平成12年9月29日 (2000.9.29)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/240,589		
(32) 優先日	平成12年10月13日 (2000.10.13)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/242,322		
(32) 優先日	平成12年10月20日 (2000.10.20)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 Gタンパク質共役受容体

## (57) 【要約】

本発明はヒトGタンパク質共役受容体 (GCRRC)、およびGCRRCを同定し、コードするポリヌクレオチドを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニストおよびアンタゴニストをも提供する。本発明はまた、GCRRCの異常発現に関連する疾患を診断、治療または予防する方法をも提供する。

## 【特許請求の範囲】

## 【請求項 1】

以下の (a) 乃至 (d) からなる群から選択した単離されたポリペプチド。

(a) SEQ ID NO: 1 - 16 (配列番号 1 乃至 16) を有する群から選択したアミノ酸配列を含むポリペプチド

(b) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列と少なくとも 90% が同一であるような天然のアミノ酸配列を含むポリペプチド

(c) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片

(d) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片 10

## 【請求項 2】

SEQ ID NO: 1 - 16 からなる群から選択された請求項 1 の単離されたポリペプチド。

## 【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

## 【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

## 【請求項 5】

SEQ ID NO: 17 - 32 を有する群から選択した請求項 4 に記載の単離されたポリヌクレオチド。 20

## 【請求項 6】

請求項 3 に記載のポリヌクレオチドに機能的に連結したプロモーター配列を含む組換えポリヌクレオチド。

## 【請求項 7】

請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

## 【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを含む遺伝形質転換体。

## 【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、 30

(a) 前記ポリペプチドの発現に好適な条件下で、請求項 1 のポリペプチドをコードするポリヌクレオチドに機能的に連結されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養する過程と、

(b) そのように発現した前記ポリペプチドを回収する過程。

## 【請求項 10】

前記ポリペプチドが、SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有することを特徴とする請求項 9 に記載の方法。

## 【請求項 11】

請求項 1 に記載のポリペプチドと特異結合するような単離された抗体。

12. 以下の (a) 乃至 (d) からなる群から選択した単離されたポリヌクレオチド。 40

(a) SEQ ID NO: 17 - 32 からなる群から選択したポリヌクレオチド配列を有するポリヌクレオチド

(b) SEQ ID NO: 17 - 32 からなる群から選択したポリヌクレオチド配列と少なくとも 90% が同一であるような天然のポリヌクレオチド配列を含むポリヌクレオチド

(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (a) ~ (d) の RNA 等価物

## 【請求項 12】

請求項 12 に記載のポリヌクレオチドの少なくとも 60 の連続したヌクレオチドを含む単 50

離されたポリヌクレオチド。

【請求項 13】

請求項 12 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

(a) 前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を有する少なくとも 20 の連続したヌクレオチドを含むプローブを用いて前記サンプルをハイブリダイズする過程と、

(b) 前記ハイブリダイゼーション複合体の存在・不存在を検出し、該複合体が存在する場合にはオプションでその量を検出する過程からなり、前記プローブと前記標的ポリヌクレオチドの間でハイブリダイゼーション複合体が形成されるような条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズすることを特徴とする方法。

10

【請求項 14】

前記プローブが少なくとも 60 の連続したヌクレオチドを含むことを特徴とする請求項 14 に記載の方法。

【請求項 15】

請求項 12 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

(a) ポリメラーゼ連鎖反応増幅を用いて前記標的ポリヌクレオチドまたはその断片を増幅する過程と、

(b) 前記標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含むことを特徴とする方法。

20

【請求項 16】

有効量の請求項 1 のポリペプチドと、薬剤として許容できる賦形剤とを有することを特徴とする成分。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有することを特徴とする請求項 17 に記載の成分。

【請求項 18】

機能性 GCRIC の発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して請求項 17 に記載の成分を投与する過程を含むことを特徴とする方法。

30

【請求項 19】

請求項 1 に記載のポリペプチドのアゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝すステップと、

(b) 前記サンプルにおいてアゴニスト活性を検出するステップとを含むことを特徴とするスクリーニング方法。

【請求項 20】

請求項 20 に記載の方法によって同定したアゴニスト化合物と、薬剤として許容できる賦形剤とを含むことを特徴とする成分。

40

【請求項 21】

機能性 GCRIC の発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して請求項 21 に記載の成分を投与する過程を含むことを特徴とする方法。

【請求項 22】

請求項 1 に記載のポリペプチドのアンタゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝すステップと、

(b) 前記サンプルにおいてアンタゴニスト活性を検出するステップとを含むことを特徴

50

とするスクリーニング方法。

【請求項 23】

請求項 23 に記載の方法によって同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを含むことを特徴とする成分。

【請求項 24】

機能性 G C R E C の過剰発現に関連する疾患又は病状の治療方法であって、そのような治療を必要とする患者に対して請求項 24 に記載の成分を投与する過程を含むことを特徴とする方法。

【請求項 25】

請求項 1 に記載のポリペプチドに特異結合する化合物をスクリーニングする方法であって 10

( a ) 適切な条件下で請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物に混合させる過程と、

( b ) 請求項 1 に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 に記載のポリペプチドに特異結合する化合物を同定する過程とを含むことを特徴とする方法。

【請求項 26】

請求項 1 に記載のポリペプチドの活性を調節する化合物をスクリーニングする方法であって、

( a ) 請求項 1 に記載のポリペプチドの活性が許容された条件下で、請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物に混合させる過程と、 20

( b ) 請求項 1 に記載のポリペプチドの活性を試験化合物の存在下で算定する過程と、

( c ) 試験化合物の存在下での請求項 1 に記載のポリペプチドの活性を、試験化合物の不存在下での請求項 1 に記載のポリペプチドの活性と比較する過程とを含み、試験化合物の存在下での請求項 1 に記載のポリペプチドの活性の変化が、請求項 1 に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。

【請求項 27】

請求項 5 の配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、

( a ) 前記標的ポリヌクレオチドの発現に好適な条件下で、前記標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、 30

( b ) 前記標的ポリヌクレオチドの変異発現を検出する過程と、

( c ) 可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程とを含むことを特徴とする方法。

【請求項 28】

試験化合物の毒性を算定する方法であって、

( a ) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、

( b ) 処理した前記生体サンプルの核酸と、請求項 12 のポリヌクレオチドの少なくとも 20 の連続するヌクレオチドを含むプローブをハイブリダイズさせるステップであって、このハイブリダイゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 12 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、 40

( c ) ハイブリダイゼーション複合体の収量を定量するステップと、

( d ) 前記処理された生物学的サンプル中の前記ハイブリダイゼーション複合体の量を、処理されていない生物学的サンプル中の前記ハイブリダイゼーション複合体の量と比較する過程とを含み、前記処理された生物学的サンプル中の前記ハイブリダイゼーション複合体の量の差が、前記試験化合物の毒性を標示することを特徴とする方法。

【請求項 29】

生物学的サンプル中の G C R E C の発現に関連する症状または疾患に対する診断試験法で 50

あって、

(a) 前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体が形成されるのに適した条件下で、前記生物学的サンプルを請求項 11 に記載の抗体と結合する過程と、

(b) 前記複合体を検出する過程とを含み、前記複合体の存在が、前記生物学的サンプル中の前記ポリペプチドの存在と相関することを特徴とする方法。

【請求項 30】

前記抗体が、

(a) キメラ抗体

(b) 単鎖抗体

(c) Fab 断片

(d) F(ab')<sub>2</sub> 断片

(e) ヒト化抗体 のいずれかであることを特徴とする請求項 11 に記載の抗体。

【請求項 31】

請求項 11 に記載の抗体と、許容できる賦形剤とを含む化合物。

【請求項 32】

被検者の GCRFC の発現に関連する病状又は疾患の診断方法であって、請求項 32 に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

【請求項 33】

前記抗体が標識されることを特徴とする請求項 32 に記載の化合物。

【請求項 34】

被検者の GCRFC の発現に関連する病状又は疾患の診断方法であって、請求項 34 に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

【請求項 35】

請求項 11 に記載の抗体の特異性を有するポリクローナル抗体を調製する方法であって、(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドまたはその免疫原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体をポリペプチドでスクリーニングし、それによって、SEQ ID NO: 1 - 16 を有する群から選択したアミノ酸配列を有するポリペプチドに特異結合するようなポリクローナル抗体を同定する過程とを含むことを特徴とする方法。

【請求項 36】

請求項 36 に記載の方法で産出したポリクローナル抗体。

【請求項 37】

請求項 37 に記載のポリクローナル抗体及び適切なキャリアを含む化合物。

【請求項 38】

請求項 11 に記載の抗体の特異性を有するモノクローナル抗体を作製する方法であって、(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドまたはその免疫原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体産出細胞を単離する過程と、

(c) 不死化の細胞を用いて前記抗体産出細胞を融合して、モノクローナル抗体を産出するハイブリドーマ細胞を形成する過程と、

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するような前記培養モノクローナル抗体から単離する過程とを含むことを特徴とする方法。

【請求項 39】

請求項 39 に記載の方法で産出したモノクローナル抗体。

10

20

30

40

50

## 【請求項 40】

請求項 40 に記載のモノクローナル抗体及び適切なキャリアを含む化合物。

## 【請求項 41】

F a b 発現ライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項 11 に記載の抗体。

## 【請求項 42】

組換え免疫グロブリンライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項 11 に記載の抗体。

## 【請求項 43】

SEQ ID NO : 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドをサンプル中から検出する方法であって、

( a ) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項 11 に記載の抗体をインキュベートする過程と、

( b ) 特異結合を検出する過程とを含み、該特異結合が、SEQ ID NO : 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドがサンプル中に存在することを標示することを特徴とする方法。

## 【請求項 44】

SEQ ID NO : 1 - 16 を有する群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、

( a ) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項 11 に記載の抗体をインキュベートする過程と、

( b ) 前記サンプルから前記抗体を分離し、SEQ ID NO : 1 - 16 を有する群から選択したアミノ酸配列を有する精製ポリペプチドを得る過程とを含むことを特徴とする方法。

## 【請求項 45】

少なくとも一つのエレメントが請求項 13 のポリヌクレオチドであるようなマイクロアレイ。

## 【請求項 46】

ポリヌクレオチドを含むサンプルの転写イメージを生成する方法であって、

a ) サンプルのポリヌクレオチドを標識する過程と、

b ) ハイブリダイゼーション複合体の形成に適した条件下で請求項 46 のマイクロアレイのエレメントを標識されたポリヌクレオチドに接触させる過程と、

c ) サンプル内のポリヌクレオチドの発現を定量化する過程からなる方法。

## 【請求項 47】

固体基板上の異なる物理的位置に固定された異なるヌクレオチド分子を含むアレイであって、該ヌクレオチド分子が標的ポリヌクレオチドの少なくとも 30 の連続したヌクレオチドと特異的にハイブリダイズする最初のオリゴヌクレオチドまたはポリヌクレオチドを含み、該標的ポリヌクレオチドが請求項 12 のポリヌクレオチドであるようなアレイ。

## 【請求項 48】

前記の最初のオリゴヌクレオチドまたはポリヌクレオチド配列が該標的ポリヌクレオチドの少なくとも 30 の連続するヌクレオチドと完全に相補的であることを特徴とする、請求項 48 に記載のアレイ。

## 【請求項 49】

前記の最初のオリゴヌクレオチドまたはポリヌクレオチド配列が該標的ポリヌクレオチドの少なくとも 60 の連続するヌクレオチドと完全に相補的であることを特徴とする、請求項 48 に記載のアレイ。

## 【請求項 50】

前記の最初のオリゴヌクレオチドまたはポリヌクレオチド配列が該標的ポリヌクレオチドと完全に相補的であることを特徴とする、請求項 48 に記載のアレイ。

## 【請求項 51】

10

20

30

40

50

マイクロアレイであることを特徴とする、請求項 4 8 に記載のアレイ。

【請求項 5 2】

前記の最初のオリゴヌクレオチドまたはポリヌクレオチド配列を含むヌクレオチド分子にハイブリダイズされた該標的ポリヌクレオチドをさらに含むことを特徴とする請求項 4 8 に記載のアレイ。

【請求項 5 3】

リンカーが該ヌクレオチド分子のうちの少なくとも一つを該固体基板に結合することを特徴とする請求項 4 8 に記載のアレイ。

【請求項 5 4】

基板上の異なる各物理的位置が複数のヌクレオチド分子を含み、任意の物理的位置にあるその複数のヌクレオチド分子が同じ配列を含み、基板上の異なる各物理的位置が基板上的他の物理的位置におけるヌクレオチド分子の配列と異なる配列を有することを特徴とする、請求項 4 8 に記載のアレイ。 10

【請求項 5 5】

嗅覚および/または味覚を調節、模倣および/または阻止する化合物を同定するための方法であって、

a) 下記からなる群から選択された嗅覚および/または味覚受容体ポリペプチドと化合物とを接触させる過程と、

i) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を含むポリペプチド、 20

ii) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および

iii) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列と少なくとも 90% の同一性を有するアミノ酸配列を有する嗅覚および/または味覚受容体。

b) 化合物が該受容体ポリペプチドに特異的に結合および/または活性に影響を与えるかどうかを同定する過程とを含む方法。

【請求項 5 6】

該受容体ポリペプチドが哺乳動物細胞の表面で発現することを特徴とする、請求項 5 6 に記載の方法。

【請求項 5 7】 30

該哺乳動物細胞が G タンパク質を発現することを特徴とする、請求項 5 7 に記載の方法。

【請求項 5 8】

該哺乳動物細胞が複数の G タンパク質を発現することを特徴とする、請求項 5 8 に記載の方法。

【請求項 5 9】

該哺乳動物細胞が別の嗅覚および/または味覚受容体ポリペプチドを発現することを特徴とする、請求項 5 9 に記載の方法。

【請求項 6 0】

該受容体ポリペプチドが別のポリペプチドに融合していることを特徴とする、請求項 5 6 に記載の方法。 40

【請求項 6 1】

SEQ ID NO: 1 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 2】

SEQ ID NO: 2 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 3】

SEQ ID NO: 3 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 4】

SEQ ID NO: 4 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 5】

SEQ ID NO: 5 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。 50

## 【請求項 66】

SEQ ID NO : 6 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 67】

SEQ ID NO : 7 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 68】

SEQ ID NO : 8 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 69】

SEQ ID NO : 9 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 70】

SEQ ID NO : 10 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

10

## 【請求項 71】

SEQ ID NO : 11 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 72】

SEQ ID NO : 12 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 73】

SEQ ID NO : 13 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 74】

SEQ ID NO : 14 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 75】

SEQ ID NO : 15 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

20

## 【請求項 76】

SEQ ID NO : 16 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

## 【請求項 77】

SEQ ID NO : 17 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

## 【請求項 78】

SEQ ID NO : 18 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

## 【請求項 79】

SEQ ID NO : 19 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

30

## 【請求項 80】

SEQ ID NO : 20 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

## 【請求項 81】

SEQ ID NO : 21 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

## 【請求項 82】

SEQ ID NO : 22 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

40

## 【請求項 83】

SEQ ID NO : 23 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

## 【請求項 84】

SEQ ID NO : 24 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

## 【請求項 85】

SEQ ID NO : 25 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

## 【請求項 86】

50

SEQ ID NO : 26 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 87】

SEQ ID NO : 27 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 88】

SEQ ID NO : 28 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 89】

SEQ ID NO : 29 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。 10

【請求項 90】

SEQ ID NO : 30 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 91】

SEQ ID NO : 31 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 92】

SEQ ID NO : 32 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。 20

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、Gタンパク質共役受容体の核酸配列及びアミノ酸配列に関し、細胞増殖異常、神経障害、心血管障害、胃腸障害、自己免疫/炎症疾患及び代謝障害並びにウイルス感染の診断、治療並びに予防におけるこれらの配列の利用と、Gタンパク質共役受容体の核酸配列及びアミノ酸配列の発現における外因性化合物の効果の算定におけるこれらの配列の利用に関する。さらに、本発明は味覚または嗅覚の調節に関与する分子を同定するために特定のGタンパク質共役受容体を使用することに関する。

【0002】

(発明の背景)

シグナル伝達は、それによって細胞が細胞外シグナルに応答するような一般的プロセスである。原形質膜でのシグナル伝達は、ホルモン、神経伝達物質または成長因子などのシグナル分子が細胞膜受容体に結合することから開始される。このように活性化された受容体が細胞内の生化学的カスケードを誘発し、このカスケードは転写因子などの細胞内標的分子の活性化で終る。シグナル伝達のこのプロセスは、細胞増殖、分化及び遺伝子転写を含む全てのタイプの細胞機能を制御する。同定されている遺伝子の最も大きいファミリーの1つによってコードされるGタンパク質共役受容体(GPCR)は、原形質膜での細胞外シグナルの伝達において中心的役割を果たす。GPCRが治療の標的として成功であることは歴史的に証明されている。 30

【0003】

GPCRは、7つの疎水性膜貫通ドメインを有することを特徴とする膜内在性タンパク質であり、それらのドメインがまとまって逆平行アルファ( )ヘリックスの束を形成している。GPCRのサイズは、400以下から1000以上のアミノ酸に及ぶ(Strosberg, A. D. (1991) Eur. J. Biochem. 196:1-10、Coughlin, S. R. (1994) Curr. Opin. Cell Biol. 6:191-197)。GPCRのアミノ末端は、細胞外にあり、長さが可変で、多くの場合グリコシル化される。カルボキシル末端は、細胞質内にあり、通常はリン酸化される。細胞外ループは、細胞内ループと交互に現れ、膜貫通ドメインに連結している。第2及び第3細胞外ループを結びつけるシステインジスルフィド架橋は、アゴニスト及び 50

アンタゴニストと相互に作用し得る。GPCRの最も保存されたドメインは、膜貫通ドメイン及び最初の2つの細胞質ループである。膜貫通ドメインは、受容体の構造的及び機能的特徴をある程度明らかにする。大抵の場合、ヘリックスの束がリガンド結合ポケットを形成する。細胞外N末端セグメント、または3つの細胞外ループの内の1つまたは複数のループが、リガンド結合にも関与し得る。リガンド結合は、受容体の細胞内部分で構造変化を誘導することにより受容体を活性化する。次に、この活性化された受容体の大きな第3の細胞内ループがヘテロ3量体であるグアニンヌクレオチド結合Gタンパク質複合体と相互作用し、このGタンパク質複合体が更に細胞内のシグナル伝達作用を仲介する。この細胞内のシグナル伝達作用には、サイクリックAMP(cAMP)、ホスホリパーゼC、およびイノシトール三リン酸などのセカンドメッセンジャーの活性化、並びに活性化されたGPCRとイオンチャネルタンパク質との相互作用が含まれる(Watson, S. および S. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, 2-6頁; Bolander, F. F. (1994) Molecular Endocrinology, Academic Press, San Diego CA, 162-176頁; Baldwin, J. M. (1994) Curr. Opin. Cell Biol. 6: 180-190を参照)。

10

#### 【0004】

GPCRには、知覚性シグナルメディエータ(例えば光及び嗅覚刺激性分子)の受容体、アデノシン、アミノ酪酸(GABA)、肝細胞成長因子、メラノコルチン、ニューロペプチドY、オピオイドペプチド、オプシン、ソマトスタチン、タキキニン、血管作用性腸管ポリペプチドファミリー及びバソプレシン、生体アミン(例えばドーパミン、エピネフリン及びノルエピネフリン、ヒスタミン、グルタミン酸(向代謝性作用)、アセチルコリン(ムスカリン様作用)及びセロトニン)、ケモカイン、炎症の脂質メディエータ(例えばプロスタグランジン及びプロスタノイド、血小板活性化因子及びロイコトリエン)及びペプチドホルモン(例えばボンベシン、ブラジキニン、カルシトニン、C5aアナフィラトキシン、エンドセリン、卵胞刺激ホルモン(FSH)、ゴナドトロピン放出ホルモン(GnRH)、ニューロキニン及び甲状腺刺激ホルモン放出ホルモン(TRH)及びオキシトシン)がある。刺激がまだ同定されていない受容体として作用するGPCRは、オーファン受容体として知られている。

20

30

#### 【0005】

GPCRファミリーの多様性は、選択的スプライシングによって更に増加する。多くのGPCR遺伝子にはイントロンが含まれ、スプライシング変異体が同定されている受容体は現在では30を超えている。変異の数が最も大きいのは、タンパク質C末端においてである。N末端及び細胞質ループの変異体もまた高頻度であるのに対して、細胞外ループまたは膜貫通ドメインでの変異体は頻度が低い。変異が発生し得るような部位を2ヶ所以上有する受容体もある。スプライシング変異体は、分布、シグナル伝達、結合、制御及びリガンド結合の特徴に観察される差異に基づき、機能的に異なるようである(Kilpatrick, G. J. ら(1999) Trends Pharmacol. Sci. 20: 294-301)。

40

#### 【0006】

GPCRは、3つの主要なサブファミリー即ちロドプシン様、セクレチン様及び代謝調節型グルタミン酸受容体サブファミリーに分けることができる。GPCRサブファミリーのメンバーは、同様の機能及び特徴的な膜7回貫通構造を共有しているが、アミノ酸配列は互いに異なる。最大のファミリーはロドプシン様GPCRを有し、ロドプシン様GPCRはホルモン、神経伝達物質及び光を含む多様な細胞外シグナルを伝送する。ロドプシンは、動物の網膜に見られる感光性GPCRである。脊椎動物では、ロドプシン分子は光受容体(杆体)細胞に見られる膜性スタックに包埋されている。各ロドプシン分子はcGMPレベルの低下を誘発し、これにより原形質膜ナトリウムチャネルが閉鎖することにより1光子の光に反応する。この方法で、可視シグナルが神経インパルスに変換される。その他

50

のロドプシン様GPCRは、神経伝達物質への応答に直接関与する。このようなGPCRには、アドレナリンに対する受容体（アドレナリン受容体）、アセチルコリンに対する受容体（ムスカリン様受容体）、アデノシンに対する受容体、ガラニンに対する受容体及びグルタミン酸に対する受容体（N-メチル-D-アスパラギン酸/NMDA受容体）がある（Watson, S. 及び S. Arkininstall (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego CA, 7-9頁, 19-22頁, 32-35頁, 130-131頁, 214-216頁, 221-222頁、Habert-Ortoli, E. ら (1994) Proc. Natl. Acad. Sci. USA 91:9780-9783に概説されている）。

10

## 【0007】

ガラニン受容体は、インスリン、アセチルコリン、セロトニン及びノルアドレナリンの分泌を阻害する神経内分泌ペプチドガラニンの活性を仲介し、プロラクチン及び成長ホルモンの放出を刺激する。ガラニン受容体は、摂食障害、痛み、うつ病及びアルツハイマー病に関与している（Kask, K. ら (1997) Life Sci. 60:1523-1533）。その他の神経系ロドプシン様GPCRには、発達及び神経病理学において役割を持つと考えられるリゾホスファチジン酸、およびその他のリゾリン脂質に対する受容体など、増大する数のファミリーが含まれる（Chun, J. ら (1999) Cell Biochem. Biophys. 30:213-242）。

## 【0008】

GPCRの最大のサブファミリーである嗅覚受容体もまた、ロドプシン様GPCRファミリーのメンバーである。この受容体は、嗅覚シグナルの伝達により機能する。異なる臭気を区別するためには、多数の異なる嗅覚受容体が必要である。各嗅覚知覚ニューロンは1種類の嗅覚受容体しか発現せず、特有の受容体を発現するニューロンが、鼻の経路中の異なる位置を占める。例えば、ラット脳ライブラリから単離したRA1c受容体は、脳の非常に特有の領域及び嗅覚上皮の特定の区域だけに発現が限定されていることを示していた（Raining, K. ら (1998) Receptors Channels 6:141-151）。しかしながら、嗅覚様受容体の発現は嗅覚組織に限定されない。例えば、典型的なGPCR特性を有する嗅覚様受容体をコードする3つのラット遺伝子は、味覚及び嗅覚組織のみならず男性生殖組織においても発現パターンを示す（Thomas, M. B. ら (1996) Gene 178:1-5）。

20

30

## 【0009】

セクレチン様GPCRサブファミリーのメンバーは、セクレチン、カルシトニン、グルカゴン、成長ホルモン放出ホルモン、副甲状腺ホルモン及び血管作用性小腸ペプチドなどのペプチドホルモンをリガンドとして有する。例えば、セクレチン受容体は、膵臓及び小腸で酵素及びイオンの分泌を刺激するペプチドホルモンであるセクレチンに対応する（前出のWatson, 278-283頁）。セクレチン受容体は、長さが約450アミノ酸であり、胃腸細胞の原形質膜に見られる。セクレチンがその受容体に結合することにより、cAMPの産出が誘発される。

## 【0010】

炎症及び免疫反応に結びつけられるセクレチン様GPCRの例には、EGFモジュール含有ムチン様ホルモン受容体（Emr1）及びCD97受容体タンパク質がある。これらのGPCRは、最近特徴付けられたEGF-TM7受容体サブファミリーのメンバーである。これら膜7回貫通ホルモン受容体は、in vivoでヘテロ二量体として存在し、3から7の潜在的カルシウム結合EGF様モチーフを含む。CD97は、主に白血球で発現し、活性化されたB及びT細胞上で顕著に上方制御される（McKnight, A. J. および S. Gordon (1998) J. Leukoc. Biol. 63:271-280）。

40

## 【0011】

第3GPCRサブファミリーは、代謝調節型グルタミン酸受容体ファミリーである。グル

50

タミン酸は、中枢神経系の主要な興奮性神経伝達物質である。代謝調節型グルタミン酸受容体は、細胞内エフェクターの活性を調節し、長期の増強作用に関与する（前出のWatson, p. 130）。 $Ca^{2+}$  検出受容体は、カルシウムイオンの細胞外濃度の変化を検出するものであり、カルシウム結合に関与し得る酸性アミノ酸のクラスターを含む大きな細胞外ドメインを有する。代謝調節型グルタミン酸受容体ファミリーには、フェロモン受容体、 $GABA_B$  受容体及び味覚受容体もある。

【0012】

その他のGPCRのサブファミリーには、線虫(*Caenorhabditis elegans* 及び *Caenorhabditis briggsae*)に見られる化学受容体遺伝子の2つの群があり、これらは哺乳動物の嗅覚受容体遺伝子に遠く関係している。細胞膜上の接合因子への反応に関与している酵母フェロモン受容体STE2及びSTE3は、細胞性粘菌(*Dictyostelium discideum*)で個々の細胞の集合を調整し、多くの発達の制御遺伝子の発現を制御すると考えられているcAMP受容体と同様に、膜7回貫通シグネチャを有する。

10

【0013】

GPCR突然変異は、機能または恒常的活性化の減少を招き得るものであり、多数のヒト疾患に関係している（前出のCoughlin）。例えば、色素性網膜炎はロドプシン遺伝子の突然変異から発生し得る。更に、甲状腺刺激ホルモン受容体の体細胞活性化突然変異は、機能亢進甲状腺アデノーマの原因となることが報告されており、恒常的活性化に感受性が高い或るGPCRが癌原遺伝子のように振舞い得ることを示唆している（Parma, J. ら (1993) *Nature* 365: 649 - 651）。以下のリガンド即ち黄体形成ホルモン（思春期早発症）、バソプレシン $V_2$ （X連鎖の腎原発性糖尿病）、グルカゴン（糖尿病及び高血圧）、カルシウム（副甲状腺機能亢進症、低カルシウム尿症（hypocalcemia）、高カルシウム血症）、副甲状腺ホルモン（短肢小人症）、 $\beta_3$ -アドレナリン受容体（肥満症、非インスリン依存型糖尿病）、成長ホルモン放出ホルモン（小人症）及び副腎皮質刺激ホルモン（糖質コルチコイド欠乏症）に対するGPCR受容体には、ヒトの疾患に関連する突然変異も含まれる（Wilson, S. ら (1998) *Br. J. Pharmacol.* 125: 1387 - 1392、Stadel, J.M. ら (1997) *Trends Pharmacol. Sci.* 18: 430 - 437）。GPCRは、抑うつ症、分裂病、不眠症、高血圧、不安、ストレス、腎不全その他幾つかの心血管障害にも関与している（Horn, F. and G. Vriend (1998) *J. Mol. Med.* 76: 464 - 468）。

20

30

【0014】

更に、過去20年間にGPCRの活性化及び阻害を誘発する数百の新薬が認められてきている。これらの薬剤の治療標的は、癌、骨粗鬆症、子宮内膜症のみならず心血管障害、胃腸障害、中枢神経系疾患を含めた幅広い疾病及び疾患に及ぶ（前出のWilson、同Stadel）。例えば、ドーパミンアゴニストのLドーパはパーキンソン病の治療に用いられ、ドーパミンアンタゴニストは精神分裂病及びハンチントン病の初期段階の治療に用いられる。アドレナリン受容体のアゴニスト及びアンタゴニストは、喘息、高血圧その他の心血管障害及び不安の治療に用いられてきた。ムスカリン様アゴニストは、緑内障及び頻脈の治療に用いられ、セロトニン $5HT1D$ アンタゴニストは片頭痛に対して用いられ、ヒスタミン $H1$ アンタゴニストはアレルギー性及びアナフィラキシー性反応、枯草熱、痒み及び動揺病に対して用いられる（前出のHorn）。

40

【0015】

最近の調査では、糖尿病、肥満症及び骨粗鬆症を含む代謝障害の治療においてGPCRを将来的に使用する可能性を示唆している。例えば、腎原発性糖尿病の原因となる突然変異体 $V2$ バソプレシン受容体は、突然変異を含む領域に及ぶC末端 $V2$ 受容体ペプチドの同時発現により*in vitro*で機能的に救出され得る。これは疾病治療の新規な手段の可能性を示唆する（Schoneberg, T. ら (1996) *EMBO J.* 15: 1283 - 1291）。メラノコルチン4受容体(MC4R)における突然変異は、ヒ

50

トの体重調節及び肥満症に結びつけられる。バソプレシンV2受容体の突然変異体と同様に、これらMC4Rの突然変異体は原形質膜への輸送中に欠陥があるので(Ho, G. and R. G. MacKenzie (1999) J. Biol. Chem. 274:35816-35822)、同様の手段で治療し得る。副甲状腺ホルモン(PTH)のための1型受容体は、血流中のカルシウムホメオスタシスのPTH依存型制御を仲介するGPCRである。PTH/受容体の相互作用の研究は、骨粗鬆症の治療のための新規なPTH受容体リガンドの開発を可能にし得る(Mannstadt, M.ら(1999) Am. J. Physiol. 277:F665-F675)。

#### 【0016】

GPCRのケモカイン受容体グループは、炎症及び感染症の治療に役立つ可能性を有する(レビューは、Locati, M. and P. M. Murphy (1999) Annu. Rev. Med. 50:425-440を参照のこと)。ケモカインは、白血球輸送、造血及び血管形成の制御において細胞内シグナルとして作用する小ポリペプチドである。マウスにおける種々のケモカイン受容体の標的破壊は、病的炎症において及び多発性硬化症などの自己免疫疾患においてこれらの受容体が役割を果たしていることを示す。ヘルペスウイルス及びヒト免疫不全症ウイルス(HIV-1)を含む病原菌、によっても、ケモカイン受容体は感染を促進するために利用される。ケモカイン受容体CCR5の短縮変種は、HIV-1によるT細胞の感染に対する補助受容体として作用するものであり、AIDSに対する抵抗力を生じさせ、CCR5のアンタゴニストがAIDSの発達予防に有用たり得ることを示唆する。

#### 【0017】

一部のGPCRは味覚および嗅覚に関与することが報告されている。ヒトその他の真核生物の感覚受容体の数多くの完全または部分配列が現在知られている。(Pilpel, Y. および D. Lancet (1999) Protein Sci. 8:969-977; Mombaerts, P. (1999) Annu. Rev. Neurosci. 22:487-509等を参照。また、EP 867508A2; US 5,874,243; WO 92/17585; WO 95/18140; WO 97/17444; および WO 99/67282等の特許を参照。)ヒトゲノムは多様な嗅覚受容体をコードする約1000個の遺伝子を含んでいると報告されている(Rouquier, S. 他(1998) Nat. Genet. 18:243-250; Trask, B. J. 他(1998) Hum. Mol. Genet. 7:2007-2020)。

#### 【0018】

新たなGタンパク質共役受容体及びそれをコードするポリヌクレオチドの発見は、細胞増殖異常、神経障害、心血管障害、胃腸障害、自己免疫/炎症疾患及び代謝障害並びにウイルス性感染症の診断、治療並びに予防と、Gタンパク質共役受容体の核酸配列及びアミノ酸配列の発現における外因性化合物の効果の算定において有用である新たな組成を提供することにより、当分野における要求を満たす。

#### 【0019】

(発明の要約)

本発明は、総称して「GCREC」、個別にはそれぞれ「GCREC-1」、「GCREC-2」、「GCREC-3」、「GCREC-4」、「GCREC-5」、「GCREC-6」、「GCREC-7」、「GCREC-8」、「GCREC-9」、「GCREC-10」、「GCREC-11」、「GCREC-12」、「GCREC-13」、「GCREC-14」、および「GCREC-15」、「GCREC-16」と呼ぶGタンパク質共役受容体である精製されたポリペプチドを提供する。或る実施態様において本発明は、(a) SEQ ID NO: 1-16からなる群から選択したアミノ酸配列からなるポリペプチド、(b) SEQ ID NO: 1-16からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-16からなる群から選択したアミノ酸配列を有するポリペプチドの生物

学的活性断片、または (d) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択した単離されたポリペプチドを提供する。一実施態様では、SEQ ID NO: 1 - 16 のアミノ酸配列を含む単離されたポリペプチドを提供する。

【0020】

また、本発明は (a) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列の生物学的活性断片、または (d) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列の免疫原性断片を含む群から選択したポリペプチドをコードするような単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドはSEQ ID NO: 1 - 16 からなる群から選択したポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO: 17 - 32 からなる群から選択される。

10

【0021】

本発明はさらに、嗅覚および/または味覚に關与するGタンパク質共役受容体を提供する。また、本発明は該Gタンパク質共役受容体をコードするポリヌクレオチド配列を提供する。

【0022】

本発明は更に、(a) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列と少なくとも90%の相同性のある天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または (d) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を有する群から選択したポリペプチドをコードするようなポリヌクレオチドと機能的に連結したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明は組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

20

【0023】

また、本発明は、(a) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および (d) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を製造する方法を提供する。製造方法は、(a) 組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、(b) そのように発現したポリペプチドを回収する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に連結したプロモーター配列を有する。

30

40

【0024】

本発明は更に、(a) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および (d) SEQ ID NO: 1 - 16 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択されたポリペプチドに特異結合するような単離された抗体を提供する。

【0025】

本発明は更に、(a) SEQ ID NO: 17 - 32 からなる群から選択したポリヌクレ

50

オチド配列を含むポリヌクレオチド、(b) SEQ ID NO: 17-32 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c)(a)に相補的なポリヌクレオチド、(d)(b)に相補的なポリヌクレオチド、および(e)(a)~(d)のRNA等価物からなる群から選択された単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドは少なくとも60の連続したヌクレオチドを有する。

【0026】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは(a) SEQ ID NO: 17-32 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b) SEQ ID NO: 17-32 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c)(a)に相補的なポリヌクレオチド、(d)(b)に相補的なポリヌクレオチド、および(e)(a)~(d)のRNA等価物からなる群から選択されたポリヌクレオチド配列を含む。検出方法は、(a)サンプル中の標的ポリヌクレオチドに相補的な配列からなる少なくとも20の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズする過程と、(b)ハイブリダイゼーション複合体の存在・不存在を検出し、複合体が存在する場合にはオプションでその量を検出する過程からなり、プローブと標的ポリヌクレオチドまたはその断片の間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは標的ポリヌクレオチドに特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチドを含む。

【0027】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO: 17-32 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b) SEQ ID NO: 17-32 からなる群から選択したポリヌクレオチド配列と少なくとも90%の相同性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c)(a)のポリヌクレオチドに相補的なポリヌクレオチド、(d)(b)のポリヌクレオチドに相補的なポリヌクレオチド、および(e)(a)~(d)のRNA等価物からなる群から選択されたポリヌクレオチド配列を有する。検出方法は、(a)ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b)標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

【0028】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む成分を提供する。有効量のポリペプチドは、(a) SEQ ID NO: 1-16 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1-16 からなる群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1-16 からなる群から選択したアミノ酸配列の生物学的活性断片、または(d) SEQ ID NO: 1-16 からなる群から選択したアミノ酸配列の免疫原性断片を含む群から選択する。一実施例では、SEQ ID NO: 1-16 からなる群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的GCRCの発現の低下に関連した疾患やその症状の治療方法を提供する。

【0029】

本発明はまた、(a) SEQ ID NO: 1-16 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1-16 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1-16 を有する群から選択したアミノ酸配列の生物学的活性断片、および(d) SEQ ID NO: 1-16 からなる群から選択したアミノ酸配列

の免疫原性断片からなる群から選択されたポリペプチドのアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)ポリペプチドを有するサンプルを化合物に曝す過程と、(b)サンプル中のアゴニスト活性を検出する過程とを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。一実施態様では、本発明は機能性GCRECの発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して成分を投与する過程を含む方法を提供する。

**【0030】**

本発明は更に、(a)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列を有するポリペプチド、(b)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を有するポリペプチド、(c)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列の生物学的活性断片、または(d)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列の免疫原性断片を含む群から選択されたポリペプチドのアнтаゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)ポリペプチドを含むサンプルを化合物に曝露する過程と、(b)サンプル中のアゴニスト活性を検出する過程とを含む。一実施態様で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む成分を提供する。別の実施態様では、機能性GCRECの過剰発現に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して成分を投与する過程を含む方法を提供する。

10

20

**【0031】**

本発明は更に、(a)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列の生物学的活性断片、または(d)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列の免疫原性断片を含む群から選択されたポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)ポリペプチドを適切な条件下で少なくとも1つの試験化合物に混合させる過程と、(b)試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

30

**【0032】**

本発明は更に、(a)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列を有するポリペプチド、(b)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を有するポリペプチド、(c)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列の生物学的活性断片、または(d)SEQ ID NO: 1-16からなる群から選択したアミノ酸配列の免疫原性断片を含む群から選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)ポリペプチドの活性が許容された条件下で、ポリペプチドを少なくとも1つの試験化合物と混合させる過程と、(b)ポリペプチドの活性を試験化合物の存在下で算定する過程と、(c)試験化合物の存在下でのポリペプチドの活性を試験化合物の不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物であることを意味する。

40

**【0033】**

本発明はさらに、嗅覚および/または味覚に関与する本発明のGタンパク質共役受容体、その生物学的に活性な断片(受容体活性を持つものを含む)、およびそれと少なくとも90%の配列同一性を持つアミノ酸配列を使って前記の受容体ポリペプチドのアゴニストまたはアンタゴニストとなる化合物を同定する方法を提供する。これらの化合物は特定の味および/または臭いの調節、阻止および/または模倣に有用である。

**【0034】**

50

本発明はまた、本発明の嗅覚および/または味覚受容体、その生物学的に活性な断片(受容体活性を持つものを含む)、およびそれと少なくとも90%の配列同一性を持つアミノ酸配列と一つ以上のその他の嗅覚および/または味覚受容体ポリペプチドを使って特定の味および/または臭いを調節、模倣および/または阻止する一つまたは複数の化合物を同定することに関する。

【0035】

本発明はまた、本発明の嗅覚または味覚受容体ポリペプチドを発現する細胞およびそうした細胞を使った細胞ベースのスクリーニングによって特定の味および/または臭いを調節、模倣および/または阻止する分子を同定することに関する。

【0036】

さらに、本発明は少なくとも本発明の一つの嗅覚または味覚Gタンパク質共役受容体ポリペプチド、一つのGタンパク質、およびオプションで一つ以上のその他の嗅覚および/または味覚Gタンパク質共役受容体ポリペプチドを同時に発現する細胞、およびそうした細胞を使った細胞ベースのスクリーニングによって特定の味および/または臭いを調節、模倣および/または阻止する分子を同定することに関する。

【0037】

更に本発明は、SEQ ID NO: 17 - 32 からなる群から選択されたポリヌクレオチド配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、(a)この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(b)この標的ポリヌクレオチドの発現の変化を検出するステップとを含む、スクリーニング方法を提供する。

【0038】

本発明は更に、(a)核酸を含む生物学的サンプルを試験化合物で処理する過程と、(b)(i)SEQ ID NO: 17 - 32 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii)SEQ ID NO: 17 - 32 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii)(i)に相補的な配列を有するポリヌクレオチド、(iv)(ii)に相補的なポリヌクレオチド、(v)(i)~(iv)のRNA等価物からなる群から選択したポリヌクレオチドの少なくとも20の連続したヌクレオチドを含むプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程とを含む試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、上記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、上記標的ポリヌクレオチドは、(i)SEQ ID NO: 17 - 32 からなる群から選択したポリヌクレオチド配列を有するポリヌクレオチド、(ii)SEQ ID NO: 17 - 32 からなる群から選択したポリヌクレオチド配列と少なくとも90%の相同性を有する天然のポリヌクレオチド配列を有するポリヌクレオチド、(iii)(i)のポリヌクレオチドに相補的なポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v)(i)~(iv)のRNA等価物を含む群から選択する。或いは、標的ポリヌクレオチドは、上記(i)~(v)を含む群から選択したポリヌクレオチド配列の断片と、(c)ハイブリダイゼーション複合体の量を定量する過程と、(d)処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量の差は、試験化合物の毒性を示す。

【0039】

【発明の記載について】

本発明のタンパク質、ヌクレオチド配列及び方法について説明するが、その前に、説明した特定の装置、材料及び方法に本発明が限定されるものではなく、改変し得ることを理解されたい。また、ここで使用する専門用語は特定の実施例を説明する目的で用いたものに過ぎず、特許請求の範囲にのみ限定される本発明の範囲を限定することを意図したもので

10

20

30

40

50

はないことも併せて理解されたい。

【0040】

本明細書及び請求の範囲において単数形を表す「或る」、「その(この等)」は、文脈で明確に示していない場合は複数形を含むことに注意されたい。従って、例えば「或る宿主細胞」は複数の宿主細胞を含み、その「抗体」は複数の抗体が含まれ、当業者には周知の等価物なども含まれる。

【0041】

本明細書中で用いる全ての専門用語及び科学用語は、特に定義されている場合を除き、当業者に一般に理解されている意味と同じ意味を有する。本明細書で説明するものと類似あるいは同等の任意の装置、材料及び方法を用いて本発明の実施または試験を行うことができるが、ここでは好適な装置、材料、方法について説明する。本発明で言及する全ての刊行物は、刊行物中で報告されていて且つ本発明に関係があるであろう細胞、プロトコル、試薬及びベクターについて説明及び開示する目的で引用しているものである。本明細書のいかなる開示内容も、本発明が先行技術の効力によってこのような開示に対して先行する権利を与えられていないことを認めるものではない。

【0042】

(定義)

用語「GCREC」は、天然、合成、半合成或いは組換え体など全ての種(特にウシ、ヒツジ、ブタ、マウス、ウマ及びヒトを含む哺乳動物)から得られる実質的に精製されたGCRECのアミノ酸配列を指す。

【0043】

用語「アゴニスト」は、GCRECの生物学的活性を強めたり、模倣する分子を指す。アゴニストの例として、タンパク質、核酸、糖質、小分子その他の任意の化合物や成分を挙げることができるが、これらはGCRECと直接相互作用することによって、或いはGCRECが関与する生物学的経路の構成エレメントに作用することによって、GCRECの活性を調節する。

【0044】

用語「対立遺伝子変異配列」は、GCRECをコードする遺伝子の別の形を指す。対立遺伝子変異体は、核酸配列における少なくとも1つの突然変異から作製し得る。また、変異RNAまたはポリペプチドからも作製し得る。ポリペプチドの構造または機能は、変異することもしないこともある。遺伝子は、天然の対立遺伝子変異体を全く有しないか、1個若しくは数個の天然の対立遺伝子変異体を有し得る。一般に対立遺伝子変異体を生じさせる通常の突然変異性変化は、ヌクレオチドの自然欠失、付加または置換に帰するものである。これら各変化は、単独或いは他の変化と共に、所定の配列内で1回若しくは数回生じ得る。

【0045】

GCRECをコードする「変異」核酸配列は、様々なヌクレオチドの欠失、挿入、或いは置換が起こっても、GCRECと同じポリペプチド或いはGCRECの機能特性の少なくとも1つを備えるポリペプチドを指す。この定義に含まれるのは、GCRECをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能或いは検出困難な多型現象と、GCRECをコードするポリヌクレオチド配列のための正常な染色体遺伝子座以外の遺伝子座に占めるような、対立遺伝子変異体に対する不適切或いは予期しないハイブリダイゼーションである。コードされたタンパク質も変異され得、サイレント変化を生じGCRECと機能的に等価となるアミノ酸残基の欠失、挿入、或いは置換を含み得る。計画的アミノ酸置換は、GCRECの生物学的または免疫学的活性が保持される限りにおいて、残基の極性、電荷、溶解度、疎水性、親水性、及び/または両親媒性特性の類似性に基づき行い得る。例えば、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似している非荷電極性側鎖を有するアミノ酸には、アスパラギンとグルタミン、セリンとスレオニンがある。親水性値が近似している非荷電側鎖を有するアミノ酸には、ロイ

10

20

30

40

50

シンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

【0046】

用語「アミノ酸」及び「アミノ酸配列」は、オリゴペプチド、ペプチド、ポリペプチド、タンパク質配列、或いはそれらの任意の断片を指し、天然の分子及び合成分子を含む。「アミノ酸配列」が天然のタンパク質分子の配列を指す場合、「アミノ酸配列」及び類似の用語は、アミノ酸配列を記載したタンパク質分子に関連する完全で元のままのアミノ酸配列に限定するものではない。

【0047】

用語「増幅」は、核酸配列の複製物を作製することに関連する。増幅は通常、当業者によく知られたポリメラーゼ連鎖反応（PCR）技術を用いて行う。 10

【0048】

用語「アンタゴニスト」は、GCRECの生物学的活性を阻害或いは減弱する分子である。アンタゴニストとしては、抗体などのタンパク質、核酸、糖質、小分子またはその他の任意の化合物や成分を挙げることができるが、これらはGCRECと直接相互作用することによって、或いはGCRECが関与する生物学的経路の構成エレメントに作用することによって、GCRECの活性を調節する。

【0049】

「抗体」の語は、抗原決定基と結合することができる、そのままの免疫グロブリンやその断片、例えばFa、F(ab')<sub>2</sub>及びFv断片を指す。GCRECポリペプチドを結合する抗体は、免疫抗原として、そのままのポリペプチド、または関心のある小ペプチドを含む断片を用いて作製可能である。動物（マウス、ラット、ウサギ等）を免疫化するために用いるポリペプチドまたはオリゴペプチドは、RNAの翻訳、または化学合成によって得られるポリペプチドまたはオリゴペプチドに由来し得るもので、好みに応じて担体タンパク質に抱合することも可能である。通常用いられる担体であってペプチドと化学結合するものは、ウシ血清アルブミン、サイログロブリン及びスカシガイのヘモシアニン（KLH）等がある。その結合ペプチドは、動物を免疫化するために用いる。 20

【0050】

用語「抗原決定基」は、特定の抗体と接触する分子の領域（即ちエピトープ）を指す。タンパク質またはタンパク質断片を用いて宿主動物を免疫化する場合、タンパク質の多数の領域が、抗原決定基（タンパク質の特定の領域または3次元構造）に特異結合する抗体の産生を誘導し得る。抗原決定基は、抗体に結合するための無損傷抗原（即ち免疫応答を誘導するために用いられる免疫原）と競合し得る。 30

【0051】

「aptamer」の語は特定の分子標的に結合する核酸またはオリゴヌクレオチド分子を指す。aptamerは*in vitro*の進化プロセスから誘導される（たとえば米国特許番号5,270,163記載のSELEX（Systematic Evolution of Ligands by Exponential Enrichment））。このプロセスでは大きな組換えライブラリから標的に特異的なaptamerの配列が選択される。aptamerの組成は二本鎖または一本鎖で、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導體またはその他のヌクレオチド様分子を含むことがある。aptamerのヌクレオチド成分の糖類を修飾することにより（たとえばリボヌクレオチドの2'-OH基を2'-Fまたは2'-NH<sub>2</sub>で置換）、核酸分解酵素への耐性を改善したり、血液中における寿命を延ばすなど、望ましい改良を加えることができる。循環系からaptamerが除去される速度を遅くするために、aptamerを高分子量担体等の分子に抱合させることができる。aptamerは、たとえば光活性化または架橋剤によって各々のリガンドと特異的に架橋させることができる。（Brody, E.N. および L. Gold (2000) J. Biotechnol. 74: 5-13等を参照。） 40

「intramer」の語は*in vivo*で発現されるaptamerを意味する。た 50

例えば、*vaccinia* ウィルスに基づく RNA 発現系を使って、白血球の細胞質で特定の RNA aptamer が高レベルで発現されている (Blind, M. 他 (1999) Proc. Natl Acad. Sci. USA 96:3606-3610)。

【0052】

「*spiegelmer*」の語は L-DNA、L-RNA その他の左旋性ヌクレオチド誘導体またはヌクレオチド様分子を含む aptamer を指す。左旋性のヌクレオチドを含む aptamer は右旋性ヌクレオチドを含む基質に対して通常作用する天然の酵素による分解に対して耐性がある。

【0053】

本明細書において「アンチセンス」は、特定の核酸配列のセンス(コーディング)鎖と塩基対を形成し得る任意の組成物を指す。アンチセンス成分には、DNA や、RNA や、ペプチド核酸(PNA)や、ホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸等の修飾されたバックボーン連鎖を有するオリゴヌクレオチドや、2'-メトキシエチル糖または2'-メトキシエトキシ糖等の修飾された糖類を有するオリゴヌクレオチドや、或いは5-メチルシトシン、2-デオキシウラシルまたは7-デアザ-2'-デオキシグアノシン等の修飾された塩基を有するオリゴヌクレオチドがある。アンチセンス分子は、化学合成または転写を含む任意の方法で製造することができる。相補的アンチセンス分子は、ひとたび細胞に導入されたら、細胞が形成した天然の核酸配列と塩基対を形成し、転写または翻訳を妨害する二重鎖を形成する。「負」または「マイナス」という表現は、ある参考DNA分子のアンチセンス鎖を意味し、「正」または「プラス」という表現は同センス鎖を意味する。

10

20

【0054】

用語「生物学的に活性」は、天然分子の構造的、調節的、或いは生化学的な機能を有するタンパク質を指す。同様に、用語「免疫学的に活性」または「免疫原性」は、天然或いは組換え体のGCRC、合成のGCRCまたはそれらの任意のオリゴペプチドが、適当な動物或いは細胞の特定の免疫応答を誘発して特定の抗体と結合する能力を指す。

【0055】

用語「相補的」は、塩基対合によってアニールする2つの一本鎖核酸配列間の関係を指す。例えば、配列「5' A - G - T 3'」は、相補配列「3' T - C - A 5'」と対を形成する。

30

【0056】

「所定のポリヌクレオチド配列を含む組成物」または「所定のアミノ酸配列を含む組成物」は広い意味で、所定のポリヌクレオチド配列若しくはアミノ酸配列を含む任意の組成物を指す。この成分には、乾燥製剤または水溶液が含まれ得る。GCRC若しくはGCRCの断片をコードするポリヌクレオチド配列を含む組成物は、ハイブリダイゼーションプローブとして使用され得る。このプローブは、凍結乾燥状態で保存可能であり、糖質などの安定化剤と結合させることが可能である。ハイブリダイゼーションにおいては、塩(例えばNaCl)、界面活性剤(例えばドデシル硫酸ナトリウム; SDS)及びその他の構成成分(例えばデンハート液、脱脂粉乳、サケの精子のDNA等)を含む水溶液中にプローブを分散させることができる。

40

【0057】

「コンセンサス配列」は、不要な塩基を分離するためにDNA配列の解析を繰り返し行い、XL-PCRキット(PE Biosystems, Foster City CA)を用いて5'及び/または3'の方向に伸長され、再度シーケンシングされた核酸配列、またはGELVIEW断片構築システム(GCG, Madison, WI)またはPhrap(University of Washington, Seattle WA)等の断片構築用のコンピュータプログラムを用いて1つ或いはそれ以上の重複するcDNAやEST、またはゲノムDNA断片から構築された核酸配列を指す。伸長及び構築の両方を行ってコンセンサス配列を決定する配列もある。

50

## 【 0 0 5 8 】

用語「保存的なアミノ酸置換」は、元のタンパク質の特性を殆ど変えない置換を指す。即ち、置換によってそのタンパク質の構造や機能が大きくは変わらず、そのタンパク質の構造、特にその機能が保存される。下表は、タンパク質中で元のアミノ酸と置換され得るアミノ酸と、保存アミノ酸置換と認められるアミノ酸を示している。

元の残基	保存的な置換	
Ala	Gly, Set	
Arg	His, Lys	10
Asn	Asp, Gln, His	
Asp	Asn, Glu	
Cys	Ala, Ser	
Gln	Asn, Glu, His	
Glu	Asp, Gln, His	
Gly	Ala	20
His	Asn, Arg, Gln, Glu	
Ile	Leu, Val	
Leu	Ile, Val	
Lys	Arg, Gln, Glu	
Met	Leu, Ile	
Phe	His, Met, Leu, Trp, Tyr	
Ser	Cys, Thr	30
Thr	Ser, Val	
Trp	Phe, Tyr	
Tyr	His, Phe, Trp	
Val	Ile, Leu, Thr	

## 【 0 0 5 9 】

保存アミノ酸置換では通常、( a ) 置換領域におけるポリペプチドのバックボーン構造、例えば シートや 螺旋構造、( b ) 置換部位における分子の電荷または疎水性、及び / または ( c ) 側鎖の大部分を保持する。

## 【 0 0 6 0 】

用語「欠失」は、1 個以上のアミノ酸残基が欠如するアミノ酸配列の変化、或いは 1 個以上のヌクレオチドが欠如する核酸配列の変化を指す。

## 【 0 0 6 1 】

用語「誘導体」は、化学修飾されたポリヌクレオチドまたはポリペプチドを指す。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチド配列の化学修飾に含まれ得る。ポリヌクレオチド誘導体は、天然分子の生物学的

40

50

または免疫学的機能を少なくとも1つは保持しているポリペプチドをコードする。ポリペプチド誘導体は、グリコシル化、ポリエチレングリコール化 (pegylation)、或いは任意の同様なプロセスであって、誘導起源のポリペプチドの少なくとも1つの生物学的若しくは免疫学的機能を保持するプロセスによって、修飾されたポリペプチドである。

【0062】

「検出可能な標識」は、測定可能な信号を発生し得る、ポリヌクレオチドやポリペプチドに共有結合或いは非共有結合するレポーター分子や酵素を指す。

【0063】

「示差発現」は少なくとも2つの異なったサンプルを比較することによって決められる、増加 (上方調節)、あるいは減少 (下方調節)、または欠損遺伝子またはタンパク発現の欠損を指す。このような比較は例えば、治療後サンプルと未治療のサンプルまたは病態のサンプルと正常サンプルの間で行われ得る。

10

【0064】

「エキソンシャフリング」は、異なるコード領域 (エキソン) の組換えを意味する。1つのエキソンがコードされたタンパク質の1つの構造的または機能的ドメインを代表し得るため、安定した基礎構造の新規な再分類を介して新しいタンパク質が組立てられることが可能であり、新しいタンパク質機能の進化を促進できる。

【0065】

用語「断片」は、G C R E CまたはG C R E Cをコードするポリヌクレオチドの固有の部分であって、その親配列 (parent sequence) と同一であるがその配列より長さが短いものを指す。断片は、画定された配列の全長から1ヌクレオチド/アミノ酸残基を差し引いた長さ以下の長さを有し得る。例えば或る断片は、5 ~ 1000の連続したヌクレオチドまたはアミノ酸残基を有し得る。プローブ、プライマー、抗原、治療用分子として、或いはその他の目的のために用いられる断片は、少なくとも5、10、15、20、25、30、40、50、60、75、100、150、250若しくは500の連続したヌクレオチド或いはアミノ酸残基長さであり得る。断片は、分子の特定領域から優先的に選択し得る。例えば、ポリペプチド断片は、所定の配列に示すような最初の250または500アミノ酸 (またはポリペプチドの最初の25%または50%) から選択された或る長さの連続したアミノ酸を有し得る。これらの長さは明らかに例として挙げているものであり、本発明の実施例では、配列表、表及び図面を含む明細書に裏付けされた任意の長さであってよい。

20

30

【0066】

SEQ ID NO: 17 - 32の断片には、固有のポリヌクレオチド配列領域が含まれる。この領域は、SEQ ID NO: 17 - 32を特異的に同定するものであり、例えばこの断片を得たゲノム中のSEQ ID NO: 4 - 6以外の配列とは異なるものである。SEQ ID NO: 17 - 32のある断片は、例えば、ハイブリダイゼーションや増幅技術、またはSEQ ID NO: 17 - 32を関連ポリヌクレオチド配列から区別する類似の方法に有用である。SEQ ID NO: 17 - 32の断片の正確な長さ及び断片に対応するSEQ ID NO: 17 - 32の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

40

【0067】

SEQ ID NO: 1 - 16の断片は、SEQ ID NO: 17 - 32の断片によってコードされる。SEQ ID NO: 1 - 16の断片には、SEQ ID NO: 1 - 16を特異的に同定する固有のアミノ酸配列領域が含まれている。例えば、SEQ ID NO: 1 - 16の断片は、SEQ ID NO: 1 - 16を特異認識する抗体を産出するための免疫原性ペプチドとして有用である。SEQ ID NO: 1 - 16の断片及び断片に対応するSEQ ID NO: 1 - 16の領域の正確な長さは、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0068】

50

「完全長」ポリヌクレオチド配列とは、少なくとも1つの翻訳開始コドン（例えばメチオニン）、オープンリーディングフレーム及び翻訳終止コドンを有する配列である。「完全長」ポリヌクレオチド配列は、「完全長」ポリペプチド配列をコードする。

【0069】

「相同性」の語は、2つ以上のポリヌクレオチド配列または2つ以上のポリペプチド配列の配列類似性、または配列同一性を意味する。

【0070】

ポリヌクレオチド配列についての用語「一致率」または「一致性%」とは、標準化されたアルゴリズムを用いてアラインメントされる、2つ以上のポリヌクレオチド配列間の一致する残基の百分率のことである。このようなアルゴリズムは、2配列間のアラインメントを最適化するために比較する配列において、標準化された再現性のある方法でギャップを挿入するので、2つの配列をより有意に比較できる。

10

【0071】

ポリヌクレオチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる。このプログラムは、LASERGENEソフトウェアパッケージ（一組の分子生物学的分析プログラム）（DNASTAR, Madison WI）の一部である。このCLUSTAL Vは、Higgins, D.G. 及び P.M. Sharp (1989) CABIOS 5:151-153、Higgins, D.G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列を2つ1組でアラインメントする際のデフォルトパラメータは、Ktuple = 2、gap penalty = 5、window = 4、「diagonals saved」= 4と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選択された。一致率は、アラインメントされたポリヌクレオチド配列間の「類似性パーセント」としてCLUSTAL Vによって報告される。

20

【0072】

或いは、一般的に用いられ且つ自由に入手できる配列比較アルゴリズム一式が、国立バイオテクノロジー情報センター（NCBI）Basic Local Alignment Search Tool（BLAST）から提供されており（Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410）、これはメリーランド州ベセスダにあるNCBI及びインターネット（<http://www.ncbi.nlm.nih.gov/BLAST/>）を含む幾つかの情報源から入手可能である。このBLASTソフトウェア一式には、既知のポリヌクレオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールが入手可能であり、2つのヌクレオチド配列の対を直接比較するために用いられる。「BLAST 2 Sequences」は、<http://www.ncbi.nlm.nih.gov/gorf/b12.html>にアクセスして、対話形式で利用ができる。「BLAST 2 Sequences」ツールは、blastn 及び blastp（以下に記載）の両方に用いることができる。BLASTプログラムは、一般的には、ギャップ及びデフォルト設定に設定された他のパラメータと共に用いる。例えば、2つのヌクレオチド配列を比較するために、デフォルトパラメータとして設定された「BLAST 2 Sequences」ツール Version 2.0.12（2000年4月21日）を用いてblastnを実行してもよい。デフォルトパラメータの設定例を以下に示す。

30

40

【0073】

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 及び Extension Gap : ペナルティ 2

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

10

#### 【0074】

一致率は、ある定義された配列の全長（例えば特定のSEQ IDナンバーで定義された配列）で測定し得る。或いは、より短い長さ、例えば、定義された、より大きな配列から得られた断片（例えば少なくとも20、30、40、50、70、100または200の連続したヌクレオチドの断片）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、一致率を測定し得る長さを説明し得ることを理解されたい。

20

#### 【0075】

高度の相同性を示さない核酸配列が、それにもかかわらず遺伝子コードの縮重が原因で類似のアミノ酸配列をコードする場合がある。この縮重を利用して核酸配列内で変化を生じさせて、全ての核酸配列が実質上同一のタンパク質をコードするような多数の核酸配列を生成し得ることを理解されたい。

#### 【0076】

ポリペプチド配列に用いられる用語「一致率」または「一致性%」とは、標準化されたアルゴリズムを用いてアラインメントされる2つ以上のポリペプチド配列間の一致する残基の百分率のことである。ポリペプチド配列アラインメントの方法は公知である。保存的アミノ酸置換を考慮するアラインメント方法もある。既に詳述したこのような保存的置換は通常、置換部位の酸性度及び疎水性を保存するので、ポリペプチドの構造を（従って機能も）保存する。

30

#### 【0077】

ポリペプチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組み込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる（既に説明したのでそれを参照されたい）。CLUSTAL Vを用いて、ポリペプチド配列を2つ1組でアラインメントする際のデフォルトパラメータは、Ktupl e = 1、gap penalt y = 3、window = 5、「diagonals saved」= 5と設定する。デフォルトの残基重み付け表としてPAM250マトリクスを選択する。ポリヌクレオチドアラインメントと同様に、アラインメントされたポリペプチド配列の対の同一性のパーセントは、「類似性パーセント」としてCLUSTAL Vによって報告される。

40

#### 【0078】

或いは、NCBI BLASTソフトウェア一式を用いてもよい。例えば、2つのポリペプチド配列を対で比較をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (Apr - 21 - 2000)でblastpを使用するであろう。デフォルトパラメータの設定例を以下に

50

示す。

【 0 0 7 9 】

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap : 1 penalty

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

10

【 0 0 8 0 】

一致率は、ある定義された配列の全長（例えば特定の SEQ ID ナンバーで定義された配列）で測定し得る。或いは、より短い長さ、例えば、定義された、より大きな配列から得られた断片（例えば少なくとも 15、20、30、40、50、70、または 150 の連続したヌクレオチドの断片）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、或る長さであってその長さに対して一致率を測定し得る長さを説明し得ることを理解されたい。

20

【 0 0 8 1 】

「ヒト人工染色体（HAC）」は、約 6 kb（キロベース）～ 10 Mb のサイズの DNA 配列を含み得る、安定した染色体複製の分離及び維持に必要な全てのエレメントを含む直鎖状の小染色体である。

【 0 0 8 2 】

用語「ヒト化抗体」は、もとの結合能力を保持しつつよりヒトの抗体に似せるために、非抗原結合領域のアミノ酸配列が変えられた抗体分子を指す。

30

【 0 0 8 3 】

「ハイブリダイゼーション」とは、所定のハイブリダイゼーション条件下で、ある一本鎖ポリヌクレオチドがある相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。特異的ハイブリダイゼーションは、2つの核酸配列が高い相同性を共有することを示すものである。特異的ハイブリダイゼーション複合体は許容されるアニーリング条件下で形成され、「洗浄」ステップ後もハイブリダイズされたままである。洗浄ステップは、ハイブリダイゼーションプロセスのストリンジェンシーを決定する際に特に重要であり、更にストリンジェントな条件では、非特異結合（即ち完全には一致しない核酸鎖間の対の結合）が減少する。核酸配列のアニーリングに対する許容条件は、本技術分野における当業者が慣例的に決定できる。許容条件はハイブリダイゼーション実験の間は一定でよいが、洗浄条件は所望のストリンジェンシーを得るように、従ってハイブリダイゼーション特異性も得るように実験中に変更することができる。アニーリングが許容される条件は、例えば、温度が 68 で、約 6 × SSC、約 1% (w/v) の SDS、並びに約 100 µg/ml のせん断して変性したサケ精子 DNA が含まれる。

40

【 0 0 8 4 】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度及び pH における特異配列の融点 (T<sub>m</sub>) より約 5 ~ 20 低くなるように選択する。この T<sub>m</sub> は、所定のイオン強度及び pH の条件下で、完全に一致するプローブに標的配列の

50

50%がハイブリダイズする温度である。T<sub>m</sub>を計算する式及び核酸のハイブリダイゼーション条件はよく知られており、Sambrook他(1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NYに記載されており、特に2巻の9章を参照されたい。

【0085】

本発明のポリヌクレオチド間の高いストリンジェンシーのハイブリダイゼーションでは、約0.2 x SSC及び約1%のSDSの存在の下、約68℃で1時間の洗浄過程を含む。別法では、65℃、60℃、55℃、42℃の温度で行う。SSC濃度は、約0.1%のSDS存在下で、約0.1~2 x SSCの範囲で変化し得る。通常は、ブロッキング剤を用いて非特異ハイブリダイゼーションを阻止する。このようなブロッキング剤には、例えば、約100~200 µg/mlの変性サケ精子DNAがある。例えばRNAとDNAのハイブリダイゼーションのような特定条件下では、有機溶剤、例えば約35~50% v/vの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド及びヌクレオチドにコードされるポリペプチドに対する類似の役割を強く示唆している。

【0086】

用語「ハイブリダイゼーション複合体」は、相補的な塩基間の水素結合によって、形成された2つの核酸配列の複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る(C<sub>0</sub>tまたはR<sub>0</sub>t解析等)。或いは、一方の核酸配列が溶解状態で存在し、もう一方の核酸配列が固体支持体(例えば紙、膜、フィルタ、チップ、ピンまたはガラススライド、或いは他の適切な基板であって細胞若しくはその核酸が固定される基板)に固定されているような2つの核酸配列間に形成され得る。

【0087】

用語「挿入」或いは「付加」は、1個以上のアミノ酸残基或いはヌクレオチドがそれぞれ追加されるアミノ酸配列或いは核酸配列の変化を指す。

【0088】

「免疫応答」は、炎症、外傷、免疫異常症、伝染性疾患または遺伝性疾患に関連する症状を指し得る。これらの症状は、細胞及び全身の防御系に作用し得る種々の因子、例えばサイトカイン、ケモカイン、その他のシグナル伝達分子の発現によって特徴づけることができる。

【0089】

「免疫原性断片」とは、哺乳動物等の生命体に導入されると免疫応答を誘発し得るようなGCRCのポリペプチドまたはオリゴペプチド断片である。「免疫原性断片」の語には、本明細書中で開示したような或いは当分野で既知であるような任意の抗体産出方法において有用なGCRCの任意のポリペプチドまたはオリゴペプチド断片も含まれる。

【0090】

用語「マイクロアレイ」は、基板上の複数のポリヌクレオチド、ポリペプチドまたはその他の化合物の構成を指す。

【0091】

用語「エレメント」または「アレイエレメント」は、マイクロアレイ上に固有の指定された位置を有する、ポリヌクレオチド、ポリペプチドまたはその他の化合物を指す。

【0092】

用語「調節」は、GCRCの活性の変化を指す。調節することによって例えば、GCRCのタンパク質活性、結合特性その他の生物学的、機能的または免疫学的特性が増大または低下し得る。

【0093】

「核酸」及び「核酸配列」の語は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチドまたはこれらの断片を指す。「核酸」及び「核酸配列」の語は、ゲノム起源または合成

10

20

30

40

50

起源のDNAまたはRNAであって一本鎖または二本鎖であるか或いはセンス鎖またはアンチセンス鎖を表し得るようなDNAまたはRNAや、ペプチド核酸(PNA)や、任意のDNA様またはRNA様物質を指すこともある。

【0094】

「機能的に連結した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に連結している。同一のリーディングフレーム内で2つのタンパク質コード領域を結合する必要がある場合、一般に、機能的に連結したDNA配列は非常に近接するか、或いは連続的に隣接し得る。

【0095】

「ペプチド核酸(PNA)」は、末端がリシンで終わるアミノ酸残基のペプチドのバックボーンに結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドを含む、アンチセンス分子または抗遺伝子剤を指す。末端のリシンは、成分に溶解性を与える。PNAは、相補的一本鎖DNAまたはRNAに優先的に結合して転写の伸長を停止するものであり、ポリエチレングリコール化して細胞におけるPNAの寿命を延長し得る。

【0096】

GCRECの「翻訳後修飾」には、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、蛋白分解性切断及びその他の当分野で既知の修飾が含まれ得る。これらのプロセスは、合成的或いは生化学的に生じ得る。生化学的修飾は、GCRECの酵素環境に依存し、細胞タイプによって変化し得る。

【0097】

「プローブ」とは、同一配列或いは対立遺伝子核酸配列、関連する核酸配列の検出に用いる、GCRECやそれらの相補配列、またはそれらの断片をコードする核酸配列のことである。プローブは、単離されたオリゴヌクレオチドまたはポリヌクレオチドであって、検出可能な標識またはレポーター分子に結合したものである。典型的な標識には、放射性アイソトープ、リガンド、化学発光試薬及び酵素がある。「プライマー」とは、相補的な塩基対を形成して標的のポリヌクレオチドにアニーリング可能な、通常はDNAオリゴヌクレオチドである短い核酸である。プライマーは次に、DNAポリメラーゼ酵素によって標的DNA鎖に沿って延長し得る。プライマー対は、例えばポリメラーゼ連鎖反応(PCR)による核酸配列の増幅(及び同定)に用い得る。

【0098】

本発明に用いるようなプローブ及びプライマーは通常、既知の配列の少なくとも15の連続したヌクレオチドを含んでいる。特異性を高めるために長めのプローブ及びプライマー、例えば開示した核酸配列の少なくとも20、25、30、40、50、60、70、80、90、100または150の連続したヌクレオチドからなるようなプローブ及びプライマーを用いてもよい。これよりもかなり長いプローブ及びプライマーもある。表、図面及び配列リストを含む本明細書に裏付けされた任意の長さのヌクレオチドを用いることができるものと理解されたい。

【0099】

プローブ及びプライマーの調製及び使用方法については、Sambrook, J. 他 (1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NY、Ausubel, F.M. 他, (1987) Current Protocols in Molecular Biology, Greene Publi. Assoc. & Wiley-Intersciences, New York NY、Innis他 (1990) PCR Protocols, A Guide to Methods and Applications Academic Press, San Diego CA等を参照されたい。PCRプライマー対は、その目的のためのコンピュータプログラム、例えばPrimer (Version 0.5, 1991, Whitehead Institute for Biomedical Research,

10

20

30

40

50

C a m b r i d g e M A ) を用いるなどして既知の配列から得ることができる。

【 0 1 0 0 】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えば O L I G O 4 . 0 6 ソフトウェアは、各 1 0 0 ヌクレオチドまでの P C R プライマー対の選択に有用であり、3 2 キロベースまでのインプットポリヌクレオチド配列からオリゴヌクレオチド及び最大 5 , 0 0 0 までの大きめのポリヌクレオチドでの分析するのにも有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組み込まれている。例えば、P r i m O U プライマー選択プログラム(テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能)は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。P r i m e r 3 プライマー選択プログラム( W h i t e h e a d I n s t i t u t e / M I T C e n t e r f o r G e n o m e R e s e a r c h (マサチューセッツ州ケンブリッジ)より入手可能)によって、ユーザーは、プライマー結合部位として避けたい配列を指定できる「非プライミングライブラリ( m i s p r i m i n g l i b r a r y )」を入力できる。P r i m e r 3 は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である。(後二者のプライマー選択プログラムのソースコードは、各自のソースから得てユーザー固有のニーズを満たすように変更してもよい。) P r i m e r G e n プログラム(英国ヒトゲノムマッピングプロジェクト-リソースセンター(英国ケンブリッジ)から一般向けに入手可能)は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラムは、固有であって保存されたオリゴヌクレオチド及びポリヌクレオチドの断片の同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチド及びポリヌクレオチドの断片は、ハイブリダイゼーション技術において、例えば P C R またはシーケンシングプライマーとして、マイクロアレイエレメントとして、或いは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

10

20

30

【 0 1 0 1 】

本明細書における「組換え核酸」は天然の配列ではなく、2つ以上の配列の離れたセグメントを人工的に組み合わせた配列である。この人為的組合せはしばしば化学合成によって達成するが、より一般的には核酸の単離セグメントの人為的操作によって、例えばの S a m b r o o k らの文献(前出)に記載されているような遺伝子工学的的手法によって達成する。組換え核酸の語は、単に核酸の一部を付加、置換または欠失した変異核酸も含む。しばしば組換え核酸には、プロモーター配列に機能的に連結した核酸配列が含まれる。このような組換え核酸は、例えばある細胞を形質転換するために使用されるベクターの一部とすることが可能である。

【 0 1 0 2 】

或いはこのような組換え核酸は、ウイルスベクターの一部であって、例えばワクシニアウイルスに基づくものであり得る。そのようなワクシニアウイルスは哺乳動物に接種され、その組換え核酸が発現されて、その哺乳動物ないで防御免疫応答を誘導するように使用することができる。

40

【 0 1 0 3 】

「調節エレメント」は、通常は遺伝子の非翻訳領域に由来する核酸配列であり、エンハンサー、プロモーター、イントロン及び 5 ' 及び 3 ' の非翻訳領域( U T R )を含む。調節エレメントは、転写、翻訳または R N A 安定性を調節する宿主タンパク質またはウイルスタンパク質と相互作用する。

【 0 1 0 4 】

「レポーター分子」は、核酸、アミノ酸または抗体の標識に用いられる化学的または生化

50

学的な部分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、阻害因子、磁気粒子及びその他の当分野で既知の成分がある。

【0105】

本明細書において、DNA配列に対する「RNA等価物」とは、基準となるDNA配列と同じ直鎖の核酸配列から構成されるが、窒素性塩基のチミンがウラシルに置換され、糖鎖のバックボーンがデオキシリボースではなくリボースからなる。

【0106】

用語「サンプル」は、その最も広い意味で用いられている。GCRECをコードする核酸若しくはその断片、またはGCREC自体を含む疑いのあるサンプルは、体液や、細胞から単離された細胞、染色体、細胞小器官または膜からの抽出物や、細胞や、溶解しているか基質に結合しているゲノムDNA、RNAまたはcDNAや、組織や、組織プリント等から構成され得る。

10

【0107】

用語「特異的結合」及び「特異的に結合する」は、タンパク質若しくはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物との間の相互作用を指す。この相互作用は、タンパク質の特定の構造（例えば抗原決定基即ちエピトープ）であって結合分子が認識するものが存在するか否かに依存していることを意味している。例えば、抗体がエピトープ「A」に対して特異的である場合、結合していない標識した「A」及び抗体を含む反応液に、エピトープAを含むポリペプチド或いは結合していない無標識の「A」が存在すると、抗体と結合する標識Aの量が減少する。

20

【0108】

用語「実質的に精製された」は、自然の環境から取り除かれてから、単離或いは分離された核酸配列或いはアミノ酸配列であって、自然に結合している組成物が少なくとも約60%除去されたものであり、好ましくは約75%以上の除去、最も好ましくは90%以上除去されたものを指す。

【0109】

「置換」とは、一つ以上のアミノ酸またはヌクレオチドをそれぞれ別のアミノ酸またはヌクレオチドに置き換えることである。

【0110】

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルター、チップ、スライド、ウエハ、ファイバー、磁気または非磁気ビード、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。基板は、壁、溝、ピン、チャネル、孔等、様々な表面形態を有することができ、基板表面にはポリヌクレオチドやポリペプチドが結合する。

30

【0111】

「転写イメージ」は、所定条件下での所定時間における特定の細胞の種類または組織による集合的遺伝子発現のパターンを指す。

【0112】

「形質転換」とは、外来DNAが受容細胞に導入されるプロセスのことである。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核宿主細胞または真核宿主細胞に挿入する任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。限定するものではないが形質転換方法には、ウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクション及び微粒子銃を用いる方法がある。「形質転換された」細胞には、導入されたDNAが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一時的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

40

【0113】

ここで用いる「遺伝形質転換体」とは任意の生物体であり、限定するものではないが動植物を含み、生物体の1個若しくは数個の細胞が、ヒトの関与によって、例えば本技術分野

50

でよく知られている形質転換技術によって導入された異種核酸を有する。核酸の細胞への導入は、直接または間接的に、細胞の前駆物質に導入することによって、計画的な遺伝子操作によって、例えば微量注射法によって或いは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種或いは *in vitro* 受精を指すものではなく、組換え DNA 分子の導入を指すものである。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌及び動植物がある。本発明の単離された DNA は、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合によって宿主に導入することができる。本発明の DNA をこのような生物体に移入する技術はよく知られており、前出の Sambrook ら (1989) 等の参考文献に与えられている。

10

## 【0114】

特定の核酸配列の「変異体」は、核酸配列 1 本全部の長さに対して特定の核酸配列と少なくとも 40% の相同性を有する核酸配列であると定義する。その際、デフォルトパラメータに設定した「BLAST 2 Sequences」ツール Version 2.0.9 (1999年5月7日) を用いて *blastn* を実行する。このような核酸対は、所定の長さに対して、例えば少なくとも 50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% またはそれ以上の相同性を示し得る。或る変異体は、例えば「対立遺伝子」変異体 (前述)、「スプライス」変異体、「種」変異体または「多型性」変異体として説明し得る。スプライス変異体は参照分子とかなりの相同性を有し得るが、mRNA プロセッシング中のエキソンの異なるスプライシングによって通常、ポリヌクレオチドがより多くまたはより少数の塩基を有することになる。対応するポリペプチドは、追加機能ドメインを有するか或いは参照分子に存在するドメインが欠落していることがある。種変異体は、種によって異なるポリヌクレオチド配列である。結果的に生じるポリペプチドは通常、相互にかなりのアミノ酸相同性を有する。多型性変異体は、与えられた種の個体間で特定の遺伝子のポリヌクレオチド配列が異なる。また、多型性変異体は、1 つのヌクレオチド塩基によってポリヌクレオチド配列が変化する「1 塩基多型性」(SNP) を含み得る。SNP の存在は、例えば特定の個体群、病状または病状性向を示し得る。

20

## 【0115】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の 1 本の長さ全体で特定のポリペプチド配列に対して少なくとも 40% の相同性を有するポリペプチド配列として画定される。ここで、デフォルトパラメータに設定した「BLAST 2 Sequence s」ツール Version 2.0.9 (1999年5月7日) を用いて *blastp* を実行する。このようなポリペプチド対は、所定の長さに対して、例えば少なくとも 50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% またはそれ以上の配列同一性を示し得る。

30

## 【0116】

(発明)

本発明は、新規なヒト G タンパク質 (GCREC) と、GCREC をコードするポリヌクレオチドの発見と、細胞増殖異常、神経障害、心血管障害、胃腸障害、自己免疫 / 炎症疾患及び代謝障害並びにウイルス性感染症の診断、治療並びに予防におけるこれらの組成物の利用に基づくものである。

40

## 【0117】

表 1 は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の命名の概略である。各ポリヌクレオチド及びその対応するポリペプチドは、1 つの *Incyte* プロジェクト識別番号 (*Incyte* プロジェクト ID) と関連する。各ポリペプチド配列は、ポリペプチド配列識別番号 (ポリペプチド SEQ ID NO) と *Incyte* ポリペプチド配列番号 (*Incyte* ポリペプチド ID) によって表示した。各ポリヌクレオチド配列は、ポリヌクレオチド配列識別番号 (ポリヌクレオチド SEQ ID NO) と *Incyte* ポリヌクレオチドコンセンサス配列番号 (*Incyte* ポリヌクレオチド ID) によって

50

表示した。と *Incyte* ポリヌクレオチド配列番号 (*Incyte* ポリヌクレオチド ID) によって表示した。

【0118】

表2は、GenBankタンパク質 (*genpept*) データベースに対するBLAST分析によって同定されたような、本発明のポリペプチドとの相同性を有する配列を示している。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号 (ポリペプチド SEQ ID NO : ) およびそれに対応する *Incyte* ポリペプチド配列番号 (*Incyte* ポリペプチド ID) を示す。列3は、GenBankの最も近い相同体のGenBankの識別番号 (*Genbank ID NO :*) を示す。列4は、各ポリペプチドとそのGenBank相同体との間の一致を表す確率スコアを示す。列5は、GwnBank相同体のアノテーションを示し、更に該当箇所には適当な引用文も示す。

10

これらを引用することを以って本明細書の一部とする。

【0119】

表3は、本発明のポリペプチドの様々な構造的特徴を示す。列1および列2はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号 (SEQ ID NO : ) およびそれに対応する *Incyte* ポリペプチド配列番号 (*Incyte* ポリペプチド ID) を示す。列3は、各ポリペプチドのアミノ酸残基数を示す。列4および列5はそれぞれ、GCG配列分析ソフトウェアパッケージのMOTIFSプログラム (*Genetics Computer Group, Madison WI*) によって決定された、リン酸化およびグリコシル化の可能性のある部位を示す。列6は、シグネチャ配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。列7は、タンパク質の構造/機能の分析のための分析方法を示し、該当箇所にはさらに分析方法に利用した検索可能なデータベースを示す。

20

【0120】

合わせて、表2及び表3は本発明のポリペプチドの特性を概略的に示している。これらの特性は、特許請求の範囲に記載されたポリペプチドがGタンパク質共役受容体であることを証明する。例えば、SEQ ID NO : 2はラット膜7回貫通Gタンパク質共役受容体 (*GenBank ID g5525078*) と39%の同一性を有することが *Basic Local Alignment Search Tool (BLAST)* によって示された。(表2参照)。BLAST確率スコアは  $1.2e^{-89}$  であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO : 2はまた、セクレチンファミリーの膜7回貫通受容体ドメインを有する。これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。(表3参照)。BLIMPSおよびBLAST分析から得たデータは、SEQ ID NO : 2がGタンパク質共役受容体であることを裏づける証拠を更に提供する。別の例において、SEQ ID NO : 4はヒトの膜7回貫通ドメイン受容体 (*GenBank ID g2117161*) に32%の同一性を有する。(表2参照)。BLAST確率スコアは  $3.2e^{-35}$  である。SEQ ID NO : 4はまた、*latrophilin/CL-1* 様GPSSドメインおよびセクレチンファミリー膜7回貫通受容体ドメインを含むが、これはHMMに基づくPFAMデータベースの検索によって決定された。(表3参照)。BLIMPSおよびMOTIFS分析から得たデータは、SEQ ID NO : 4が膜7回貫通Gタンパク質共役受容体であることを裏づける証拠を更に提供する。別の例においてSEQ ID NO : 5は、BLASTにより決定されるマウスのオキシトシン受容体 (*GenBank ID g1902964*) に34%の同一性を有する。(表2参照)。BLAST確率スコアは  $7.1e^{-21}$  である。SEQ ID NO : 5はまた、膜7回貫通受容体ドメインを含むが、これはHMMに基づくPFAMデータベースの検索によって決定された。(表3参照)。BLIMPS分析から得たデータは、SEQ ID NO : 5がGタンパク質共役受容体であることを裏づける証拠を更に提供する。別の例において、SEQ

30

40

50

I D N O : 6 はヒトのシステニルロイコトリエン受容体 ( G e n B a n k I D g 5 3 5 9 7 1 8 ) に 2 3 % の同一性を有するが、これは B L A S T によって決定された。(表 2 参照)。B L A S T 確率スコアは  $2.5 \times 10^{-21}$  である。B L I M P S 分析から得たデータは、S E Q I D N O : 6 が G タンパク質共役受容体であることを裏づける証拠を更に提供する。別の例において、S E Q I D N O : 8 はメキシコのテトラ ( 盲目洞窟魚として知られる ) のオプシン ( G e n B a n k I D g 4 4 0 6 2 6 ) に 3 2 % の同一性を有するが、これは B L A S T によって決定された。(表 2 参照)。B L A S T 確率スコアは  $2.2 \times 10^{-22}$  である。S E Q I D N O : 8 はまた、ロドプシンファミリーの膜 7 回貫通受容体ドメインを含むが、これは H M M に基づく P F A M データベースの検索によって決定された。(表 3 参照)。B L I M P S、M O T I F S 及び P R O F I L E S C A N 分析から得たデータは、S E Q I D N O : 8 がロドプシンファミリーの G タンパク質共役受容体であることを裏づける証拠を更に提供する。別の例において、S E Q I D N O : 9 は推定上のヒトの神経伝達物質受容体 ( G e n B a n k I D g 2 4 6 5 4 3 2 ) に 4 4 % の同一性を有するが、これは B L A S T によって決定された。B L A S T 確率スコアは  $5.2 \times 10^{-69}$  である。S E Q I D N O : 9 はまた、ロドプシンファミリーの膜 7 回貫通受容体ドメインを含むが、これは H M M に基づく P F A M データベースの検索によって決定された。(表 3 参照)。B L I M P S および B L A S T 分析から得たデータは、S E Q I D N O : 9 がロドプシンファミリーの G タンパク質共役受容体であることを裏づける証拠を更に提供する。別の例において、S E Q I D N O : 1 1 は B L A S T によって同定されるマーモットの嗅覚受容体 ( G e n B a n k I D g 5 9 0 1 4 8 8 ) に 8 2 % の同一性を有する。(表 2 参照)。B L A S T 確率スコアは  $1.5 \times 10^{-101}$  である。S E Q I D N O : 1 1 はまた、ロドプシンファミリーの膜 7 回貫通受容体ドメインを含むが、これは H M M に基づく P F A M データベースの検索によって決定された。(表 3 参照)。B L I M P S、M O T I F S 及び P R O F I L E S C A N 分析から得たデータは、S E Q I D N O : 1 1 が G タンパク質共役受容体であることを裏づける証拠を更に提供する。別の例において、S E Q I D N O : 1 2 はヒトの嗅覚受容体 ( G e n B a n k I D g 2 7 9 2 0 1 8 ) に 6 4 % の同一性を有するが、これは B L A S T によって決定された。(表 2 参照)。B L A S T 確率スコアは  $8.4 \times 10^{-99}$  である。S E Q I D N O : 1 2 はまた、ロドプシンファミリーの膜 7 回貫通受容体ドメインを含むが、これは H M M に基づく P F A M データベースの検索によって決定された。(表 3 参照)。B L I M P S、M O T I F S 及び P R O F I L E S C A N 分析から得たデータは、S E Q I D N O : 1 2 が G タンパク質共役受容体であることを裏づける証拠を更に提供する。別の例において、S E Q I D N O : 1 5 はニワトリの嗅覚受容体 4 ( G e n B a n k I D g 1 2 4 6 5 3 4 ) に 5 7 % の同一性を有するが、これは B L A S T によって決定された。(表 2 参照)。B L A S T 確率スコアは  $1.1 \times 10^{-91}$  である。S E Q I D N O : 1 5 はまた、膜 7 回貫通受容体ドメイン ( ロドプシンファミリー ) を含むが、これは H M M に基づく P F A M データベースの検索によって決定された。(表 3 参照)。B L I M P S、M O T I F S、及び P R O F I L E S C A N 解析よりのデータは、S E Q I D N O : 1 5 が嗅覚受容体である、さらに実証的な証拠を提供する。S E Q I D N O : 1、S E Q I D N O : 3、S E Q I D N O : 7、S E Q I D N O : 1 0、S E Q I D N O : 1 3 - 1 4 および S E Q I D N O : 1 6 については、同様の方法で分析し、注釈を付けた。S E Q I D N O : 1 - 1 6 の解析のためのアルゴリズム及びパラメータが表 7 で記述されている。

#### 【 0 1 2 1 】

表 4 に示すように、本発明の完全長ポリヌクレオチド配列は、c D N A 配列またはゲノム D N A 由来のコード ( エキソン ) 配列を用いて、或いはこれら 2 種類の配列を任意に組み合わせ構築した。列 1 および列 2 はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号 ( ポリヌクレオチド S E Q I D N O : ) およびそれに対応する I n c y t e ポリヌクレオチドコンセンサス配列番号 ( I n c y t e ポリヌクレオチド I D ) を示す。列 3 は、塩基対における各ポリヌクレオチド配列の長さを示す。列 4 は

、例えば、SEQ ID NO: 17-32を同定するため、或いはSEQ ID NO: 17-32と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列5はcDNA配列、ゲノムDNAから予想されたコード配列(エキソン)及び/またはcDNA及びゲノムDNAを共に有する配列集合に対応する識別番号を示している。これらの配列は、本発明の完全長ポリヌクレオチド配列を構築するのに用いた。表4の列6および列7はそれぞれ、列5の配列に対応するcDNA配列およびゲノム配列の開始ヌクレオチド(5')位置および終了ヌクレオチド(3')位置を示す。

#### 【0122】

表4の列5の識別番号は、特に例えばIncyte cDNAとそれに対応するcDNAライブラリを指す場合もある。例えば、2536292F6はIncyte cDNA配列の識別番号であり、BRAINOT18はそれが由来するcDNAライブラリの識別番号である。cDNAライブラリが示されていないIncyte cDNAは、プールされているcDNAライブラリ(例えば、72051732V1)に由来する。または、列5の識別番号は、完全長ポリヌクレオチド配列の組み立てに用いたGenBankのcDNAすなわちEST(例えばg3738039)の識別番号の場合もある。さらに、列5の識別番号は、ENSEMBL(The Sanger Centre、(英国ケンブリッジ))データベースから由来した配列を同定し得る(例えば、「ENST」命名を含む配列)。或いは、列5の識別番号は、NCBI RefSeq Nucleotide Sequence Records データベースから由来する場合もあり(即ち「NM」または「NT」の命名を含む配列)、またNCBI RefSeq Protein Sequence Recordsから由来する場合もある(即ち「NP」の命名を含む配列)。または列5の識別番号は、「エキソンスティック(exon-stitching)」アルゴリズムにより結び合わせたcDNA及びGenscan予想エキソンの両方からなる群を意味する場合がある(実施例5を参照)。例えば、FL\_XXXXXX\_N<sub>1</sub>\_N<sub>2</sub>\_YYYYY\_N<sub>3</sub>\_N<sub>4</sub>は「ステッチされた」配列であり、その内、XXXXXXはアルゴリズムが適用される配列のクラスターの識別番号であり、YYYYYはアルゴリズムにより生み出される予測の数であり、N<sub>1</sub>, N<sub>2</sub>, N<sub>3</sub>, ...がある場合には、解析中に手で編集された特定のエキソンを表す(実施例5参照)。または、列5の識別番号は「エキソストレッチ(exon-stretching)」アルゴリズムにより結び合わせたエキソンの集合を指す場合もある。例えば、FLXXXXXX\_gAAAAA\_gBBBBB\_1\_Nは「ストレッチ」配列の識別番号である。ここでXXXXXはIncyteプロジェクト識別番号、gAAAAAは「エキソストレッチ」アルゴリズムを適用したヒトゲノム配列のGenBank識別番号、gBBBBBは一番近いGenBankタンパク質相同体のGenBank識別番号、またはNCBI RefSeq識別番号、Nは特定のエキソンである(実施例5を参照)。あるRefSeq配列が「エキソストレッチ」アルゴリズムのためのタンパク質相同体として使用された場合には、RefSeq識別番号(「NM」、「NP」、または「NT」によって表される)が、GenBank識別(即ち、gBBBBB)の代わりに使用される場合もある。

#### 【0123】

或いは、接頭コードは、手で編集された構成配列、ゲノムDNA配列から予測された構成配列、または組み合わされた配列解析方法から由来する構成配列を同定する。以下の表はコンポーネント配列の接頭コードおよびその接頭コードに対応する配列解析法の例を示している(実施例4および実施例5を参照。)

10

20

30

40

接頭コード	解析タイプやプログラムの例
GNN, GFG, ENST	例えば、GENSCAN (Stanford University, CA, USA) または FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)を用いたゲノム配列からのエクソン予測
GBI	手動で編集されたゲノム配列の解析
FL	スティッチまたはストレッチゲノム配列(実施例5参照)
INCY	EST 配列のゲノムへのマッピングからの全長転写とエクソンの予想エクソンと転写を予想するために、ゲノム位置と EST 構成データが組み合わされる。

10

20

## 【0124】

場合によっては、最終コンセンサスポリヌクレオチド配列を確認するための列5に示すような配列の適用範囲と重複する Incyte cDNA の適用範囲が得られたが、関連する Incyte cDNA 識別番号は示さなかった。

30

## 【0125】

表5は、Incyte cDNA 配列を用いて構築された完全長ポリヌクレオチド配列のための代表的な cDNA ライブラリを示している。代表的な cDNA ライブラリは、上記のポリヌクレオチド配列を構築及び確認するために用いられる Incyte cDNA 配列によって、最も頻繁に代表される Incyte cDNA ライブラリである。cDNA ライブラリを作製するために用いた組織及びベクターを表5に示し、表6で説明している。

## 【0126】

本発明には、GCRECの変異体も含まれる。好適なGCRECの変異体のアミノ酸配列は、GCRECアミノ酸配列と少なくとも約80%、または少なくとも約90%、或いは少なくとも約95%もの一致率を有し、GCRECの機能的若しくは構造的特徴を少なくとも1つ有するような変異体である。

40

## 【0127】

本発明には、GCRECをコードするポリヌクレオチドも含まれる。或る例では、GCRECをコードするSEQ ID NO: 17-32からなる群から選択された配列を有するポリヌクレオチド配列が本発明に含まれている。SEQ ID NO: 17-32のポリヌクレオチド配列は、配列表に示されているように等価RNA配列と同等の価値を有しているが、窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくてリボースから構成されている。

50

## 【0128】

本発明には、G C R E Cをコードするポリヌクレオチド配列の変異配列も含まれる。具体的には、そのようなポリヌクレオチド配列の変異配列は、G C R E Cをコードするポリヌクレオチド配列と少なくとも約70%、或いは少なくとも約85%、または少なくとも約95%もの一致率を有する。本発明の特定の実施形態は、S E Q I D N O : 1 7 - 3 2 からなる一群から選択された核酸配列と少なくとも70%のポリヌクレオチド配列同一性、或いは少なくとも85%のポリヌクレオチド配列同一性、更には少なくとも95%ものポリヌクレオチド配列同一性を有するS E Q I D N O : 1 7 - 3 2 からなる群から選択された配列を含むポリヌクレオチド配列の変異配列を提供する。上記したポリヌクレオチド変異配列は何れも、G C R E Cの機能的或いは構造的特徴の少なくとも1つを有するアミノ酸配列をコードする。 10

## 【0129】

当業者であれば、遺伝暗号の縮重の結果、G C R E Cをコードする多数のポリヌクレオチド配列（既知の遺伝子または天然の遺伝子のポリヌクレオチド配列と最低限の類似性しか有しないポリヌクレオチド配列もある）を産出し得ることは理解できよう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得るようなありとあらゆる可能性のあるポリヌクレオチド配列変異体を網羅し得る。これらの組合せは、天然のG C R E Cのポリヌクレオチド配列に適用されるような標準トリプレット遺伝暗号を基に作られるものであり、このような変異は全て明確に開示されているものと考えられる。 20

## 【0130】

G C R E C及びその変異体をコードするヌクレオチド配列は通常、好適に選択されたストリンジェントな条件下で天然のG C R E Cのヌクレオチド配列とハイブリダイズ可能であるが、G C R E Cまたはその誘導体であって実質上異なるコドンの使用法があるもの、例えば天然に存在しないコドンの封入があるものをコードするヌクレオチド配列を作り出すことは有益であろう。宿主が特定のコドンを利用する頻度に基づいて、特定の真核宿主又は原核宿主に発生するペプチドの発現率を高めるようにコドンを選択することが可能である。コードされたアミノ酸配列を変えなくG C R E C及びその誘導体をコードするヌクレオチド配列を実質上変更する別の理由には、天然の配列から作製される転写物より好ましい、例えば半減期が長いなどの特性を有するRNA転写物の作製がある。 30

## 【0131】

本発明には、G C R E C、G C R E C誘導体及びこれらの断片をコードするDNA配列又はそれらの断片を完全に合成化学によって作製することも含まれる。作製後、当分野でよく知られている試薬を用いて、この合成配列を任意の様々な入手可能な発現ベクター及び細胞系中に挿入し得る。更に、合成化学を用いてG C R E Cまたはその任意の断片をコードする配列に突然変異を誘導し得る。 30

## 【0132】

更に本発明には、種々のストリンジェントな条件下で、請求項に記載されたポリヌクレオチド配列、特に、S E Q I D N O : 1 7 - 3 2 及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、W a h l , G . M . 及びS . L . B e r g e r ( 1 9 8 7 ) M e t h o d s E n z y m o l . 1 5 2 : 3 9 9 - 4 0 7、K i m m e l , A . R . ( 1 9 8 7 ) M e t h o d s E n z y m o l . 1 5 2 : 5 0 7 - 5 1 1等を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。 40

## 【0133】

DNAシーケンシングの方法は当分野では公知であり、本発明のいずれの実施例もDNAシーケンシング方法を用いて実施可能である。DNAシーケンシング方法には酵素を用いることができ、例えばDNAポリメラーゼIのクレノウ断片、S E Q U E N A S E ( U S B i o c h e m i c a l , C l e v e l a n d O H )、T a qポリメラーゼ ( A p p l i e d B i o s y s t e m s )、熱安定性T7ポリメラーゼ ( A m e r s h a m , P h a r m a c i a B i o t e c h , P i s c a t a w a y N J )を用いること 50

ができる。或いは、例えばELONGASE増幅システム(Life Technologies, Gaithersburg MD)において見られるように、ポリメラーゼと校正エキソヌクレアーゼを併用することができる。好適には、MICROLAB 2200液体転移システム(Hamilton, Reno, NV)、PTC 200サーマルサイクラー(MJ Research, Watertown MA)及びABI CATALYST 800サーマルサイクラー(Applied Biosystems)等の装置を用いて配列の準備を自動化する。次に、ABI 373 或いは 377 DNAシーケンシングシステム(Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム(Molecular Dynamics, Sunnyvale CA)または当分野でよく知られている他の方法を用いてシーケンシングを行う。結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する。(Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7、Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, 856 - 853頁等を参照)。

#### 【0134】

GRECをコードする核酸配列を、部分的ヌクレオチド配列を利用し且つ当分野でよく知られているPCR法をベースにした種々の方法を用いて伸長させ、プロモーター及び調節要素等の上流配列を検出することができる。例えば、使用し得る方法の一つである制限部位PCR法は、ユニバーサルプライマー及びネステッドプライマーを用いてクローニングベクター内のゲノムDNAから未知の配列を増幅する方法である(例えば、Sarkar, G. (1993) PCR Methods Applic. 2: 318322を参照。)別の方法に逆PCR法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知の配列を増幅する方法である。鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限酵素断片から得る(例えば、Triglia, T. 他 (1988) Nucleic Acids Res. 16: 8186。)第3の方法としてキャプチャPCR法があり、これはヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片をPCR増幅する方法に参与している。(Lagerstrom, M.ら(1991) PCR Methods Applic 1: 111-119等を参照) 30  
この方法では、PCRを行う前に複数の制限酵素の消化及びライゲーション反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。また、未知の配列を検索するために用い得る別の方法については当分野で知られている(Parker, J.D. 他 (1991) Nucleic Acids Res. 19: 3055-3060等を参照)。更に、PCR、ネステッドプライマー及びPromoter Finder (商標)ライブラリ(Clontech, Palo Alto CA)を用いてゲノムDNAをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を見付けるのに有用である。全てのPCRベースの方法に対して、市販されているソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア(National Biosciences, Plymouth 40  
MN) 或いは別の好適なプログラムを用いて、長さが約22~30ヌクレオチド、GC含有率が約50%以上、温度約68~72で鋳型に対してアニーリングするようにプライマーを設計し得る。

#### 【0135】

完全長cDNAをスクリーニングする際は、より大きなcDNAを含むようにサイズ選択されたライブラリを用いるのが好ましい。更に、ランダムプライマーのライブラリは、しばしば遺伝子の5'領域を有する配列を含み、オリゴd(T)ライブラリが完全長cDNAを作製できない状況に対して好適である。ゲノムライブラリは、5'非転写調節領域への配列の伸長に有用であろう。

#### 【0136】

市販のキャピラリー電気泳動システムを用いて、シーケンシングまたはPCR産物のサイズを分析し、またはそのヌクレオチド配列を確認することができる。具体的には、キャピラリーシーケンシングは、電気泳動による分離のための流動性ポリマーと、4つの異なるヌクレオチドに特異的であるような、レーザーで活性化される蛍光色素と、発光された波長の検出に利用するCCDカメラとを有し得る。出力/光の強度は、適切なソフトウェア(Applied Biosystems社のGENOTYPER、SEQUENCE NAVIGATOR等)を用いて電気信号に変換し得る。サンプルのロードからコンピュータ分析及び電子データ表示までの全プロセスがコンピュータ制御可能である。キャピラリー電気泳動法は、特定のサンプルに少量しか存在しないようなDNA小断片のシーケンシングに特に適している。

10

## 【0137】

本発明の別の実施例では、GCRECをコードするポリヌクレオチド配列またはその断片を、適切な宿主細胞内でGCREC、GCRECの断片またはその機能的等価物を発現させるような組換えDNA分子内でクローニングし得る。遺伝暗号固有の宿重に起因して、実質的同一或いは機能的等価のアミノ酸配列をコードするような別のDNA配列を作製してGCRECの発現に利用し得る。

## 【0138】

種々の目的でGCRECがコードする配列を変えるために、本発明のヌクレオチド配列を当分野で通常知られている方法を用いて組み換えることができる。ここで目的には、限定するものではないが遺伝子産物のクローニング、プロセッシング、発現の調節がある。遺伝子断片及び合成オリゴヌクレオチドのランダムなフラグメンテーション及びPCR再アセンブリによるDNAシャッフリングを用い、ヌクレオチド配列を組み換えることが可能である。例えば、オリゴヌクレオチドを介した部位特異的変異誘導を利用して、新規な制限部位の作製、グリコシル化パターンの変更、コドン優先の変更、スプライス変異体の生成等を起こす突然変異を導入し得る。

20

## 【0139】

本発明のヌクレオチドを、MOLECULAR BREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他(1999) Nat. Biotechnol. 17:793-797; Christians, F.C. 他(1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他(1996) Nat. Biotechnol. 14:315-319)などのDNAシャッフリング技術を用いてシャッフリングして、SECPの生物学的または酵素的な活性、或いは他の分子や化合物と結合する能力などのSECPの生物学的特性を変更或いは改良することができる。DNAシャッフリングは、遺伝子断片のPCRを介する組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、所望の特性を有する遺伝子変異体を同定するための選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してDNAシャッフリング及び選択/スクリーニングを行ってもよい。従って、人工的な育種及び急速な分子の進化によって多様な遺伝子が作られる。例えば、ランダムポイント突然変異を有する単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングすることができる。或いは、所定の遺伝子を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子と組み換え、それによって天然に存在する複数の遺伝子の遺伝多様性を、指図された制御可能な方法で最大化させることができる。

30

40

## 【0140】

別の実施例によれば、当分野でよく知られている化学的方法を用いて、GCRECをコードする配列の全部或いは一部を合成し得る(Caruthers, M.H.ら(1980) Nucleic Acids Symp. Ser. 7:215-223、Horn, T.ら(1980) Nucleic Acids Symp. Ser. 7:225-232等を参照)。或いは、化学的方法を用いてGCRECそれ自体またはその断片

50

を合成し得る。例えば、種々の液相または固相技術を用いてペプチド合成を行うことができる (Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, 55-60頁、Roberge, J. Y. ら (1995) Science 269:202204等を参照)。自動合成はABI 431Aペプチドシンセサイザ (Applied Biosystems) を用いて達成し得る。更にGCRECのアミノ酸配列またはその任意の一部は、直接合成する間、または他のタンパク質から得た配列またはその任意の一部と結合する間、或いはその両方を行っている間に変更し、変異型ポリペプチドを生成することが可能である。

【0141】

ペプチドは、分離用高速液体クロマトグラフィーを用いて実質上精製可能である (Chiez, R. M. および F. Z. Regnier (1990) Methods Enzymol. 182:392-421等を参照)。合成ペプチドの組成は、アミノ酸分析またはシーケンシングによって確認することができる (前出のCreighton, 28-53頁等を参照)。

【0142】

生物学的に活性なGCRECを発現させるために、GCRECまたはその誘導体をコードするヌクレオチド配列を好適な発現ベクターに挿入することができる。好適な発現ベクターとは即ち好適な宿主内で挿入されたコーディング配列の転写及び翻訳の調節に必要な要素を含むベクターである。これらの必要な要素には、ベクター及びGCRECをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型プロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。このような要素は、長さ及び特異性が様々である。固有開始シグナルを用いて、GCRECをコードする配列をより効果的に翻訳することも可能である。このようなシグナルには、ATG開始コドンと、コザック配列などの近傍の配列が含まれる。GCRECをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームのATG開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳要素及び開始コドンは、様々な天然物及び合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である (例えば、Scharf, D. 他 (1994) Results Probl. Cell Differ. 201-18-162.を参照)。

【0143】

当業者によく知られている方法を用いて、GCRECをコードする配列と、好適な転写及び翻訳調節要素とを含む発現ベクターを構築することが可能である。これらの方法には、in vitro組換えDNA技術、合成技術、及びin vivo遺伝子組換え技術が含まれる (例えば、Sambrook, J. 他. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章及び8章、及び16-17章; Ausubel, F. M. 他. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, 9章、13章、16章を参照)。

【0144】

種々の発現ベクター/宿主系を利用して、GCRECをコードする配列を保持及び発現し得る。限定するものではないがこのような発現ベクター/宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター (例えばバキュロウイルス) に感染した昆虫細胞系や、ウイルス発現ベクター (例えばカリフラワーモザイクウイルスCaMVまたはタバコモザイクウイルスTMV) または細菌発現ベクター (例えば

10

20

30

40

50

Ti または pBR322 プラスミド) で形質転換させた植物細胞系、動物細胞系などの微生物等がある。(前出の Sambrook、前出の Ausubel、Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509、; Engelhard, E.K. ら (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. ら (1996) Hum. Gene Ther. 7:1937-1945、Takamatsu, N. (1987) EMBOJ. 6:307-311、; 『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196 ページ、Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J.J. ら (1997) Nat. Genet. 15:345-355 等を参照) レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる (Di Nicola, M. ら (1998) Cancer Gene Ther. 5(6):350-356、Yu, M. ら (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344、Buller, R.M. ら (1985) Nature 317(6040):813-815; McGregor, D.P. ら (1994) Mol. Immunol. 31(3):219-226、Verma, I.M. and N. Somia (1997) Nature 389:239-242 等を参照)。本発明は使用される宿主細胞によって限定されるものではない。

#### 【0145】

細菌系では、多数のクローニングベクター及び発現ベクターが、GCREC をコードするポリヌクレオチド配列の使用目的に応じて選択可能である。例えば、GCREC をコードするポリヌクレオチド配列の日常的なクローニング、サブクローニング及び増殖には、PBLUESCRIPT (Stratagene, La Jolla CA) または PSPORT1 プラスミド (Life Technologies) などの多機能大腸菌ベクターを用いることができる。GCREC をコードする配列の、ベクターの多数のクローニング部位への連結反応によって、lacZ 遺伝子が破壊され、組換え分子を含む形質転換された細菌を同定するための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列における in vitro 転写、ジデオキシのシークエンシング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失の生成にも有用であろう (例えば、Van Heeke, G. および S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509 を参照)。多量の GCREC が必要な場合、例えば抗体を生成する場合などには、GCREC の発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導 T5 バクテリオファージプロモーターまたは誘導 T7 バクテリオファージプロモーターを含むベクターが使用できる。

#### 【0146】

酵母の発現系を使用して GCREC を生成し得る。因子、アルコールオキシダーゼ、PGH プロモーター等の構成型或いは誘導型のプロモーターを含む多数のベクターが、出芽酵母菌 (Saccharomyces cerevisiae) またはピキア酵母 (Pichia pastoris) に使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む。(例えば、Ausubel, 1995, 前出、Bitter, G.A. ら (1987) Methods Enzymol. 153:516-544、及び Scorer, C.A. ら (1994) Bio/Technology 12:1-181-184. を参照)。

#### 【0147】

10

20

30

40

50

植物系を使用してGCRECを発現することも可能である。GCRECをコードする配列の転写は、ウイルスプロモーター、例えば、CaMV由来の35S及び19Sプロモーターを単独で、或いは、TMV(Takamatsu, N. (1987) EMBO J 6:307-311)由来のオメガリーダー配列と組み合わせて用いて促進し得る。或いは、RUBISCOの小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい(例えば、Coruzzi, G. ら (1984) EMBO J. 3:1671-1680; Broglie, R. ら (1984) Science 224:838-843; および Winter, J. ら (1991) Results Probl. Cell Differ. 17:85-105を参照)これらの構成物は、直接DNA形質転換または病原体を媒介とする形質移入によって、植物細胞内に導入可能である(『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196ページ等を参照)。

#### 【0148】

哺乳動物細胞においては、多数のウイルスベースの発現系を利用し得る。発現ベクターとしてアデノウイルスを用いる場合、GCRECをコードする配列は、後発プロモーター及び3連リーダー配列からなるアデノウイルス転写/翻訳複合体に連結され得る。可欠E1またはE3領域へウイルスのゲノムを挿入し、宿主細胞でGCRECを発現する感染ウイルスを得ることが可能である。(例えば、Logan, J. および T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659を参照)。更に、ラウス肉腫ウイルス(RSV)エンハンサー等の転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。SV40またはEBVをベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

#### 【0149】

ヒト人工染色体(HAC)を用いて、プラスミドに含まれ且つプラスミドから発現するものより大きなDNAの断片を輸送することもできる。治療のために約6kb~10MbのHACsを作製し、従来の送達方法(リポソーム、ポリカチオンアミノポリマー、またはベシクル)で供給する(例えば、Harrington, J. J. 他 (1997) Nat Genet. 15:345-355.を参照)。

#### 【0150】

長期にわたり哺乳動物系内で組換えタンパク質を生成するためには、細胞株内のGCRECの安定発現が望ましい。例えば、発現ベクターであって複製発現因子、内在性発現因子、或いはその両者のウイルス起源を含むものと、同一或いは別のベクター上の選択可能マーカー遺伝子とを用いて、GCRECをコードする配列を細胞株に形質転換することが可能である。ベクターの導入後、選択培地に移す前に強化培地で約1~2日間細胞を増殖させることができる。選択可能マーカーの目的は選択培地への抵抗性を与えることであり、選択可能マーカーが存在することにより、導入された配列をうまく発現するような細胞の成長及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

#### 【0151】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、tk<sup>-</sup>単純細胞のために用いられる単純ヘルペスウイルスチミジンキナーゼ遺伝子と、apr<sup>-</sup>細胞のために用いられる単純ヘルペスウイルスのアデニンホスホリボシルトランスフェラーゼ遺伝子がある(例えば、Wigler, M. 他 (1977) Cell 11:223-232; 及びLowy, I. 他 (1980) Cell 22:817-823を参照)。また、選択の基礎として代謝拮抗物質、抗生物質或いは除草剤への耐性を用いることができる。例えばdhfrはメトトレキサートに対する耐性を与え、neoはアミノグリコシッドネオマイシン及びG-418に対する耐性を与え、alsはクロルスルフロンに対する耐性を、patはホスフィトリシンアセチルトランスフェラーゼに対する耐性を各々与える(Wigler, M. 他 (1980)

Proc. Natl. Acad. Sci. USA 77:35673570; Colbere Garapin, F. 他 (1981) J. Mol. Biol. 150:114 等を参照。)この他の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変える *trpB* 及び *hisD* は、文献に記載されている (Hartman, S. C. および R. C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:80478051 を参照。)可視マーカー、例えばアントシアニン、緑色蛍光タンパク質 (GFP; Clontech)、グルクロニダーゼ及びその基質 グルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である (Rhodes, C. A. (1995) Methods Mol. Biol. 55:121131 等を参照)。

10

## 【0152】

マーカー遺伝子発現の存在/不存在によって目的の遺伝子の存在が示唆されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、GCRECをコードする配列をマーカー遺伝子配列内に挿入すると、GCRECをコードする配列を含む形質転換細胞は、マーカー遺伝子機能の欠落により同定することが可能である。または単一プロモーター制御下で、GCRECをコードする配列とタンデムにマーカー遺伝子を配置することも可能である。誘導または選択に応答したマーカー遺伝子の発現は通常、タンデム遺伝子の発現も示す。

20

## 【0153】

一般に、GCRECをコードする核酸配列を含み且つGCRECを発現する宿主細胞は、当業者によく知られている種々の方法を用いて同定することが可能である。限定するものではないが当業者によく知られている方法には、DNA-DNA或いはDNA-RNAハイブリダイゼーション、PCR法、核酸或いはタンパク質の検出、定量、或いはその両方を行うための膜系、溶液ベース或いはチップベースの技術を含むタンパク質バイオアッセイまたはイムノアッセイ技術がある。

## 【0154】

特異的ポリクローナル抗体または特異的モノクローナル抗体を用いてGCRECの発現の検出及び計測を行うための免疫学的方法は、当分野で公知である。このような技術の例としては、酵素に結合したイムノソルベントアッセイ (ELISA)、ラジオイムノアッセイ (RIA)、蛍光活性化細胞選別 (FACS) などが挙げられる。GCREC上の2つの非干渉エピトープに反応するモノクローナル抗体を用いた、2部位モノクローナルベースのイムノアッセイ (two-site, monoclonal-based immunoassay) が好ましいが、競合結合アッセイも用いることもできる。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である (Hampton, R. ら (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV, Coligan, J. E. 他 (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY, Pound, J. D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ 等を参照)。

30

40

## 【0155】

多岐にわたる標識方法及び抱合方法が、当業者に知られており、様々な核酸アッセイおよびアミノ酸アッセイにこれらの方法を用い得る。GCRECをコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いはPCRプローブを産出する方法には、オリゴ標識化、ニックトランスレーション、末端標識化、または標識ヌクレオチドを用いるPCR法がある。或いは、GCRECをコードする配列またはその任意の断片を、mRNAプローブを産出するためのベクターにクローニ

50

ングすることも可能である。このようなベクターは、当分野において知られており、市販もされており、T7、T3またはSP6等の好適なRNAポリメラーゼ及び標識されたヌクレオチドを加えて、in vitroでRNAプローブの合成に用いることができる。このような方法は、例えばAmersham Pharmacia Biotech、Promega (Madison WI)、U.S. Biochemical等から市販されている種々のキットを用いて実行することができる。検出を容易にするために用い得る好適なレポーター分子或いは標識には、基質、補助因子、インヒビター、磁気粒子のほか、放射性核種、酵素、蛍光剤、化学発光剤、発色剤等がある。

#### 【0156】

GCRECをコードするヌクレオチド配列を用いて形質転換した宿主細胞は、細胞培地からのタンパク質の回収及び発現に適した条件下で培養し得る。形質転換細胞から製造されたタンパク質が分泌されるか細胞内に留まるかは、使用される配列、ベクター、或いはその両者に依存する。当業者であれば理解し得るように、GCRECをコードするポリヌクレオチドを含む発現ベクターを、原核細胞膜または真核細胞膜を透過するGCRECの直接分泌を誘導するシグナル配列を含むように設計し得る。 10

#### 【0157】

更に、宿主細胞株の選択は、挿入した配列の発現を調節する能力または発現したタンパク質を所望の形に処理する能力によって行い得る。限定するものではないがこのようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化及びアシル化がある。タンパク質の「プレプロ」または「プロ」形を切断する翻訳後のプロセッシングを利用して、標的タンパク質、折りたたみ及び/または活性を特定することが可能である。翻訳後の活性のための固有の細胞装置及び特徴のある機構を有する種々の宿主細胞(例えばCHO、HeLa、MDCK、MEK293、WI38等)は、American Type Culture Collection (ATCC, Manassas, VA)から入手可能であり、外来タンパク質の正しい修飾及び処理を確実にするよう 20  
に選択し得る。

#### 【0158】

本発明の別の実施例では、GCRECをコードする天然の核酸配列、修飾核酸配列または組換え核酸配列を、上記任意の宿主系において融合タンパク質の翻訳をもたらす異種配列に連結反応させることができる。例えば、市販されている抗体を用いて認識可能な異種部分を含むキメラGCRECタンパク質は、GCREC活性阻害剤に対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分及び異種ペプチド部分も、市販されている親和性基質を用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ(GST)、マルトース結合タンパク質(MBP)、チオレドキシン(Trx)、カルモジュリン結合ペプチド(CBP)、6-His、FLAG、c-myc、赤血球凝集素(HA)がある。GSTは固定化グルタチオン上で、MBPはマルトース上で、Trxはフェニルアルシンオキシド上で、CBPはカルモジュリン上で、そして6-Hisは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。FLAG、c-myc及び赤血球凝集素(HA)は、これらのエピトープ標識を特異的に認識する市販されているモノクローナル抗体及びポリクローナル抗体を用いて、融合タンパク質の免疫親和性精製を可能にする。また、GCRECが精製後に異種部分から切断できるように、GCRECコード配列と異種タンパク質配列の間にタンパク質分解性切断部位を含めるように融合タンパク質を遺伝子操作することもできる。融合タンパク質の発現及び精製方法は、前出のAusubel (1995) 10章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現及び精製を促進することもできる。 30  
40

#### 【0159】

本発明の更に別の実施例では、TNTウサギ網状赤血球可溶化液またはコムギ胚芽抽出系(Promega)を用いて、放射能標識したGCRECの合成がin vitroで可能である。これらの系は、T7、T3またはSP6プロモーターと機能的に連結したタン 50

パク質コード配列の転写及び翻訳をカップルさせる。翻訳は、例えば<sup>35</sup>Sメチオニンのような放射能標識したアミノ酸前駆体の存在下で起こる。

【0160】

本発明のGCRECまたはその断片を用いて、GCRECに特異結合する化合物をスクリーニングすることができる。少なくとも1つまたは複数の試験化合物を用いて、GCRECへの特異的な結合をスクリーニングすることが可能である。試験化合物の例には、抗体、オリゴヌクレオチド、タンパク質（例えば受容体）または小分子が挙げられる。

【0161】

一実施例では、このように同定された化合物は、例えばリガンドやその断片などのGCRECの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している（Coligan, J. E. 他（1991）Current Protocols in Immunology 1（2）の5章等を参照）。同様に、化合物は、GCRECが結合する天然受容体、或いは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、このような化合物に対するスクリーニングには、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方としてGCRECを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョバエ、大腸菌からの細胞が含まれる。GCRECを発現する細胞またはGCRECを含有する細胞膜断片を試験化合物と接触させて、GCRECまたは化合物の何れかの結合、刺激または阻害を分析する。

【0162】

あるアッセイは、単に試験化合物をポリペプチドに実験的に結合させ、結合を、蛍光色素、放射性同位体、酵素抱合体またはその他の検出可能な標識により検出することができる。例えば、このアッセイは、少なくとも1つの試験化合物を、溶液中の或いは固体支持物に固定されたGCRECと結合させるステップと、GCRECとこの化合物との結合を検出するステップを含み得る。別法では、標識された競合物の存在下での試験化合物の結合の検出及び測定を行うことができる。更にこのアッセイでは、無細胞再構成標本、化学ライブラリまたは天然の生成混合物を用いて実施することができ、試験化合物は、溶液中で遊離させるか固体支持体に固定させる。

【0163】

本発明のGCRECまたはその断片を用いて、GCRECの活性を調整する化合物をスクリーニングすることが可能である。このような化合物には、アゴニスト、アンタゴニスト、或るいは部分的または逆アゴニスト等が含まれる。一実施例では、GCRECが少なくとも1つの試験化合物と結合する、GCRECの活性が許容される条件下でアッセイを実施し、試験化合物の存在下でのGCRECの活性が試験化合物不在下でのGCRECの活性と比較する。試験化合物の存在下でのGCRECの活性の変化は、GCRECの活性を調整する化合物の存在を示唆する。別法では、試験化合物をGCRECの活性に適した条件下でGCRECを含む*in vitro*または細胞遊離系と結合させてアッセイを実施する。これらアッセイの何れかにおいて、GCRECの活性を調整する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも1つから複数の試験化合物をスクリーニングすることができる。

【0164】

別の実施例では、胚性幹細胞（ES細胞）における相同組換えを用いて動物モデル系内で、GCRECまたはその哺乳動物相同体をコードするポリヌクレオチドを「ロックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号及び第5,767,337号等を参照）。例えば129/SvJ細胞株等のマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo: Capocchi, M. R. (1989) Science 244: 1288-1292）等のマーカー遺伝子で破壊した目的の遺伝子を含むベクターで形質転換す

る。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をノックアウトする(Martha, J. D. (1996) Clin. Invest. 97:1999-2002; Wagner, K. U. 他(1997) Nucleic Acids Res. 25:4323-4330)。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、可能性のある治療薬や毒性薬剤で検査することができる。

## 【0165】

10

GCRECをコードするポリヌクレオチドを*in vitro*でヒト胚盤胞由来のES細胞において操作することが可能である。ヒトES細胞は、内胚葉、中胚葉及び外胚葉の細胞の種類を含む少なくとも8つの別々の細胞系統に分化する可能性を有する。これらの細胞系統は、例えば神経細胞、造血系統及び心筋細胞に分化する(Thomson, J. A. 他(1998) Science 282:1145-1147)。

## 【0166】

GCRECをコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ノックイン」ヒト化動物(ブタ)または遺伝子組換え動物(マウスまたはラット)を作製することが可能である。ノックイン技術を用いて、GCRECをコードするポリヌクレオチドの或る領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、可能性のある医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばGCRECを乳汁内に分泌するなどGCRECを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る(Janne, J. 他(1998) Biotechnol. Annu. Rev. 4:55-74)。

20

## (治療)

化学的及び構造的類似性、例えば配列及びモチーフとの関連における類似性が、GCRECの領域とGタンパク質共役受容体との間に存在する。更に、GCRECの発現は、脳組織に密接に関連する。従ってGCRECは、細胞増殖異常、神経障害、心血管障害、胃腸障害、自己免疫/炎症性疾患及び代謝障害並びにウイルス感染において或る役割を果たすものと考えられる。GCRECの発現または活性の増大に関連する疾患の治療においては、GCRECの発現または活性を低下させることが望ましい。また、GCRECの発現または活性の低下に関連する疾患の治療においては、GCRECの発現または活性を増大させることが望ましい。

30

## 【0167】

従って、或る実施例において、GCRECの発現または活性の低下に関連した疾患の治療または予防のために、患者にGCRECまたはその断片や誘導体を投与することが可能である。限定するものではないがこのような疾患の例として細胞増殖異常が含まれ、その中には日光性角化症、動脈硬化症、アテローム性動脈硬化症、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症と、腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、奇形癌を含む癌、具体的には副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌等が含まれ、神経障害が含まれ、その中には癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病その他の錐体外路障害、筋萎縮性側索硬化その他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症その他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患と、クールー、クロイツフェルトヤコブ病及びガストマンストラウスラーシャインカー症候群を含むプリオン病

40

50

と、致死性家族性不眠症、神経系の栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜性血管芽腫症 (cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、精神薄弱その他の中枢神経系発達障害、脳性麻痺、神経骨格異常、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィーその他の神経筋疾患、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分障害、不安障害及び精神分裂病を含む精神障害と、季節型感情障害 (SAD) と、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、トゥレット病、進行性核上麻痺、皮質基底核変性症 (corticobasal degeneration)、家族性前頭側頭骨痴呆が含まれ、また心血管障害も含まれ、その中には動静脈瘤、アテローム性動脈硬化症、高血圧、脈管炎、レイノー病、静脈奇形、動脈解離、静脈瘤、血栓静脈炎及び静脈血栓、血管の腫瘍、血栓崩壊の合併症、バルーン血管形成術 (balloon angioplasty)、血管置換術、大動脈冠動脈バイパス術移植手術 (coronary artery bypass graft surgery) などの血管疾患と、うっ血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪状石灰化 (mitral annular calcification)、僧帽弁脱出、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性紅斑性狼瘡の心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心臓移植の合併症などの心疾患が含まれ、また胃腸障害も含まれ、その中には嚥下障害、消化性食道炎、食道痙攣、食道狭窄、食道癌、消化不良、消化障害、胃炎、胃癌、食欲不振、悪心、嘔吐、胃不全麻痺、洞または幽門の浮腫、腹部アングナ、胸焼け、胃腸炎、イレウス、腸管感染、消化性潰瘍、胆石症、胆嚢炎、胆汁うっ滞、膵臓炎、膵臓癌、胆道疾患、肝炎、高ビリルビン血症、硬変症、肝臓の受動性うっ血、ヘパトーム、感染性大腸炎、潰瘍性大腸炎、潰瘍性直腸炎、クローン病、ホイップル病、マロリー ヴァイス症候群、結腸癌、結腸閉塞、過敏性腸症候群、短小腸症候群、下痢、便秘、胃腸出血、後天性免疫不全症候群 (AIDS) 腸症、黄疸、肝性脳症、肝腎症候群、肝炎、血色素症、ウィルソン病、アンチトリプシン欠損症、ライ症候群、原発性硬化性胆管炎、肝梗塞、門脈循環閉塞及び血栓、小葉中心壊死、肝臓紫斑病、肝静脈血栓、肝静脈閉塞症、子癇前症、子癇、妊娠性急性肝脂肪、妊娠性肝臓内胆汁うっ滞と、結節性再生及び腺腫、癌腫を含む肝癌が含まれ、また自己免疫/炎症疾患も含まれ、その中には後天性免疫不全症候群 (AIDS)、アジソン病、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫多発性内分泌腺障害症 (APCED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜の炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性紅斑性狼瘡、全身性硬化症、血小板減少症、潰瘍性大腸炎、ブドウ膜炎、ウェルナー症候群、癌の合併症、血液透析、体外循環、ウイルス感染症、細菌性感染症、真菌性感染症、寄生虫感染症、原虫感染症、蠕虫の感染症及び外傷が含まれ、また代謝障害も含まれ、その中には糖尿病、肥満症及び骨粗鬆症が含まれ、またアデノウイルス、アレナウイルス、ブニヤウイルス、カリチウイルス、コロナウイルス、フィロウイルス、ヘパドナウイルス、ヘルペスウイルス、フラビウイルス、オルソミクソウイルス、パルボウイルス、パポバウイルス、パラミクソウイルス、ピコルナウイルス、ポックスウイルス、レオウイルス、レトロウイルス、ラプトウイルス及びトンガウイルス (tongavirus) として分類されるウイルス性病原体による感染症も含まれる。

【0168】

別の実施例では、G C R E C またはその断片や誘導体を発現し得るベクターを患者に投与して、限定するものではないが上記した疾患を含む G C R E C の発現または活性の低下に関連した疾患を治療または予防することも可能である。

【0169】

更に別の実施例では、実質的に精製された G C R E C を含む成分を好適な医薬用担体と共に患者に投与して、限定するものではないが上記した疾患を含む G C R E C の発現または活性の低下に関連した疾患を治療または予防することも可能である。

【0170】

更に別の実施例では、G C R E C の活性を調節するアゴニストを患者に投与して、限定するものではないが上記した疾患を含む G C R E C の発現または活性の低下に関連した疾患を治療または予防することも可能である。

10

【0171】

更に別の実施例では、患者に G C R E C のアンタゴニストを投与して、G C R E C の発現または活性の増大に関連した疾患を治療または予防することが可能である。限定するものではないがこのような疾患の例には、上記した細胞増殖異常、神経障害、心血管障害、胃腸障害、自己免疫/炎症疾患及び代謝異常並びにウイルス感染がある。一実施態様においては、アンタゴニストとして直接的に、或いは G C R E C を発現する細胞または組織に薬剤を輸送するターゲティングまたは輸送機構として間接的に G C R E C と特異結合する抗体を用いることができる。

【0172】

別の実施例では、G C R E C をコードするポリヌクレオチドの相補体を発現するベクターを患者に投与して、限定するものではないが上記した疾患を含む G C R E C の発現または活性の増大に関連した疾患を治療または予防することも可能である。

20

【0173】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補配列、またはベクターを、別の好適な治療薬と組み合わせて投与することもできる。併用療法で用いる好適な治療薬は、当業者が従来の医薬原理に従って選択し得る。治療薬と組み合わせることにより、上記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法を用いることにより少量の各薬剤で医薬効果をあげることが可能となり、それによって副作用の可能性を低減し得る。

30

【0174】

G C R E C のアンタゴニストは、本技術分野で一般的に知られている方法を用いて製造し得る。具体的には、精製された G C R E C を用いて抗体を作るか、治療薬のライブラリをスクリーニングして、G C R E C と特異結合するものを同定することが可能である。G C R E C の抗体も、本技術分野で一般的に知られている方法を用いて製造することが可能である。限定するものではないがこのような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖抗体、F a b 断片及び F a b 発現ライブラリによって作られた断片が含まれ得る。中和抗体（即ち二量体の形成を阻害する抗体）は通常、治療用に好適である。

【0175】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒト、その他を含む種々の宿主は、G C R E C、または G C R E C の任意の断片またはオリゴペプチドであって免疫原性の特性を有するものを注入することによって免疫化することができる。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。限定するものではないがこのようなアジュバントには、フロイントアジュバントと、水酸化アルミニウム等のミネラルゲルアジュバントと、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、油性乳剤、スカシガイのヘモシニアン、ジニトロフェノール等の界面活性剤とがある。ヒトに用いられるアジュバントの中では、B C G（カルメット ゲラン桿菌）及びコリネバクテリウム パルヴム（C o r y n e b a c t e r i u m p a r v u m）が特に好ましい。

40

50

## 【0176】

GCRECに対する抗体を誘導するために用いるオリゴペプチド、ペプチドまたは断片は、少なくとも約5のアミノ酸からなり、一般的には少なくとも約10のアミノ酸からなるアミノ酸配列を有するものが好ましい。これらのオリゴペプチド、ペプチドまたは断片は、天然のタンパク質のアミノ酸配列の一部と同一であり且つ小さな天然の分子の全アミノ酸配列を含むことが望ましい。GCRECアミノ酸の短い伸長部は、別のタンパク質、例えばスカシガイのヘモシニアンの配列と融合し、キメラ分子に対する抗体が産生され得る。

## 【0177】

GCRECに対するモノクローナル抗体は、抗体分子を産生する任意の技術を用いて、培地内の連続した細胞株によって作製し得る。限定するものではないがこのような技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術及びEBV-ハイブリドーマ技術がある(Kohler, G. 他(1975) Nature 256:495-497、Kozbor, D. 他(1985) J. Immunol. Methods 81:31-42、Cote, R.J. 他(1983) Proc. Natl. Acad. Sci. USA 80:2026-2030、Cole, S.P. 他(1984) Mol. Cell Biol. 62:109-120等を参照)。

10

## 【0178】

更に、ヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの「キメラ抗体」作製のために発達した技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる(例えば、Morrisson, S.L. 他(1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. 他(1984) Nature 312:604-608; Takeda, S. 他(1985) Nature 314:452, 454を参照)。或いは、一本鎖抗体を産生するために説明された技術を適用し、本技術分野で知られている方法を用いて、GCREC特異性一本鎖抗体を産生し得る。関連特異性を有するガイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる(Burton D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137等を参照)。

20

## 【0179】

抗体の産生は、リンパ球集団における *in vivo* 産生の誘導によって、或いは免疫グロブリンライブラリのスクリーニングまたは文献に開示されているような高特異結合試薬のパネルのスクリーニングによっても行い得る(Orlandi, R. 他(1989) Proc. Natl. Acad. Sci. USA 86:3833-3837、Winter, G. 他(1991) Nature 349:293-299等を参照)。

30

## 【0180】

GCRECのための特異結合部位を有する抗体を産生することもできる。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製されるF(ab')<sub>2</sub>断片と、F(ab')<sub>2</sub>断片のジスルフィド架橋を還元することによって作製されるFab断片とがある。或いは、Fab発現ライブラリを作製することによって、モノクローナルFab断片を所望の特異性で迅速且つ容易に同定することが可能となる(Huse, W.D. ら(1989) Science 256:1275-1281等を参照)。

40

## 【0181】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる免疫放射線アッセイまたは競合結合アッセイに対する数々のプロトコルは、本技術分野において公知である。このようなイムノアッセイは通常、GCRECとその特異性抗体間の複合体形成の計測に関与している。2つの非干渉性GCRECエピト

50

ープに反応するモノクローナル抗体を用いるような、2部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合結合アッセイを利用してもよい(前出のPoundの文献)。

【0182】

ラジオイムノアッセイ技術と共に様々な方法、例えばスキッチャード分析を用いて、GCRECに対する抗体の親和性を評価する。親和性は結合定数 $K_a$ で表す。 $K_a$ は、平衡状態においてGCREC抗体複合体のモル濃度を遊離抗体と遊離抗原のモル濃度で除した値であると定義する。ポリクローナル抗体は多様なGCRECエピトープに対する親和性が不均一であり、ポリクローナル抗体試薬のために決定した $K_a$ は、GCREC抗体の平均親和性または結合活性を表す。モノクローナル抗体は特定のGCRECエピトープに対して単一特異的であり、モノクローナル抗体試薬のために決定した $K_a$ は、親和性の真の測定値を表す。 $K_a$ 値が約 $10^9 \sim 10^{12}$  L/molの範囲にあるような高親和性抗体試薬は、GCREC抗体複合体が激しい操作に耐えなければならないイムノアッセイに用いるのが好ましい。 $K_a$ 値が約 $10^6 \sim 10^7$  L/molの範囲にあるような低親和性抗体試薬は、GCRECが抗体から最終的に活性化状態で解離する必要がある免疫精製及び類似の処理に用いるのが好ましい(Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC, Liddell, J. E. and Cyster, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0183】

ポリクローナル抗体試薬の抗体価及び結合活性を更に評価して、後に使う或る適用例に対するこのような試薬の品質及び適性を決定することができる。例えば、少なくとも $1 \sim 2$  mg/ml、好ましくは $5 \sim 10$  mg/mlの特異抗体を含むポリクローナル抗体試薬は、GCREC抗体複合体を沈殿させる必要がある処理において通常用いられる。抗体の特異性、抗体価、結合活性、様々な適用例における抗体の品質や使用に対する指針については、一般に入手可能である。(前出のCattyの文献、同Coliganらの文献等を参照)。

【0184】

本発明の別の実施例では、GCRECをコードするポリヌクレオチド、GCRECの任意の断片または相補配列を治療目的で使用することができる。ある実施態様では、GCRECをコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子(DNA及びRNA、修飾ヌクレオチド)を設計して遺伝子発現を変更することができる。このような技術は当分野では周知であり、センスまたはアンチセンスオリゴヌクレオチドまたは大きな断片が、GCRECをコードする配列の制御領域から、またはコード領域に沿ったさまざまな位置から設計可能である。(Agrawal, S.他(1996) Antisense Therapeutics, Humana Press Inc., Totawa NJを参照)。

【0185】

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を発現する発現プラスミドの形で細胞内に輸送することが可能である。(Slater, J. E.他(1998) J. Allergy Clin. Immunol. 102(3): 469-475及びScanlon, K. J.他(1995) 9(13): 1288-1296.等を参照。)アンチセンス配列はまた、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる(Miller, A. D. (1990) Blood 76: 271、前出のAusubel, Uckert, W. and Walther (1994) Pharmacol. Ther. 63(3): 32

10

20

30

40

50

3 - 347等を参照。)その他の遺伝送達機構には、リポソーム系、人工的なウイルスエンベロープ及び当分野で公知のその他のシステムが含まれる(Rossi, J. J. (1995) Br. Med. Bull. 51(1): 217-225; Boado, R. J. 他(1998) J. Pharm. Sci. 87(11): 1308-1315、Morris, M. C. 他(1997) Nucleic Acids Res. 25(14): 2730-2736. 等を参照)。

#### 【0186】

本発明の別の実施例では、GCRECをコードするポリヌクレオチドを、体細胞若しくは生殖細胞遺伝子治療に用いることが可能である。遺伝子治療を行うことにより、(i) 遺伝子欠損症(例えばX染色体鎖遺伝(Cavazzana-Calvo, M. 他(2000) Science 288: 669-672)により特徴付けられる重度の複合型免疫欠損(SCID) - X1の場合)、先天性アデノシンデアミナーゼ(ADA)欠損症に関連する重度の複合型免疫欠損(Blaese, R. M. 他(1995) Science 270: 475-480、Bordignon, C. 他(1995) Science 270: 470-475)、嚢胞性繊維症(Zabner, J. 他(1993) Cell 75: 207-216; Crystal, R. G. 他(1995) Hum. Gene Therapy 6: 643-666、Crystal, R. G. 他(1995) Hum. Gene Therapy 6: 667-703)、サラセミア(thalassaemia)、家族性高コレステロール血症、第VIII因子若しくは第IX因子欠損に起因する血友病(Crystal, R. G. (1995) Science 270: 404-410、Verma, I. M. および N. Somia (1997) Nature 389: 239-242)を治療し、(ii) 条件的致死性遺伝子産物を発現させ(例えば制御不能な細胞増殖に起因する癌の場合)、(iii) 細胞内の寄生虫(例えばヒト免疫不全ウイルス(HIV)(Baltimore, D. (1988) Nature 335: 395-396、Poeschl, E. 他(1996) Proc. Natl. Acad. Sci. USA. 93: 11395-11399)、B型若しくはC型肝炎ウイルス(HBV、HCV)、Candida albicans及びParacoccidioides brasiliensis等の真菌寄生虫、並びにPlasmodium falciparum及びTrypanosoma cruzi等の原虫寄生体に対する防御機能を有するタンパク質を発現させることができる。GCRECの発現若しくは調節に必要な遺伝子の欠損が疾患を発生させる場合、形質導入した細胞の好適な集団からGCRECを発現することにより、遺伝子欠損に起因する症状の発現を緩和し得る。

#### 【0187】

本発明の更なる実施例では、GCRECをコードする哺乳動物発現ベクターを作製し、これらのベクターを機械的手段によってGCREC欠損細胞に導入することによって、GCRECの欠損による疾患や異常症を治療する。in vivo 或いは ex vitro の細胞に用いる機械的導入技術には、(i) 個々の細胞内への直接的なDNA微量注射法、(ii) 遺伝子銃、(iii) リポソームを介した形質移入、(iv) 受容体を介した遺伝子導入、及び(v) DNAトランスポソンの使用がある(Morgan, R. A. 及び W. F. Anderson (1993) Annu. Rev. Biochem. 62: 191-217、Ivics, Z. (1997) Cell 91: 501-510; Boulay, J-L. 及び H. Recipon (1998) Curr. Opin. Biotechnol. 9: 445-450)。

#### 【0188】

GCRECの発現に効果的である発現ベクターには、限定するものではないが、PCDNA3.1、EPITAG、PRCCMV2、PREP、PVAX、PCR2-TOPOTAベクター(Invitrogen, Carlsbad CA)、PCMVSCRIPT、PCMV-TAG、PEGSH/PERV(Stratagene, La Jolla CA) および PTETOFF、PTETON、PTRE2、PTRE2

LUC、PTKHYG (Clontech, Palo Alto CA)が含まれる。GCRECは、(i)恒常的に活性化プロモーター(例えばサイトメガロウイルス(CMV)、ラウス肉腫ウイルス(RSV)、SV40ウイルス、チミジンキナーゼ(TK)またはアクチン遺伝子)、(ii)誘導性プロモーター(例えばテトラサイクリン調節性プロモーター(Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551、Gossen, M. 他(1995) Science 268:1766-1769、Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456)、市販のInvitrogen社のT-REXプラスミドに含まれる)、エクジソン誘導性プロモーター(Invitrogen社のプラスミドPVGRXR及びPINDから得られる)、FK506/ラパマイシン誘導性プロモーターまたはRU486/ミフェプリストーン誘導性プロモーター(前出のRossi, F.M.V. and H.M. Blau)、または(iii)正常個体由来の、GCRECをコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いて、発現させることができる。

【0189】

市販のリポソーム形質転換キット(例えばInvitrogen社のPerfect Lipid Transfection Kit)を用いれば、当業者は経験にそれほど頼らないでもポリヌクレオチドを培養中の標的細胞に導入することが可能になる。別法では、リン酸カルシウム法(Graham, F.L. 及び A.J. Eb (1973) Virology 52:456-467)若しくは電気穿孔法(Neumann, B. 他(1982) EMBO J. 1:841-845)を用いて形質転換を行う。初代培養細胞にDNAを導入するためには、標準化された哺乳動物の形質移入プロトコルの修飾が必要である。

【0190】

本発明の別の実施例では、GCRECの発現に関連する遺伝子欠損によって起こる疾患や異常症は、(i)レトロウイルス末端反復配列(LTR)プロモーターまたは独立プロモーターの制御下でGCRECをコードするポリヌクレオチドと、(ii)好適なRNAパッケージングシグナルと、(iii)追加レトロウイルス・シス作用性RNA配列及び効率的なベクターの増殖に必要なコード配列を伴うRev応答性エレメント(RRE)とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター(例えばPFB及びPFBNEO)はStratagene社から市販されており、刊行データ(Riviere, I. 他(1995) Proc. Natl. Acad. Sci. U.S.A. 92:6733-6737)に基づいている。上記データを引用することをもって本明細書の一部とする。ベクターは、好適なベクター産生細胞系(VPCL)において増殖され、VPCLは、標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子またはVSVG等の汎親和性エンベロープタンパク質を発現する(Armentano, D. ら(1987) J. Virol. 61:1647-1650、Bender, M.A. ら(1987) J. Virol. 61:1639-1646、Adam, M.A. および A.D. Miller (1988) J. Virol. 62:3802-3806、Dull, T. 他(1998) J. Virol. 72:8463-8471、Zufferey, R. 他(1998) J. Virol. 72:9873-9880)。RIGGに付与された米国特許第5,910,434号(「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」)において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団(例えばCD4<sup>+</sup>T細胞)の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている(Ranga, U. 他

(1997) *J. Virol.* 71:7020-7029、Bauer, G. 他 (1997) *Blood* 89:2259-2267、Bonyhadi, M. L. (1997) *J. Virol.* 71:4707-4716、Ranga, U. 他 (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:1201-1206、Su, L. (1997) *Blood* 89:2283-2290)。

【0191】

別法では、アデノウイルス系遺伝子治療の輸送系を用いて、GCRECの発現に関連する1若しくは複数の遺伝子異常を有するような細胞にGCRECをコードするポリヌクレオチドを輸送する。アデノウイルス系ベクターの作製及びパッケージングについては、当業者に公知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を脾臓の無損傷の脾島内に導入するために多面的に有用であることが証明された (Csete, M. E. 他 (1995) *Transplantation* 27:263-268)。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第5,707,618号 ("Adenovirus vectors for gene therapy") に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについては、Antinozzi, P. A. 他 (1999) *Annu. Rev. Nutr.* 19:511-544 及び Verma, I. M. 及び N. Somia (1997) *Nature* 18:389:239-242も参照されたい。両文献は、引用することをもって本明細書の一部とする。

10

20

【0192】

別法では、ヘルペス系遺伝子治療の輸送系を用いて、GCRECの発現に関連する1若しくは複数の遺伝子異常を有する標的細胞にGCRECをコードするポリヌクレオチドを輸送する。HSVが親和性を有するような中枢神経系の細胞にGCRECを導入する際には、単純ヘルペスウイルス(HSV)系のベクターの使用は特に役立つ。ヘルペス系ベクターの作製及びパッケージングは、当業者に公知である。複製適格性単純ヘルペスウイルス(HSV)I型系のベクターは、レポーター遺伝子を霊長類の眼に送達するために用いられてきた (Liu, X. 他 (1999) *Exp. Eye Res.* 169:385-395)。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号 ("Herpes simplex virus swains for gene transfer") に開示されており、該特許の引用をもって本明細書の一部とする。米国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも1つの外在性遺伝子を有するゲノムを含む組換えHSV d92についての記載がある。上記特許はまた、ICP4、ICP27及びICP22のために除去される組換えHSV系統の作製及び使用について開示している。HSVベクターについては、Goins, W. F. 他 (1999) *J. Virol.* 73:519-532 及び Xu, H. 他 (1994) *Dev. Biol.* 163:152-161も参照されたい。両文献は、引用をもって本明細書の一部とする。クローン化ヘルペスウイルス配列の操作、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。

30

40

【0193】

別法では、ウイルス(正の一本鎖RNAウイルス)ベクターを用いてGCRECをコードするポリヌクレオチドを標的細胞に輸送する。プロトタイプのウイルスであるセムリキ森林熱ウイルス(Semliki Forest Virus, SFV)の生物学的研究が広範に行われており、遺伝子導入ベクターがSFVゲノムに基づいていることが分かった (Garoff, H. 及び K. - J. Li (1998) *Cun. Opin. Biotech.* 9:464-469)。ウイルスRNAの複製中に、通常はウイルスのキャプシッドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲ

50

ノムRNAは、完全長のゲノムRNAより高いレベルに複製されるため、酵素活性（例えばプロテアーゼ及びポリメラーゼ）を有するウイルスタンパク質に比べてキャプシドタンパク質が過剰産生される。同様に、GCRECに対するコード配列をカプシドコード領域のウイルスゲノムに導入することにより、ベクター導入細胞において多数のGCRECコードRNAが産生され、高レベルのGCRECが合成される。通常はウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドビスウイルス（SIN）の変異体を有するハムスター正常腎臓細胞（BHK-21）の持続的な感染を確立する能力は、ウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している（Dryga, S. A. 他（1997）*Virology* 228 : 74 - 83）。ウイルスの宿主の範囲が広いことにより、様々な細胞タイプへのGCRECの導入が可能になる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。ウイルスの感染性cDNAクローンの処置方法、ウイルスのcDNA及びRNAの形質移入方法及びウイルスの感染方法は、当業者に公知である。

#### 【0194】

転写開始部位由来のオリゴヌクレオチドを用いて遺伝子発現を阻害することも可能である。転写開始部位とは例えば開始部位から数えて約-10と約+10の間である。同様に、三重らせん塩基対の形成方法を用いて阻害が可能となる。三重らせん塩基対形成は、ポリメラーゼ、転写因子または調節分子の結合のために十分に開くような二重らせんの能力を阻害するので、三重らせん塩基対形成は有用である。三重らせんDNAを用いる最近の治療の進歩については文献に記載がある（Gee, J. E. ら（1994）*in: Huber, B. E. 及び B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY, 163 - 177* 頁等を参照）。相補配列またはアンチセンス分子もまた、転写物がリボソームに結合するのを阻止することによってmRNAの翻訳を阻止するべく設計することができる。

#### 【0195】

リボザイムは酵素性RNA分子であり、RNAの特異的切断を触媒するためにリボザイムを用いることもできる。リボザイム作用のメカニズムは、相補的標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションおよびその後起こる内ヌクレオチド鎖切断に参与している。例えば、遺伝子操作で作られたハンマーヘッド型リボザイム分子は、GCRECをコードする配列の内ヌクレオチド鎖分解性の切断を特異的且つ効果的に触媒する可能性がある。

#### 【0196】

任意のRNA標的内の特異的リボザイム切断部位は、GUA、GUU、GUC配列を含めたりボザイム切断部位に対して標的分子をスキャンすることによって先ず同定される。一度同定されると、切断部位を含む標的遺伝子の領域に対応する15~20リボヌクレオチドの短いRNA配列が、そのオリゴヌクレオチドを機能不全にするような2次構造の特徴をもっていないかを評価することが可能になる。候補標的の適合性の評価も、リボヌクレアーゼ保護アッセイを用いて相補的オリゴヌクレオチドとのハイブリダイゼーションのアクセス可能性をテストすることによって行うことができる。

#### 【0197】

本発明の相補リボ核酸分子及びリボザイムは、核酸分子合成のために当分野でよく知られている任意の方法を用いて作製し得る。任意の方法には、固相フォスフォアミダイト化学合成等のオリゴヌクレオチドを化学的に合成する方法がある。或いは、GCRECをコードするDNA配列の*in vitro*及び*in vivo*転写によってRNA分子を産出し得る。このようなDNA配列は、T7やSP6等の好適なRNAポリメラーゼプロモーターを用いて多様なベクター内に取り込むことが可能である。或いは、相補的RNAを構成的或いは誘導的に合成するようなこれらcDNA産物を、細胞系、細胞または組織内に導入することができる。

#### 【0198】

10

20

30

40

50

細胞内の安定性を高め、半減期を長くするためにRNA分子を修飾することができる。限定するものではないが可能な修飾には、分子の5'末端、3'末端、あるいはその両方においてフランキング配列を追加したり、分子の主鎖内においてホスホジエステラーゼ結合ではなくホスホオチオネートまたは2'-O-メチルを使用したりすることが含まれる。この概念は、PNAの産出に固有のものであり、これら全ての分子に拡大することができる。それには、内在性エンドヌクレアーゼによって容易には認識されないアデニン、シチジン、グアニン、チミン、及びウリジンにアセチル-、メチル-、チオ-及び同様の修飾をしたものの他、非従来型塩基、例えばイノシン、クエオシン(queosine)、ワイプトシン(wybutosine)等を加えることでできる。

#### 【0199】

本発明の追加実施例は、GCRECをコードするポリヌクレオチドの変異発現に有効な化合物をスクリーニングする方法を含む。限定するものではないが特異ポリヌクレオチドの発現変化を起こすのに有効な化合物には、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチド、転写因子その他のポリペプチド転写制御因子、及び特異ポリヌクレオチド配列と相互作用し得る非高分子化学的実体がある。有効な化合物は、ポリヌクレオチド発現のインヒビターまたはエンハンサーのいずれかとして作用することによりポリヌクレオチド発現を変異し得る。従って、GCRECの発現または活性の増加に関連する疾患の治療においては、GCRECをコードするポリヌクレオチドの発現を特異的に阻害する化合物が治療上有用であり、GCRECの発現または活性の低下に関連する疾患の治療においては、GCRECをコードするポリヌクレオチドの発現を特異的に促進する化合物が治療上有用であり得る。

#### 【0200】

特異ポリヌクレオチドの変異発現における有効性に対して、少なくとも1個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を変異させる場合と、既存の、市販のまたは専売の、天然または非天然の化合物ライブラリから選択する場合と、標的ポリヌクレオチドの化学的及び/または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効であることが知られているような化合物の化学修飾がある。GCRECをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝され、このように得られる。サンプルには例えば、無傷細胞、透過化処理した細胞、in vitro無細胞系または再構成生化学系があり得る。GCRECをコードするポリヌクレオチドの発現における変化は、当分野で通常知られている任意の方法でアッセイする。通常、GCRECをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特異ヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量し、それによって1つ若しくは複数の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現変化の検出は、ポリヌクレオチドの発現を変異する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えば分裂酵母(Schizosaccharomyces pombe)遺伝子発現系(Atkins, D. 他(1999)米国特許第5,932,435号、Arndt, G.M. 他(2000)Nucleic Acids Res. 28:E15)またはHeLa細胞等のヒト細胞系(Clark, M.L. 他(2000)Biochem. Biophys. Res. Commun. 268:8-13)を用いてスクリーニングを実行する。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性のためのオリゴヌクレオチド(デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド)の組み合わせライブラリをスクリーニングすることに関与している(Bruice, T.W. 他(1997)米国特許第5,686,242号、Bruice, T.W. 他(2000)米国特許第6,022,691号)。

10

20

30

40

50

## 【0201】

ベクターを細胞または組織に導入する多数の方法が利用可能であり、*in vivo*、*in vitro*及び*ex vivo*の使用に対して同程度に適している。*ex vivo*治療の場合、ベクターを患者から採取した幹細胞内に導入し、クローニング増殖して同一患者に自家移植で戻すことができる。トランスフェクション、リボソーム注入またはポリカチオンアミノポリマーによる送達は、当分野でよく知られている方法を用いて実行することができる (Goldman, C. K. 他 (1997) *Nat. Biotechnol.* 15: 462 - 466. 等を参照)。

## 【0202】

上記の治療方法はいずれも、例えば、ヒト、イヌ、ネコ、ウシ、ウマ、ウサギ、サル等の哺乳動物を含めて治療が必要な全ての対象に適用できる。 10

## 【0203】

本発明の追加実施例は、通常薬剤として許容できる賦形剤で処方される活性成分を有する成分の投与に関連する。賦形剤には例えば、糖、でんぷん、セルロース、ゴム及びタンパク質がある。様々な処方が通常知られており、詳細は Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA) の最新版に記載されている。このような成分は、GC REC、GC REC に対する抗体、擬態、アゴニスト、アンタゴニスト、またはGC RECインヒビターから構成し得る。

## 【0204】

本発明に用いられる成分は、任意の数の経路によって投与することができ、限定するものではないが経路には、経口、静脈内、筋肉内、動脈内、骨髄内、クモ膜下腔内、心室内、肺、経皮、皮下、腹腔内、鼻腔内、腸内、局所、舌下または直腸がある。 20

## 【0205】

肺から投与する成分は、液状または乾燥粉末状で調製し得る。このような成分は通常、患者が吸入する直前にエアロゾル化する。小分子 (例えば伝統的な低分子量有機薬) の場合には、速効剤のエアロゾル送達は当分野で公知である。高分子 (例えばより大きなペプチド及びタンパク質) の場合には、当該分野において肺の肺泡領域を介しての肺送達が最近向上したことにより、インスリン等の薬剤を実質的に血液循環へ輸送することを可能にした (Patton, J. S. 他, 米国特許第 5, 997, 848号等を参照)。肺送達は、針注射なしに投与する点で優れており、有毒な可能性のある浸透エンハンサーの必要性をなくす。 30

## 【0206】

本発明での使用に適した成分には、所定の目的を達成するために必要なだけの量の活性成分を含有する成分が含まれる。有効投与量の決定は、当業者の能力の範囲内で行う。

## 【0207】

特殊形状の成分は、GC RECまたはその断片を含む高分子を直接細胞内輸送するために調製される。例えば、細胞不透過性高分子を含むリボソーム製剤は、細胞融合及び高分子の細胞内送達を促進し得る。或いは、GC RECまたはその断片を HIV Tat-1 タンパク質から陽イオン N 末端部に結合することもできる。このようにして生成された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入することがわかっている (Schwarze, S. R. 他 (1999) *Science* 285: 1569 - 1572)。 40

## 【0208】

任意の化合物に対して、細胞培養アッセイ、例えば新生物性細胞の細胞培養アッセイにおいて、或いは、動物モデル、例えばマウス、ウサギ、イヌまたはブタ等において、先ず治療の有効投与量を推定することができる。動物モデルはまた、好適な濃度範囲及び投与経路を決定するためにも用い得る。このような情報を用いて、次にヒトに対する有益な投与量及び投与経路を決定することができる。

## 【0209】

治療上の有効投与量は、症状や容態を回復させるような活性成分量を参考にする。そのような活性成分の例としては、G C R E C またはその断片、G C R E C の抗体、G C R E C のアゴニスト、アンタゴニストまたはインヒビターがある。薬用有効度及び毒性は、細胞培養または動物実験における標準的な薬剤手法によって、例えば E D<sub>50</sub>（集団の 50% の医薬的有効量）または L D<sub>50</sub>（集団の 50% の致死量）を測定するなどして決定することができる。毒性効果の治療効果に対する投与量の比は、治療指数であり、L D<sub>50</sub> / E D<sub>50</sub> 比として表すことができる。高い治療指数を示すような成分が望ましい。細胞培養アッセイ及び動物実験から得られたデータは、ヒトに用いるための投与量の範囲を調剤するのに用いられる。このような組成物が含まれる投与量は、毒性を殆ど或いは全く含まず、E D<sub>50</sub> を含むような血中濃度の範囲にあることが好ましい。用いられる投与形態、患者の感受性及び投与の経路によって、投与量はこの範囲内で様々に変わる。

10

**【0210】**

正確な投与量は、治療が必要な被験者に関する要素を考慮して、現場の医者が決定することになる。効果的なレベルの活性成分を与え、或いは所望の効果を維持するべく、投与量及び投与を調節する。被験者に関する要素としては、疾患の重症度、患者の通常の状態、患者の年齢、体重及び性別、投与の時間及び頻度、薬剤の配合、反応感受性及び治療に対する応答等を考慮する。作用期間が長い成分は、特定の製剤の半減期及びクリアランス率によって 3 ~ 4 日毎に 1 度、1 週間に 1 度、或いは 2 週間に 1 度の間隔で投与し得る。

**【0211】**

通常投与量は、投与の経路にもよるが約 0.1 ~ 100,000 µg であり、合計で約 1 g までとする。特定の投与量及び送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。当業者は、タンパク質またはインヒビターに対する処方とは異なる、ヌクレオチドに対する処方を利用することになる。同様に、ポリヌクレオチドまたはポリペプチドの送達は、特定の細胞、状態、位置等に特異的なものとなる。

20

**【0212】**

（診断）

別の実施例では、G C R E C の発現によって特徴付けられる疾患の診断のために、或いは G C R E C や G C R E C のアゴニスト、アンタゴニストまたは阻害剤で治療を受けている患者をモニターするためのアッセイにおいて、G C R E C を特異的に結合する抗体が用いられることがある。診断目的に有用な抗体は、上記の治療の箇所で記載した方法と同じ方法で調合される。G C R E C の診断アッセイには、抗体及び標識を利用してヒトの体液において或いは細胞や組織のエキスにおいて G C R E C を検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

30

**【0213】**

G C R E C を測定するための様々なプロトコル、例えば E L I S A、R I A、F A C S 等が本技術分野において知られており、G C R E C 発現の修正レベル或いは異常レベルを診断する基準を提供する。複合体の形成に適した条件下でヒト対象等の正常な哺乳動物対象から採取した体液または細胞と G C R E C に対する抗体とを結合させることにより、G C R E C 発現の正常値または標準値が決定される。標準複合体形成量は、種々の方法、例えば測光法で定量できる。対象内で発現した G C R E C の量、制御、検体からの病変サンプルを標準値と比較する。標準値と被験者との偏差が疾患を診断するパラメータとなる。

40

**【0214】**

別の実施例によれば、G C R E C をコードするポリヌクレオチドを診断目的に用いることもできる。用いられることができるポリヌクレオチドには、オリゴヌクレオチド配列、相補的 R N A 及び D N A 分子、そして P N A が含まれる。ポリヌクレオチドは、検体における G C R E C の発現が疾患と相関し得るような該検体における遺伝子発現の検出及び定量

50

に用いることができる。診断アッセイは、G C R E Cの不在、存在及び過剰発現を測定するために、そして治療インターベンション中にG C R E Cレベルの調製をモニターするために用いることができる。

【0215】

ある実施形態では、G C R E Cをコードする核酸配列を同定するために、G C R E Cまたは近縁の分子をコードする、ゲノム配列を含むポリヌクレオチド配列を検出可能なP C Rプローブとのハイブリダイゼーションを用いることができる。プローブが同定するのはG C R E C、突然変異体または関連配列をコードするような、天然に存在する配列のみであるか否かは、プローブの特異性及びハイブリダイゼーション或いは増幅のストリンジェンシーによって決定されることになる。ここで、プローブの特異性とは、プローブが高特異領域（例えば5'調節領域）からなるのか、低特異領域（例えば保存されたモチーフ）からなるのかということである。

【0216】

プローブはまた、関連する配列の検出にも用いることができ、その配列はG C R E Cをコードする任意の配列と少なくとも50%の相同性を有し得る。目的の本発明のハイブリダイゼーションプローブには、D N AあるいはR N Aが可能であり、S E Q I D N O : 17-32の配列、或いはG C R E C遺伝子のプロモーター、エンハンサー、イントロンを含むゲノム配列に由来し得る。

【0217】

G C R E CをコードするD N Aに対する特異的ハイブリダイゼーションプローブを作製する方法には、G C R E CまたはG C R E C誘導体をコードするポリヌクレオチド配列を、m R N Aプローブを作製するためのベクターにクローニングする方法が含まれる。m R N Aプローブ作製のためのベクターは、当業者に知られており、市販されており、好適なR N Aポリメラーゼ及び好適な標識されたヌクレオチドを加えることによって、i n v i t r oでR N Aプローブを合成するために用いられ得る。ハイブリダイゼーションプローブは、種々のレポーターの集団によって標識され得る。レポーター集団の例としては、<sup>32</sup>Pまたは<sup>35</sup>S等の放射性核種、或いはアビジン/ビオチン結合系を介してプローブに結合されたアルカリホスファターゼ等の酵素標識などが挙げられる。

【0218】

G C R E Cをコードするポリヌクレオチド配列は、G C R E Cの発現に関係する疾患の診断の為に用い得る。限定するものではないがこのような疾患の例として細胞増殖異常が含まれ、その中には日光性角化症、動脈硬化症、アテローム性動脈硬化症、滑液包炎、硬変、肝炎、混合型結合組織病（M C T D）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症と、腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、奇形癌を含む癌、具体的には副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌等が含まれ、神経障害が含まれ、その中には癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病その他の錐体外路障害、筋萎縮性側索硬化その他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症その他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患と、クールー、クロイツフェルト ヤコブ病及びガスマン ストラウスラー シャインカー症候群を含むプリオン病と、致死性家族性不眠症、神経系の栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜性血管芽腫症（c e r e b e l l o r e t i n a l h e m a n g i o b l a s t o m a t o s i s）、脳3叉神経血管症候群、精神薄弱その他の中枢神経系発達障害、脳性麻痺、神経骨格異常、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィーその他の神経筋疾患、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分障害、不安障害及び精神分裂病を含む精神障害と、季節型感情障害（S A D）と、静座不能、健忘症、

10

20

30

40

50

緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、トゥレット病、進行性核上麻痺、皮質基底核変性症 (c o r t i c o b a s a l d e g e n e r a t i o n)、家族性前頭側頭骨痲呆が含まれ、また心血管障害も含まれ、その中には動静脈瘤、アテローム性動脈硬化症、高血圧、脈管炎、レイノー病、静脈奇形、動脈解離、静脈瘤、血栓静脈炎及び静脈血栓、血管の腫瘍、血栓崩壊の合併症、バルーン血管形成術 ( b a l l o o n a n g i o p l a s t y)、血管置換術、大動脈冠動脈バイパス術移植手術 ( c o r o n a r y a r t e r y b y p a s s g r a f t s u r g e r y)などの血管疾患と、うっ血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪状石灰化 ( m i t r a l a n n u l a r c a l c i f i c a t i o n)、僧帽弁脱出、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性紅斑性狼瘡の心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心臓移植の合併症などの心疾患が含まれ、また胃腸障害も含まれ、その中には嚥下障害、消化性食道炎、食道痙攣、食道狭窄、食道癌、消化不良、消化障害、胃炎、胃癌、食欲不振、悪心、嘔吐、胃不全麻痺、洞または幽門の浮腫、腹部アングナ、胸焼け、胃腸炎、イレウス、腸管感染、消化性潰瘍、胆石症、胆嚢炎、胆汁うっ滞、膵臓炎、膵臓癌、胆道疾患、肝炎、高ビリルビン血症、硬変症、肝臓の受動性うっ血、ヘパトーム、感染性大腸炎、潰瘍性大腸炎、潰瘍性直腸炎、クローン病、ホイップル病、マロリー ヴァイス症候群、結腸癌、結腸閉塞、過敏性腸症候群、短小腸症候群、下痢、便秘、胃腸出血、後天性免疫不全症候群 ( A I D S )腸症、黄疸、肝性脳症、肝腎症候群、肝炎、血色素症、ウィルソン病、アンチトリプシン欠損症、ライ症候群、原発性硬化性胆管炎、肝梗塞、門脈循環閉塞及び血栓、小葉中心壊死、肝臓紫斑病、肝静脈血栓、肝静脈閉塞症、子癇前症、子癇、妊娠性急性肝脂肪、妊娠性肝臓内胆汁うっ滞と、結節性再生及び腺腫、癌腫を含む肝癌が含まれ、また自己免疫/炎症疾患も含まれ、その中には後天性免疫不全症候群 ( A I D S)、アジソン病、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫多発性内分泌腺障害症 ( A P E C E D)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレープス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜の炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性紅斑性狼瘡、全身性硬化症、血小板減少症、潰瘍性大腸炎、ブドウ膜炎、ウェルナー症候群、癌の合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫の感染症及び外傷が含まれ、また代謝障害も含まれ、その中には糖尿病、肥満症及び骨粗鬆症が含まれ、またアデノウイルス、アレナウイルス、プニヤウイルス、カリチウイルス、コロナウイルス、フィロウイルス、ヘパドナウイルス、ヘルペスウイルス、フラビウイルス、オルソミクソウイルス、パルボウイルス、パポバウイルス、パラミクソウイルス、ピコルナウイルス、ポックスウイルス、レオウイルス、レトロウイルス、ラウドウイルス及びトンガウイルス ( t o n g a v i r u s)として分類されるウイルス性病原体による感染症も含まれる。G C R E Cをコードするポリヌクレオチド配列は、サザン法、ノーザン法、ドットプロット法やその他の膜ベースの技術と、P C R法と、ディップスティック ( d i p s t i c k)法、ピン及びマルチフォーマット E L I S A様アッセイと、変異G C R E Cの発現を検出するために患者から採取した体液または組織を利用するマイクロアレイとにおいて使用し得る。このような定性方法または定量方法は、当分野で公知である。

#### 【0219】

或る形態では、関連する疾患、特に上記した疾患を検出するアッセイにおいて、G C R E Cをコードするヌクレオチド配列が有用であり得る。G C R E Cをコードするヌクレオチ

ド配列は標準法で標識化され、ハイブリダイゼーション複合体の形成に好適な条件下で、患者から採取した体液または組織のサンプルに添加することができる。好適なインキュベーション期間が経過したらサンプルを洗浄し、シグナルを定量して標準値と比較する。患者サンプルのシグナル量が制御サンプルと比べて著しく変化している場合は、サンプル内のGCRCをコードするヌクレオチド配列の変異レベルは関連する疾患の存在を示している。このようなアッセイは、動物実験、臨床試験における特定の治療効果を推定するため、或いは個々の患者の治療をモニターするために用いることもできる。

#### 【0220】

GCRCの発現に関連する疾患の診断基準を提供するために、発現のための正常あるいは標準プロフィールを確立する。これは、ハイブリダイゼーション或いは増幅に好適な条件下で、動物或いはヒトの正常な対象から抽出した体液或いは細胞を、GCRCをコードする配列またはその断片と結合させることにより達成され得る。実質的に精製されたポリヌクレオチドを既知量用いて行った実験から得た値を正常な対象から得た値と比較することにより、標準ハイブリダイゼーションを定量することができる。このようにして得た標準値は、疾患の徴候を示す患者から得たサンプルから得た値と比較することができる。標準値からの偏差を用いて疾患の存在を確定する。

10

#### 【0221】

疾患の存在が確定されて治療プロトコルが開始されると、患者の発現レベルが正常な被検者に観察されるレベルに近づき始めたかどうかを測定するため、ハイブリダイゼーションアッセイを通常ペースで繰り返し得る。連続アッセイから得られた結果を用いて、数日から数ヶ月の期間にわたる治療の効果を示し得る。

20

#### 【0222】

癌に関しては、個体からの生体組織における異常な量の転写物（過少発現または過剰発現）の存在は、疾患の発生素質を示したり、実際に臨床的症状が現れる前に疾患を検出する方法を提供したりし得る。この種のより明確な診断により、医療の専門家が予防方法または積極的な治療法を早くから利用し、それによって癌の発生または更なる進行を防止することが可能となる。

#### 【0223】

GCRCをコードする配列から設計されたオリゴヌクレオチドを追加的に診断上利用することは、PCRの利用に關与し得る。これらのオリゴマーは、化学的に合成するか、酵素により生産するか、或いは*in vitro*で産出し得る。オリゴマーは、好ましくはGCRCをコードするポリヌクレオチドの断片、或いはGCRCをコードするポリヌクレオチドと相補的ポリヌクレオチドの断片を含み、最適条件下で特定の遺伝子や条件を識別するべく利用される。また、オリゴマーは、やや緩いストリンジェント条件下で、近縁のDNA或いはRNA配列の検出、定量、或いはその両方のため用いることが可能である。

30

#### 【0224】

或る態様において、GCRCをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて単1塩基多型性(SNP)を検出し得る。SNPは、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入及び欠失である。限定するものではないがSNPの検出方法には、SSCP(single-stranded conformation polymorphism)及び蛍光SSCP(fSSCP)がある。SSCPでは、GCRCをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、ポリメラーゼ連鎖反応法(PCR)を用いたDNAの増幅を行う。DNAは例えば、病変組織または正常組織、生検サンプル、体液その他に由来し得る。DNA内のSNPは、一本鎖形状のPCR生成物の2次及び3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSSCPでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってDNAシーケンシング機などの高処理機器でアンプリマー(amplimer)の検出が可能になる。更に、インシリコSNP(in silico SNP, isSNP)と呼ばれる配列デ

40

50

ータベース分析法は、一般的なコンセンサス配列に配列されるような個々の重畳するDNA断片の配列を比較することにより、多型性を同定し得る。これらのコンピュータベースの方法は、DNAの実験室での調整及び統計モデル及びDNA配列クロマトグラムの自動分析を用いたシーケンシングのエラーに起因する配列の変異をフィルタリングして除去する。別の態様では、例えば高処理MASSARRAYシステム(Sequenom, Inc., San Diego CA)を用いた質量分析によりSNPを検出し、特徴付ける。

#### 【0225】

GCRECの発現を定量するために用い得る方法には、ヌクレオチドの放射標識またはビオチン標識、調節核酸の相互増幅(coamplification)及び標準曲線から得た結果の補間もある(例えば、Melby, P.C. 他(1993) J. Immunol. Methods 159:235244; Duplaa, C. 他(1993) Anal. Biochem. 212:229236を参照。)目的のオリゴマーが種々の希釈液中に存在し、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットのアッセイを行うことによって、複数のサンプルの定量速度を加速することができる。

#### 【0226】

更に別の実施例では、本明細書に記載した任意のポリヌクレオチド配列由来のオリゴヌクレオチドまたはより長い断片を、マイクロアレイにおけるエレメントとして用いることができる。多数の遺伝子の関連発現レベルを同時にモニターする転写イメージング技術にマイクロアレイを用いることが可能である。これについては、以下に記載する。マイクロアレイはまた、遺伝変異体、突然変異及び多型性の同定に用いることができる。この情報を用いることで、遺伝子機能を決定し、疾患の遺伝的根拠を理解し、疾患を診断し、遺伝子発現の機能としての疾病の進行/後退をモニターし、疾病治療における薬剤の活性を開発及びモニターすることができる。特に、患者にとって最もふさわしく、有効的な治療法を選択するために、この情報を用いて患者の薬理ゲノムプロファイルを開発することができる。例えば、患者の薬理ゲノムプロファイルに基づき、患者に対して高度に効果的で副作用を殆ど示さない治療薬を選択することができる。

#### 【0227】

別の実施例では、GCRECに特異的な抗体、GCRECまたはその断片をマイクロアレイ上で要素として用いることができる。マイクロアレイを用いて、上記のようなタンパク質-タンパク質相互作用、薬剤-標的相互作用及び遺伝子発現プロファイルをモニターまたは測定することが可能である。

#### 【0228】

或る実施例は、或る組織または細胞タイプの転写イメージを生成するような本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプにより遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される(Seilliamer 他、米国特許第5,840,484号の"Comparative Gene Transcript Analysis"を参照。この特許に言及することを以って本明細書の一部とする)。従って、特定の組織または細胞タイプの転写または逆転写全体に本発明のポリヌクレオチドまたはその補体をハイブリダイズすることにより、転写イメージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはその補体がマイクロアレイ上のエレメントのサブセットを複数含むような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロファイルを提供し得る。

#### 【0229】

転写イメージは、組織、株化細胞、生検またはその生物学的サンプルから単離した転写物を用いて生成し得る。転写イメージは従って、組織または生検サンプルの場合には in vivo、または株化細胞の場合には in vitroでの遺伝子発現を反映する。

10

20

30

40

50

## 【0230】

本発明のポリヌクレオチドの発現のプロフィールを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、*in vitro*モデル系及び薬剤の前臨床評価に関連して使用し得る。全ての化合物は、作用及び毒性のメカニズムを暗示する、しばしば分子フィンガープリントまたは毒性シグネチャと称されるような特徴的な遺伝子発現パターンを惹起する(Nuwaysir, E. F. 他(1999) Mol. Carcinog. 24:153-159、Steiner, S. 及び N. L. Anderson (2000) Toxicol. Lett. 112-113:467-471、該文献は特に引用することを以って本明細書の一部となす)。試験化合物が、既知の毒性を有する化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子及び遺伝子ファミリーからの発現情報を含んでいる場合には、最も有用且つ正確である。理想的には、発現のゲノム全域にわたる測定が最高品質のシグネチャを提供する。たとえ、発現が任意の試験された化合物によって変化しない遺伝子があったとしても、それらの発現レベルを残りの発現データを標準化するために使用できるため、それらの遺伝子は重要である。標準化手順は、異なる化合物で処理した後の発現データの比較に有用である。毒性シグネチャの要素への遺伝子機能を割り当てることで毒性メカニズムの阻止に役立つが、毒性の予測につながるシグネチャの統計的に一致させるのに遺伝子機能の知識は必要とされない。(例えば2000年2月29日にNational Institute of Environmental Health Sciencesより発行されたPress Release 00-02を参照されたい。これについては<http://www.niehs.nih.gov/oc/news/toxchip.htm>で入手可能である)。従って、毒物学的スクリーニングの際に毒性シグネチャを用いて、全ての発現した遺伝子配列を含めることは重要且つ望ましいことである。

## 【0231】

或る実施例では、核酸を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。処理した生物学的サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1つ若しくは複数のプローブでハイブリダイズし、それによって本発明のポリヌクレオチドに対応する転写レベルを定量し得る。処理した生物学的サンプル中の転写レベルを、未処理生物学的サンプル中のレベルと比較する。両サンプルの転写レベルの差は、処理されたサンプル中で試験化合物が引き起こす毒性反応を示す。

## 【0232】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞タイプのプロテオームを分析することに関連する。プロテオームの語は、特定の組織または細胞タイプでのタンパク質発現の包括的パターンを指す。プロテオームの各タンパク質成分は、個々に更に分析の対象とすることができる。プロテオーム発現パターン即ちプロフィールは、所与の条件下で所与の時間に発現したタンパク質の数及び相対存在量を定量することにより分析し得る。従って細胞のプロテオームのプロフィールは、特定の組織または細胞タイプのポリペプチドを分離及び分析することにより作成し得る。或る実施例では、1次元等電点電気泳動によりサンプルからタンパク質を分離し、2次元ドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に応じて分離するような2次元ゲル電気泳動により分離が達成される(前出のSteiner および Anderson)。タンパク質は、通常クーマシーブルーまたはシルバーまたは蛍光染色などの物質を用いてゲルで染色することにより、分散した、独特な位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または未処理のいずれかの生物学的サンプルから得られるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシーケンシングする。スポット内のタンパク質の素性は、その部分配列を、好適には少なくとも5個の連続するアミノ酸残基を、本発明

のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

【0233】

タンパク質の ( p r o t e o m i c ) プロフィールは、G C R E C に特異的な抗体を用いて G C R E C 発現レベルを定量することによっても作成し得る。或る実施例では、マイクロアレイ上でエレメントとして抗体を用い、マイクロアレイをサンプルに曝して各アレイ要素へのタンパク質結合レベルを検出することによりタンパク質発現レベルを定量する ( L u e k i n g , A . 他 ( 1 9 9 9 ) *Anal. Biochem.* 270 : 103 - 111、M e n d o z e , L . G . 他 ( 1 9 9 9 ) *Biotechniques* 27 : 778 - 788 )。検出は当分野で既知の様々な方法で行うことができ、例えば、チオールまたはアミノ反応性蛍光化合物を用いてサンプル中のタンパク質を反応させ、各アレイのエレメントにおける蛍光結合の量を検出し得る。

10

【0234】

プロテオームレベルでの毒性シグネチャも中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行に分析するべきである。或る組織の或るタンパク質に対しては、転写とタンパク質の存在量の相関が乏しいこともあるので ( A n d e r s o n , N . L . 及び J . S e i l h a m e r ( 1 9 9 7 ) *Electrophoresis* 18 : 533 - 537 )、転写イメージにはそれ程影響しないがタンパク質のプロフィールを変化させるような化合物の分析においてプロテオーム毒性シグネチャは有用たり得る。更に、体液中での転写の分析は m R N A 急速な分解により困難であるので、タンパク質のプロフィール作成はこのような場合により信頼でき、情報価値がある。

20

【0235】

別の実施例では、タンパク質を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。処理された生物学的サンプル中で発現したタンパク質は、各タンパク質の量を定量し得るように分離する。各タンパク質の量を、未処理生物学的サンプル中の対応するタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する反応を示す。個々のタンパク質は、個々のタンパク質のアミノ酸残基をシークエンシングし、これら部分配列を本発明のポリペプチドと比較することにより同定する。

【0236】

別の実施例では、タンパク質を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。生物学的サンプルから得たタンパク質は、本発明のポリペプチドに特異的な抗体を用いてインキュベートする。抗体により認識されたタンパク質の量を定量する。処理された生物学的サンプル中のタンパク質の量を、未処理生物学的サンプル中のタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する毒性反応を示す。

30

【0237】

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして分析する ( B r e n n a n , T . M . 他 ( 1 9 9 5 ) の米国特許第 5 , 4 7 4 , 7 9 6 号、S c h e n a , M . 他 ( 1 9 9 6 ) *Proc. Natl. Acad. Sci. USA* 93 : 10614 - 10619、B a l d e s c h w e i l e r 他 ( 1 9 9 5 ) P C T 出願第 W O 9 5 / 2 5 1 1 1 6 号、S h a l o n , D . 他 ( 1 9 9 5 ) P C T 出願第 W O 9 5 / 3 5 5 0 5 号、H e l l e r , R . A . 他 ( 1 9 9 7 ) *Proc. Natl. Acad. Sci. USA* 94 : 2150 - 2155、H e l l e r , M . J . 他 ( 1 9 9 7 ) 米国特許第 5 , 6 0 5 , 6 6 2 号等を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, 編集。(1999) Oxford University Press, London に記載されている。該文献は、特に引用することを以って本明細書の一部となす。

40

【0238】

50

本発明の別の実施例では、天然のゲノム配列をマッピングする際に有用なハイブリダイゼーションプローブを作製するために、GCRECをコードする核酸配列を用いることが可能である。コード配列または非コード配列のいずれかを用いることができ、或る例では、コード配列よりも非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。核酸配列は、特定の染色体、染色体の特定領域または人工形成の染色体、例えば、ヒト人工染色体(HAC)、酵母人工染色体(YAC)、細菌人工染色体(BAC)、細菌P1産物、或いは単一染色体cDNAライブラリに対してマッピングされる。(例えば、Harrington, J. J. 他(1997) Nat. Genet. 15:345-355; Price, C. M. (1993) Blood Rev. 7:127134; および Trask, B. J. (1991) Trends Genet. 7:149154を参照。)一度マッピングすると、本発明の核酸配列を用いて例えば病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型(RFLP)と関連させるような遺伝子連鎖地図を作製できる。(例えば、Lander, E. S. および D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照。)

#### 【0239】

蛍光原位ハイブリッド形成法(FISH)は、他の物理的及び遺伝地図データと関連し得る(前出のHeinz-Ulrich, ら(1995) in Meyers, 965-968頁等を参照。)遺伝地図データの例は、種々の科学雑誌あるいはOnline Mendelian Inheritance in Man(OMIM)のウェブサイトに見ることができる。物理的染色体地図上のGCRECをコードする遺伝子の位置と特定の疾患との相関性或いは特定の疾患に対する素因は、その疾患に関係するDNAの領域を画定するのに役立つものであり、従って更に位置クローニングする試みとなり得る。

#### 【0240】

確定した染色体マーカーを用いた連鎖分析等の物理的マッピング技術及び染色体標本原位ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳動物の染色体上に遺伝子を配置することにより、特定のヒト染色体の数或いはアームが分かっていない場合でも関連するマーカーを明らかにし得る。この情報は、位置クローニングその他の遺伝子発見技術を用いて遺伝的疾患を調査する研究者にとって価値がある。いったん、疾患または症候群の原因となる遺伝子が、血管拡張性失調症の11q22-23領域等、特定の遺伝子領域への遺伝的結合によって大まかに位置決めがなされると、該領域に対するいかなるマッピングも、更なる調査のための関連遺伝子或いは調節遺伝子を表すことができる(Gatti, R. A. 他(1988) Nature 336:577-580等を参照)転座、反転等に起因する、健常者、保有者、感染者の三者間における染色体位置の相違を発見するために本発明のヌクレオチド配列を用いてもよい。

#### 【0241】

本発明の別の実施例では、種々の薬剤スクリーニング技術を以って化合物のライブラリをスクリーニングするために、GCREC、GCRECの触媒作用断片、免疫原断片、またはそのオリゴペプチドを用いることができる。薬剤スクリーニングに用いる断片は、溶液中に遊離しているか、固体支持物に固定されるか、細胞表面上に保持されるか、細胞内に位置することになる。GCRECとテストされる薬剤との結合複合の形成は計測できる。

#### 【0242】

別の薬剤スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化合物を高い処理能力でスクリーニングするために用いられる(Geyesen, 他(1984) PCT application WO84/03564等を参照。)この方法においては、多数の異なる小さな試験用化合物を固体基板上で合成する。試験用化合物は、GCREC或いはその断片と反応させ、洗浄する。次に、本技術分野でよく知られている方法で、結合したGCRECを検出する。精製したGCRECはまた、上記した薬剤の

スクリーニング技術において用いるプレート上で直接コーティングすることもできる。別法では、非中和抗体を用いてペプチドを捕捉し、ペプチドを固体支持物に固定することもできる。

【0243】

別の実施例では、競合薬スクリーニングアッセイを用いることができる。このアッセイでは、GCRECを結合できる中和抗体が、GCRECを結合するための試験化合物と特異的に競合する。この方法では、抗体が、1若しくは数個の抗原決定因子をGCRECと共有するペプチドの存在を検出する。

【0244】

別の実施例では、新規技術が現在知られているヌクレオチド配列の特性（限定するものではないがトリプレット遺伝暗号及び特異的塩基対の相互作用等を含む）に依存するのであれば、まだ発展途上である分子生物学技術のいずれにおいても、GCRECをコードするヌクレオチド配列を用いることができる。 10

【0245】

更に詳細説明をしなくとも、当業者であれば以上の説明を以って本発明を最大限に利用できるであろう。従って、これ以下に記載する実施例は単なる例示目的にすぎず、いかようにも本発明を限定するものではない。

【0246】

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第60/236,546号、第60/240,589号、第60/242,223号、第60/242,322号、第60/245,855号、第60/245,900、第60/247,587および第60/249,343は、言及することをもって本明細書の一部となす。 20

【0247】

(実施例)

1 cDNAライブラリの作製

Incyte cDNAは、LIFESEQ GOLDデータベース(Incyte Genomics, Palo Alto CA)に記載されたcDNAライブラリに由来するものであり、表4の列5に列記した。ホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解した組織もあり、また、ホモジナイズしてフェノールまたは好適な変性剤の混合液に溶解した組織もある。変性剤の混合液は、例えばフェノールとグアニジニウムイソチオシアネートの単相溶液であるTRIZOL(Life Technologies)等である。結果として得られた溶解物は、塩化セシウムにおいて遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、或いは別の方法を用いて、溶解物からRNAを沈殿させた。 30

【0248】

RNAの純度を高めるため、RNAのフェノールによる抽出及び沈殿を必要な回数繰り返した。場合によっては、DNアーゼでRNAを処理した。殆どのライブラリでは、オリゴd(T)連結常磁性粒子(Promega)、OLIGOTEXラテックス粒子(QIAGEN, Valencia CA)またはOLIGOTEX mRNA精製キット(QIAGEN)を用いて、ポリ(A+) RNAを単離した。別法では、別のRNA単離キット、例えばPOLY(A)PURE mRNA精製キット(Ambion, Austin TX)を用いて組織溶解物からRNAを直接単離した。 40

【0249】

場合によってはStratagene社へのRNA提供を行い、対応するcDNAライブラリをStratagene社が作製することもあった。そうでない場合は、UNIZAPベクターシステム(Stratagene)またはSUPERSCRIPTプラスミドシステム(Life Technologies)を用いて本技術分野で公知の推奨方法または類似の方法でcDNAを合成し、cDNAライブラリを作製した(前出のAusubel, 1997, unit 5.1-6.6等を参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本 50

鎖 cDNA に連結反応させ、好適な制限酵素で cDNA を消化した。殆どのライブラリに対して、cDNA のサイズ (300 ~ 1000 bp) 選択は、SEPHACRYL S1000、SEPHAROSE CL2B または SEPHAROSE CL4B カラムクロマトグラフィー (Amersham Pharmacia Biotech)、或いは調製用アガロースゲル電気泳動法を用いて行った。合成オリゴヌクレオチドアダプターを二本鎖 cDNA に連結反応させ、好適な制限酵素または酵素で cDNA を消化した。好適なプラスミドは、例えば PBLUESCRIPT プラスミド (Stratagene)、PSPORT1 プラスミド (Life Technologies)、PCDNA2.1 プラスミド (Invitrogen, Carlsbad CA)、PBK-CMV プラスミド (Stratagene)、PCR2-TOPOTA プラスミド (Invitrogen)、PCMV-ICIS プラスミド (Stratagene)、pIGEN (Incyte Genomics, Palo Alto CA)、pINCY (Incyte Genomics) 等およびその誘導體である。組換えプラスミドは、Stratagene 社の XL1-Blue、XL1-Blue MRF または SOLR、或いは Life Technologies 社の DH5、DH10B または ELECTROMAX DH10B を含む大腸菌細胞に形質転換した。

10

20

30

40

50

#### 【0250】

##### 2 cDNA クローンの単離

UNIZAP ベクターシステム (Stratagene) を用いた *in vivo* 切除によって、或いは細胞溶解によって、実施例 1 のようにして得たプラスミドを宿主細胞から回収した。Magic または WIZARD Minipreps DNA 精製システム (Promega)、AGTC Miniprep 精製キット (Edge Biosystems, Gaithersburg MD)、QIAGEN 社の QIAWELL 8 Plasmid、QIAWELL 8 Plus Plasmid 及び QIAWELL 8 Ultra Plasmid 精製システム、R.E.A.L. Prep 96 プラスミドキットの中から少なくとも 1 つを用いて、プラスミドを精製した。沈殿させた後、0.1 ml の蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥せずに、4 で保管した。

#### 【0251】

別法では、高処理フォーマットにおいて直接結合 PCR 法を用いて宿主細胞溶解物からプラスミド DNA を増幅した (Rao, V. B. (1994) Anal. Biochem. 216: 1-14)。宿主細胞の溶解及び熱サイクリング過程は、単一反応混合液中で行った。サンプルを処理し、それを 384 穴プレート内で保管し、増幅したプラスミド DNA の濃度を PICO GREEN 色素 (Molecular Probes, Eugene OR) 及び FLUOROSKANI I 蛍光スキャナ (Lab systems Oy, Helsinki, Finland) を用いて蛍光分析的に定量した。

#### 【0252】

##### 3 シークエンシング及び分析

実施例 2 に記載したようにプラスミドから回収した Incyte cDNA を、以下に示すようにシークエンシングした。cDNA のシークエンス反応は、標準的方法或いは高処理装置、例えば ABI CATALYST 800 サーマルサイクラー (Applied Biosystems) または PTC-200 サーマルサイクラー (MJ Research) を HYDRA マイクロディスペンサー (Robbins Scientific) または MICROLAB 2200 (Hamilton) 液体転移システムと併用して処理した。cDNA のシークエンス反応は、Amersham Pharmacia Biotech 社が提供する試薬、または ABI シークエンシングキット、例えば ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems) の試薬を用いて準備した。cDNA のシークエンス反応の電気泳動的分離及び標識したポリヌクレオチドの検出には、MEGABACE 1000 DNA シークエンシングシステム (Molecular Dynamics) か、標準 ABI プロトコル及び塩基対呼び出しソフトウ

エアを用いるABI PRISM 373または377シーケンシングシステム (Applied Biosystems) か、或いはその他の本技術分野でよく知られている配列解析システムを用いた。cDNA配列内のリーディングフレームは、標準的方法 (前出のAusubel, 1997, ユニット7.7に概説) を用いて決定した。cDNA配列の幾つかを選択して、実施例8に記載した方法で配列を伸長させた。

#### 【0253】

Incyte cDNA配列に由来するポリヌクレオチド配列は、ベクター、リンカー及びポリ(A)配列を除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、BLAST、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びプログラムを用いた。公共のデータベース、例えばGenBankの霊長類及びげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM、及びPFAM等のような隠れマルコフモデル(HMM)ベースのタンパク質ファミリーデータベースの選択に対して、Incyte cDNA配列またはその翻訳を問い合わせた(HMMは、遺伝子ファミリーのコンセンサス1次構造を分析する確率的アプローチである。例えば、Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365を参照のこと)。問合せは、BLAST、FASTA、BLIMPS及びHMMERに基づくプログラムを用いて行った。Incyte cDNA配列は、完全長のポリヌクレオチド配列を産出するように構築された。或いは、GenBank cDNA、GenBank EST、ステッチされた配列、ストレッチされた配列またはGenscan予測コード配列 (実施例4及び5を参照) を用いてIncyte cDNAの集団を完全長まで伸長させた。Phred、Phrap及びConsedに基づくプログラムを用いて構築し、GenMark、BLAST及びFASTAに基づくプログラムを用いてcDNAの集団をオープンリーディングフレームに対してスクリーニングした。対応する完全長ポリペプチド配列を誘導するべく完全長ポリヌクレオチド配列を翻訳した。或いは、本発明のポリペプチドは完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。引き続き、GenBankタンパク質データベース(genpept)、SwissProt、BLOCKS、PRINTS、DOMO、PRODOM及びPrositate等のデータベース、PFAM等の隠れマルコフモデル(HMM)ベースのタンパク質ファミリーデータベースに対する問合せによって完全長ポリペプチド配列を分析した。完全長ポリヌクレオチド配列はまた、MACDNASIS PROソフトウェア(日立ソフトウェアエンジニアリング, South San Francisco CA)及びLASERGENEソフトウェア(DNASTAR)を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算するMEGALIGNマルチシーケンシングアラインメントプログラム(DNASTAR)に組み込まれているようなCLUSTALアルゴリズムによって特定されるデフォルトパラメータを用いて生成する。

#### 【0254】

Incyte cDNA及び完全長配列の分析及びアセンブリに利用したツール、プログラム及びアルゴリズムの概略と、適用可能な説明、引用文献、閾値パラメータを表7に示す。用いたツール、プログラム及びアルゴリズムを表5の列1に、それらの簡単な説明を列2に示す。列3は好適な引用文献であり、全ての文献はそっくりそのまま引用を以って本明細書の一部となす。適用可能な場合には、列4は2つの配列が一致する強さを評価するために用いたスコア、確率値その他のパラメータを示す(スコアが高ければ高いほど2配列間の相同性が高くなる)。

#### 【0255】

完全長ポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO: 17-32のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20~約4000ヌクレオチドの断片を表4の列4に示した。

## 【0256】

Genscan 遺伝子同定プログラムを公衆のゲノム配列データベース（例えば gbpr i 及び gbhtg）に対して実行することにより、推定上の G タンパク質共役受容体を先 10  
 ず同定する。Genscan は、様々な生物からゲノム DNA 配列を分析する汎用遺伝子  
 同定プログラムである（Burge, C. 及び S. Karlin (1997) J.  
 Mol. Biol. 268:78-94、Burge, C. 及び S. Karli  
 n (1998) Cuff. Opin. Struct. Biol. 8:346-35  
 4 参照）。プログラムは予測エキソンを連結し、メチオニンから終止コドンに及ぶ構築さ  
 れた cDNA 配列を形成する。Genscan の出力は、ポリヌクレオチド及びポリペプ  
 チド配列の FASTA データベースである。Genscan が一度に分析する配列の最大  
 範囲は、30 kb に設定した。いずれの Genscan 予測 cDNA 配列が G タンパク質  
 共役受容体をコードするかを決定するために、コードされたポリペプチドを G タンパク質  
 共役受容体のための PFAM モデルに対して問い合わせることにより分析した。Incy  
 te cDNA 配列の相同体を G タンパク質共役受容体として注釈を付けてきたことによ  
 り、潜在的 G タンパク質共役受容体も同定した。こうして選択された Genscan 予測  
 配列は、次に BLAST 分析により公共データベース gbpr i 及び gbhtg と比較し  
 た。必要であれば、genpept からヒットしたトップの BLAST と比較すること  
 により Genscan 予測配列を編集し、余分なまたは取り除かれたエキソンなどの Gen  
 scan により予測された配列のエラーを修正する。BLAST 分析はまた、任意の In  
 cyte cDNA または Genscan 予測配列の公共 cDNA 適用範囲の発見に用い 20  
 られるので、転写の証拠を提供する。Incyte cDNA 適用範囲が利用できる場合  
 には、この情報を用いて Genscan 予測配列を修正または確認した。完全長ポリヌク  
 レオチド配列は、実施例 3に記載された構築プロセスを用いて、Incyte cDNA  
 配列及び/または公共の cDNA 配列で Genscan 予測コード配列を構築することに  
 より得た。或いは、完全長ポリヌクレオチド配列は編集または非編集の Genscan 予  
 測コード配列に完全に由来する。

## 【0257】

5 cDNA 配列データを使ったゲノム配列データの構築  
ステッチ配列 (Stitched Sequence)

部分 cDNA 配列は、実施例 4に記載の Genscan 遺伝子同定プログラムにより予測 30  
 されたエキソンを用いて伸長させた。実施例 3に記載されたように構築された部分 cDN  
 A は、ゲノム DNA にマッピングし、関連する cDNA 及び 1 つ若しくは複数のゲノム配  
 列から予測された Genscan エキソンを含むクラスタに分解した。cDNA 及びゲノ  
 ム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて  
 各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するよう  
 な潜在的スプライシング変異体を生み出した。区間全体の長さがクラスタ中の 2 以上の配列  
 に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えら  
 れた。例えば、1 つの cDNA 及び 2 つのゲノム配列に区間が存在する場合、3 つの区間  
 は全て等しいと考えられる。このプロセスは、無関係であるが連続したゲノム配列を cD  
 NA 配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列 (p  
 arent sequence) に沿って現われるようにステッチアルゴリズムで縫い合 40  
 わせ、可能な最も長い配列および変異配列を作製する。1 種類の親配列に沿って発生した  
 間隔と間隔との連鎖 (cDNA - cDNA またはゲノム配列 - ゲノム配列) は、親の種類  
 を変える連鎖 (cDNA - ゲノム配列) に優先させた。結果として得られるステッチ配列  
 は、BLAST 分析により公共データベース genpept 及び gbpr i に翻訳されて  
 比較された。Genscan により予測された不正確なエキソンは、genpept から  
 ヒットしたトップの BLAST と比較することにより修正した。必要な場合には、追加 c  
 DNA 配列を用いるかゲノム DNA の検査により配列を更に伸長させた。

## 【0258】

ストレッチ配列 (Stretched Sequence)

部分DNA配列は、BLAST分析に基づくアルゴリズムにより完全長まで伸長された。まず、BLASTプログラムを用いて、GenBankの霊長類、げっ歯類、哺乳動物、脊椎動物及び真核生物のデータベースなどの公共データベースに対し、実施例3に記載されたように構築された部分cDNAを問い合わせた。次に、最も近いGenBankタンパク質相同体をBLAST分析によりIncyte cDNA配列または実施例4に記載のGenScanエキソン予測配列のいずれかと比較した。結果として得られる高スコアリングセグメント対(HSP)を用いてキメラタンパク質を産出し、翻訳した配列をGenBankタンパク質相同体上にマッピングした。元のGenBankタンパク質相同体に関連して、キメラタンパク質内で挿入または削除が起こり得る。GenBankタンパク質相同体、キメラタンパク質またはその両方をプローブとして用い、公共のヒトゲノムデータベースから相同ゲノム配列を検索した。このようにして、部分的なDNA配列を相同ゲノム配列の付加によりストレッチすなわち伸長した。結果として得られるストレッチ配列を検査し、完全遺伝子を含んでいるか否かを決定した。

10

【0259】

6 GCRECをコードするポリヌクレオチドの染色体マッピング  
SEQ ID NO: 17-32を構築するために用いた配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、Incyte LIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 17-32と一致するこれらのデータベースの配列を、Phrap(表7)などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。スタンフォード・ヒトゲノムセンター(SHGC)、ホワイトヘッド・ゲノム研究所(WIGR)、Genethon等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスタ化された配列が前もってマッピングされたかを決定した。マッピングされた配列がクラスタに含まれている場合は、個々の配列番号を含めてそのクラスタの全配列が地図上の位置に割り当てられた。

20

【0260】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体のpアームの末端に関連して測定する。(センチモルガン(cM)は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1cMは、ヒト中のDNAの1メガベース(Mb)にほぼ等しい。尤も、この値は、組換えのホットスポット及びコールドスポットに起因して広範囲に変化する。)cM距離は、配列が各クラスタ内に含まれるような放射線ハイブリッドマーカーに対して境界を提供するようなGenethonによってマッピングされた遺伝マーカーに基づく。NCBI「GeneMap99」(<http://www.ncbi.nlm.nih.gov/genemap>)などの一般個人が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、上記した区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

30

7 ポリヌクレオチド発現の分析

ノーザン分析は、転写された遺伝情報の存在を検出するために用いられる実験技術であり、特定の細胞種または組織からのRNAが結合される膜への標識されたヌクレオチド配列のハイブリダイゼーションに関与している(前出のSambrook, 7章、前出のAusubel, F.M. 他, 4章及び16章等を参照)。

40

【0261】

BLASTに適用する類似のコンピュータ技術を用いて、GenBankやLifeSeq(Incyte Pharmaceuticals)等のヌクレオチドデータベースにおいて同一または関連分子を検索する。ノーザン分析は、多数の膜系ハイブリダイゼーションよりも非常に速い。更に、特定の同一を厳密な或いは相同的なものとして分類するかどうかを決定するため、コンピュータ検索の感度を変更することができる。検索の基準は積スコアであり、次式で定義される。

【数1】

## (BLASTスコア×配列一致率)

---

5 × (長さ(配列1), 長さ(配列2))の最小値

## 【0262】

積スコアは、2つの配列間の類似度及び配列が一致する長さの両方を考慮している。積スコアは、0～100の規準化された値であり、次のようにして求める。BLASTスコアにヌクレオチドの配列一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。高スコアリングセグメント対(HSP)に一致する各塩基に+5のスコアを割り当て、各不適性塩基対に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る(ギャップにより隔離され得る)。2以上のHSPがある場合には、最高BLASTスコアの塩基対を用いて積スコアを計算する。積スコアは、断片的重畳とBLASTアラインメントの質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合のみ得られる。積スコア70は、一端が100%一致し、70%重畳しているか、他端が88%一致し、100%重畳しているかのいずれかの場合に得られる。積スコア50は、一端が100%一致し、50%重畳しているか、他端が79%一致し、100%重畳しているかのいずれかの場合に得られる。

10

## 【0263】

或いは、GCRECをコードするポリヌクレオチド配列を、ポリヌクレオチド配列が派生した組織源に関連して分析した。例えば或る完全長配列は、Incyte cDNA配列(実施例3を参照)と少なくとも一部は重畳するように構築される。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、以下の生物/組織カテゴリー即ち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液及び免疫系、肝、筋骨格系、神経系、脾臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性/混合性または尿路の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患/病状カテゴリー即ち癌、細胞系、発達、炎症、神経性、外傷、心血管、鬱血、その他の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの

20

30

## 【0264】

## 8 GCRECをコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列もまた、完全長分子の適切な断片から設計したオリゴヌクレオチドプライマーを用いて該断片を伸長させて生成した。一方のプライマーは既知の断片の5'伸長を開始するべく合成し、他方のプライマーは既知の断片の3'伸長を開始するべく合成した。開始プライマーは、長さが約22～30ヌクレオチド、GC含有率が約50%以上となり、約68～72の温度で標的配列にアニーリングするように、OLIGO 4.06ソフトウェア(National Biosciences)或いは別の適切なプログラムを用いて、cDNAから設計した。ヘアピン構造及びプライマー-プライマー二量体を生ずるようなヌクレオチドの伸長は全て回避した。

40

## 【0265】

配列を伸長するために、選択されたヒトcDNAライブラリを用いた。2段階以上の伸長が必要または望ましい場合には、付加的プライマー或いはプライマーのネステッドセットを設計した。

## 【0266】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた

50

。PCRは、PTC-200サーマルサイクラー(MJ Research, Inc.)を用いて96穴プレート内で行った。反応混合液は、鋳型DNA及び200 nmolの各プライマー、 $Mg^{2+}$ と $(NH_4)_2SO_4$ と-メルカプトエタノールを含むバッファー、Taq DNAポリメラーゼ(Amersham Pharmacia Biotech)、ELONGASE酵素(Life Technologies)、Pfu DNAポリメラーゼ(Stratagene)を含む。プライマーの組、PCI AとPCI Bに対して以下のパラメータで増幅を行った。ステップ1: 94 で3分間 ステップ2: 94 で15秒 ステップ3: 60度で1分間 ステップ4: 68 で2分間 ステップ5: ステップ2、3、4を20回繰り返す ステップ6: 68 で5分間 ステップ7: 4 で保存。プライマー対T7、SK+に対しては、上記パラメータに代えて以下のパラメータを用いた。ステップ1: 94 で3分間 ステップ2: 94 で15秒 ステップ3: 57 で1分間 ステップ4: 68 で2分間 ステップ5: ステップ2、3、4を20回繰り返す ステップ6: 68 で5分間 ステップ7: 4 で保存。

10

20

30

40

50

**【0267】**

各ウェルのDNA濃度は、1X TE及び0.5  $\mu$ lの希釈していないPCR産物に溶解した100  $\mu$ lのPICOGREEN定量試薬(0.25 (v/v) PICOGREEN; Molecular Probes, Eugene OR)を不透明な蛍光光度計プレート(Coming Costar, Acton MA)の各ウェルに分配してDNAが試薬と結合できるようにして測定する。サンプルの蛍光を計測してDNAの濃度を定量するべくプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンした。反応混合物のアリコート5~10  $\mu$ lを1%アガロースミニゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを決定した。

**【0268】**

伸長させたヌクレオチドは、脱塩及び濃縮して384穴プレートに移し、CviJIコラウイルスエンドヌクラーゼ(Molecular Biology Research, Madison WI)を用いて消化し、pUC 18ベクター(Amersham Pharmacia Biotech)への再連結反応前に音波処理またはせん断した。シヨットガン・シークエンシングのために、消化したヌクレオチドを低濃度(0.6~0.8%)のアガロースゲル上で分離し、断片を切除し、寒天をAgar ACE(Promega)で消化した。伸長させたクローンをT4リガーゼ(New England Biolabs, Beverly MA)を用いてpUC 18ベクター(Amersham Pharmacia Biotech)に再連結し、Pfu DNAポリメラーゼ(Stratagene)で処理して制限部位の張出部(overhang)を満たし、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB/2Xカルベニシリン培養液の384ウェルプレートに37で一晩培養した。

**【0269】**

細胞を溶解して、Taq DNAポリメラーゼ(Amersham Pharmacia Biotech)及びPfu DNAポリメラーゼ(Stratagene)を用いて以下の手順でDNAをPCR増幅した。ステップ1 94 で3分間 ステップ2 94 で15秒 ステップ3 60 で1分間 ステップ4 72 で2分間 ステップ5 ステップ2、3、及び4を29回繰り返す ステップ6 72 で5分間 ステップ7 4 で保管。上記したようにPICOGREEN試薬(Molecular Probes)でDNAを定量した。DNAの回収率が低いサンプルは、上記と同一の条件を用いて再増幅した。サンプルは20%ジメチルスルホキシド(1:2, v/v)で希釈し、DYENAMIC エネルギートランスファー シークエンシングプライマー、及びDYENAMIC DIRECT kit(Amersham Pharmacia Biotech)またはABI PRISM BIGDYE ターミネーターサイクル シークエンシング反応キット(Terminator cycle sequencing ready reaction k

it) (Applied Biosystems) を用いてシーケンシングした。

【0270】

同様に、上記手順を用いて完全長ヌクレオチド配列を検証し、或いはそのような伸長のために設計されたオリゴヌクレオチド及び適切なゲノムライブラリを用いて5'調節配列を得る。

【0271】

9 個々のハイブリダイゼーションプローブの標識及び使用

SEQ ID NO: 17-32 から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しては事実上同一の手順が用いられる。オリゴヌクレオチドは、OLIGO 4.06 ソフトウェア (National Biosciences) 等の最新ソフトウェアを用いて設計し、各オリゴマー50 pmol と、[<sup>32</sup>P]アデノシン3リン酸 (Amersham Pharmacia Biotech) 250 μCi と、T4ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) を混合することにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25 超細繊維分子サイズ排除デキストラン ビードカラム (Amersham Pharmacia Biotech) を用いて十分に精製する。Ase I、Bgl II、Eco RI、Pst I、Xba I または Pvu II (DuPont NEN) のいずれか1つのエンドヌクレアーゼで消化されたヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において、毎分10<sup>7</sup> カウントの標識されたプローブを含むアリコットを用いる。

【0272】

各消化物から得たDNAは、0.7%アガロースゲル上で分画してナイロン膜 (Nyttran Plus, Schleicher & Schuell, Durham NH) に移す。ハイブリダイゼーションは、40°C で16時間行う。非特異的シグナルを除去するため、例えば0.1Mクエン酸ナトリウム食塩水及び0.5%ドデシル硫酸ナトリウムに一致する条件下で、プロットを室温で順次洗浄する。オートラジオグラフィまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

【0273】

10 マイクロアレイ

マイクロアレイの表面上でアレイエレメントの連鎖または合成は、フォトリソグラフィ、圧電印刷 (インクジェット印刷、前出の Baldeschweiler 等を参照)、機械的マイクロスポッティング技術及びこれらから派生したものを用いて達成することが可能である。上記各技術において基板は、均一且つ非多孔性の固体とするべきである (Schedena (1999) 前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。或いは、ドットプロット法またはスロットプロット法に類似のアレイを利用して、熱的、紫外線的、化学的または機械的結合手順を用いて基板の表面にエレメントを配置及び結合させてもよい。通常のアレイは、手作業で、または利用可能な方法や機械を用いて作製でき、任意の適正数のエレメントを有し得る (Schedena, M. 他 (1995) Science 270: 467-470、Shalton, D. 他 (1996) Genome Res. 6: 639-645、Marshall, A. および J. Hodgson (1998) Nat. Biotechnol. 16: 27-31. を参照)。

【0274】

完全長cDNA、発現配列タグ (EST)、またはその断片またはオリゴマーは、マイクロアレイのエレメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、レーザGENEソフトウェア (DNASTAR) 等の本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメントは、生物学的サンプル中でポリヌクレオチドを用いてハイブリダイズされる。生物学的サンプル中のポリヌクレオチド

は、検出を容易にするために蛍光標識またはその他の分子タグに抱合される。ハイブリダイゼーション後、生物学的サンプルからハイブリダイズされていないヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントにおいてハイブリダイゼーションを検出する。或いは、レーザ脱離及び質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のエレメントにハイブリダイズする各ポリヌクレオチドの相補性の度合及び相対存在度は、算定し得る。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。

#### 【0275】

##### 組織または細胞サンプルの準備

グアニジウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ(dT)セルロース法を用いてポリ(A)<sup>+</sup> RNAを精製する。各ポリ(A)<sup>+</sup> RNAサンプルを、MMLV逆転写酵素、0.05 pg/μlのオリゴ(dT)プライマー(21mer)、1×第一鎖合成バッファー、0.03 unit/μlのRNAアーゼ阻害因子、500 μMのdATP、500 μMのdGTP、500 μMのdTTP、40 μMのdCTP、40 μMのdCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。逆転写反応は、GEMBRIGHTキット(Incyte)を用いてポリ(A)<sup>+</sup> RNA含有の25体積ml内で行う。特異的制御ポリ(A)<sup>+</sup> RNAは、非コード酵母ゲノムDNAから*in vitro*転写により合成する。各反応サンプル(1つはCy3、もう1つはCy5標識)は、2.5 mlの0.5 M水酸化ナトリウムで処理し、85 °Cで20分間インキュベートし、反応を停止させてRNAを分解させる。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピカラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。結合後、2つの反応サンプルは、1 mlのグリコーゲン(1 mg/ml)を用いて析出させたエタノール、60 mlの酢酸ナトリウム及び300 mlの100%エタノールである。サンプルは次に、Speed VAC(Savant Instruments Inc., Holbrook NY)を用いて乾燥して仕上げ、14 μl 5×SSC/0.2% SDS中で再懸濁する。

#### 【0276】

##### マイクロアレイの準備

本発明の配列を用いて、アレイエレメントを作製する。各アレイエレメントは、クローン化cDNAインサートによりベクター含有細菌性細胞から増幅する。PCR増幅は、cDNAインサートの側面に位置するベクター配列に相補的なプライマーを用いる。30サイクルのPCRによって、1~2 ngの初期量から5 μgを超える最終量までアレイエレメントを増幅する。増幅されたアレイエレメントは、SEPHACRYL-400(Amersham Pharmacia Biotech)を用いて精製される。

#### 【0277】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス(Corning)は、処理中及び処理後に0.1%のSDS及びアセトン中で超音波をかけ、蒸留水で非常に良く洗って洗浄する。スライドガラスは、4%フッ化水素酸(VWR Scientific Products Corporation (VWR), West Chester PA)中でエッチングし、蒸留水中で非常に良く洗浄し、95%エタノール中で0.05%アミノプロピルシラン(Sigma)を用いてコーティングする。コーティングしたスライドガラスは、110 °Cのオープンで硬化させる。

#### 【0278】

米国特許第5,807,522号で説明されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。該特許は、引用を以って本明細書の一部となす。平均濃度が100 ng/μlのアレイエレメントDNA 1 μlを高速機械装置により開放型キャピラリープリンティングエレメント(open capillary print

10

20

30

40

50

ing element) に充填する。装置はここで、スライド毎に約 5 n l のアレイエレメントサンプルを加える。

#### 【0279】

マイクロアレイには、STRATALINKER UV 架橋剤 (Stratagene) を用いて UV 架橋する。マイクロアレイは、室温において 0.2% SDS で 1 度洗浄し、蒸留水で 3 度洗浄する。リン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) 中の 0.2% カゼイン中において 60 で 30 分間マイクロアレイをインキュベートした後、前に行ったように 0.2% SDS 及び蒸留水で洗浄することにより、非特異結合部位をブロックする。

#### 【0280】

##### ハイブリダイゼーション

ハイブリダイゼーション反応は、5 x SSC, 0.2% SDS ハイブリダイゼーション緩衝液中の Cy 3 及び Cy 5 標識した cDNA 合成生成物を各 0.2 µg 含む 9 µl のサンプル混合体を有する。サンプル混合体は、65 まで 5 分間加熱し、マイクロアレイ表面上で等分して 1.8 cm<sup>2</sup> のカバーガラスで覆う。アレイは、顕微鏡用スライドより僅かに大きい空洞を有する防水チェンバーに移す。チェンバーのコーナーに 140 µl の 5 x SSC を加えることにより、チェンバー内部を湿度 100% に保持する。アレイを含むチェンバーは、60 で約 6.5 時間インキュベートする。アレイは、第 1 洗浄緩衝液中 (1 x SSC, 0.1% SDS) において 45 で 10 分間洗浄し、第 2 洗浄緩衝液中 (0.1 x SSC) において 45 で 10 分間各々 3 度洗浄して乾燥させる。

#### 【0281】

##### 検出

レポーター標識ハイブリダイゼーション複合体は、Cy 3 の励起のためには 488 nm、Cy 5 の励起のためには 632 nm でスペクトル線を発生し得る Innova 70 混合ガス 10 W レーザ (Coherent, Inc., Santa Clara CA) を備えた顕微鏡で検出する。20 x 顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いて、アレイ上に励起レーザ光の焦点を当てる。アレイを含むスライドを顕微鏡のコンピュータ制御の X-Y ステージに置き、対物レンズを通過してラスタースキャンする。本実施例で用いた 1.8 cm x 1.8 cm のアレイは、20 µm の解像度でスキャンした。

#### 【0282】

2 つの異なるスキャンで、混合ガスマルチラインレーザは 2 つの蛍光色素を連続的に励起する。発光された光は、波長に基づき分離され、2 つの蛍光色素に対応する 2 つの光電子増倍管検出器 (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) に送られる。アレイと光電子増倍管間に設置された好適なフィルターを用いて、シグナルをフィルターする。用いる蛍光色素の最大発光の波長は、Cy 3 では 565 nm、Cy 5 では 650 nm である。装置は両方の蛍光色素からのスペクトルを同時に記録し得るが、レーザ源において好適なフィルターを用いて、蛍光色素 1 つにつき 1 度スキャンし、各アレイを通常 2 度スキャンする。

#### 【0283】

スキャンの感度は通常、既知濃度のサンプル混合体に添加される cDNA 対照種により生成されるシグナル強度を用いて較正する。アレイ上の特定の位置には相補的 DNA 配列が含まれ、その位置におけるシグナルの強度をハイブリダイジング種の重量比 1:100,000 に相関させる。異なる源泉 (例えば試験される細胞及び対照細胞など) からの 2 つのサンプルを、各々異なる蛍光色素で標識し、他と異なって発現した遺伝子を同定するために単一のアレイにハイブリダイズする場合には、その較正を、2 つの蛍光色素で較正する cDNA のサンプルを標識し、ハイブリダイゼーション混合体に各々等量を加えることによって行う。

#### 【0284】

光電子増倍管の出力は、IBM コンパチブル PC コンピュータにインストールされた 12

10

20

30

40

50

ビットRTI-835Hアナログ-デジタル(A/D)変換ボード(Analog Devices, Inc., Norwood MA)を用いてデジタル化される。デジタル化されたデータは、青色(低シグナル)から赤色(高シグナル)までの擬似カラー範囲へのリニア20色変換を用いてシグナル強度がマッピングされたようなイメージとして表示される。データは、定量的にも分析される。2つの異なる蛍光色素を同時に励起及び測定する場合には、各蛍光色素の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク(発光スペクトルの重なりに起因する)を補正する。

#### 【0285】

グリッドが蛍光シグナルイメージ上に重ねられ、それによって各スポットからのシグナルはグリッドの各エレメントに集められる。各エレメント内の蛍光シグナルは統合され、シグナルの平均強度に応じた数値が得られる。シグナル分析に用いるソフトウェアは、GEMTOOLS遺伝子発現分析プログラム(Incyte)である。

#### 【0286】

##### 11 相補的ポリヌクレオチド

GCRECをコードする配列或いはその任意の一部に対して相補的配列は、天然のGCRECの発現を検出し、低下させ、または阻害するために用いられる。約15~30塩基対を含むオリゴヌクレオチドの使用について記すが、これより小さな或いは大きな配列の断片の場合でも本質的に同じ方法を用いることができる。OLIGO4.06ソフトウェア(National Biosciences)とGCRECをコードする配列とを用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な5'配列から相補的オリゴヌクレオチドを設計し、これを用いてプロモーターがコーディング配列に結合するのを阻害する。翻訳を阻害するためには、GCRECをコードする転写物にリボソームが結合しないように相補的オリゴヌクレオチドをデザインする。

#### 【0287】

##### 12 GCRECの発現

GCRECの発現及び精製は、細菌またはウイルスをベースにした発現系を用いて行うことができる。細菌でGCRECを発現するために、抗生物質耐性遺伝子及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクローニングする。このようなプロモーターには、lacオペレーター調節エレメントに関連するT5またはT7バクテリオファージプロモーター及びtrp-lac(tac)ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、BL21(DE3)等の好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル-Dチオガラクトピラノシド(IPTG)で誘導されるとGCRECを発現する。真核細胞でのGCRECの発現は、一般にバキュロウイルスとして知られているAutographica californica核多面性ウイルス(ACMNPV)を昆虫細胞株または哺乳動物細胞株に感染させて行う。バキュロウイルスの可欠ポリヘドリン遺伝子を、相同的組換え、或いは転移プラスミドの仲介に關与する細菌仲介遺伝子転移のどちらかによって、GCRECをコードするcDNAと置換する。ウイルスの感染力は維持され、強い多角体プロモーターによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合はSpodoptera frugiperda(Sf9)昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる(Engelhard, E. K. 他(1994) Proc. Natl. Acad. Sci. USA 91:3224-3227, Sandig, V. 他(1996) Hum. Gene Ther. 7:1937-1945.等を参照)。

#### 【0288】

殆どの発現系では、融合タンパク質としてGCRECを合成するのに例えばグルタチオンSトランスフェラーゼ(GST)またはペプチドエピトープ標識、例えばFLAGや6-Hisを用いる。これらを用いることにより、未精製細胞溶解物から組換え融合タンパク質の親和性ベースの精製を迅速に1ステップで行うことができる。GSTは日本住血吸

虫からの26 kDaの酵素であり、タンパク質の活性及び抗原性を維持した状態で、固定化グルタチオン上で融合タンパク質の精製を可能とする(Amersham Pharmacia Biotech)。精製後、GSTの部分特定の開発部位においてGCRCからタンパク質的に切断することが可能である。FLAGは8アミノ酸のペプチドであり、市販されているモノクローナル及びポリクローナル抗FLAG抗体(Eastman Kodak)を用いて免疫親和性精製を可能にする。6ヒスチジン残基が連続して伸長した6-Hisは、金属キレート樹脂(QIAGEN)上での精製を可能にする。タンパク質の発現及び精製の方法は、前出のAusubel(1995)10章、16章に記載されている。これらの方法で精製したGCRCを直接用いて、適用可能な場合には実施例16、17及び18のアッセイを行うことができる。

10

## 【0289】

## 13 機能的アッセイ

GCRC機能は、哺乳動物細胞培養系において生理学的に高められたレベルでのGCRCをコードする配列の発現によって評価する。cDNAを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにcDNAをサブクローニングする。選択されるベクターには、PCMV SPORTプラスミド(Life Technologies)及びPCR3.1(Invitrogen)が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤或いは電気穿孔法を用いて、5~10 µgの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に一時的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2 µgのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質(GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー(FCM)を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するか或いは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるDNA染色によって計測される核DNA内容物の変化、前方散乱光と90°側方散乱光によって計測される細胞サイズと顆粒状性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞内におけるタンパク質の発現の変化、及び蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。フローサイトメトリー法については、Ormerod, M. G. (1994) Flow Cytometry Oxford, New York, NY. に記述がある。

20

30

## 【0290】

遺伝子発現におけるGCRCの影響は、GCRCをコードする配列とCD64またはCD64-GFPのいずれかが形質移入された高度に精製された細胞集団を用いて評価することができる。CD64またはCD64-GFPは、形質転換された細胞表面で発現し、ヒト免疫グロブリンG(IgG)の保存領域と結合する。形質転換細胞と非形質転換細胞は、ヒトIgGまたはCD64に対する抗体(DYNAL, Lake Success, NY)で覆われた磁気ビーズを用いて有効に分離することができる。mRNAは、本技術分野で公知の方法で細胞から精製することができる。GCRCその他の目的の遺伝子をコードするmRNAの発現は、ノーザン分析或いはマイクロアレイ技術で分析することができる。

40

## 【0291】

## 14 GCRC特異抗体の産生

ポリアクリルアミドゲル電気泳動法(PAGE; Harrington, M. G. (1990) Methods Enzymol. 182: 488-495等を参照)または

50

他の精製技術を用いて実質上精製されたGCRECを用いて、標準プロトコルでウサギを免疫化して抗体を産出する。

【0292】

或いは、レーザGENEソフトウェア(DNASTAR)を用いてGCRECアミノ酸配列を解析し、免疫原性の高い領域を決定する。そして対応するオリゴペプチドを合成し、このオリゴペプチドを用いて当業者によく知られている方法で抗体を生成する。適切なエピトープ、例えばC末端付近或いは隣接する親水性領域にあるエピトープの選択については、当分野で公知である(前出のAusubel, 1995, 11章等を参照)。

【0293】

通常は、長さ約15残基のオリゴペプチドを、Fmocケミストリを用いるABI 431Aペプチドシンセサイザ(Applied Biosystems)を用いて合成し、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル(MBS)を用いた反応によってKLH(Sigma-Aldrich, St. Louis MO)に結合させて、免疫原性を高める(前出のAusubel, 1995等を参照)。完全フロイントアダプタントにおいてオリゴペプチド-KLH複合体を用いてウサギを免疫化する。得られた抗血清の抗ペプチド活性及び抗GCREC活性を検査するには、ペプチドまたはGCRECを基質に結合し、1%BSAを用いてブロックし、ウサギ抗血清と反応させて洗浄し、さらに放射性ヨウ素で標識したヤギ抗ウサギIgGと反応させる。

【0294】

15 特異抗体を用いた天然のGCRECの精製  
天然または組換えGCRECを、GCREC特異抗体を用いたイムノアフィニティークロマトグラフィにより実質的に精製する。イムノアフィニティークラムは、抗GCREC抗体を活性化クロマトグラフィ用樹脂、例えばCNBr活性化セファロース(Amersham Pharmacia Biotech)と共有結合させることにより構築する。結合後に、製造者の使用説明書に従って樹脂をブロックし、洗浄する。

【0295】

GCRECを含む培養液をイムノアフィニティークラムに通し、GCRECを優先的に吸着する条件下(例えば洗浄剤が存在する高イオン強度緩衝液)でカラムを洗浄する。抗体とGCRECの結合を破壊する条件(例えばpH2~3の緩衝液、或いは尿素またはチオシアン酸塩イオン等の高濃度のカオトロップ剤)でカラムを溶出させ、GCRECを回収する。

【0296】

16 GCRECと相互作用する分子の同定  
GCRECと相互作用する分子は、Gタンパク質などのシグナル伝達に関与する分子以外にもアゴニスト及びアンタゴニストを含み得る。GCRECまた断片を、<sup>125</sup>Iボルトンハンター試薬で標識する。(例えばBolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539を参照。)GCRECの断片には、例えば3つの細胞外ループのうち1つ若しくは複数のループ、細胞外N末端領域または第3細胞内ループがある。マルチウェルプレートの穴の中に予め配列しておいた候補分子を、標識したGCRECと共にインキュベートして洗浄し、標識したGCREC複合体を有する全ての穴をアッセイする。GCREC濃度を変えて得たデータを用いて、候補分子とのGCRECの数、親和性及び会合の値を計算する。

【0297】

或いは、GCRECと相互作用する分子は、Fields, S. 及びO. Song (1989) Nature 340:245-246に記載されているような酵母2ハイブリッドシステムを用いて分析するか、またはMATCHMAKERシステム(Clontech)等の2ハイブリッドシステムに基づく市販のキットを用いて分析する。

【0298】

GCRECはまた、ハイスループットな方法で酵母2ハイブリッドシステムを使用するPATHCALLINGプロセス(CuraGen Corp., New Haven CT

）に用いて、遺伝子の2大ライブラリにコードされるタンパク質間の全ての相互作用を決定することができる（Nandabalan, K. ら（2000）米国特許第6,057,101号）。

#### 【0299】

潜在的GCRECアゴニストまたはアンタゴニストは、実施例17及び18に記載のアッセイを用いてGCREC受容体活性の活性化または阻害を試験し得る。候補分子は、既知のGPCRアゴニストまたはアンタゴニスト、ペプチドライブラリまたは組み合わせの化学ライブラリから選択し得る。

#### 【0300】

GCRECとGタンパク質などの細胞内シグナル伝達分子の相互作用を検出する方法は、  
 オルファンGタンパク質共役膜7回貫通受容体から得た内部セグメントまたは細胞質ドメインが、既知のGタンパク質共役膜7回貫通受容体の相似ドメインと交換され、Gタンパク質及びオルファン受容体ドメインに活性化された下流シグナル伝達経路を同定するために用いられるという前提に基づく（Kobilka, B. K. ら（1988）Science 240:1310-1316）。相似方法においては、オルファン受容体のドメインが融合タンパク質の一部としてクローニングされ、結合アッセイで用いて特定のGタンパク質との相互作用を実証し得る。研究によれば、Gタンパク質共役膜7回貫通受容体の第3細胞内ループがGタンパク質相互作用及びシグナル伝達に重要であることが示されてきた（Conklin, B. R. ら（1993）Cell 73:631-641）。例えば、GCRECの第3細胞内ループに対応するDNA断片は、ポリメラーゼ連鎖反応法（PCR）により増幅し、pGEX（Pharmacia Biotech）などの融合ベクターにサブクローニングし得る。構成体は誘導された適切な細菌性宿主に形質転換され、融合タンパク質はグルタチオンセファロース4B（Pharmacia Biotech）アフィニティークロマトグラフィーにより細胞溶解産物から精製する。

#### 【0301】

in vitro 結合アッセイの場合、Gタンパク質を含む細胞抽出物は、50mM Tris、pH7.8、1mM EGTA、5mMのMgCl<sub>2</sub>、20mMのCHAPS、20%グリセロール、各10μgのアプロチニン及びロイペプチンと、20μlの50mMフッ化フェニルメチルスルフォニルで抽出することにより準備する。溶解産物は絶えず攪拌しながら氷上で45分間インキュベートし、23,000g、4 で15分間遠心分離して上澄みを集める。750μgの細胞抽出物をグルタチオンSトランスフェラーゼ（GST）融合タンパク質ビーズを用いて4 で2時間インキュベートする。GSTビーズは、リン酸緩衝食塩水で5回洗浄する。結合Gサブユニットは、百日咳毒素またはコレラ毒素を用いた[<sup>32</sup>P]ADPリボシル化により検出する。SDSサンプル緩衝液（4.6%（w/v）SDS、10%（v/v）メルカプトエタノール、20%（w/v）グリセロール、95.2mM Tris-HCl、pH6.8、0.01%（w/v）ブロムフェノールブルー）を添加することにより反応を停止させる。[<sup>32</sup>P]ADP標識されたタンパク質は、10%SDS-PAGEゲルで分離し、オートラジオグラフを行う。ゲル中で分離したタンパク質はニトロセルロースペーパーに移し、プロット（blot to）（5%脱脂粉乳、50mMのTris-HCl（pH8.0）、2mMのCaCl<sub>2</sub>、80mMのNaCl、0.02%のNaN<sub>3</sub>、及び0.2%ノニデットP-40）を用いて室温で1時間遮断し、G サブタイプ選択的抗体（1:500; Calbiochem-Novabiochem）で1.5時間インキュベートする。3洗浄後に、プロットはホースラディッシュペルオキシダーゼ（HRP）共役ヤギ抗ウサギ免疫グロブリン（1:2000, Cappel, Westchester PA）でインキュベートし、化学ルミネセンススペースのECL法（Amersham Corp）で可視化する。

#### 【0302】

##### 17 GCREC活性の実証

GCREC活性に対するアッセイは、細胞表面におけるGCRECの発現を測定する。

GCRECをコードするcDNAは、好適な哺乳動物細胞系に形質移入する。細胞表面タンパク質は、記載されているようにビオチンで標識する(de la Fuente, M. A. ら (1997) Blood 90:2398-2405)。GCREC特異抗体を用いて免疫沈降を実行し、ドデシル硫酸ナトリウムポリアクリルアミドゲル電気泳動(SDS-PAGE)及び免疫ブロット技術を用いて免疫沈降サンプルを分析する。標識された免疫沈降と未標識免疫沈降の比は、細胞表面に発現したGCRECの量に比例する。

#### 【0303】

或いは、GCREC活性のためのアッセイは、リガンド/受容体が仲介する細胞増殖の調節のための原型アッセイに基づく。このアッセイでは、スイスマウス3T3細胞におけるDNA合成の速度を測定する。当分野で公知の形質移入方法を用いて、GCRECをコードするポリヌクレオチドを含むプラスミドを静止状態の3T3培養細胞に加える。一過性に形質移入された細胞は次に、放射性DNA前駆分子である $[^3\text{H}]$ チミジンの存在下でインキュベートする。次に、培養した細胞にGCRECリガンドの変化する量を加える。 $[^3\text{H}]$ チミジンの酸沈殿可能DNAへの取り込みは、ラジオアイソトープカウンターを用いて適切な時間間隔で測定する。取り込み量は、新たに合成されたDNAの量に正比例する。少なくとも100倍のGCRECリガンド密度範囲に対する投与-反応の1次曲線は、受容体活性を示す。ミリリットル当りの活性の1単位は、50%の反応レベルを産出するGCRECの密度として画定され、100%であれば $[^3\text{H}]$ チミジンを酸沈殿可能DNAに最大限取り込むことを表す(McKay, I. and I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY, 73頁)。

#### 【0304】

或いは、GCREC活性のためのアッセイは、GPCRファミリータンパク質がGタンパク質活性化第2メッセンジャーシグナル伝達経路(例えばcAMP; Gaucelin, P. ら (1998) J. Biol. Chem. 273:4990-4996)を調整する能力に基づく。当分野で既知の方法を用いて、完全長GCRECをコードするプラスミドを哺乳動物細胞系(例えばチャイニーズハムスター卵巣(CHO)またはヒト胎児腎臓(HEK-293)細胞系)に形質移入する。培地中の12穴トレーで形質移入された細胞を成長させ、培養液を廃棄して、接着した細胞をPBSで軽く洗浄する。リガンドを入れた培養液とリガンドを入れない培養液で中で細胞を30分間インキュベートし、次に培養液を除去し、細胞を1Mの過塩素酸で処理して溶解する。溶解産物中のcAMPレベルは、当分野で既知の方法を用いて放射免疫アッセイにより測定する。リガンドに曝された細胞から得た溶解産物中のcAMPレベルをリガンドを入れなかった場合と比較したときの変化は、形質移入された細胞に存在するGCRECの量に比例する。

#### 【0305】

イノシトールリン酸レベルの変化を測定するためには、 $1 \times 10^5$  細胞/穴を含む24穴プレートで細胞を成長させ、イノシトールを含まない培地及び $[^3\text{H}]$ ミオイノシトールを用いて $2 \mu\text{Ci}$ /穴で48時間インキュベートする。培養液を除去し、10mMのLiClを含む緩衝液で細胞を洗浄し、その後リガンドを添加する。反応は、過塩素酸の添加により停止する。イノシトールリン酸は、Dowex AG1-X8(Bio-Rad)陰イオン交換樹脂及び液体シンチレーションにより計数された全ての標識されたイノシトールリン酸で抽出及び分離される。リガンドに曝された細胞から得た標識されたイノシトールリン酸のレベルをリガンドなしのものと比較したときの変化は、形質移入された細胞に存在するGCRECの量に比例する。

#### 【0306】

##### 18 GCRECリガンドの同定

GCRECは、CHO(チャイニーズハムスター卵巣)またはHEK(ヒト胎児腎臓)293などの真核細胞系において発現する。これらの細胞系は、GPCR発現過程が良好であり、発現したGCRECを下流エフェクターに機能的に結合させることができるよう

な広範囲のGタンパク質を含む。候補リガンドの存在下で発現した受容体の活性化に対し、形質転換した細胞をアッセイする。活性は、cAMPまたはCa<sup>2+</sup>などの細胞内セカンドメッセンジャーの変化により測定する。当分野で公知の標準的な方法を用いるか、活性化受容体によるタンパク質キナーゼCの刺激に反応して発光タンパク質(例えばホタルルシフェラーゼまたは緑色蛍光タンパク質)がプロモーターの転写調節下にあるようなレポーター遺伝子アッセイを用いて直接測定する(Milligan, C.ら(1996) Trends Pharmacol. Sci. 17:235-237)。アッセイ技術は、これらセカンドメッセンジャーシステムの両方に利用可能であり、アデニリルシクラーゼ活性化Flash Plate Assay(NEN Life Sciences Products)などのマルチウェルプレートフォーマットまたはFluo-4 AM(Molecular Probes)などの蛍光Ca<sup>2+</sup>指示薬でFLIPR蛍光定量的プレート読出しシステム(Molecular Devices)を併用して高処理読出しを可能にする。生理的関連性のあるセカンドメッセンジャー経路が知られていない場合、GCRECは、ホスホリパーゼC及びCa<sup>2+</sup>可動化に関与する経路を介してGCRECのシグナル伝達を通すために、広範囲のGタンパク質と共役することが実証されているようなGタンパク質G<sub>15/16</sub>と同時発現し得る(Offerrnanns, S. and M.I. Simon(1995) J. Biol. Chem. 270:15175-15180)。或いは、内在性GPCRが不足しているように遺伝子操作された酵母系にGCRECを発現し、それによってOCREC活性化スクリーニングに対してバックグラウンドが存在しない利点を提供し得る。これらの酵母系は、内在性酵母フェロモン受容体経路の対応する構成要素をヒトGPCR及びGタンパク質で置換する。シグナルに対する正常な酵母反応を、選択的培地の正の成長またはレポーター遺伝子発現に変換するように、下流シグナル伝達経路も変更する(Broach, J.R. and J. Thorner(1996) Nature 384(supp.):14-16)。既知のGPCRリガンド及びその他天然の生理活性分子を含む推定上のリガンドに対して受容体をスクリーニングする。組織、生物学的液体及び細胞上澄みから抽出した生物学的抽出物もスクリーニングする。

#### 【0307】

当業者は、本発明の範囲及び精神から逸脱することなく本発明の記載した方法及びシステムの種々の改変を行い得る。本発明について説明するにあたり特定の好適実施例に関連して説明を行ったが、本発明の範囲が、そのような特定の実施例に不当に制限されるべきではないことを理解されたい。実際に、分子生物学または関連分野の専門家には明らかな、本明細書に記載されている本発明の実施方法の様々な改変は、特許請求の範囲内にあるものとする。

(表の簡単な説明)

表1は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の命名法の概略を示す。

#### 【0308】

表2は、GenBank識別番号及び本発明のポリペプチドに最も近いGenBank相同体の注釈を示す。各ポリペプチドとそのGenBank相同体が一致する確率スコアも併せて示す。

#### 【0309】

表3は、予測されるモチーフ及びドメインを含む本発明のポリヌクレオチド配列の構造的特徴を、ポリペプチドの分析に用いるための方法、アルゴリズム及び検索可能なデータベースと共に示す。

#### 【0310】

表4は、本発明のポリヌクレオチド配列を構築するために用いたcDNAやゲノムDNA断片を、ポリヌクレオチド配列の選択した断片と共に示す。

#### 【0311】

表5は、本発明のポリヌクレオチドの代表的なcDNAライブラリを示す。

10

20

30

40

50

【 0 3 1 2 】

表 6 は、表 5 に示した c D N A ライブラリの作製に用いた組織及びベクターを説明する付表である。

【 0 3 1 3 】

表 7 は、本発明のポリヌクレオチドとポリペプチドの分析に用いたツール、プログラム、アルゴリズムを、適用可能な説明、引用文献及び閾値パラメータと共に示す。

【 表 1 】

表 1

Incyte プロジェクト ID	Incyte ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	Incyte ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID
2536292	1	2536292CD1	17	2536292CB1
7477708	2	7477708CD1	18	7477708CB1
7474823	3	7474823CD1	19	7474823CB1
644692	4	644692CD1	20	644692CB1
3837054	5	3837054CD1	21	3837054CB1
6157025	6	6157025CD1	22	6157025CB1
55012817	7	55012817CD1	23	55012817CB1
7475061	8	7475061CD1	24	7475061CB1
7477374	9	7477374CD1	25	7477374CB1
7479890	10	7479890CD1	26	7479890CB1
7482825	11	7482825CD1	27	7482825CB1
7483087	12	7483087CD1	28	7483087CB1
7483134	13	7483134CD1	29	7483134CB1
7478550	14	7478550CD1	30	7478550CB1
7483142	15	7483142CD1	31	7483142CB1
7483151	16	7483151CD1	32	7483151CB1

10

20

30

40

【 表 2 】

表 2

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	GenBank ID NO:	種 属 ス ト ク	GenBank 相 同 体
1	2536292CD1	G1039470	8.1E-07	[マウ] プロトン受容体 VNI (Dulac, C. および R. Axel (1996) Cell 83:195-206)
2	7477708CD1	G5525078	1.2E-89	[ラット] 糖7回貫通受容体 (Abe, J. 他 (1999) J. Biol. Chem. 274:19957-19964)
3	74748223CD1	G10241847	6.0E-79	[ヒト] ヒスタミン H4 受容体 (Oda, J. 他 (2000) J. Biol. Chem. 275:36781-36786)
4	644692CD1	G2117161	3.2E-35	[ヒト] 糖7回貫通ドメイン受容体 (Osterhoff, C. 他 (1997) DNA Cell Biol. 16:379-389)
5	3837054CD1	G1902964	7.1E-21	[マウス] オキントン受容体 (Kubota, Y. (1996) Mol. Cell. Endocrinol. 124:25-32)
6	6157025CD1	G5359718	2.5E-21	[ヒト] システイニル/ロイコトリエン受容体 (Sara, H.M. 他 (1999) Mol. Pharmacol. 56:657-663)
7	55012817CD1	G6006811	8.6E-60	[マウス] 錠行受容体 (Hsu, S.Y. 他 (2000) Mol. Endocrinol. 14:1257-1271)
8	7475061CD1	G440626	2.2E-22	[マウス] のろろ(Astyanax mexicanus) オプシン (Horowsky, B. 他 (2001) Proc. Natl. Acad. Sci. USA 98:8966-8971)
9	7477374CD1	G14600082	0.0	[ヒト] トレースマミン受容体 3 (Hirose, M.B. et al. (1999) Cell 96:713-723)
10	7479890CD1	G10441732	0.0	[ヒト] ロイコトリエン受容体含有Gタンパク質共役受容体 (Hsu, S.Y. 他 (2000) Mol. Endocrinol. 14:1257-1271)
11	7482825CD1	G4680254	1.3E-68	[マウス] 嗅覚受容体 S1 (Hirose, M.B. et al. (1999) Cell 96:713-723)
12	7483087CD1	G5901488 G2792018	1.5E-101 8.4E-99	[マウス] 嗅覚受容体 [ヒト] 嗅覚受容体 (Vanderhaeghen, P. 他 (1997) Biochem. Biophys. Res. Commun. 237:283-287)
13	7483134CD1	G6178008	8.3E-92	[マウス] 臭気物質受容体 MOR18 (Tsuboi, A. 他 (1999) J. Neurosci. 19:8409-8418)
14	7478550CD1	G3983394	1.2E-53	[マウス] 嗅覚受容体 F7 (Krautwurst, D. 他 (1998) Cell 95:917-926)
15	7483142CD1	G1246534	1.1E-91	[マウス] 嗅覚受容体 4 (Leibovici, M. 他 (1996) Dev. Biol. 175:118-131)
16	7483151CD1	G1246532	1.9E-78	[マウス] 嗅覚受容体 3 (Leibovici, M. 他 (1996) Dev. Biol. 175:118-131)

【表 3】

表 3-1

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基	潜在的な酸化部位	潜在的なシリンヒル化部位	シグネチャ配列、ドメインおよびモチーフ	分析手法およびデータベース
1	253622CD1	217	S183 S53 T172 T206 T45 Y128	N117 N181	膜貫通ドメイン: L63-L88	HMMER
2	7477708CD1	578	S179 S196 S224 S248 S255 S570 T185 T190 T346 T46 T499 T539 T97 S4 T16 T51	N122 N130 N175 N183 N228 N90	膜貫通ドメイン: Y218-V341, L400-I419, I479-V497 膜7回貫通受容体(セクレチンファミリー): S313-V560 Latrophilin/OL-1 様 GPS ドメイン: R260-L311 セクレチン様 GPCR スーパーファミリー: I518-A538, S376-F397, Y318-W342, T9-V33, A390-I413 Gタンパク質共受容体 PD000752: S307-Q559 EMR1 ホルモン受容体: DM05221 I37225 347-738: C264-Q559 DM05221 P48960 347-738: C265-Q559 DM05221 A57172 465-886: V231-F558	HMMER HMMER HMMER-PFAM HMMER-PFAM BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
3	7474823CD1	441	S174 S209 S215 S232 S314 S322 S359 S415 T129 T320 T363 T428 T433	N324 N393 N6	シグナルペプチド M1-A51 膜貫通ドメイン: Y226-N247 膜7回貫通受容体(ロドジンファミリー): F131-W255 Gタンパク質共受容体モチーフ: A153-V169 Gタンパク質共受容体シグネチャ: D147-V194 Gタンパク質共受容体 BL00237: W133-A172, F236-N247, T177-M203 ロドジン様 GPCR スーパーファミリー-PR00237: K184-L205, L428-Y251, V182-Y206, W225-Y251, K184-E208, D147-V169	SPScan HMMER HMMER-PFAM MOTIFS ProfileScan BLIMPS-BLOCKS BLIMPS-PRINTS

10

20

30

【表 4】

表3-2

SEQ ID NO.	Inocyte ポリペプチド ID	アミノ酸残基	潜在的な酸化部位	潜在的な酸化部位	潜在的グリコシ化部位	シグネチャ配列、ドメインおよびモチーフ	Analytical Methods and Databases
3						Gタンパク質共役受容体: DM00013 P30546 22-383: V119-H259 DM00013 P22270 103-599: S118-R257 DM00013 S58868 45-481: S118-R257 DM00013 P31390 21-315: I120-G281	BLAST-DOMO
4	644692CD1	797	S249 S3 3344 S349 S350 S368 S389 S466 S492 S499 S647 S660 S665 S666 S696 S746 S781 S788 T192 T225 T233 T294 T338 T423 T459 T47 T558 T661 T721 T768 T786	N159 N178 N191 N247 N261 N312 N316 N387 N413 N657 N709 N82	Signal peptide: M1-G26 膜貫通ドメイン: N8-W30, V470-I489, I567-Q592, V632-W652, S672-L689 膜7回貫通受容体(セクレチンファミリー) 7tm_2: P431-V726 Lactrophilin/CL-1様 GPS ドメインGPS:: Y379-D427 Gタンパク質共役受容体 BL00649: S389-T416, G441-V486, C526-L551, G574-Q598, W626-N655, S666-A687, T710-L735 線虫の膜内在性タンパク質 Srb シグネチャ PR00699E: V467-F487 ホリモン: EWR1: 白血球; 抗原: DM05221 A57172 465-886: C526-Y720 DM05221 I37225 347-738: Y603-T751 Gタンパク質共役受容体 PD000752: N511-E730	MOTIFS HMNER HMNER-PFAM HMNER-PFAM	
5	3837054CD1	434	S41 S131 S293	N15 N27 N60	Gタンパク質共役受容体: DM00013 P47901 28-353: R43-V182, W260-A357, W227-L256 Gタンパク質共役受容体 BL00237: A115-F154, R288-A314, N344-C360 膜貫通ドメイン: R45-C65 膜7回貫通受容体(ロドプシノプファミリー) 7tm_1): G59-F181, H228-Y352 シグナル切断: M1-A58	BLAST-DOMO BLAST-DOMO BLIMPS- BLOCKS HMNER HMNER-PFAM SPSCAN	

【表5】

表3-3

SEQ ID NO.	Inocyte ポリペプチド ID	アミノ酸残基	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチン配列、ドメインおよびモチーフ	分析手法およびデータベース
6	6157025CD1	339	T6 S10 Y175 S217 S276 S332	N5 N9 N308	Gタンパク質共役受容体: DM00013 E34993 36-327; L20-F285 DM00013 E30872 52-338; L20-F285 Gタンパク質共役受容体 BL00237: R218-F244, N9-F25, W81-C120 膜貫通ドメイン: I21-W45, V43-S63, M99-I116, M139-Y155, I185-V208, N227-L249 シグナルペプチド: M1-G20 シグナルペプチド: M1-E22 膜貫通ドメイン: G499-I523, I477-F495, H436-F458, Y378-L405, A274-I293 膜7回貫通受容体(セクレソファミリ): V265-S528	BLAST-DOMO BLIMPS-BLOCKS HMMER SPScan HMMER HMMER HMMER-PPAM
7	55012817CD1	549	S19 S57 T103 T153 T224 T3 T407 T415 T537 T84	N144 N210 N413 N98	Gタンパク質共役受容体 BL00649: M280-V325, C338-L363, G386-D410, Y429-F458, N472-A493 セクレソチン様 GPCR スーパーファミリ- PR00249: R270-R294, A340-L363, F379-G404, L492-I517, K475-F495, V503-L524 EMR1 白血球抗原: DM05221 I37225 347-738; P207-I517 DM05221 P48960 347-738; P207-I517 DM05221 A57172 465-886; L177-I517 Gタンパク質共役受容体 PD000752 : I271-L524	BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DOMO
8	7475061CD1	188	S163 S179 S64 T159 T174		シグナルペプチド: M1-I50 膜貫通ドメイン: F34-F52, M90-F112	BLAST-PRODOM SPScan HMMER

【表 6】

表3-4

SEQ ID NO:	Incyte 抗体ID	アミノ酸残基	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法およびデータベース
8					膜7回貫通受容体(ロドプシンファミリー): I38-Y141 視覚色素(オプシン)網膜結合部: A130-F148 視覚色素(オプシン)網膜結合部: A111-G166 Gタンパク質共役受容体 BL00237: P13-F52, L43-Y54, V80-V106, A133-A149 ロドプシン様GPCRスーパーファミリー-PR00237: L36-K60, G31-F52, A27-V49, L91-F112, I35-I58, L85-W109, S123-A149 Gタンパク質共役受容体: DM00013   P51472   36-326: P2-C151 DM00013   S39028   36-326: P2-C151 DM00013   P32312   31-322: P2-C151 DM00013   P32310   35-326: P2-C151 膜貫通ドメイン: S33-I55, F187-V207 膜7回貫通受容体(ロドプシンファミリー): G49-Y131, V183-Y295 Gタンパク質共役受容体 BL00237: W98-W137, F190-Y201, K235-D261, N287-G303 ロドプシン様GPCRスーパーファミリー-PR00237: T67-F88, D112-G134, W182-F205, A240-I264, Y277-G303, I34-L58 Gタンパク質共役受容体: DM00013   S55549   13-327: M132-V310 DM00013   P32251   21-399: P26-A240 DM00013   I49480   45-449: Y27-A322 DM00013   P18825   45-452: Y27-I215	HMMER-PFAM MOTIFS ProfileScan BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DMO
9	7477374CD1	332	S224 S228 S311 S6 T109 T90 Y99	N19 N4		HMMER HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DMO

10

20

30

【表7】

表3-5

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基	潜在的リン酸化部位	潜在的グリコシル化部位	シグナチーフ配列、ドメインおよびモチーフ	分析手法およびデータベース
9					Gタンパク質共役受容体 PD000009 : K61-M132	BLAST-PRODOM
10	7479890CD1	948	S107 S144 S19 S488 S489 S666 S688 S643 S855 T309 Y456 Y504	N169	シグナルペプチド: M1-G24 シグナルペプチド: M1-A25 膜貫通ドメイン: A548-G571, V756-P783 ロイシンリッチリピート: R239-P262, L263-P286, K287-T309, S310-P333, R334-Q355, K356-S379, S380-H403, S404-M427, S144-F167, A168-T191, S192-H215, N216-G238 ロイシンリッチリピート N 末端ドメイン: A34-D65 ロイシンジッパーモチーフ: L57-L78	SPScan HMMEP HMMEP HMMEP-PFAM
11	7482825CD1	315	S234 S54 S69 T10 T292 T6	N67 N8	Gタンパク質共役受容体 BL00237: R616-V655, P800-R816 糖タンパク質ホルモン受容体 PR00373: F533-W550, F610-C623, C623-S637, W748-L766 Gタンパク質共役受容体: DM00013   P14763   407-693; P535-L619 DM00013   P35376   355-641; P535-D818 DM00013   P22888   352-638; P535-D818 DM00013   P35409   519-807; I545-L619 Gタンパク質共役受容体 HG38: PD175529; H428-F568 PD168963; M259-P333 PD168277; A168-H215 Gタンパク質共役受容体: DM00013   P22270   18-311; L25-L306	MOTIFS BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DOMO BLAST-PRODOM BLAST-DOMO

【表 8】

10

20

30

表3-6

SEQ ID NO:	Inxite ポリペプチド ID	アミノ酸残基	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析 方法およびデータベース
11					受容体 嗅覚タンパク質 受容体様Gタンパク質 共役膜貫通糖タンパク質 多重遺伝子ファミリー PD000921:L167-L246 Gタンパク質共役受容体タンパク質 BL00237: T283-H299, A91-P130, L208-Y219, R236-A262 嗅覚受容体シグネチャ PR00245: M61-K82, F178-D192, F239-G254, F275-F286, T292-L306 ロドプシン様GPCRスーパーファミリー-シグネチャ PR00237: L28-V52, M61-K82, F105-I127, T200-V223, K273-H299 Gタンパク質共役受容体シグネチャ G protein_receptor.prf: F103-T149 膜貫通ドメイン: F27-V51, F102-D122, M143-A165, L201-T217 膜7回貫通受容体(ロドプシンファミリー) 7tm_1: G43-Y251	BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PRINTS PROFILESAN HMMER HMMER-PFAM
12	7483087CD1	312	S152 S267 S268 S291 S67	N5 N93	Gタンパク質共役受容体: DM00013 P23265 17-306: E22-I305 DM00013 P23268 18-307: D20-I305 DM00013 P30955 18-305: D20-I305 DM00013 S29707 18-306: F21-I301 受容体 嗅覚タンパク質 受容体様Gタンパク質 共役膜貫通糖タンパク質 多重遺伝子ファミリー PD000921: L166-I246 嗅覚受容体タンパク質 受容体様Gタンパク質 共役膜貫通糖タンパク質 多重遺伝子ファミリー PD149621: V247-R307	BLAST-PRODOM BLAST-PRODOM

【表9】

表3-7

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析 方法およびデータベース
12					Gタンパク質共役受容体タンパク質 BL00237: T282-L298, Q90-F129 嗅覚受容体シグネチャ PR00245: M59-K80, F177-D191, F238-G253, A274-L285, S291-I305 膜貫通ドメイン: L30-I46, Q100-M118, L143-M162, I197-F216 膜7回貫通受容体(ロトシリンファミリー) 7tm_1: G41-Y290 Gタンパク質受容体: L110-I126 Gタンパク質共役受容体シグネチャ g protein receptor.prf: C102-V151	BLIMPS- BLOCKS BLIMPS- PRINTS HMMER HMMER-PFAM MOTIFS PROFILESKAN
13	7483134CD1	309	S227 S305 S65 T286 T52 T83	N6	Gタンパク質共役受容体: DM00013 S297I0 15-301; F26-F300 DM00013 P23256 17-306; E20-L299 DM00013 S29708 18-306; E20-M296 DM00013 P23274 18-306; E20-M296 受容体 嗅覚タンパク質 受容体様Gタンパク質 共役 貫通 糖タンパク質 多量遺伝子ファミリー PD000921: L164-I243 Gタンパク質共役受容体タンパク質 BL00237: K88-P127, P277-K293 ロトシリン様 GPCR スーパーファミリー シグネチャ PR00237: T24-N48, M57-K78, A102-L124, K138-L159, V197-L220, K267-K293 嗅覚受容体シグネチャ PR00245: M57-K78, F175-D189, V235-M250, T286-F300 膜貫通ドメイン: V28-I45, I204-N222	BLAST- PRODOM BLIMPS- BLOCKS BLIMPS- PRINTS BLIMPS- PRINTS HMMER

10

20

30

表 3-8

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基	潜在的 リン酸化 部位	潜在的 グリコシル 部位	シグネチャ配列、 ドメインおよびモチーフ	分析 方法および データベース
13					膜7回貫通受容体(ロドプシンファミリー) 7cm_1:G39-Y285 Gタンパク質受容体-T108-I124 Gタンパク質共役受容体シグネチャ g_protein_receptor.prf: F100-A145 シグナル切断: M1-G39	HMMER_Pfam MOTIFS PROFILESSCAN
14	7478550CD1	309	S107 T21 T26 T3 T48	N46 N61	膜7回貫通受容体(ロドプシンファミリー): G81-Y292 Gタンパク質共役受容体シグネチャ: F142-G186 Gタンパク質共役受容体 BL00237: K130-S169 (E-value<0.018) 嗅覚受容体シグネチャ PR00245: M99-K120, F217-D231, F278-A293 ロドプシン様 GPCR スーパーファミリー-シグネチャ PRO0237: F144-I166, M180-V201, V239-L262, A277-R301, F66-H90, M99-K120 受容体 嗅覚 Gタンパク質 共役 膜貫通 権タンパク質 多重塩基子ファミリー PD000921: I206-L285 Gタンパク質共役受容体 DM00013 F23266 17-306: L67-L307 DM00013 S51356 18-307: L67-S305 DM00013 S29707 18-306: L58-T308 DM00013 E23274 18-306: L58-T308	HMMER L68-I86, M238-S256 HMMER_Pfam PROFILESSCAN BLIMPS- BLOCKS BLIMPS- PRINTS BLIMPS- PRINTS BLAST- PRODOM BLAST-DMO

【表 1 1】

10

20

30

表3-9

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基	潜在的リン酸化部位	潜在的グリコシル化部位	シグナル配列、ドメインおよびモチーフ	分析 方法および データベース
15	7483142CD1	315	S191 S294 S70	N68 N8	シグナル切断: M1-G54 膜貫通ドメイン: F31-L51 膜7回貫通受容体(ロトシグナファミリー): M44-Y293 Gタンパク質共役受容体: BL00237: K93-F132, M238-R264, I285-K301 膜受容体シグネチャ: PR00245: F180-D194, C241-G256, V277-L288, S294-L308, M62-N83 GPR: オープン受容体シグネチャ: PR0644: I52-Y63, S294-W305 Gタンパク質共役受容体シグネチャ: F105-V153 受容体 膜 Gタンパク質 共役 膜貫通 糖タンパク質 多重膜伝子ファミリー: PD000921: L169-M249 PD149621: V250-K311 Gタンパク質共役受容体: DM00013   P37067   17-306; T21-L304 DM00013   S29709   11-299; T21-L308 DM00013   P23266   17-306; P24-L308 DM00013   S51356   18-307; T21-L304 Gタンパク質共役受容体シグネチャ: S113-I129	SPSCAN HMMEP HMMEP-PPAM BLIMPS- BLOCKS BLIMPS- PRINTS BLIMPS- PRINTS PROFILESSCAN BLAST- PRODOM BLAST- DOMO MOTIFS

【表 1 2】

10

20

30

表3-10

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析手法およびデータベース
16	7483151CD1	307	S230 S261 S266 S291 T78	N107	膜貫通ドメイン: E28-A47, I196-I215 膜7回貫通受容体(ロトブジンファミリー): G41-Y290 Gタンパク質共役受容体 BL00237: H90-F129, I282-I298 Gタンパク質共役受容体シグネチャ: Y102-F150 嗅覚受容体シグネチャ PR00245: S291-M305, I59-K80, F177-N191, F238-G253, Y274-L285 ロトブジン様GPCRスーパーファミリー-シグネチャ PR00237: I59-K80, F104-I126, I201-A224, K272-I298, P226-W50 受容体 嗅覚 Gタンパク質 共役 膜貫通 糖タンパク質 多量輸送子ファミリー: PD000921: L166-L246 PD149621: V248-M305 Gタンパク質共役受容体 DM00013 P37067 17-306: L17-I304 DM00013 S51356 18-307: L17-R303 DM00013 S29709 11-299: T18-M305 DM00013 P23274 18-306: Q24-W305 Gタンパク質共役受容体シグネチャ: T110-I126	HMMER HMMER-PFAM BLIMPS- BLOCKS PROFILESAN BLIMPS- PRINTS BLIMPS- PRINTS BLAST- PRODOM BLAST- DOMO MOTIFS

【表 1 3】

10

20

30

表4-1

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	配列 長さ	選択された 断片	配列 断片	5' 位置	3' 位置
17	2536292CB1	2422	288-1079, 2271-2422, 105-150, 1835-1898	72051732V1 2536292P6 (BRAINOT18) FL2536392_96007890_000006_9499 5709 55062753H1 55062751J1 55062763H1 72049620V1 55062754J1 55094142J1	787 836 301 1 480 648 1488 658 1	2422 1690 1175 636 2422 2422 2422 2422 639
18	747708CB1	1912	1798-1912, 1-1741, 1304-1481	FL747708_98217793_000019_9552 5078 1971428H1 (UCMCL5T01) GBI-97547159_000017_000014.edi t GNN-97547159_000014_004 GNN-97106195_000002_002.edi	338 1 1028 144	1912 1255 782
19	7474823CB1	1326	1-185, 231-1326, 261-396, 626-1326	GNN-97547159_000014_004 GNN-97106195_000002_002.edi	517 1	1326 516
20	644692CB1	3058	1-1633, 2354-3058, 2369-2504, 1-2201	71737648V1 55052474J1 71735994V1 55067945J1 71907834V1 71736314V1 693242R8 (SINTMR02) 53738039	1922 370 1490 1 2612 2293 910 1088	1326 2535 1195 2079 671 3058 2985 1676 1541
21	3837054CB1	1993	1-1063, 1540-1714, 1-87, 1543-1993, 447-829	FL3837054_98575884_g1150834_1_2 GNN-98575884_004 56005250J1 3837054F6 (DEMTNT01) 2814920H1 (OVARNOF10)	849 359 1649 1371 1	1349 1420 1993 1935 187

【表 1 4】

10

20

30

表4-2

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	配列 長	選択された 断片	配列 断片	5' 位置	3' 位置
21				3406743H1 (PROSTUS08) GNN.G5686520.000054.002.edit	103	354
22	6157025CB1	1499	1-81, 248-1499, 850-920, 1265-1499	6834181F6 (BRSTN002) FL6157025.G10567930.000001.953 53887.1.1 7169316F8 (MCLRNOC01)	826	1499
23	55012817CB1	2455	1397-1482, 318-704, 1893-2171, 1-1014, 1619-2182	4425914H1 (BRAPDI101) 71691523V1 71691480V1 71691220V1 71690572V1 71691571V1 71691384V1 71688567V1	32	676
24	7475061CB1	2056	1-688, 1207-2056, 1736-1917, 1207-1680	72488727D1 72489896D1 72487236D1 72492196D1	841	1769
25	7477374CB1	999	1-138, 264-999	FL7477374.G9930948.000007.6739 496	1	999
26	7479890CB1	3429	1-2724, 2850-2900, 1-2300, 2367-2732	4021537F6 (BRANNOF02) GBI.G9967464.000017.000007.000 001.edit 55017814H1 58013229J1 8042178J1 (OVARFUE01) GNN.G9967464.000015.002 71702678V1 6609077H2 (PLACFEC01) 7720776H1 (THYRDIE01) 7726057J1 (THYRDIE01)	1427	2110
					1	588
					737	1280
					2532	3429
					2032	2692
					729	2847
					2160	2786
					366	877
					880	1291
					1242	1974

【表15】

表 4-3

ポリシフト SEQ ID NO:	Incyte ポリシフト ID	配列 長	選択された 断片	配列 断片	5' 位置	3' 位置
27	7482825CB1	948	1-87, 210-299, 923-948	GNN:G10045182_000002_006	1	948
38	7483087CB1	939	259-939, 1-86, 912-939	GNN:G6015288_000044_010	1	939
29	7483134CB1	930	1-327, 1-100, 889-930, 400-930	GNN:G7143464_000018_002	1	930
30	7478550CB1	1161	112-335, 116-358, 925-1161	7077972FB (BRAUTDR04) 7077972FB (BRAUTDR04)	1 416	926 1161
31	7483142CB1	948	1-36, 605-782, 913-948	FL7483142_g6086488_000017_g374 6443_1_1	1	948
32	7483151CB1	924	876-924	FL7483151_g9438337_000003_g151 4480_1_1-2	1	924

【表 16】

10

20

30

表 5

ポリヌクレオチド SEQ ID NO:	Incyte プロシクド ID	代表的 ライブラリ
17	2536292CB1	BRAINOT18
19	7474823CB1	UCMCL5T01
20	644692CB1	SINTTMR02
21	3837054CB1	DENDTNT01
22	6157025CB1	MGNCTXN05
23	55012817CB1	BRAPDT01
26	7479890CB1	LUNGNON07
30	7478550CB1	BRAUIDR04

10

20

30

【表 17】



表6-2

ライブラリ	ベクター	ライブラリの説明
DENDRYT01	pINCY	ライブラリは、抹消血液の処理済樹状細胞から単離された RNA を用いて作製した。
LJUNG007	pINCY	標準化した肺組織ライブラリは、肺組織ライブラリの 510 万個の独立性クローンから作製した。開始 RNA は肺組織から単離された RNA を使って作製された。このライブラリは、極めて長時間 (48 時間/一回) の再アニーリングによるハイブリダイゼーションを用いたことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を適用した条件を用いて 2 回におよびノーマライズした。
MONOXYM05	pINCY	標準化した処理済単球細胞組織ライブラリは、単球細胞ライブラリの 103 万個の独立性クローンから作製した。開始 RNA は 4 2 歳の女性から採られた抹消血液の処理済単球細胞から単離された RNA を使って作製された。細胞はインターフェロン 10 (IL-10) およびリポ多糖 (LPS) で処理された。このライブラリは、極めて長時間 (48 時間/一回) の再アニーリングによるハイブリダイゼーションを用いたことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を適用した条件を用いて 2 回におよびノーマライズした。
SINTPMR02	FCDNA2.1	このランダムにプライミングしたライブラリは 59 才男性から抽出した小腸組織から単離した RNA から作製した。一致する腸管組織の検理検査では、標数 (9) のカルチノイド腫瘍 (グレード 1) が小腸に見られた。最大の腫瘍は大きな腸管の塊を伴っていた。腸曲した腫瘍の一部分が腫瘍に転写していた。一つ (13 のうちの 1) の局所リンパ節が悪性度極性であった。腫瘍の生検では局所的な脂肪動脈硬化が見られた。
UCMCL5T01	PBLUESCRIPT	ライブラリは、12 人の献血者の腫瘍の血液から得られた単核細胞から単離された RNA を使って作製された。細胞を IL-5 と共に 12 日培養してから、RNA をブールされた溶解物から得た。

10

20

30

表 7-1

プログラム	説明	引用文献	パラメータ関連値
ABI FACTURA	核酸配列においてベクター配列を除去して、不定の塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finderは、アミノ酸または核酸配列の比較、注釈付けに有用である。	Applied Biosystems, Foster City, CA Paracel Inc., Pasadena, CA	不一致<50%
ABI AutoAssembler	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA	
BLAST	Basic Local Alignment Search Tool はアミノ酸および核酸配列の配列類似性検索に有用であり、BLAST には5つの機能がある：blastp、blastn、blastx、tblastnおよびtblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) (1997) Nucleic Acids Res. 25:3389-3402.	E <sub>ST</sub> s 確率値=1.0E-8 以下 完全長配列確率値= 1.0E-10以下
FASTA	問合せ配列と同種の配列群配との類似性を検索する Pearson および Lipman アルゴリズム。FASTA は最低5つの機能からなる：fasta、tfasta、fastx、tfastx および ssearch。	Pearson, W.R. 及び D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; および Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	E <sub>ST</sub> s fasta E 値=1.06E-6 構築された E <sub>ST</sub> s fasta 同一性=95%以上 一致した長さ=200塩基以上 fastx E 値=1.0E-8 以下 完全長配列 fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOM および PFAM データベースの配列と対応させて遺伝子ファミリー、配列相同性および構造的インゲージング領域を検索する BLocks IMProved Searcher。	Henikoff, S. 及び J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. 及び S. Henikoff (1996) Methods Enzymol. 266:88-105; Atwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値= 1.0E-3 以下
HMMER	PFAM のようなタンパク質ファミリー コンセンサス配列の隠れマルコフモデル (HMM) に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol. 235:1501-1531; Sommarhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322; Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, 1-350 ページ。1-350.	PFAM ヒット 確率値=1.0E-3 以下 シグナルペプチド ヒット: スコア =0 以上

表7-2

プログラム	説明	引用文献	パラメータ関連
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造のモチーフおよび配列モチーフを検索するアルゴリズム。	Gribakov, M. 他 (1988) CABIOS 4:61-66; Gribakov, M. 他 (1989) Methods Enzymol. 183:146-159; Bairoch, A. 他. (1997) Nucleic Acids Res. 25:217-221.	標準化された質のスコアとその特定のPrositeモチーフに対するGCG指定”HIGH”値 通常、スコア=1.4-2.1.
Phred	高い感度と精度で自動配列決定機のトレースを塩基読み出しアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づくSWATやCrossMatchを含むPhrap Revised Assembly プログラムで、配列相同性の検索やDNA配列の構築に有用である。	Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. および M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap で構築したものの表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンして、分泌シグナルペプチドの存在を調べる重み付けマトリックス解析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M. 及び S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	加重マトリックスを用いて蛋白配列での膜貫通セグメントの描写および配向を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れMarkov モデル(HMM)を使って蛋白配列の膜貫通セグメントを描写し、配向を決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc.Sixth Intl. Conf. on Intelligent Systems for Mol.Biol., Glasgow 他, eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182頁	
Motifs	Prositeで定義された配列と一致したパターンのあるノード配列を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, 第9版, M51-59頁, Genetics Computer Group, Madison, WI	

10

20

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  
4 April 2002 (04.04.2002)

PCT

(10) International Publication Number  
WO 02/26825 A2

(51) International Patent Classification: C07K 14/705

(21) International Application Number: PCT/US01/30661

(22) International Filing Date:  
28 September 2001 (28.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/236,546 29 September 2000 (29.09.2000) US  
60/240,589 13 October 2000 (13.10.2000) US  
60/242,322 20 October 2000 (20.10.2000) US  
60/242,223 20 October 2000 (20.10.2000) US  
60/245,855 3 November 2000 (03.11.2000) US  
60/245,900 3 November 2000 (03.11.2000) US  
60/247,587 9 November 2000 (09.11.2000) US  
60/249,343 15 November 2000 (15.11.2000) US

(71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US); GRAUL, Richard, C. [US/US]; 682-29th Avenue, San Francisco, CA 94121 (US); WALLA, Narinder, K. [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US); GANDHI, Ameena, R. [US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025 (US); HAFALIA, April, J. A. [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US); RAMKUMAR, Jayalaxmi [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US); TRIBOULEY, Catherine, M. [FR/US]; 1121 Tennessee Street, #5, San Francisco, CA 94107 (US); THORNTON, Michael [US/US]; 9 Madway Road, Woodside, CA 94062 (US); KALLICK, Deborah, A. [US/US]; 900 Olive Street, Menlo Park, CA 94025 (US); YAO, Monique, G. [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US); ELLIOTT, Vicki, S. [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US); BURFORD, Neil [GB/US]; 105 Wildwood Circle, Durham, CT 06422

(US). KHAN, Farrah, A. [IN/US]; 3617 Central Road #102, Glenview, IL 60025 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LU, Yan [CN/US]; 3885 Corrina Way, Palo Alto, CA 94303 (US). ARVIZU, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). ROOPA, Reddy [IN/US]; 1233 W. McKinley Avenue, #3, Sunnyvale, CA 94086 (US). NGUYEN, Dannie, B. [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). LEE, Ernestine, A. [US/US]; 624 Kains Street, Albany, CA 94706 (US). LU, Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). ISON, Craig, H. [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). WALSH, Roderick, T. [IE/GB]; 8 Boundary Court, St. Lawrence Road, Canterbury, Kent CT1 3EZ (GB). POLICKY, Jennifer, L. [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, 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, TH, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian 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, LI, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BI, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

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

(54) Title: G-PROTEIN COUPLED RECEPTORS

(57) Abstract: The invention provides human G-protein coupled receptors (GCREC) and polynucleotides which identify and encode GCREC. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of GCREC.



WO 02/26825 A2

WO 02/26825

PCT/US01/30661

**G-PROTEIN COUPLED RECEPTORS****TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of G-protein coupled receptors and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of G-protein coupled receptors. The present invention further relates to the use of specific G-protein coupled receptors to identify molecules that are involved in modulating taste or olfactory sensation.

**BACKGROUND OF THE INVENTION**

Signal transduction is the general process by which cells respond to extracellular signals. Signal transduction across the plasma membrane begins with the binding of a signal molecule, e.g., a hormone, neurotransmitter, or growth factor, to a cell membrane receptor. The receptor, thus activated, triggers an intracellular biochemical cascade that ends with the activation of an intracellular target molecule, such as a transcription factor. This process of signal transduction regulates all types of cell functions including cell proliferation, differentiation, and gene transcription. The G-protein coupled receptors (GPCRs), encoded by one of the largest families of genes yet identified, play a central role in the transduction of extracellular signals across the plasma membrane. GPCRs have a proven history of being successful therapeutic targets.

GPCRs are integral membrane proteins characterized by the presence of seven hydrophobic transmembrane domains which together form a bundle of antiparallel alpha ( $\alpha$ ) helices. GPCRs range in size from under 400 to over 1000 amino acids (Strosberg, A.D. (1991) *Eur. J. Biochem.* 196:1-10; Coughlin, S.R. (1994) *Curr. Opin. Cell Biol.* 6:191-197). The amino-terminus of a GPCR is extracellular, is of variable length, and is often glycosylated. The carboxy-terminus is cytoplasmic and generally phosphorylated. Extracellular loops alternate with intracellular loops and link the transmembrane domains. Cysteine disulfide bridges linking the second and third extracellular loops may interact with agonists and antagonists. The most conserved domains of GPCRs are the transmembrane domains and the first two cytoplasmic loops. The transmembrane domains account, in part, for structural and functional features of the receptor. In most cases, the bundle of  $\alpha$  helices forms a ligand-binding pocket. The extracellular N-terminal segment, or one or more of the three extracellular loops, may also participate in ligand binding. Ligand binding activates the receptor by inducing a conformational change in intracellular portions of the receptor. In turn, the large, third intracellular loop of the activated receptor interacts with a heterotrimeric guanine nucleotide binding

WO 02/26825

PCT/US01/30661

(G) protein complex which mediates further intracellular signaling activities, including the activation of second messengers such as cyclic AMP (cAMP), phospholipase C, and inositol triphosphate, and the interaction of the activated GPCR with ion channel proteins. (See, e.g., Watson, S. and S. Arkininstall (1994) *The G-protein Linked Receptor Facts Book*, Academic Press, San Diego CA, pp. 2-6;

- 5 Bolander, F.F. (1994) *Molecular Endocrinology*, Academic Press, San Diego CA, pp. 162-176;  
Baldwin, J.M. (1994) *Curr. Opin. Cell Biol.* 6:180-190.)

GPCRs include receptors for sensory signal mediators (e.g., light and olfactory stimulatory molecules); adenosine,  $\gamma$ -aminobutyric acid (GABA), hepatocyte growth factor, melanocortins, neuropeptide Y, opioid peptides, opsins, somatostatin, tachykinins, vasoactive intestinal polypeptide family, and vasopressin; biogenic amines (e.g., dopamine, epinephrine and norepinephrine, histamine, glutamate (metabotropic effect), acetylcholine (muscarinic effect), and serotonin); chemokines; lipid mediators of inflammation (e.g., prostaglandins and prostanoids, platelet activating factor, and leukotrienes); and peptide hormones (e.g., bombesin, bradykinin, calcitonin, C5a anaphylatoxin, endothelin, follicle-stimulating hormone (FSH), gonadotropic-releasing hormone (GnRH), neurokinin, 15 thyrotropin-releasing hormone (TRH), and oxytocin). GPCRs which act as receptors for stimuli that have yet to be identified are known as orphan receptors.

The diversity of the GPCR family is further increased by alternative splicing. Many GPCR genes contain introns, and there are currently over 30 such receptors for which splice variants have been identified. The largest number of variations are at the protein C-terminus. N-terminal and 20 cytoplasmic loop variants are also frequent, while variants in the extracellular loops or transmembrane domains are less common. Some receptors have more than one site at which variance can occur. The splice variants appear to be functionally distinct, based upon observed differences in distribution, signaling, coupling, regulation, and ligand binding profiles (Kilpatrick, G.J. et al. (1999) *Trends Pharmacol. Sci.* 20:294-301).

25 GPCRs can be divided into three major subfamilies: the rhodopsin-like, secretin-like, and metabotropic glutamate receptor subfamilies. Members of these GPCR subfamilies share similar functions and the characteristic seven transmembrane structure, but have divergent amino acid sequences. The largest family consists of the rhodopsin-like GPCRs, which transmit diverse extracellular signals including hormones, neurotransmitters, and light. Rhodopsin is a photosensitive 30 GPCR found in animal retinas. In vertebrates, rhodopsin molecules are embedded in membranous stacks found in photoreceptor (rod) cells. Each rhodopsin molecule responds to a photon of light by triggering a decrease in cGMP levels which leads to the closure of plasma membrane sodium channels. In this manner, a visual signal is converted to a neural impulse. Other rhodopsin-like GPCRs are directly involved in responding to neurotransmitters. These GPCRs include the receptors 35 for adrenaline (adrenergic receptors), acetylcholine (muscarinic receptors), adenosine, galanin, and

WO 02/26825 PCT/US01/30661  
glutamate (N-methyl-D-aspartate/NMDA receptors). (Reviewed in Watson, S. and S. Arkininstall  
(1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 7-9, 19-22,  
32-35, 130-131, 214-216, 221-222; Habert-Ortoli, E. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9780-  
9783.)

5 The galanin receptors mediate the activity of the neuroendocrine peptide galanin, which  
inhibits secretion of insulin, acetylcholine, serotonin and noradrenaline, and stimulates prolactin and  
growth hormone release. Galanin receptors are involved in feeding disorders, pain, depression, and  
Alzheimer's disease (Kask, K. et al. (1997) Life Sci. 60:1523-1533). Other nervous system  
rhodopsin-like GPCRs include a growing family of receptors for lysophosphatidic acid and other  
10 lysophospholipids, which appear to have roles in development and neuropathology (Chan, J. et al.  
(1999) Cell Biochem. Biophys. 30:213-242).

The largest subfamily of GPCRs, the olfactory receptors, are also members of the rhodopsin-  
like GPCR family. These receptors function by transducing odorant signals. Numerous distinct  
olfactory receptors are required to distinguish different odors. Each olfactory sensory neuron  
15 expresses only one type of olfactory receptor, and distinct spatial zones of neurons expressing distinct  
receptors are found in nasal passages. For example, the RA1c receptor, which was isolated from a  
rat brain library, has been shown to be limited in expression to very distinct regions of the brain and a  
defined zone of the olfactory epithelium (Raming, K. et al. (1998) Receptors Channels 6:141-151).  
However, the expression of olfactory-like receptors is not confined to olfactory tissues. For example,  
20 three rat genes encoding olfactory-like receptors having typical GPCR characteristics showed  
expression patterns not only in taste and olfactory tissue, but also in male reproductive tissue (Thomas,  
M.B. et al. (1996) Gene 178:1-5).

Members of the secretin-like GPCR subfamily have as their ligands peptide hormones such as  
secretin, calcitonin, glucagon, growth hormone-releasing hormone, parathyroid hormone, and  
25 vasoactive intestinal peptide. For example, the secretin receptor responds to secretin, a peptide  
hormone that stimulates the secretion of enzymes and ions in the pancreas and small intestine  
(Watson, supra, pp. 278-283). Secretin receptors are about 450 amino acids in length and are found in  
the plasma membrane of gastrointestinal cells. Binding of secretin to its receptor stimulates the  
production of cAMP.

30 Examples of secretin-like GPCRs implicated in inflammation and the immune response include  
the EGF module-containing, mucin-like hormone receptor (*Emr1*) and CD97 receptor proteins. These  
GPCRs are members of the recently characterized EGF-TM7 receptors subfamily. These seven  
transmembrane hormone receptors exist as heterodimers in vivo and contain between three and seven  
potential calcium-binding EGF-like motifs. CD97 is predominantly expressed in leukocytes and is  
35 markedly upregulated on activated B and T cells (McKnight, A.J. and S. Gordon (1998) J. Leukoc.

WO 02/26825  
Biol. 63:271-280).

PCT/US01/30661

The third GPCR subfamily is the metabotropic glutamate receptor family. Glutamate is the major excitatory neurotransmitter in the central nervous system. The metabotropic glutamate receptors modulate the activity of intracellular effectors, and are involved in long-term potentiation (Watson, *supra*, p.130). The  $Ca^{2+}$ -sensing receptor, which senses changes in the extracellular concentration of calcium ions, has a large extracellular domain including clusters of acidic amino acids which may be involved in calcium binding. The metabotropic glutamate receptor family also includes pheromone receptors, the GABA<sub>B</sub> receptors, and the taste receptors.

Other subfamilies of GPCRs include two groups of chemoreceptor genes found in the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*, which are distantly related to the mammalian olfactory receptor genes. The yeast pheromone receptors STE2 and STE3, involved in the response to mating factors on the cell membrane, have their own seven-transmembrane signature, as do the cAMP receptors from the slime mold *Dictyostelium discoideum*, which are thought to regulate the aggregation of individual cells and control the expression of numerous developmentally-regulated genes.

GPCR mutations, which may cause loss of function or constitutive activation, have been associated with numerous human diseases (Coughlin, *supra*). For instance, retinitis pigmentosa may arise from mutations in the rhodopsin gene. Furthermore, somatic activating mutations in the thyrotropin receptor have been reported to cause hyperfunctioning thyroid adenomas, suggesting that certain GPCRs susceptible to constitutive activation may behave as protooncogenes (Parma, J. et al. (1993) *Nature* 365:649-651). GPCR receptors for the following ligands also contain mutations associated with human disease: luteinizing hormone (precocious puberty); vasopressin V<sub>2</sub> (X-linked nephrogenic diabetes); glucagon (diabetes and hypertension); calcium (hyperparathyroidism, hypocalcemia, hypercalcemia); parathyroid hormone (short limbed dwarfism);  $\beta_2$ -adrenoceptor (obesity, non-insulin-dependent diabetes mellitus); growth hormone releasing hormone (dwarfism); and adrenocorticotropin (glucocorticoid deficiency) (Wilson, S. et al. (1998) *Br. J. Pharmacol.* 125:1387-1392; Stadel, J.M. et al. (1997) *Trends Pharmacol. Sci.* 18:430-437). GPCRs are also involved in depression, schizophrenia, sleeplessness, hypertension, anxiety, stress, renal failure, and several cardiovascular disorders (Horn, F. and G. Vriend (1998) *J. Mol. Med.* 76:464-468).

In addition, within the past 20 years several hundred new drugs have been recognized that are directed towards activating or inhibiting GPCRs. The therapeutic targets of these drugs span a wide range of diseases and disorders, including cardiovascular, gastrointestinal, and central nervous system disorders as well as cancer, osteoporosis and endometriosis (Wilson, *supra*; Stadel, *supra*). For example, the dopamine agonist L-dopa is used to treat Parkinson's disease, while a dopamine antagonist is used to treat schizophrenia and the early stages of Huntington's disease. Agonists and

WO 02/26825 PCT/US01/30661  
antagonists of adrenoceptors have been used for the treatment of asthma, high blood pressure, other cardiovascular disorders, and anxiety; muscarinic agonists are used in the treatment of glaucoma and tachycardia; serotonin 5HT<sub>1D</sub> antagonists are used against migraine; and histamine H<sub>1</sub> antagonists are used against allergic and anaphylactic reactions, hay fever, itching, and motion sickness (Horn, 5 *supra*).

Recent research suggests potential future therapeutic uses for GPCRs in the treatment of metabolic disorders including diabetes, obesity, and osteoporosis. For example, mutant V<sub>2</sub> vasopressin receptors causing nephrogenic diabetes could be functionally rescued *in vitro* by co-expression of a C-terminal V<sub>2</sub> receptor peptide spanning the region containing the mutations. This result suggests a possible novel strategy for disease treatment (Schöneberg, T. et al. (1996) EMBO J. 15:1283-1291). 10 Mutations in melanocortin-4 receptor (MC4R) are implicated in human weight regulation and obesity. As with the vasopressin V<sub>2</sub> receptor mutants, these MC4R mutants are defective in trafficking to the plasma membrane (Ho, G. and R.G. MacKenzie (1999) J. Biol. Chem. 274:35816-35822), and thus might be treated with a similar strategy. The type I receptor for parathyroid hormone (PTH) is a 15 GPCR that mediates the PTH-dependent regulation of calcium homeostasis in the bloodstream. Study of PTH/receptor interactions may enable the development of novel PTH receptor ligands for the treatment of osteoporosis (Mannstadt, M. et al. (1999) Am. J. Physiol. 277:F665-F675).

The chemokine receptor group of GPCRs have potential therapeutic utility in inflammation and infectious disease. (For review, see Locati, M. and P.M. Murphy (1999) Annu. Rev. Med. 50:425- 20 440.) Chemokines are small polypeptides that act as intracellular signals in the regulation of leukocyte trafficking, hematopoiesis, and angiogenesis. Targeted disruption of various chemokine receptors in mice indicates that these receptors play roles in pathologic inflammation and in autoimmune disorders such as multiple sclerosis. Chemokine receptors are also exploited by infectious agents, including herpesviruses and the human immunodeficiency virus (HIV-1) to facilitate infection. A truncated 25 version of chemokine receptor CCR5, which acts as a coreceptor for infection of T-cells by HIV-1, results in resistance to AIDS, suggesting that CCR5 antagonists could be useful in preventing the development of AIDS.

The involvement of some GPCRs in taste and olfactory sensation has been reported. Complete or partial sequences of numerous human and other eukaryotic sensory receptors are 30 currently known. (See, e.g., Pilpel, Y. and D. Lancet (1999) Protein Sci. 8:969-977; Mombaerts, P. (1999) Annu. Rev. Neurosci. 22:487-509. See also, e.g., patents EP 867508A2; US 5,874,243; WO 92/117585; WO 95/18140; WO 97/17444; and WO 99/67282.) It has been reported that the human genome contains approximately one thousand genes that encode a diverse repertoire of olfactory receptors (Rouquier, S. et al. (1998) Nat. Genet. 18:243-250; Trask, B.J. et al. (1998) Hum. Mol. 35 Genet. 7:2007-2020).

WO 02/26825

PCT/US01/30661

The discovery of new G-protein coupled receptors, and the polynucleotides encoding them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of G-protein coupled receptors.

#### SUMMARY OF THE INVENTION

The invention features purified polypeptides, G-protein coupled receptors, referred to collectively as "GCREC" and individually as "GCREC-1," "GCREC-2," "GCREC-3," "GCREC-4," "GCREC-5," "GCREC-6," "GCREC-7," "GCREC-8," "GCREC-9," "GCREC-10," "GCREC-11," "GCREC-12," "GCREC-13," "GCREC-14," "GCREC-15," and "GCREC-16." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-16.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-16. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:17-32.

The invention additionally provides G-protein coupled receptors that are involved in olfactory and/or taste sensation. The invention further provides polynucleotide sequences that encode said G-protein coupled receptors.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting

WO 02/26825 PCT/US01/30661  
of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least  
90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a  
biologically active fragment of a polypeptide having an amino acid sequence selected from the group  
5 consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino  
acid sequence selected from the group consisting of SEQ ID NO:1-16. In one alternative, the  
invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the  
invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group  
consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting  
10 of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least  
90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a  
biologically active fragment of a polypeptide having an amino acid sequence selected from the group  
consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino  
acid sequence selected from the group consisting of SEQ ID NO:1-16. The method comprises a)  
15 culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is  
transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a  
polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a  
polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence  
20 selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally  
occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the  
group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an  
amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic  
fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ  
25 ID NO:1-16.

The invention further provides an isolated polynucleotide selected from the group consisting of  
a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ  
ID NO:17-32, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least  
90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32,  
30 c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to  
the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide  
comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample,  
said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of  
35 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ

WO 02/26825 PCT/US01/30661  
ID NO:17-32, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional GCREC, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected

WO 02/26825 PCT/US01/30661  
from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional GCREC, comprising administering to a patient in need of such treatment the composition.

10 Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional GCREC, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a

WO 02/26825 PCT/US01/30661  
naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected  
from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide  
having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an  
immunogenic fragment of a polypeptide having an amino acid sequence selected from the group  
5 consisting of SEQ ID NO:1-16. The method comprises a) combining the polypeptide with at least one  
test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity  
of the polypeptide in the presence of the test compound, and c) comparing the activity of the  
polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of  
the test compound, wherein a change in the activity of the polypeptide in the presence of the test  
10 compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides methods of using G-protein coupled receptors of the invention  
involved in olfactory and/or taste sensation, biologically active fragments thereof (including those  
having receptor activity), and amino acid sequences having at least 90% sequence identity therewith,  
to identify compounds that agonize or antagonize the foregoing receptor polypeptides. These  
15 compounds are useful for modulating, blocking and/or mimicking specific tastes and/or odors.

The present invention also relates to the use of olfactory and/or taste receptors of the  
invention, biologically active fragments thereof (including those having receptor activity), and  
polypeptides having at least 90% sequence identity therewith, in combination with one or more other  
olfactory and/or taste receptor polypeptides, to identify a compound or plurality of compounds that  
20 modulate, mimic, and/or block a specific olfactory and/or taste sensation.

The invention also relates to cells that express an olfactory or taste receptor polypeptide of the  
invention, a biologically active fragment thereof (including those having receptor activity), or a  
polypeptide having at least 90% sequence identity therewith, and the use of such cells in cell-based  
screens to identify molecules that modulate, mimic, and/or block specific olfactory or taste sensations.  
25

Still further, the invention relates to a cell that co-expresses at least one olfactory or taste G-  
protein coupled receptor polypeptide of the invention, and a G-protein, and optionally one or more other  
olfactory and/or taste G-protein coupled receptor polypeptides, and the use of such a cell in screens to  
identify molecules that modulate, mimic, and/or block specific olfactory and/or taste sensations.

The invention further provides a method for screening a compound for effectiveness in  
30 altering expression of a target polynucleotide, wherein said target polynucleotide comprises a  
polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, the method  
comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting  
altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said  
35 method comprising a) treating a biological sample containing nucleic acids with the test compound; b)

WO 02/26825 PCT/US01/30661  
hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20  
contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide  
comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, ii) a  
polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a  
5 polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, iii) a polynucleotide  
having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of  
ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific  
hybridization complex is formed between said probe and a target polynucleotide in the biological  
sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising  
10 a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, ii) a  
polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a  
polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, iii) a polynucleotide  
complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of  
ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment  
15 of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the  
amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated  
biological sample with the amount of hybridization complex in an untreated biological sample, wherein  
a difference in the amount of hybridization complex in the treated biological sample is indicative of  
toxicity of the test compound.

20

#### BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide  
sequences of the present invention.

25 Table 2 shows the GenBank identification number and annotation of the nearest GenBank  
homolog for polypeptides of the invention. The probability score for the match between each  
polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including  
predicted motifs and domains, along with the methods, algorithms, and searchable databases used for  
analysis of the polypeptides.

30 Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble  
polynucleotide sequences of the invention, along with selected fragments of the polynucleotide  
sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

35 Table 6 provides an appendix which describes the tissues and vectors used for construction of  
the cDNA libraries shown in Table 5.

WO 02/26825

PCT/US01/30661

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

#### DESCRIPTION OF THE INVENTION

5 Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

10 It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

15 Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the  
20 cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

#### DEFINITIONS

"GCREC" refers to the amino acid sequences of substantially purified GCREC obtained from  
25 any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of GCREC. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of GCREC either by directly interacting with  
30 GCREC or by acting on components of the biological pathway in which GCREC participates.

An "allelic variant" is an alternative form of the gene encoding GCREC. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to  
35 allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides.

WO 02/26825

PCT/US01/30661

Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding GCREC include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as GCREC or a polypeptide with at least one functional characteristic of GCREC. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding GCREC, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding GCREC. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent GCREC. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of GCREC is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where “amino acid sequence” is recited to refer to a sequence of a naturally occurring protein molecule, “amino acid sequence” and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

“Amplification” relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term “antagonist” refers to a molecule which inhibits or attenuates the biological activity of GCREC. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of GCREC either by directly interacting with GCREC or by acting on components of the biological pathway in which GCREC participates.

The term “antibody” refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind GCREC polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or

WO 02/26825 PCT/US01/30661  
oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize  
5 the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on  
10 the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No.  
15 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH<sub>2</sub>), which may improve a desired property, e.g.,  
20 resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) *J. Biotechnol.* 74:5-13.)

The term "intramer" refers to an aptamer which is expressed *in vivo*. For example, a vaccinia  
25 virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) *Proc. Natl Acad. Sci. USA* 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on  
30 substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified  
35 sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having

WO 02/26825 PCT/US01/30661  
modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic GCREC, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding GCREC or fragments of GCREC may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhard's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
5	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
10	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
15	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
20	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be

WO 02/26825 PCT/US01/30661  
assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of GCREC or the polynucleotide encoding GCREC which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:17-32 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:17-32, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:17-32 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:17-32 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:17-32 and the region of SEQ ID NO:17-32 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-16 is encoded by a fragment of SEQ ID NO:17-32. A fragment of SEQ ID NO:1-16 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-16. For example, a fragment of SEQ ID NO:1-16 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-16. The precise length of a fragment of SEQ ID NO:1-16 and the region of SEQ ID NO:1-16 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a

WO 02/26825 PCT/US01/30661  
standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/b2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62  
Reward for match: 1  
Penalty for mismatch: -2  
Open Gap: 5 and Extension Gap: 2 penalties  
Gap x drop-off: 50  
Expect: 10  
Word Size: 11  
Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example,

WO 02/26825 PCT/US01/30661  
over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

```
Matrix: BLOSUM62
Open Gap: 11 and Extension Gap: 1 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 3
Filter: on
```

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for

WO 02/26825 PCT/US01/30661  
instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

5 "Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

10 "Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 25 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance,

WO 02/26825 PCT/US01/30661  
sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as  
formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances,  
such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily  
apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency  
5 conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is  
strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid  
sequences by virtue of the formation of hydrogen bonds between complementary bases. A  
hybridization complex may be formed in solution (e.g., C<sub>0</sub>t or R<sub>0</sub>t analysis) or formed between one  
10 nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid  
support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate  
to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide  
sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

15 "Immune response" can refer to conditions associated with inflammation, trauma, immune  
disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression  
of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect  
cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of GCREC which is  
20 capable of eliciting an immune response when introduced into a living organism, for example, a  
mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of  
GCREC which is useful in any of the antibody production methods disclosed herein or known in the  
art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides,  
25 polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other  
chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of GCREC. For example, modulation  
may cause an increase or a decrease in protein activity, binding characteristics, or any other biological,  
30 functional, or immunological properties of GCREC.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide,  
polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or  
synthetic origin which may be single-stranded or double-stranded and may represent the sense or the  
antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

35 "Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a

WO 02/26825 PCT/US01/30661  
functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

5 "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

10 "Post-translational modification" of an GCREC may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of GCREC.

"Probe" refers to nucleic acid sequences encoding GCREC, their complements, or fragments  
15 thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target  
20 DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100,  
25 or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold  
30 Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge  
35 MA).

WO 02/26825

PCT/US01/30661

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions

WO 02/26825 PCT/US01/30661  
(UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radiomulchides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing GCRC, nucleic acids encoding GCRC, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods

WO 02/26825 PCT/US01/30661  
well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), *supra*.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The

WO 02/26825 PCT/US01/30661  
presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

10

#### THE INVENTION

The invention is based on the discovery of new human G-protein coupled receptors (GCREC), the polynucleotides encoding GCREC, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections.

15

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

20

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

25

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS

30

WO 02/26825 PCT/US01/30661  
program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI).  
Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7  
shows analytical methods for protein structure/function analysis and in some cases, searchable  
databases to which the analytical methods were applied.

5 Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these  
properties establish that the claimed polypeptides are G-protein coupled receptors. For example, SEQ  
ID NO:2 is 39% identical to rat seven transmembrane G-protein coupled receptor (GenBank ID  
g5525078) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The  
BLAST probability score is  $1.2e-89$ , which indicates the probability of obtaining the observed  
10 polypeptide sequence alignment by chance. SEQ ID NO:2 also contains a secretin family 7-  
transmembrane receptor domain as determined by searching for statistically significant matches in the  
hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See  
Table 3.) Data from BLIMPS and BLAST analyses provide further corroborative evidence that SEQ  
ID NO:2 is a G-protein coupled receptor. In an alternative example, SEQ ID NO:4 is 32% identical to  
15 human seven transmembrane-domain receptor (GenBank ID g2117161) as determined by BLAST.  
(See Table 2.) The BLAST probability score is  $3.2e-35$ . SEQ ID NO:4 also contains a  
latrophilin/CL-1-like GPS domain and a secretin family 7-transmembrane receptor domain as  
determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS and  
MOTIFS analyses provide further corroborative evidence that SEQ ID NO:4 is a 7-transmembrane  
20 G-protein coupled receptor. In an alternative example, SEQ ID NO:5 is 34% identical to murine  
oxytocin receptor (GenBank ID g1902964) as determined by BLAST. (See Table 2.) The BLAST  
probability score is  $7.1e-21$ . SEQ ID NO:5 also contains a 7-transmembrane receptor domain as  
determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS  
analyses provide further corroborative evidence that SEQ ID NO:5 is a G-protein coupled receptor.  
25 In an alternative example, SEQ ID NO:6 is 23% identical to a human cysteinyl leukotriene receptor  
(GenBank ID g5359718) as determined by BLAST. (See Table 2.) The BLAST probability score is  
 $2.5e-21$ . Data from BLIMPS analyses provide further corroborative evidence that SEQ ID NO:6 is a  
G-protein coupled receptor. In an alternative example, SEQ ID NO:8 is 32% identical to an opsin  
from the Mexican tetra, a blind cave fish, (GenBank ID g440626) as determined by BLAST. (See  
30 Table 2.) The BLAST probability score is  $2.2e-22$ . SEQ ID NO:8 also contains a rhodopsin family 7-  
transmembrane receptor domain as determined by searching the HMM-based PFAM database. (See  
Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further  
corroborative evidence that SEQ ID NO:8 is a rhodopsin family G-protein coupled receptor. In an  
alternative example, SEQ ID NO:9 is 44% identical to a putative human neurotransmitter receptor  
35 (GenBank ID g2465432) as determined by BLAST. The BLAST probability score is  $5.2e-69$ . SEQ

WO 02/26825 PCT/US01/30661  
ID NO:9 also contains a rhodopsin family 7-transmembrane receptor domain as determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS and BLAST analyses provide further corroborative evidence that SEQ ID NO:9 is a rhodopsin family G-protein coupled receptor. In an alternative example, SEQ ID NO:11 is 82% identical to Marmota marmota olfactory receptor (GenBank ID g5901488) as determined by BLAST. (See Table 2.) The BLAST probability score is 1.5e-101. SEQ ID NO:11 also contains a rhodopsin family 7-transmembrane receptor domain as determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:11 is a G-protein coupled receptor. In an alternative example, SEQ ID NO:12 is 64% identical to Homo sapiens olfactory receptor (GenBank ID g2792018) as determined by BLAST. (See Table 2.) The BLAST probability score is 8.4e-99. SEQ ID NO:12 also contains a rhodopsin family 7-transmembrane receptor domain as determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:12 is a G-protein coupled receptor. In an alternative example, SEQ ID NO:15 is 57% identical to chicken olfactory receptor 4 (GenBank ID g1246534) as determined by BLAST. (See Table 2.) The BLAST probability score is 1.1e-91. SEQ ID NO:15 also contains a 7-transmembrane receptor (rhodopsin family) domain as determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:15 is an olfactory receptor. SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:10, SEQ ID NO:13-14, and SEQ ID NO:16 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-16 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:17-32 or that distinguish between SEQ ID NO:17-32 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective

WO 02/26825  
full length sequences.

PCT/US01/30661

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 2536292F6 is the identification number of an Incyte cDNA sequence, and BRAINOT18 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 72051732V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g3738039) which contributed to the assembly of the full length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the identification numbers in column 5 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, FL\_XXXXXX\_N<sub>1</sub>\_N<sub>2</sub>\_YYYYY\_N<sub>3</sub>\_N<sub>4</sub> represents a "stitched" sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYYY is the number of the prediction generated by the algorithm, and N<sub>1,2,3,4</sub>, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the identification numbers in column 5 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, FLXXXXXX\_gAAAAA\_gBBBBB\_1\_N is the identification number of a "stretched" sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
--------	--

WO 02/26825

PCT/US01/30661

GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
5 INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

10 Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

15 The invention also encompasses GCREC variants. A preferred GCREC variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the GCREC amino acid sequence, and which contains at least one functional or structural characteristic of GCREC.

20 The invention also encompasses polynucleotides which encode GCREC. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:17-32, which encodes GCREC. The polynucleotide sequences of SEQ ID NO:17-32, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

25 The invention also encompasses a variant of a polynucleotide sequence encoding GCREC. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding GCREC. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:17-  
30 32 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ

WO 02/26825 PCT/US01/30661  
ID NO:17-32. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of GCREC.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding GCREC, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring GCREC, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode GCREC and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring GCREC under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding GCREC or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding GCREC and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode GCREC and GCREC derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding GCREC or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:17-32 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE

WO 02/26825 PCT/US01/30661  
amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA  
5 sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

10 The nucleic acid sequences encoding GCRFC may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.)

15 Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et  
20 al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo  
25 Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about  
30 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence  
35 into 5' non-transcribed regulatory regions.

WO 02/26825

PCT/US01/30661

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

10 In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode GCREC may be cloned in recombinant DNA molecules that direct expression of GCREC, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express GCREC.

15 The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter GCREC-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. 25 Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of GCREC, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These 30 preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene 35 family, either from the same or different species, thereby maximizing the genetic diversity of multiple

WO 02/26825  
naturally occurring genes in a directed and controllable manner.

PCT/US01/30661

In another embodiment, sequences encoding GCREC may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.) Alternatively, GCREC itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) *Proteins, Structures and Molecular Properties*, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of GCREC, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, *supra*, pp. 28-53.)

In order to express a biologically active GCREC, the nucleotide sequences encoding GCREC or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding GCREC. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding GCREC. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding GCREC and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding GCREC and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques,

WO 02/26825 PCT/US01/30661  
and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

5 A variety of expression vector/host systems may be utilized to contain and express sequences encoding GCREC. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); 10 plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New 15 York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 20 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding GCREC. For example, routine cloning, 25 subcloning, and propagation of polynucleotide sequences encoding GCREC can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding GCREC into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for 30 *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of GCREC are needed, e.g. for the production of antibodies, vectors which direct high level expression of GCREC may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

35 Yeast expression systems may be used for production of GCREC. A number of vectors

WO 02/26825 PCT/US01/30661  
containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*;  
5 Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of GCREC. Transcription of sequences encoding GCREC may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 10 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill,  
15 New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding GCREC may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain  
20 infective virus which expresses GCREC in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of  
25 DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of  
30 GCREC in cell lines is preferred. For example, sequences encoding GCREC can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a  
35 selective agent, and its presence allows growth and recovery of cells which successfully express the

WO 02/26825 PCT/US01/30661  
introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apv* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding GCRC is inserted within a marker gene sequence, transformed cells containing sequences encoding GCRC can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding GCRC under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding GCRC and that express GCRC may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of GCRC using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on GCRC is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See,

WO 02/26825 PCT/US01/30661  
e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

5 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding GCREC include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding GCREC, or any fragments thereof, may be cloned into a vector  
10 for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for  
15 ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding GCREC may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence  
20 and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode GCREC may be designed to contain signal sequences which direct secretion of GCREC through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of  
25 the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture  
30 Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding GCREC may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric GCREC protein  
35 containing a heterologous moiety that can be recognized by a commercially available antibody may

WO 02/26825 PCT/US01/30661  
facilitate the screening of peptide libraries for inhibitors of GCREC activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the GCREC encoding sequence and the heterologous protein sequence, so that GCREC may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled GCREC may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, <sup>35</sup>S-methionine.

GCREC of the present invention or fragments thereof may be used to screen for compounds that specifically bind to GCREC. At least one and up to a plurality of test compounds may be screened for specific binding to GCREC. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of GCREC, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) *Current Protocols in Immunology* 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which GCREC binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express GCREC, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing GCREC or cell membrane fractions which contain GCREC are then contacted with a test compound and binding, stimulation, or inhibition of activity of either GCREC or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the

WO 02/26825 PCT/US01/30661  
assay may comprise the steps of combining at least one test compound with GCREC, either in solution or affixed to a solid support, and detecting the binding of GCREC to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor.

Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.  
5 GCREC of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of GCREC. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for GCREC activity, wherein GCREC is combined with at least one test compound, and the activity of GCREC in the presence of a test compound is compared with the activity of GCREC in the absence of the test compound. A change in the activity of GCREC in the presence of the test compound is indicative of a compound that modulates the activity of GCREC. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising GCREC under conditions suitable for GCREC activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of  
10 GCREC may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding GCREC or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of  
20 human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by  
25 homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and  
30 the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding GCREC may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate  
35 into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al.

WO 02/26825  
(1998) Science 282:1145-1147).

PCT/US01/30661

Polynucleotides encoding GCRC can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding GCRC is injected into animal ES cells, and the injected sequence  
5 integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress GCRC, e.g., by secreting GCRC in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

#### 10 THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of GCRC and G-protein coupled receptors. In addition, the expression of GCRC is closely associated with brain tissue. Therefore, GCRC appears to play a role in cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and  
15 viral infections. In the treatment of disorders associated with increased GCRC expression or activity, it is desirable to decrease the expression or activity of GCRC. In the treatment of disorders associated with decreased GCRC expression or activity, it is desirable to increase the expression or activity of GCRC.

Therefore, in one embodiment, GCRC or a fragment or derivative thereof may be  
20 administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCRC. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia,  
25 lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease,  
30 Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including  
35 kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial

WO 02/26825 PCT/US01/30661  
insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous  
sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation  
and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal  
disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular  
5 dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis  
and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic  
paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective  
disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias,  
paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy,  
10 corticobasal degeneration, and familial frontotemporal dementia; a cardiovascular disorder such as  
arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial  
dissections, varicose veins, thrombophlebitis and phlebotrombosis, vascular tumors, complications of  
thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery,  
congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive  
15 heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid  
aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart  
disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus  
erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart  
disease, congenital heart disease, and complications of cardiac transplantation; a gastrointestinal  
20 disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal  
carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis,  
antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of  
the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic  
carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver,  
25 hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease,  
Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short  
bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency  
syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic  
steatosis, hemochromatosis, Wilson's disease, alpha<sub>1</sub>-antitrypsin deficiency, Reye's syndrome, primary  
30 sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis,  
peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty  
liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular  
hyperplasias, adenomas, and carcinomas; an autoimmune/inflammatory disorder such as acquired  
immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies,  
35 ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia,

WO 02/26825 PCT/US01/30661  
autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a metabolic disorder such as diabetes, obesity, and osteoporosis; and an infection by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, and togavirus.

15 In another embodiment, a vector capable of expressing GCREC or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCREC including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified GCREC in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCREC including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of GCREC may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCREC including, but not limited to, those listed above.

25 In a further embodiment, an antagonist of GCREC may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GCREC. Examples of such disorders include, but are not limited to, those cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections described above. In one aspect, an antibody which specifically binds GCREC may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express GCREC.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding GCREC may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GCREC including, but not limited to, those described above.

35 In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary

WO 02/26825 PCT/US01/30661  
sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of GCREC may be produced using methods which are generally known in the art. In particular, purified GCREC may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind GCREC. Antibodies to GCREC may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with GCREC or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to GCREC have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of GCREC amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to GCREC may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc.*

WO 02/26825 PCT/US01/30661  
Nat. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce GCREC-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for GCREC may also be generated. For example, such fragments include, but are not limited to, F(ab)<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab)<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between GCREC and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering GCREC epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for GCREC. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of GCREC-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple GCREC epitopes, represents the average affinity, or avidity, of the antibodies for GCREC. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular GCREC epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the GCREC-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of GCREC, preferably in active form, from the

WO 02/26825 PCT/US01/30661  
antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC;  
Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons,  
New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine  
5 the quality and suitability of such preparations for certain downstream applications. For example, a  
polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg  
specific antibody/ml, is generally employed in procedures requiring precipitation of GCREC-antibody  
complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for  
antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and  
10 Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding GCREC, or any  
fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications  
of gene expression can be achieved by designing complementary sequences or antisense molecules  
(DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene  
15 encoding GCREC. Such technology is well known in the art, and antisense oligonucleotides or larger  
fragments can be designed from various locations along the coding or control regions of sequences  
encoding GCREC. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc.,  
Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense  
20 sequences into appropriate target cells can be used. Antisense sequences can be delivered  
intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence  
complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g.,  
Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995)  
9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral  
25 vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood*  
76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other  
gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other  
systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et  
al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.*  
30 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding GCREC may be used for  
somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency  
(e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-  
linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined  
35 immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency

WO 02/26825 PCT/US01/30661  
 (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475),  
 cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene  
 Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassemias, familial  
 hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal,  
 5 R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii)  
 express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated  
 cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g.,  
 against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988)  
 Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399),  
 10 hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides  
brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the  
 case where a genetic deficiency in GCREC expression or regulation causes disease, the expression of  
 GCREC from an appropriate population of transduced cells may alleviate the clinical manifestations  
 caused by the genetic deficiency.

15 In a further embodiment of the invention, diseases or disorders caused by deficiencies in  
 GCREC are treated by constructing mammalian expression vectors encoding GCREC and introducing  
 these vectors by mechanical means into GCREC-deficient cells. Mechanical transfer technologies for  
 use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic  
 gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and  
 20 (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem.  
 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J.-L. and H. Récipon (1998) Curr. Opin.  
 Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of GCREC include, but are not  
 limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors  
 25 (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA),  
 and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). GCREC  
 may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous  
 sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or  $\beta$ -actin genes), (ii) an inducible promoter  
 (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci.  
 30 USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau  
 (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen));  
 the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the  
 FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V.  
 and Blau, H.M. supra), or (iii) a tissue-specific promoter or the native promoter of the endogenous

WO 02/26825  
gene encoding GCREC from a normal individual.

PCT/US01/30661

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to GCREC expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding GCREC under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and FEBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4<sup>+</sup> T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding GCREC to cells which have one or more genetic abnormalities with respect to the expression of GCREC. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are

WO 02/26825 PCT/US01/30661  
described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Somia (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

5 In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding GCREC to target cells which have one or more genetic abnormalities with respect to the expression of GCREC. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing GCREC to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with  
10 ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92  
15 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned  
20 herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to  
25 deliver polynucleotides encoding GCREC to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA,  
30 resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for GCREC into the alphavirus genome in place of the capsid-coding region results in the production of a large number of GCREC-coding RNAs and the synthesis of high levels of GCREC in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a  
35 persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN)

WO 02/26825 PCT/US01/30661  
indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of GCREC into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of  
5 manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can  
10 be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.B. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A  
15 complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example,  
20 engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding GCREC.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides,  
25 corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared  
30 by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding GCREC. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA  
35 constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell

WO 02/26825  
lines, cells, or tissues.

PCT/US01/30661

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding GCREC. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased GCREC expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding GCREC may be therapeutically useful, and in the treatment of disorders associated with decreased GCREC expression or activity, a compound which specifically promotes expression of the polynucleotide encoding GCREC may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding GCREC is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding GCREC are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding GCREC. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to

WO 02/26825 PCT/US01/30661  
a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

10 Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

15 Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which, generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of GCREC, antibodies to GCREC, and mimetics, agonists, antagonists, or inhibitors of GCREC.

20 The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

25 Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without

WO 02/26825 PCT/US01/30661  
needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

5 Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising GCREC or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, GCREC or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to  
10 transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and  
15 route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example GCREC or fragments thereof, antibodies of GCREC, and agonists, antagonists or inhibitors of GCREC, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be  
20 determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population) or LD<sub>50</sub> (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD<sub>50</sub>/ED<sub>50</sub> ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are  
25 used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED<sub>50</sub> with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the  
30 subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or  
35 biweekly depending on the half-life and clearance rate of the particular formulation.

WO 02/26825

PCT/US01/30661

Normal dosage amounts may vary from about 0.1  $\mu\text{g}$  to 100,000  $\mu\text{g}$ , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

#### DIAGNOSTICS

In another embodiment, antibodies which specifically bind GCREC may be used for the diagnosis of disorders characterized by expression of GCREC, or in assays to monitor patients being treated with GCREC or agonists, antagonists, or inhibitors of GCREC. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for GCREC include methods which utilize the antibody and a label to detect GCREC in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring GCREC, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of GCREC expression. Normal or standard values for GCREC expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to GCREC under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of GCREC expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding GCREC may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of GCREC may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of GCREC, and to monitor regulation of GCREC levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding GCREC or closely related molecules may be used to identify nucleic acid sequences which encode GCREC. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the

WO 02/26825 PCT/US01/30661  
probe identifies only naturally occurring sequences encoding GCREC, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the GCREC encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:17-32 or from genomic sequences including promoters, enhancers, and introns of the GCREC gene.

Means for producing specific hybridization probes for DNAs encoding GCREC include the cloning of polynucleotide sequences encoding GCREC or GCREC derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as <sup>32</sup>P or <sup>35</sup>S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding GCREC may be used for the diagnosis of disorders associated with expression of GCREC. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic

WO 02/26825 PCT/US01/30661  
 paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cardiovascular disorder such as  
 5 arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid  
 10 aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal  
 15 carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease,  
 20 Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha<sub>1</sub>-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis,  
 25 peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia,  
 30 autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxicins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome,  
 35 multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis,

WO 02/26825

PCT/US01/30661

osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a metabolic disorder such as diabetes, obesity, and osteoporosis; and an infection by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, and tongavirus. The polynucleotide sequences encoding GCREC may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered GCREC expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding GCREC may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding GCREC may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding GCREC in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of GCREC, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding GCREC, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several

WO 02/26825  
days to months.

PCT/US01/30661

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding GCREC may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding GCREC, or a fragment of a polynucleotide complementary to the polynucleotide encoding GCREC, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding GCREC may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding GCREC are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of GCREC include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplax, C.

WO 02/26825 PCT/US01/30661  
et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

5 In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene  
10 function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and  
15 display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, GCREC, fragments of GCREC, or antibodies specific for GCREC may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

20 A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No.  
25 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The  
30 resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*, as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention  
35 may also be used in conjunction with *in vitro* model systems and preclinical evaluation of

WO 02/26825 PCT/US01/30661  
pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental  
compounds. All compounds induce characteristic gene expression patterns, frequently termed  
molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity  
(Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000)  
5 Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a  
signature similar to that of a compound with known toxicity, it is likely to share those toxic properties.  
These fingerprints or signatures are most useful and refined when they contain expression information  
from a large number of genes and gene families. Ideally, a genome-wide measurement of expression  
provides the highest quality signature. Even genes whose expression is not altered by any tested  
10 compounds are important as well, as the levels of expression of these genes are used to normalize the  
rest of the expression data. The normalization procedure is useful for comparison of expression data  
after treatment with different compounds. While the assignment of gene function to elements of a  
toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not  
necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for  
15 example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released  
February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is  
important and desirable in toxicological screening using toxicant signatures to include all expressed  
gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample  
20 containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated  
biological sample are hybridized with one or more probes specific to the polynucleotides of the present  
invention, so that transcript levels corresponding to the polynucleotides of the present invention may be  
quantified. The transcript levels in the treated biological sample are compared with levels in an  
untreated biological sample. Differences in the transcript levels between the two samples are  
25 indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present  
invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global  
pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome  
can be subjected individually to further analysis. Proteome expression patterns, or profiles, are  
30 analyzed by quantifying the number of expressed proteins and their relative abundance under given  
conditions and at a given time. A profile of a cell's proteome may thus be generated by separating  
and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is  
achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by  
isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl  
35 sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins

WO 02/26825 PCT/US01/30661  
are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for GCREC to quantify the levels of GCREC expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lucking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendozze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological

WO 02/26825 PCT/US01/30661  
sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding GCREC may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding GCREC on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse,

WO 02/26825 PCT/US01/30661  
may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

10 In another embodiment of the invention, GCREC, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between GCREC and the agent being tested may be measured.

15 Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with GCREC, or fragments thereof, and washed. Bound GCREC is then detected by methods well known in the art. Purified GCREC can also be coated directly onto plates for use in the aforementioned drug screening techniques.

20 Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding GCREC specifically compete with a test compound for binding GCREC. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with GCREC.

25 In additional embodiments, the nucleotide sequences which encode GCREC may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

30 Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/236,546, U.S. Ser. No. 60/240,589, U.S. Ser. No. 60/242,223, U.S. Ser.

WO 02/26825 PCT/US01/30661  
No. 60/242,322, U.S. Ser. No. 60/245,855, U.S. Ser. No. 60/245,900, U.S. Ser. No. 60/247,587, and  
U.S. Ser. No. 60/249,343, are expressly incorporated by reference herein.

#### EXAMPLES

##### 5 I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)<sup>+</sup> RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPRT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPRT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XLI-Blue, XLI-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

##### 35 II. Isolation of cDNA Clones

WO 02/26825

PCT/US01/30661

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, 5 QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4 °C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) *Anal. Biochem.* 216:1-14). Host cell lysis and thermal 10 cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

### III. Sequencing and Analysis

15 Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared 20 using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI 25 protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing 30 vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family 35 databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary

WO 02/26825 PCT/US01/30661  
structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.)  
The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER.

The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences.

Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or

5 Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA  
assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and  
Coused, and cDNA assemblages were screened for open reading frames using programs based on  
GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive  
the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may  
10 begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide  
sequences were subsequently analyzed by querying against databases such as the GenBank protein  
databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden  
Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide  
sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering,  
15 South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide  
sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as  
incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also  
calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of  
20 Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold  
parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second  
column provides brief descriptions thereof, the third column presents appropriate references, all of  
which are incorporated by reference herein in their entirety, and the fourth column presents, where  
applicable, the scores, probability values, and other parameters used to evaluate the strength of a  
25 match between two sequences (the higher the score or the lower the probability value, the greater the  
identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and  
polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID  
NO:17-32. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and  
30 amplification technologies are described in Table 4, column 4.

#### IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative G-protein coupled receptors were initially identified by running the Genscan gene  
identification program against public genomic sequence databases (e.g., gbpr1 and gbhtg). Genscan is  
a general-purpose gene identification program which analyzes genomic DNA sequences from a  
35 variety of organisms (See Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and

WO 02/26825 PCT/US01/30661  
S. Karlin (1998) *Curr. Opin. Struct. Biol.* 3:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode G-protein coupled receptors, the encoded polypeptides were analyzed by querying against PFAM models for G-protein coupled receptors. Potential G-protein coupled receptors were also identified by homology to Incyte cDNA sequences that had been annotated as G-protein coupled receptors. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

#### V. Assembly of Genomic Sequence Data with cDNA Sequence Data

##### 20 "Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent

WO 02/26825 PCT/US01/30661  
type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

5 **“Stretched” Sequences**

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore “stretched” or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

15 **VI. Chromosomal Mapping of GCREC Encoding Polynucleotides**

The sequences which were used to assemble SEQ ID NO:17-32 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:17-32 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome’s p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI “GeneMap’99” World Wide Web site

WO 02/26825 PCT/US01/30661  
 (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

#### VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel (1995) *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding GCREC are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia,

WO 02/26825 PCT/US01/30661  
male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system;  
pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or  
urinary tract. The number of libraries in each category is counted and divided by the total number of  
libraries across all categories. Similarly, each human tissue is classified into one of the following  
5 disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma,  
cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided  
by the total number of libraries across all categories. The resulting percentages reflect the tissue- and  
disease-specific expression of cDNA encoding GCREC. cDNA sequences and cDNA library/tissue  
information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

10 VIII. Extension of GCREC Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate  
fragment of the full length molecule using oligonucleotide primers designed from this fragment. One  
primer was synthesized to initiate 5' extension of the known fragment, and the other primer was  
synthesized to initiate 3' extension of the known fragment. The initial primers were designed using  
15 OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30  
nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target  
sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would  
result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one  
20 extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR  
was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction  
mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ ,  
and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE  
25 enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters  
for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min;  
Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage  
at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C,  
3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4  
30 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN  
quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE  
and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar,  
Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II  
35 (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the

WO 02/26825 PCT/US01/30661  
concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

#### IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:17-32 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10<sup>7</sup> counts per minute of the labeled probe is used in a typical membrane-based

WO 02/26825 PCT/US01/30661  
hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase  
I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon  
membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16  
5 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature  
under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate.  
Hybridization patterns are visualized using autoradiography or an alternative imaging means and  
compared.

#### X. Microarrays

10 The linkage or synthesis of array elements upon a microarray can be achieved utilizing  
photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical  
microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned  
technologies should be uniform and solid with a non-porous surface (Schna (1999), supra).  
Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a  
15 procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface  
of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may  
be produced using available methods and machines well known to those of ordinary skill in the art and  
may contain any appropriate number of elements. (See, e.g., Schna, M. et al. (1995) Science  
270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998)  
20 Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may  
comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be  
selected using software well known in the art such as LASERGENE software (DNASTAR). The  
array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the  
25 biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection.  
After hybridization, nonhybridized nucleotides from the biological sample are removed, and a  
fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser  
desorption and mass spectrometry may be used for detection of hybridization. The degree of  
complementarity and the relative abundance of each polynucleotide which hybridizes to an element on  
30 the microarray may be assessed. In one embodiment, microarray preparation and usage is described  
in detail below.

#### Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and  
poly(A)<sup>+</sup> RNA is purified using the oligo-(dT) cellulose method. Each poly(A)<sup>+</sup> RNA sample is  
35 reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ $\mu$ l oligo-(dT) primer (21mer), 1X first

WO 02/26825 PCT/US01/30661  
strand buffer, 0.03 units/ $\mu$ l RNase inhibitor, 500  $\mu$ M dATP, 500  $\mu$ M dGTP, 500  $\mu$ M dTTP, 40  $\mu$ M dCTP, 40  $\mu$ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)<sup>+</sup> RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)<sup>+</sup> RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14  $\mu$ l 5X SSC/0.2% SDS.

#### Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5  $\mu$ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1  $\mu$ l of the array element DNA, at an average concentration of 100 ng/ $\mu$ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

#### Hybridization

Hybridization reactions contain 9  $\mu$ l of sample mixture consisting of 0.2  $\mu$ g each of Cy3 and

WO 02/26825 PCT/US01/30661  
Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140  
5  $\mu$ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

#### Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an  
10 Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a  
15 resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The  
20 emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on  
25 the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and  
30 adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high  
35 signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and

WO 02/26825 PCT/US01/30661  
measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Lucyte).

#### XI. Complementary Polynucleotides

Sequences complementary to the GCREC-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring GCREC. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of GCREC. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the GCREC-encoding transcript.

#### XII. Expression of GCREC

Expression and purification of GCREC is achieved using bacterial or virus-based expression systems. For expression of GCREC in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*lac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express GCREC upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of GCREC in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding GCREC by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, GCREC is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step,

WO 02/26825 PCT/US01/30661  
affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from GCREC at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ansel (1995, *supra*, ch. 10 and 16). Purified GCREC obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII, where applicable.

### XIII. Functional Assays

GCREC function is assessed by expressing the sequences encoding GCREC at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10  $\mu$ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu$ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) *Flow Cytometry*, Oxford, New York NY.

The influence of GCREC on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding GCREC and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions

WO 02/26825 PCT/US01/30661  
of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding GCREC and other genes of interest can be analyzed by northern analysis or microarray techniques.

#### XIV. Production of GCREC Specific Antibodies

GCREC substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the GCREC amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-GCREC activity by, for example, binding the peptide or GCREC to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

#### XV. Purification of Naturally Occurring GCREC Using Specific Antibodies

Naturally occurring or recombinant GCREC is substantially purified by immunoaffinity chromatography using antibodies specific for GCREC. An immunoaffinity column is constructed by covalently coupling anti-GCREC antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing GCREC are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of GCREC (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/GCREC binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and GCREC is collected.

#### XVI. Identification of Molecules Which Interact with GCREC

Molecules which interact with GCREC may include agonists and antagonists, as well as

WO 02/26825 PCT/US01/30661  
molecules involved in signal transduction, such as G proteins. GCREC, or a fragment thereof, is labeled with <sup>125</sup>I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) A fragment of GCREC includes, for example, a fragment comprising one or more of the three extracellular loops, the extracellular N-terminal region, or the third intracellular loop.

5 Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled GCREC, washed, and any wells with labeled GCREC complex are assayed. Data obtained using different concentrations of GCREC are used to calculate values for the number, affinity, and association of GCREC with the candidate ligand molecules.

Alternatively, molecules interacting with GCREC are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech). GCREC may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Potential GCREC agonists or antagonists may be tested for activation or inhibition of GCREC receptor activity using the assays described in sections XVII and XVIII. Candidate molecules may be selected from known GPCR agonists or antagonists, peptide libraries, or combinatorial chemical libraries.

20 Methods for detecting interactions of GCREC with intracellular signal transduction molecules such as G proteins are based on the premise that internal segments or cytoplasmic domains from an orphan G protein-coupled seven transmembrane receptor may be exchanged with the analogous domains of a known G protein-coupled seven transmembrane receptor and used to identify the G-proteins and downstream signaling pathways activated by the orphan receptor domains (Kobilka, B.K. et al. (1988) *Science* 240:1310-1316). In an analogous fashion, domains of the orphan receptor may be cloned as a portion of a fusion protein and used in binding assays to demonstrate interactions with specific G proteins. Studies have shown that the third intracellular loop of G protein-coupled seven transmembrane receptors is important for G protein interaction and signal transduction (Conklin, B.R. et al. (1993) *Cell* 73:631-641). For example, the DNA fragment corresponding to the third intracellular  
30 loop of GCREC may be amplified by the polymerase chain reaction (PCR) and subcloned into a fusion vector such as pGEX (Pharmacia Biotech). The construct is transformed into an appropriate bacterial host, induced, and the fusion protein is purified from the cell lysate by glutathione-Sepharose 4B (Pharmacia Biotech) affinity chromatography.

For *in vitro* binding assays, cell extracts containing G proteins are prepared by extraction with  
35 50 mM Tris, pH 7.8, 1 mM EGTA, 5 mM MgCl<sub>2</sub>, 20 mM CHAPS, 20% glycerol, 10 μg of both

WO 02/26825 PCT/US01/30661  
aprotinin and leupeptin, and 20  $\mu$ l of 50 mM phenylmethylsulfonyl fluoride. The lysate is incubated on ice for 45 min with constant stirring, centrifuged at 23,000 g for 15 min at 4°C, and the supernatant is collected. 750  $\mu$ g of cell extract is incubated with glutathione S-transferase (GST) fusion protein beads for 2 h at 4°C. The GST beads are washed five times with phosphate-buffered saline. Bound G protein subunits are detected by [<sup>32</sup>P]ADP-ribosylation with pertussis or cholera toxins. The reactions are terminated by the addition of SDS sample buffer (4.6% (w/v) SDS, 10% (v/v)  $\beta$ -mercaptoethanol, 20% (w/v) glycerol, 95.2 mM Tris-HCl, pH 6.8, 0.01% (w/v) bromophenol blue). The [<sup>32</sup>P]ADP-labeled proteins are separated on 10% SDS-PAGE gels, and autoradiographed. The separated proteins in these gels are transferred to nitrocellulose paper, blocked with blotto (5% nonfat dried milk, 50 mM Tris-HCl (pH 8.0), 2 mM CaCl<sub>2</sub>, 80 mM NaCl, 0.02% NaN<sub>3</sub>, and 0.2% Nonidet P-40) for 1 hour at room temperature, followed by incubation for 1.5 hours with G $\alpha$  subtype selective antibodies (1:500; Calbiochem-Novabiochem). After three washes, blots are incubated with horseradish peroxidase (HRP)-conjugated goat anti-rabbit immunoglobulin (1:2000, Cappel, Westchester PA) and visualized by the chemiluminescence-based ECL method (Amersham Corp.).

#### 15 XVII. Demonstration of GCREC Activity

An assay for GCREC activity measures the expression of GCREC on the cell surface. cDNA encoding GCREC is transfected into an appropriate mammalian cell line. Cell surface proteins are labeled with biotin as described (de la Fuente, M.A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using GCREC-specific antibodies, and immunoprecipitated samples are analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of GCREC expressed on the cell surface.

In the alternative, an assay for GCREC activity is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the rate of DNA synthesis in Swiss mouse 3T3 cells. A plasmid containing polynucleotides encoding GCREC is added to quiescent 3T3 cultured cells using transfection methods well known in the art. The transiently transfected cells are then incubated in the presence of [<sup>3</sup>H]thymidine, a radioactive DNA precursor molecule. Varying amounts of GCREC ligand are then added to the cultured cells. Incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a radioisotope counter, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold GCREC ligand concentration range is indicative of receptor activity. One unit of activity per milliliter is defined as the concentration of GCREC producing a 50% response level, where 100% represents maximal incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA (McKay, I. and I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY, p. 73.)

WO 02/26825

PCT/US01/30661

In a further alternative, the assay for GCREC activity is based upon the ability of GPCR family proteins to modulate G protein-activated second messenger signal transduction pathways (e.g., cAMP; Gaudin, P. et al. (1998) *J. Biol. Chem.* 273:4990-4996). A plasmid encoding full length GCREC is transfected into a mammalian cell line (e.g., Chinese hamster ovary (CHO) or human embryonic kidney (HEK-293) cell lines) using methods well-known in the art. Transfected cells are grown in 12-well trays in culture medium for 48 hours, then the culture medium is discarded, and the attached cells are gently washed with PBS. The cells are then incubated in culture medium with or without ligand for 30 minutes, then the medium is removed and cells lysed by treatment with 1 M perchloric acid. The cAMP levels in the lysate are measured by radioimmunoassay using methods well-known in the art. Changes in the levels of cAMP in the lysate from cells exposed to ligand compared to those without ligand are proportional to the amount of GCREC present in the transfected cells.

To measure changes in inositol phosphate levels, the cells are grown in 24-well plates containing  $1 \times 10^5$  cells/well and incubated with inositol-free media and [ $^3$ H]myoinositol, 2  $\mu$ Ci/well, for 48 hr. The culture medium is removed, and the cells washed with buffer containing 10 mM LiCl followed by addition of ligand. The reaction is stopped by addition of perchloric acid. Inositol phosphates are extracted and separated on Dowex AG1-X8 (Bio-Rad) anion exchange resin, and the total labeled inositol phosphates counted by liquid scintillation. Changes in the levels of labeled inositol phosphate from cells exposed to ligand compared to those without ligand are proportional to the amount of GCREC present in the transfected cells.

#### XVIII. Identification of GCREC Ligands

GCREC is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK (Human Embryonic Kidney) 293 which have a good history of GPCR expression and which contain a wide range of G-proteins allowing for functional coupling of the expressed GCREC to downstream effectors. The transformed cells are assayed for activation of the expressed receptors in the presence of candidate ligands. Activity is measured by changes in intracellular second messengers, such as cyclic AMP or  $Ca^{2+}$ . These may be measured directly using standard methods well known in the art, or by the use of reporter gene assays in which a luminescent protein (e.g. firefly luciferase or green fluorescent protein) is under the transcriptional control of a promoter responsive to the stimulation of protein kinase C by the activated receptor (Milligan, G. et al. (1996) *Trends Pharmacol. Sci.* 17:235-237). Assay technologies are available for both of these second messenger systems to allow high throughput readout in multi-well plate format, such as the adenylyl cyclase activation FlashPlate Assay (NEN Life Sciences Products), or fluorescent  $Ca^{2+}$  indicators such as Fluo-4 AM (Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In cases where the physiologically relevant second messenger pathway is not

WO 02/26825 PCT/US01/30661  
known, GCREC may be coexpressed with the G-proteins  $G_{\alpha 1516}$  which have been demonstrated to couple to a wide range of G-proteins (Offermanns, S. and M.I. Simon (1995) J. Biol. Chem. 270:15175-15180), in order to funnel the signal transduction of the GCREC through a pathway involving phospholipase C and  $Ca^{2+}$  mobilization. Alternatively, GCREC may be expressed in  
5 engineered yeast systems which lack endogenous GPCRs, thus providing the advantage of a null background for GCREC activation screening. These yeast systems substitute a human GPCR and  $G_{\alpha}$  protein for the corresponding components of the endogenous yeast pheromone receptor pathway. Downstream signaling pathways are also modified so that the normal yeast response to the signal is converted to positive growth on selective media or to reporter gene expression (Broach, J.R. and J.  
10 Thorne (1996) Nature 384 (supp.):14-16). The receptors are screened against putative ligands including known GPCR ligands and other naturally occurring bioactive molecules. Biological extracts from tissues, biological fluids and cell supernatants are also screened.

Various modifications and variations of the described methods and systems of the invention  
15 will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the  
20 following claims.

WO 02/26825

PCT/US01/30661

Table 1

Incyte Project ID	Polypeptide SEO ID NO:	Incyte Polypeptide ID	Polynucleotide SEO ID NO:	Incyte Polynucleotide ID
2536292	1	2536292CD1	17	2536292CB1
7477708	2	7477708CD1	18	7477708CB1
7474823	3	7474823CD1	19	7474823CB1
644692	4	644692CD1	20	644692CB1
3837054	5	3837054CD1	21	3837054CB1
6157025	6	6157025CD1	22	6157025CB1
55012817	7	55012817CD1	23	55012817CB1
7475061	8	7475061CD1	24	7475061CB1
7477374	9	7477374CD1	25	7477374CB1
7479890	10	7479890CD1	26	7479890CB1
7482825	11	7482825CD1	27	7482825CB1
7483087	12	7483087CD1	28	7483087CB1
7483134	13	7483134CD1	29	7483134CB1
7478550	14	7478550CD1	30	7478550CB1
7483142	15	7483142CD1	31	7483142CB1
7483151	16	7483151CD1	32	7483151CB1

WO 02/26825

PCT/US01/30661

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide	GenBank ID No:	Probability Score	GenBank Homolog
1	2534292CD1	G1039470	8.1E-07	[Rattus norvegicus] pheromone receptor VII (Datta, C. and R. Axel (1996) Cell 83:185-206)
2	7477708CD1	G5525078	1.2E-89	[Rattus norvegicus] seven transmembrane receptor (Abe, J. et al. (1999) J. Biol. Chem. 274:19957-19964)
3	7474623CD1	G10241847	6.0E-79	[Homo sapiens] histamine H4 receptor (Oda, T. et al. (2000) J. Biol. Chem. 275:36781-36786)
4	644692CD1	G2117161	3.2E-35	[Oryzopsis sp.] seven transmembrane-domain receptor (Osterhof, C. et al. (1997) DNA Cell Biol. 16:1379-1389)
5	3837054CD1	G1902964	7.1E-21	[Mus sp.] oxytocin receptor (Rubota, Y. (1996) Mol. Cell. Endocrinol. 124:25-32)
6	6157025CD1	G5339718	2.5E-21	[Homo sapiens] cysteinyl leukotriene receptor (Sara, H.M. et al. (1999) Mol. Pharmacol. 56:657-663)
7	55012817CD1	G6006811	8.6E-60	[Mus musculus] serpentine receptor
8	7475061CD1	G430626	2.2E-22	[Aplysia mexicanus] opsin
9	7477374CD1	G14600082	0.0	[Homo sapiens] trace amine receptor 3 (Borowsky, B. et al. (2001) Proc. Natl. Acad. Sci. USA 98:8966-8971)
10	7479890CD1	G10441732	0.0	[Homo sapiens] leucine-rich repeat-containing G protein-coupled receptor 6 (Heu, S.Y. et al. (2000) Mol. Endocrinol. 14:1257-1271)
11	7482825CD1	G4680254	1.3E-68	[Mus musculus] odorant receptor S1 (Hirono, M.B. et al. (1999) Cell 96:713-721)
12	7483087CD1	G5901488 G2792018	1.5E-101 8.4E-99	[Marmota marmota] olfactory receptor (Homo sapiens) olfactory receptor (Vanderhaeghen, P. et al. (1997) Biochem. Biophys. Res. Commun. 237:283-287)
13	7483134CD1	G6178008	8.3E-52	[Mus musculus] odorant receptor MOR18 (Grabol, A. et al. (1999) J. Neurosci. 19:8409-8418)
14	7478550CD1	G3983394	1.2E-53	[Mus musculus] olfactory receptor F7 (Krautwurst, D. et al. (1998) Cell 95:917-926)
15	7483142CD1	G1246534	1.1E-91	[Gallus gallus] olfactory receptor 4 (Leibovici, M. et al. (1996) Dev. Biol. 175:118-131)
16	7483151CD1	G1246532	1.9E-78	[Gallus gallus] olfactory receptor 3 (Leibovici, M. et al. (1996) Dev. Biol. 175:118-131)

Table 3

SEQ ID NO.	Inocyte Polypeptide ID	Amino Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods Databases
1	253623CD1	217	S183 S53 W172 Y206 T45 Y128	N117 N181	Transmembrane domain: L63-L68	HMMER
2	7477708CD1	578	S179 S136 S224 S248 S255 S570 T185 T190 T346 T46 T499 T539 T97 S4 T16 T51	N122 N130 N175 N183 N228 N90	Transmembrane domains: Y318-Y341 L400-L419 L479-V497 7 transmembrane receptor (secretin family): S313-Y560 Leucokinin/GL-1-like GFS domain: R260-E311 Secretin-like GPCR super family: I518-S538 S576-S597, E318-K342, T9-V33 K390-I413 G-protein coupled receptor PD00752: S307-S559 EPR1 heptahelical receptor: DM05211 T3725 I347-738: C364-S559 DM05211 R48920 I347-738: C365-S559 DM05221 R51172 L465-896: V231-E558	HMMER HMMER-PTAM HMMER-PTAM BLIMPS-PRINTS BLAST- PRODOM BLAST- DOMO
3	7474823CD1	441	S174 S209 S215 S222 S316 S372 S359 S423 T429 T428 T433 T428 T433	M324 M393 M6	Signal peptide: M1-R51 Transmembrane domain: Y226-M247 Family: F131-W235 G-protein coupled receptor motif: A153-V169 G-protein coupled receptors signature: D147-V194 G-protein coupled receptor BL00237: W133-A172, F236-M247, T177-M203 Rhodopsin-like GPCR superfamily PR00237: K184-L205, L228-Y251, V182-V206, W225-Y251, K184-E208, D147-V169	SFScan HMMER HMMER-PTAM MOTIFS ProfileScan BLIMPS- BLOCKS BLIMPS- PRINTS

WO 02/26825

PCT/US01/30661

Table 3 (cont.)

SEQ ID NO.	Leucine Zipper Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Biological Methods and Databases
3					G-protein coupled receptor: DM00013 P30546 22-383: V119-R259 DM00013 P22270 103-599: S118-R257 DM00013 S58868 45-481: S118-R257 DM00013 P31390 21-315: L120-Q281	BLAST-DBO
4	64469ZCD1	797	S249 S3 S344 S349 S350 S368 S389 S466 S492 S499 S647 S660 S665 S666 S696 S746 S781 S788 T192 T225 T233 T294 T338 T423 T459 T47 T458 T661 T721 T768 T786	N159 N178 N191 N247 N261 N312 N316 N387 N413 N657 N709 N82	Signal peptide: M1-G26 Transmembrane domain: N8-W30, V470-I489, I567-Q592, V632-W652, S672-I689 7 transmembrane receptor (Secretin family) 7tm.2: P431-V726 Tetrahelin/CU-1-like GFS domain GFS: Y379-D427 G-protein coupled receptor BL00649: S389-W416, G461-Y485, C526-L551, G571-Q598, W626-N655, S666-A687, P710-E735 C. elegans integral membrane protein SFB signature PR06699: V467-F487 HORMONE, ENRI, GSDOCCTRE, ANTIGEN: DM02211 A5712 469-866: C526-Y720 DM02211 E37225 347-738: F603-F731 PROTEIN COUPLED TRANSMEMBRANE RECEPTOR ED000724: N311-E730	MOTIFS HMMER HMMER HMMER-PFAM HMMER-PFAM
5	3837054CD1	434	S41 S131 S293	N15 N27 N60	G-PROTEIN COUPLED RECEPTORS: DM00013 P47901 28-353: N43-V182, W260-A357, W271-L256 G-protein coupled receptor BL00237: A119-F134, E288-A314, N344-C360 Transmembrane domain: R45-C65 7-transmembrane receptor (Rhodopsin family, 7tm.1): G59-F181, H228-V352 Signal cleavage: M1-A58	BLAST-DBO PRODOM BLAST-DBO BLIAPS- BLOCKS HMMER HMMER-PFAM SPSCAN

WO 02/26825

PCT/US01/30661

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Programs
6	6157025CD1	339	96 S10 Y175 S217 S276 S332	N5 N9 R308	G-PROTEIN COUPLED RECEPTORS: DM00013 P34993 36-327: L20-F285 DM00013 P30872 52-338: L20-F285 G-PROTEIN COUPLED RECEPTOR BL00237: R218-F244 N9-F25 N81-C120 Transmembrane domain: T21-M45, Y43-S63 M89-T116 M139-Y155, I185-V208 R227-F249	BLAST-DJMO
7	55012817CD1	549	S19 S57 T103 T153 T224 T3 T407 T415 T537 T84	M144 N210 M413 N98	Signal peptide: M1-C20 Signal peptide: M1-S22 Transmembrane domains: G439-L523, L477-F495, H436-F456, F378-L405, A274-L293 7 transmembrane receptor (Secretin family): V265-S528	SPScan HAMER HAMER
8	7475061CD1	188	S163 S179 S64 T159 T174		7 transmembrane receptor (Secretin family): V265-S528 Secretin-like GPCR superfamily R270-R294, A340-L363, F379-G404, L437-I517, K475-F495, V503-L524 MR1 leucocyte antigen: DM05221 I37225 347-738: P207-I517 DM05221 P48960 347-738: P207-I517 DM05221 A57172 463-866: E177-I517 G-protein coupled receptor PD000752: L271-L524 Signal peptide: M1-I50 Transmembrane domains: F34-F52, N90-F112	BLIMP8-BLOCKS BLIMP8-PRINTS BLAST-DJMO BLAST-PRODOM SPScan HAMER



Table 3 (cont.)

SEQ ID NO	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence, Domains and Motifs	Analytical Method and Database
9					G-protein coupled receptor PD000009: K61-N132	BLAST- PRODOM
10	7479890CD1	948	S107 S174 S19 S488 S489 S666 S688 S843 S855 T309 Y456 Y504	N189	Signal peptide: M1-G24 Signal peptide: M1-N25 Transmembrane domains: A548-G571, V756-F783 Leucine rich repeats: K287-T309, R239-F262, R265-F286, K356-S379, S380-F383, K384-Q397, S444-N467, S488-N501, S494-N507, S444-N467, A169-T191, S192-H215, N216-G238 Leucine rich repeat N-terminal domain: A34-D65 Leucine zipper pattern: L57-L78 G-protein coupled receptor BM00237: R16-V655, F800-R816 Glycoprotein hormone receptor PR00373: F533-W550, F610-C623, C623-S637, W746-L766 G-protein coupled receptors: DM00013 F14763 607-693: F535-L619 DM00013 F55376 355-641: F535-D618 DM00013 F22888 352-638: F535-D618 DM00013 F35409 519-807: I545-L619 Orphan G-protein coupled receptor HG38: FD17529: H428-F568 FD16963: M259-F333 FD166277: A168-H215 G-PROTEIN COUPLED RECEPTORS: DM00013 F23270 18-311: L23-I306	BLAST- PRODOM BLAST- PRODOM BLAST- PRODOM BLAST- PRODOM BLAST- DOMO BLAST- PRODOM BLAST- DOMO
11	7482825CD1	315	S234 S54 S69 T10 T292 T6	S67 N8		BLAST- DOMO

WO 02/26825

PCT/US01/30661

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Prostaglandin-Endothelin-1 Receptor Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Software
11					RECEPTOR OLFACTORY PROTEIN RECEPTOR-LIKE GPROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY ED000921: L167-L246 G-protein coupled receptors proteins BL00237: F283-H288, A91-F130, L208-V219, R236-A262 G-protein coupled receptors signature motifs: F109-L192, F239-G254, F273-F286, F292-L296 Rhodopsin-like GPCR superfamily signature F823-L849, M1-K82, F109-L177, F200-L223, K273-H299 G-protein coupled receptors signature motifs: F109-F149 Transmembrane domain: F271-F31, F102-D122, M143-A165, L201-F217, 7 transmembrane receptor (rhodopsin family) 7tm.1: G43-L291 G-Protein Receptor: All-IL127 G-PROTEIN COUPLED RECEPTORS: DM00013 F23265 I7-306: E22-I305 DM00013 F23268 I8-307: D20-I305 DM00013 F30955 I8-305: D20-I305 DM00013 S29707 I8-306: F21-L301 RECEPTOR OLFACTORY PROTEIN RECEPTOR-LIKE GPROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY ED000921: L166-L246 OLFACTORY RECEPTOR PROTEIN RECEPTOR-LIKE GPROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY ED149621: V247-F307	BLAST- PRODOM BLIMPS- BLOCKS BLIMPS- PRINTS PRINTS PROFILESCAN HMMER HMMER-PPM MOTIFS BLAST-DOMO
12	7483087CD1	312	S152 S267 S268 S291 S67	NS N93		BLAST- PRODOM BLAST- PRODOM

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Packages
12					G-protein coupled receptors proteins BL00237: F282-E298, Q90-E143 OLFACTORY RECEPTOR SIGNATURES PR00245: M59-K80, F177-F191, F238-G253, K274-E286, S291-E305 Transmembrane domain: I130-I146 G109-M118, E143-M162, I197-F216 7 transmembrane receptor (Rhodopsin family) / GPCRs: G110-E126 G-Protein Receptor: G110-E126 G-protein coupled receptors signature B-protein_receptor: PF: C102-V151	BLIMPS- BLOCKS BLIMPS- PRINTS HMMER HMMER-PPM MOTIFS PROFILESSCAN
13	7483134CD1	309	S227 S305 S65 M6 T286 T52 R83		G-PROTEIN COUPLED RECEPTORS: DM00013 S297I015-301: F26-F300 DM00013 F232S0617-306: E20-E299 DM00013 S2970818-306: E20-E296 DM00013 F2327418-306: E20-M296 RECEPTOR OLFACTORY PROTEIN RECEPTORLIKE GPROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY PD000921: L164-I243 G-protein coupled receptors proteins BL00237: K88-F127, F277-E293 Rhodopsin-like GPCR superfamily signature PR00237: T24-M48, M57-K78, A102-L124, K138-L159, V197-L220, K267-E293 Olfactory receptor signature PR00245: M57-K78, E175-D189, V235-E250, T286-F300 transmembrane domains: V28-E45, I204-M222	ELAST-DOMO BLIMPS- BLOCKS BLIMPS- PRINTS BLIMPS- PRINTS HMMER

Table 3 (cont.)

SEQ ID NO	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
13					7 transmembrane receptor (rhodopsin family), 7tm.1: G39-Y285 G Protein Receptor: H108-F124 G-protein coupled receptor signature G-protein_receptor_psf: F100-R145 Signal cleavage: M1-G39	HMMER_Pfam MOTIFS PROFILESCAN
14	7478550CD1	309	S107 T21 T26 R3 T46	N46 M51	Transmembrane domains: 168-386, M238-S256 transmembrane receptor (rhodopsin family): G61-I232 G-protein coupled receptors signature: F142-G186 G-protein coupled receptor BL00237: K130-S169 (E-value=0.018) Olfactory receptor signature PR00245: M99-K120, F217-D231, F278-A293 Rhodopsin-like GPCR superfamily signature PR00237: F144-I166, M180-V201, V239-L262, A277-R301, F66-H90, M99-K120 RECEPTOR OLFACATORY G PROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY ED000921: I206-L285 G-PROTEIN COUPLED RECEPTORS DM00013 E232466 17-306: L67-L107 DM00013 S51356 18-307: L67-S105 DM00013 S29707 18-306: L58-T108 DM00013 E23274 18-306: L58-T108 G-protein coupled receptors signature: S150-I166	HMMER_Pfam PROFILESCAN BLIMPS- BLOCKS BLIMPS- FRIMPS BLIMPS- FRIMPS ELAST- PRODOM ELAST-DOMO

WO 02/26825

PCT/US01/30661

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Residues	Potential Phosphorylation Sites	Potential Ubiquitination Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15	7483142CD1	315	S191 S294 S70	N68 N9	Signal cleavage: M1-G54 Transmembrane domain: F31-I51 7 transmembrane receptor (rhodopsin family): M4-Y293 G-protein coupled receptor SM00237: Y93-P132 K218-K264 F285-K301 Olfactory receptor signatures PR00245: F180-D184, C241-C256, V277-L286, S284-L308, M62-M83 GPCR (rhodopsin)-signature PR00644: I52-Y63, S294-M305 G-protein coupled receptors SIGNATURE: F109-V153 RECEPTOR OLFACTORY G PROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MOLTIGENE FAMILY: PD000241: L169-K249 PD14921: V250-K311 G-PROTEIN COUPLED RECEPTORS: DM00013 F37067 I7-306: T21-L504 DM00013 S29709 I1-299: T21-L508 DM00013 F23266 I7-306: F24-L508 DM00013 S11556 I8-307: T21-L504 G-protein coupled receptors signature: S113-I129	HMMER-PPAM HMMER-PPAM BLIPS-BLOCKS BLIPS-PRINTS BLIPS-PRINTS PROFILES-SCAN BLAST-FRODOO BLAST-DOMO MOTIFS

Table 3 (cont.)

SEQ ID NO	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Addresses
16	748313CD1	307	S230 S251 S266 S291 278	N107	<p>Transmembrane domains:            F28-A47, I196-I215            7 transmembrane receptor (rhodopsin family), G41-Y290            G-protein coupled receptor E100237;            H90-E129, I285-F298            G-protein coupled receptors            signature: I102-F150            Olfactory receptor signature PR00245:            S291-R305, I39-R60, F177-W191,            F238-S253, I274-L285            Rhodopsin-like GPCR superfamily            signature PR00237:            I39-R80, F104-I126, I201-A224,            S272-L288, E26-R50            RECEPTOR OLFACTORY G PROTEIN COUPLED            TRANSMEMBRANE GLYCOPROTEIN MULTIGENE            FAMILY:            PD000221: L166-L246            PDL9621: V248-R305            G-PROTEIN COUPLED RECEPTORS            DM00013 F37067 I7-306: L17-F104            DM00013 S1356 I8-307: L17-R303            DM00013 S29709 I1-299: T18-R305            DM00013 F23274 I8-306: Q24-R305            G-protein coupled receptors            signature: T110-I128</p>	<p>HMMER            HMMER-PTAM            ELIIPS-BLOCKS            PROFILSCAN            ELIIPS-PRINTS            ELIIPS-PRINTS            ELAST-PRODOM            ELAST-DOMO            MOTIFS</p>

Table 4

Polynucleotide Seq ID NO:	Incyte Polynucleotide	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
17	2536292CB1	2422	298-1079, 2211-2422, 105-150, 1835-1898	72051732VI, 2536292F6 (BRANINCHI) FL2536292_g6007890_000006_g499 2709 55062753HI 55062751JI 55062765HI 72049620VI 55062754VI 55094142VI	787 836 301	2422 1800 1175
18	7477708CB1	1912	1798-1912, 1-1741, 1304-1481	FL7477708_g8217793_000019_g552 5078	1 480 648 1488 658 1	636 2422 2422 2422 639 1912
19	7474823CB1	1326	1-185, 231-1326, 261-396, 626-1326	1971428HI (UCMCL5T01) GBL_57347159_000017_000014.edi t GNN_57547159_000014_004 GNN_57106155_000002_002.edi t GNN_57547159_000014_004 GNN_57106155_000002_002.edi t GNN_57547159_000014_004 GNN_57106155_000002_002.edi t GNN_57547159_000014_004 GNN_57106155_000002_002.edi t	1028 144 517 1 1922 370 1490 1 2612 2293 910 1088 849	782 782 1326 516 2535 1195 2079 671 3058 2985 1678 1541 1349
20	644692CB1	3058	1-1633, 2354-3058, 2369-2504, 1-2201	1-1633, 2354-3058, 2369-2504, 1-2201	1 1 1 1	1326 516 2535 1195 2079 671
21	3837054CB1	1993	1-1063, 1540-1714, 1-87, 1543-1993, 447-829	93738039 FL3837054_g8575884_g1150834_1 2 GNN_48575884_004 56008250L 3837054F6 (DENPNT01) 2814920HI (OVARNOCT10)	1088 849 359 1649 1371 1	1541 1349 1420 1993 1933 187

WO 02/26825

PCT/US01/30661

Table 4 (cont).

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence length	Selected Fragments	Sequence Fragments	5' Position	3' Position
21				3406743H1 (PROSMIS08)	103	354
22	6157025CB1	1499	1-81,	GNN C5686570_000054_002.edit	116	565
			248-1499,	FN6157025_910567930_000001_953	836	1499
			850-970,	53887_1		
			1265-1499	7169157V1 (NCEINOC01)	1	676
				4425316H1 (BRAPDT01)	32	288
23	55012817CB1	2455	1397-1482,	71691480V1	1746	2401
			318-704,	71691523V1	1405	2082
			1693-2171,	71691220V1	1	646
			1-1014,	71690572V1	654	1377
			1619-2182	71691571V1	608	1338
24	7475061CB1	2056		71688567V1	1351	2058
			1-688,	72488727D1	2162	2455
			1207-2056,	72489896D1	841	1769
			1736-1917,	72487236D1	1	953
			1207-1680	72492196D1	1046	2056
25	7477374CB1	999	1-138,	FL747374_99930948_000007_6739	1200	2056
			264-999	496	1	999
			1-2724,	4021537F6 (ERRAXNOT02)	1427	2110
			2850-2900,	GBL G9567464_000017_000007_000	1	588
			1-2300,	001.edit		
26	7479890CB1	3429	2367-2732	55017814H1	737	1280
				58013229J1	2532	3429
				8042178J1 (OVARVUE01)	2032	2692
				GNN G9567464_000015_002	729	2847
				71702678V1	2160	2786
				6609077H2 (PLACFRC01)	365	877
				7720776H1 (PHYRDE01)	880	1291
				7726057J1 (PHYRDE01)	1242	1974

WO 02/26825

PCT/US01/30661

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
27	7483282CB1	948	1-87, 210-289, 923-948,	CGN: g10045182_000002_016	1	948
38	7483087CB1	939	259-939, 1-86,	CGN: g6015288_000044_010	1	939
29	7483134CB1	930	912-939 1-927, 1-100, 889-930, 400-930,	CGN: g7143464_000018_002	1	930
30	7478550CB1	1161	112-335, 116-388, 925-1161	707797288 (ERR07804) 707797278 (ERR07804)	1	926 1161
31	7483142CB1	948	1-25, 605-782, 913-948,	WT:7483142_g8086488_000017_g374 6443_1_1	1	948
32	7483151CB1	924	876-924	WT:7483151_g9438337_000003_g151 4460_1_1-2	1	924

WO 02/26825

PCT/US01/30661

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
17	2536292CB1	BRA180718
19	7474823CB1	DCMCL5701
20	644692CB1	SIMVPM02
21	3837054CB1	DENDY7F01
22	6157025CB1	MONOTM05
23	55012817CB1	BRAPD1F01
26	7479890CB1	LUNGR007
30	7478550CB1	BRAUDR04

Table 6

Library	Vector	Library Description
BRA1N0T18	P1N1C1	Library was constructed using RNA isolated from left temporal lobe brain tissue removed from a 34-year-old Caucasian male during cerebral meninges lesion excision. Pathology for the associated tumor tissue indicated metastatic malignant melanoma. Neoplastic cells strongly expressed HMB-45. Patient history included malignant melanoma of skin of the trunk. Family history included liver cancer, acute myocardial infarction, atherosclerotic coronary artery disease, and cerebrovascular disease.
BRA1P1T01	P1N1C1	Library was constructed using RNA isolated from diseased pons tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Serology was negative. Patient history included Huntington's disease, emphysema, and tobacco abuse.
BRA1V1D1R04	P1C1M1A2.1	Library was constructed using RNA isolated from striatum, dorsal caudate nucleus, dorsal putamen, and ventral nucleus accumbens tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated no diagnostic abnormalities in the brain or intracranial vessels. There was mild meningeal fibrosis predominately over the convexities. Special stains showed no evidence of amyloid plaques or metastatic lesions. There were scattered axonal spheroids in the white matter of the cingulate cortex and thalamus. There were a few scattered neurofibrillary tangles in the entorhinal cortex and periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor, surrounded by foci of bile lakes beneath the hepatic surface scar. The liver had extensive surface scarring, congestion, cholestasis, hemorrhage, necrosis, and chronic inflammation. The patient presented with nausea, vomiting, dehydration, malnutrition, oliguria, and acute renal failure. Patient history included post-operative Budd-Chiari syndrome, biliary ascites, acute bilateral bronchopneumonia with microabscesses, hemothorax, and bilateral leg pitting edema. Previous surgeries included cholecystectomy, liver resection, hysterectomy, bilateral salpingo-oophorectomy, and portocaval shunt. The patient was treated with a nasogastric feeding tube, biliary drainage stent, paracentesis, pleurodesis, and abdominal ultrasound. Patient medications included Ampicillin, niacin, furosemide, Albuterone, Benadryl, and morphine.

WO 02/26825

PCT/US01/30661

Table 6 (cont.)

Library	Vector	Library Description
DEMD7MT01	pINCY	Library was constructed using RNA isolated from treated dendritic cells from peripheral blood.
LUMENON07	pINCY	This normalized lung tissue library was constructed from 5.1 million independent clones from a lung tissue library. Starting RNA was made from RNA isolated from lung tissue. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
MOROTXN05	pINCY	This normalized treated monocyte cell tissue library was constructed from 1.03 million independent clones from a monocyte tissue library. Starting RNA was made from RNA isolated from treated monocytes from peripheral blood removed from a 42-year-old female. The cells were treated with interleukin-10 (IL-10) and lipopolysaccharide (LPS). The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996) 791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
SLNTTMR02	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 59-year-old male. Pathology for the matched tumor tissue indicated multiple (9) carcinoid tumors, grade 1, in the small bowel. The largest tumor was associated with a large mesenteric mass. Multiple convoluted segments of bowel were adhered to the tumor. A single (1 of 13) regional lymph node was positive for malignancy. The peritoneal biopsy indicated focal fat necrosis.
UCMCL5T01	PBLUESCRIPT	Library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of 12 individuals. The cells were cultured for 12 days with IL-5 before RNA was obtained from the pooled lysates.

Table 7

Program	Description	Reference	Parameter Threshold
ABI/FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL.FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, flfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.0E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=1.00 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLocks, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Atwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sommadossi, J.P. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM hit: Probability value= 1.0E-3 or less Signal peptide hit: Score= 0 or greater

WO 02/26825

PCT/US01/30661

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Batroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores>GCC-specified "HIGHP" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phis Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match lengths= 50 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SFScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Clavette, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Batroch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, pages M51-59, Genetics Computer Group, Madison, WI.	

WO 02/26825  
What is claimed is:

PCT/US01/30661

1. An isolated polypeptide selected from the group consisting of:
- 5 a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16,
- b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16,
- 10 c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and
- d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-16.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
- 20 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:17-32.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
- 30 9. A method of producing a polypeptide of claim 1, the method comprising:
- a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
- 35 b) recovering the polypeptide so expressed.

WO 02/26825

PCT/US01/30661

10. A method of claim 9, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-16.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

5

12. An isolated polynucleotide selected from the group consisting of:

- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32,
- b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32,
- c) a polynucleotide complementary to a polynucleotide of a),
- d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).

10

15

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide

20 having a sequence of a polynucleotide of claim 12, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

25

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

30

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment

35

WO 02/26825

PCT/US01/30661

thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

5

18. A composition of claim 17, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-16.

19. A method for treating a disease or condition associated with decreased expression of functional GCREC, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- 15 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and  
b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

20

22. A method for treating a disease or condition associated with decreased expression of functional GCREC, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- 25 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and  
b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

30

25. A method for treating a disease or condition associated with overexpression of functional GCREC, comprising administering to a patient in need of such treatment a composition of claim 24.

35

WO 02/26825

PCT/US01/30661

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- 5 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

- 10 a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- 15 c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- 25 b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

- 30 a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,
- 35

WO 02/26825

PCT/US01/30661

- c) quantifying the amount of hybridization complex, and  
d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.
- 5
30. A diagnostic test for a condition or disease associated with the expression of GCREC in a biological sample, the method comprising:
- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and  
b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
- 10
31. The antibody of claim 11, wherein the antibody is:
- a) a chimeric antibody,  
b) a single chain antibody,  
c) a Fab fragment,  
d) a F(ab')<sub>2</sub> fragment, or  
e) a humanized antibody.
- 15
- 20
32. A composition comprising an antibody of claim 11 and an acceptable excipient.
33. A method of diagnosing a condition or disease associated with the expression of GCREC in a subject, comprising administering to said subject an effective amount of the composition of claim 32.
- 25
34. A composition of claim 32, wherein the antibody is labeled.
- 30
35. A method of diagnosing a condition or disease associated with the expression of GCREC in a subject, comprising administering to said subject an effective amount of the composition of claim 34.
36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim

WO 02/26825

PCT/US01/30661

11, the method comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- 5 b) isolating antibodies from said animal, and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16.

10 37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim

15 11, the method comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- 20 c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- d) culturing the hybridoma cells, and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ
- 25 ID NO:1-16.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

30

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant

P2WO 02/26825

PCT/US01/30661

immunoglobulin library.

44. A method of detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16 in a sample, the method comprising:

- 5 a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16 in the sample.
- 10

45. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16 from a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- 15 b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

20

47. A method of generating a transcript image of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- 25 b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
- c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

30

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is

PCT/US01/30661

PCT/US01/30661

completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

5

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

10

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

15

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

20

56. A method of identifying a compound that modulates, mimics and/or blocks an olfactory and/or taste sensation, the method comprising:

25

a) contacting the compound with an olfactory and/or taste receptor polypeptide selected from the group consisting of:

i) a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16,

ii) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and

30

iii) an olfactory and/or taste receptor having an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16.

b) identifying whether the compound specifically binds to and/or affects the activity of

WO 02/26825

PCT/US01/30661

said receptor polypeptide.

57. The method of claim 56, wherein said receptor polypeptide is expressed on the surface of a mammalian cell.

5

58. The method of claim 57, wherein said mammalian cell expresses a G-protein.

59. The method of claim 58, wherein said mammalian cell expresses a plurality of G-protein coupled receptors.

10

60. The method of claim 59, wherein said mammalian cell expresses another olfactory and/or taste receptor polypeptide.

61. The method of claim 56, wherein said receptor polypeptide is fused to another polypeptide.

15

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

20

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

25

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

30

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

WO 02/26825

PCT/US01/30661

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
- 5 73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
- 10 76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
- 15 78. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:17.
79. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:18.
80. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:19.
- 20 81. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:20.
82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:21.
- 25 83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:22.
84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:23.
85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:24.
- 30 86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:25.
87. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:26.

WO 02/26825

PCT/US01/30661

- 88. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:27.
- 89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:28.
- 5 90. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:29.
- 91. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:30.
- 92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:31.
- 10 93. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:32.

WO 02/26825

PCT/US01/30661

PI-0236 PCT

<110> INCYTE GENOMICS, INC.  
 BAUGHN, Mariah R.  
 GRAUL Richard C.  
 WALIA, Narinder K.  
 GANDHI, Ameena R.  
 HAFALIA, April J.A.  
 RAMKUMAR, Jaysalaxmi  
 TRIBOULEY, Catherine M.  
 THORNTON, Michael  
 KALLICK, Deborah A.  
 YAO, Monique G.  
 ELLIOTT, Vicki S.  
 BURFORD, Neil  
 KHAN, Farrah A.  
 YUE, Henry  
 LU, Yan  
 ARVIZU, Chandra  
 ROOPA, Reddy  
 NGUYEN, Dannel B.  
 LEE, Ernestine A.  
 LU, Dyung Aina M.  
 ISON, Craig H.  
 WALSH, Roderick T.  
 POLICKY, Jennifer L.

<120> G-PROTEIN COUPLED RECEPTORS

<130> PI-0236 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/236,546; 60/242,223; 60/245,900; 60/247,587; 60/249,343; 60/240,589;  
 60/242,322; 60/245,855

<151> 2000-09-29; 2000-10-20; 2000-11-03; 2000-11-09; 2000-11-15; 2000-10-13;  
 2000-10-20; 2000-11-03

<160> 32

<170> PERL Program

<210> 1

<211> 217

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2536292CD1

<400> 1

Met Lys Ser Phe Leu Pro Gly Thr Cys Ile Leu Leu Cys Ser Ala  
 1 5 10 15

WO 02/26825

PCT/US01/30661

PI-0236 PCT

Phe Asn Leu Met Phe Phe Ser Leu Phe Arg Leu Lys Tyr Asn Ile  
 20 25 30  
 Cys Ile Ile Leu Arg Ala Cys Asn Thr Met Leu Ser Ser Asn Thr  
 35 40 45  
 Ile Met Glu Ile Phe Phe Leu Ser His Ile Asp Ile Gly Ile Trp  
 50 55 60  
 Arg Asn Leu Leu Leu Leu Met Pro Ile Tyr Thr Phe Leu Ile  
 65 70 75  
 Cys Pro Gln Gln Lys Lys Pro Met Gly Leu Leu Phe Leu His Leu  
 80 85 90  
 Ser Val Ala Asn Thr Met Thr Leu Leu Arg Lys Val Ile Pro Leu  
 95 100 105  
 Ala Val Lys Ser Phe Asn Thr Lys Asn Leu Leu Asn Tyr Thr Gly  
 110 115 120  
 Cys Arg Glu Phe Glu Phe Leu Tyr Arg Val Ser Trp Gly Leu Pro  
 125 130 135  
 Leu Cys Thr Thr Tyr Leu Leu Ser Met Val Gln Ala Leu Arg Gly  
 140 145 150  
 Ser Pro Ser Lys Ser Arg Val Ile Asn Ser Leu Ile Tyr Ile Lys  
 155 160 165  
 Leu Val Pro Phe Val Asp Thr Ile Lys Tyr Gly Ser Val Thr Lys  
 170 175 180  
 Asn Leu Ser Ile Lys Met Cys Leu Ala Thr Pro His Met Gly Asn  
 185 190 195  
 Thr Ile Ala Val Ser His Thr Ser Val Ile Thr Phe Gln Asp Leu  
 200 205 210  
 Ile Phe Leu Val Leu Met Ser  
 215

&lt;210&gt; 2

&lt;211&gt; 578

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7477708CD1

&lt;400&gt; 2

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

WO 02/26825

PCT/US01/30661

PI-0236 PCT

95 100 105  
 Ser Tyr Ser Thr Ile Ala Asn His Ile Leu Asn Ser Lys Ser Ile  
 110 115 120  
 Ser Asn Trp Thr Phe Ile Pro Asp Arg Asn Ser Ser Tyr Ile Leu  
 125 130 135  
 Leu His Ser Val Asn Ser Phe Ala Arg Arg Leu Phe Ile Asp Asn  
 140 145 150  
 Ile Pro Val Asp Ile Ser Asp Val Phe Ile His Thr Met Gly Thr  
 155 160 165  
 Thr Ile Ser Gly Asp Asn Ile Gly Lys Asn Phe Thr Phe Ser Met  
 170 175 180  
 Arg Ile Asn Asp Thr Ser Asn Glu Val Thr Gly Arg Val Leu Ile  
 185 190 195  
 Ser Arg Asp Glu Leu Arg Lys Val Pro Ser Pro Ser Gln Val Ile  
 200 205 210  
 Ser Ile Ala Phe Pro Thr Ile Gly Ala Ile Leu Glu Ala Ser Leu  
 215 220 225  
 Leu Glu Asn Val Thr Val Asn Gly Leu Val Leu Ser Ala Ile Leu  
 230 235 240  
 Pro Lys Glu Leu Lys Arg Ile Ser Leu Ile Phe Glu Lys Ile Ser  
 245 250 255  
 Lys Ser Glu Glu Arg Thr Gln Cys Val Gly Trp His Ser Val  
 260 265 270  
 Glu Asn Arg Trp Asp Gln Gln Ala Cys Lys Met Ile Gln Glu Asn  
 275 280 285  
 Ser Gln Gln Ala Val Cys Lys Cys Arg Pro Ser Lys Leu Phe Thr  
 290 295 300  
 Ser Phe Ser Ile Leu Met Ser Pro His Ile Leu Glu Ser Leu Ile  
 305 310 315  
 Leu Thr Tyr Ile Thr Tyr Val Gly Leu Gly Ile Ser Ile Cys Ser  
 320 325 330  
 Leu Ile Leu Cys Leu Ser Ile Glu Val Leu Val Trp Ser Gln Val  
 335 340 345  
 Thr Lys Thr Glu Ile Thr Tyr Leu Arg His Val Cys Ile Val Asn  
 350 355 360  
 Ile Ala Ala Thr Leu Leu Met Ala Asp Val Trp Phe Ile Val Ala  
 365 370 375  
 Ser Phe Leu Ser Gly Pro Ile Thr His His Lys Gly Cys Val Ala  
 380 385 390  
 Ala Thr Phe Phe Val His Phe Phe Tyr Leu Ser Val Phe Phe Trp  
 395 400 405  
 Met Leu Ala Lys Ala Leu Leu Ile Leu Tyr Gly Ile Met Ile Val  
 410 415 420  
 Phe His Thr Leu Pro Lys Ser Val Leu Val Ala Ser Leu Phe Ser  
 425 430 435  
 Val Gly Tyr Gly Cys Pro Leu Ala Ile Ala Ala Ile Thr Val Ala  
 440 445 450  
 Ala Thr Glu Pro Gly Lys Gly Tyr Leu Arg Pro Glu Ile Cys Trp  
 455 460 465  
 Leu Asn Trp Asp Met Thr Lys Ala Leu Leu Ala Phe Val Ile Pro  
 470 475 480  
 Ala Leu Ala Ile Val Val Val Asn Leu Ile Thr Val Thr Leu Val

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

      485          490          495
Ile Val Lys Thr Gln Arg Ala Ala Ile Gly Asn Ser Met Phe Gln
      500          505          510
Glu Val Arg Ala Ile Val Arg Ile Ser Lys Asn Ile Ala Ile Leu
      515          520          525
Thr Pro Leu Leu Gly Leu Thr Trp Gly Phe Gly Val Ala Thr Val
      530          535          540
Ile Asp Asp Arg Ser Leu Ala Phe His Ile Ile Phe Ser Leu Leu
      545          550          555
Asn Ala Phe Gln Val Ser Pro Asp Ala Ser Asp Gln Val Gln Ser
      560          565          570
Glu Arg Ile His Glu Asp Val Leu
      575

```

```

<210> 3
<211> 441
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7474823CD1

```

```

<400> 3
Met Val Leu Gly Lys Asn Val Ser Met Ser Gly Pro Arg Pro Ala
      1           5           10           15
Ser Trp Gln Ser His Pro Gln Gly Leu Glu Leu Val Phe Gly Lys
      20           25           30
Trp Pro Cys Arg Cys Ser Tyr Ser Ala Val Leu Val Ile Ser Ser
      35           40           45
Ile Ser Ser Ser Gly Ala Gly Asp Ile Pro Asp Gln Asp Ser Gly
      50           55           60
Gln Tyr Trp Phe Leu Met Arg Ala Val Phe Leu Ala Cys Arg Arg
      65           70           75
Leu Pro Ser Thr Cys Val Leu Lys Arg Pro Phe Ser Glu Cys Ala
      80           85           90
Gln Arg Glu Arg Thr Asn Leu Val Leu Met Lys Lys Trp Glu Phe
      95          100          105
Leu Glu Val Pro Asp Thr Phe Glu Val Thr Gln Gln Ser Val Ile
      110          115          120
Ser Ile Pro Leu Tyr Ile Pro His Thr Leu Phe Glu Trp Asp Phe
      125          130          135
Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp Tyr Leu Leu
      140          145          150
Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr Asp Arg
      155          160          165
Tyr Leu Ser Val Ser Asn Ala Val Ser Tyr Arg Thr Gln His Thr
      170          175          180
Gly Val Leu Lys Ile Val Thr Leu Met Val Ala Val Trp Val Leu
      185          190          195
Ala Phe Leu Val Asn Gly Pro Met Ile Leu Val Ser Glu Ser Trp
      200          205          210

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

Lys Asp Glu Gly Ser Glu Cys Glu Pro Gly Phe Phe Ser Glu Trp  
 215 220 225  
 Tyr Ile Leu Ala Ile Thr Ser Phe Leu Glu Phe Val Ile Pro Val  
 230 235 240  
 Ile Leu Val Ala Tyr Phe Asn Met Asn Ile Tyr Trp Ser Leu Trp  
 245 250 255  
 Lys Arg Asp His Leu Arg Leu Gly His Pro Lys Gly Trp Gly Gln  
 260 265 270  
 Leu Val Leu Arg Leu Pro His Gly Val Glu Gly Gln Pro Trp Arg  
 275 280 285  
 Leu Gln Leu Val Pro Arg Met Gly Tyr Ile Glu Val Gly Gly Leu  
 290 295 300  
 Leu Cys Thr Ala Ala Gly Glu Met Ser Thr His Ala Arg Ser Ala  
 305 310 315  
 Lys Leu Leu Ser Thr Gly Ser Glu Asn Asp Thr Leu Pro Val Pro  
 320 325 330  
 Ser Leu Ala Ser Arg Ser Leu Cys Pro Ser Val Leu Ser Leu Gly  
 335 340 345  
 Ser Phe Pro Ser Cys Gln Ser Cys Leu Ser Asp Gln Met Ser Gln  
 350 355 360  
 Cys Asp Thr Glu Pro Glu Arg Lys Ser Phe Leu Ser Met Met Gln  
 365 370 375  
 Gly Thr Gln His Phe Asp Asn Pro Asp Gly Met Trp Ser Ser His  
 380 385 390  
 Gly Arg Asn Val Ser Ser Gly Gly Leu His Asn His Cys Ile Leu  
 395 400 405  
 Gln Met Gly Thr Gly Ser Ala Gly Ala Ser His Pro Glu Gly Pro  
 410 415 420  
 Arg Gly Gly Gln Gly Gln Val Thr Thr Arg Ala Thr Thr Gln Lys  
 425 430 435  
 Arg Val Ala Ala Ser Gly  
 440

&lt;210&gt; 4

&lt;211&gt; 797

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 644692CD1

&lt;400&gt; 4

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

WO 02/26825

PCT/US01/30661

PI-0236 PCT

	65		70		75									
Ile	Ala	Asn	Phe	Cys	Glu	Asn	Ser	Thr	Tyr	Met	Gly	Phe	Thr	Phe
	80								85					90
Ala	Arg	Ile	Pro	Val	Gly	Arg	Tyr	Gly	Pro	Ser	Leu	Gln	Thr	Cys
	95								100					105
Gly	Lys	Asp	Thr	Pro	Asn	Ala	Gly	Asn	Pro	Met	Ala	Val	Arg	Leu
	110								115					120
Cys	Ser	Leu	Ser	Leu	Tyr	Gly	Glu	Ile	Glu	Leu	Gln	Lys	Val	Thr
	125								130					135
Ile	Gly	Asn	Cys	Asn	Glu	Asn	Leu	Glu	Thr	Leu	Glu	Lys	Gln	Val
	140								145					150
Lys	Asp	Val	Thr	Ala	Pro	Leu	Asn	Asn	Ile	Ser	Ser	Glu	Val	Gln
	155								160					165
Ile	Leu	Thr	Ser	Asp	Ala	Asn	Lys	Leu	Thr	Ala	Glu	Asn	Ile	Thr
	170								175					180
Ser	Ala	Thr	Arg	Val	Val	Gly	Gln	Ile	Phe	Asn	Thr	Ser	Arg	Asn
	185								190					195
Ala	Ser	Pro	Glu	Ala	Lys	Lys	Val	Ala	Ile	Val	Thr	Val	Ser	Gln
	200								205					210
Leu	Leu	Asp	Ala	Ser	Glu	Asp	Ala	Phe	Gln	Arg	Val	Ala	Ala	Thr
	215								220					225
Ala	Asn	Asp	Asp	Ala	Leu	Thr	Thr	Leu	Ile	Glu	Gln	Met	Glu	Thr
	230								235					240
Tyr	Ser	Leu	Ser	Leu	Gly	Asn	Gln	Ser	Val	Val	Glu	Pro	Asn	Ile
	245								250					255
Ala	Ile	Gln	Ser	Ala	Asn	Phe	Ser	Ser	Glu	Asn	Ala	Val	Gly	Pro
	260								265					270
Ser	Asn	Val	Arg	Phe	Ser	Val	Gln	Lys	Gly	Ala	Ser	Ser	Ser	Leu
	275								280					285
Val	Ser	Ser	Ser	Thr	Phe	Ile	His	Thr	Asn	Val	Asp	Gly	Leu	Asn
	290								295					300
Pro	Asp	Ala	Gln	Thr	Glu	Leu	Gln	Val	Leu	Leu	Asn	Met	Thr	Lys
	305								310					315
Asn	Tyr	Thr	Lys	Thr	Cys	Gly	Phe	Val	Val	Tyr	Gln	Asn	Asp	Lys
	320								325					330
Leu	Phe	Gln	Ser	Lys	Thr	Phe	Thr	Ala	Lys	Ser	Asp	Phe	Ser	Gln
	335								340					345
Lys	Ile	Ile	Ser	Ser	Lys	Thr	Asp	Glu	Asn	Glu	Gln	Asp	Gln	Ser
	350								355					360
Ala	Ser	Val	Asp	Met	Val	Phe	Ser	Pro	Lys	Tyr	Asn	Gln	Lys	Glu
	365								370					375
Phe	Gln	Leu	Tyr	Ser	Tyr	Ala	Cys	Val	Tyr	Trp	Asn	Leu	Ser	Ala
	380								385					390
Lys	Asp	Trp	Asp	Thr	Tyr	Gly	Cys	Gln	Lys	Asp	Lys	Gly	Thr	Asp
	395								400					405
Gly	Phe	Leu	Arg	Cys	Arg	Cys	Asn	His	Thr	Thr	Asn	Phe	Ala	Val
	410								415					420
Leu	Met	Thr	Phe	Lys	Lys	Asp	Tyr	Gln	Tyr	Pro	Lys	Ser	Leu	Asp
	425								430					435
Ile	Leu	Ser	Asn	Val	Gly	Cys	Ala	Leu	Ser	Val	Thr	Gly	Leu	Ala
	440								445					450
Leu	Thr	Val	Ile	Phe	Gln	Ile	Val	Thr	Arg	Lys	Val	Arg	Lys	Thr

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

455          460          465
Ser Val Thr Trp Val Leu Val Asn Leu Cys Ile Ser Met Leu Ile
470          475          480
Phe Asn Leu Leu Phe Val Phe Gly Ile Glu Asn Ser Asn Lys Asn
485          490          495
Leu Gln Thr Ser Asp Gly Asp Ile Asn Asn Ile Asp Phe Asp Asn
500          505          510
Asn Asp Ile Pro Arg Thr Asp Thr Ile Asn Ile Pro Asn Pro Met
515          520          525
Cys Thr Ala Ile Ala Ala Leu Leu His Tyr Phe Leu Leu Val Thr
530          535          540
Phe Thr Trp Asn Ala Leu Ser Ala Ala Gln Leu Tyr Tyr Leu Leu
545          550          555
Ile Arg Thr Met Lys Pro Leu Pro Arg His Phe Ile Leu Phe Ile
560          565          570
Ser Leu Ile Gly Trp Gly Val Pro Ala Ile Val Val Ala Ile Thr
575          580          585
Val Gly Val Ile Tyr Ser Gln Asn Gly Asn Asn Pro Gln Trp Glu
590          595          600
Leu Asp Tyr Arg Gln Glu Lys Ile Cys Trp Leu Ala Ile Pro Glu
605          610          615
Pro Asn Gly Val Ile Lys Ser Pro Leu Leu Trp Ser Phe Ile Val
620          625          630
Pro Val Thr Ile Ile Leu Ile Ser Asn Val Val Met Phe Ile Thr
635          640          645
Ile Ser Ile Lys Val Leu Trp Lys Asn Asn Gln Asn Leu Thr Ser
650          655          660
Thr Lys Lys Val Ser Ser Met Lys Lys Ile Val Ser Thr Leu Ser
665          670          675
Val Ala Val Val Phe Gly Ile Thr Trp Ile Leu Ala Tyr Leu Met
680          685          690
Leu Val Asn Asp Asp Ser Ile Arg Ile Val Phe Ser Tyr Ile Phe
695          700          705
Cys Leu Phe Asn Thr Thr Gln Gly Leu Gln Ile Phe Ile Leu Tyr
710          715          720
Thr Val Arg Thr Lys Val Phe Gln Ser Glu Ala Ser Lys Val Leu
725          730          735
Met Leu Leu Ser Ser Ile Gly Arg Arg Lys Ser Leu Pro Ser Val
740          745          750
Thr Arg Pro Arg Leu Arg Val Lys Met Tyr Asn Phe Leu Arg Ser
755          760          765
Leu Pro Thr Leu His Glu Arg Phe Arg Leu Leu Glu Thr Ser Pro
770          775          780
Ser Thr Glu Glu Ile Thr Leu Ser Glu Ser Asp Asn Ala Lys Glu
785          790          795
Ser Ile

```

<210> 5  
<211> 434  
<212> PRT  
<213> Homo sapiens

WO 02/26825

PCT/US01/30661

PI-0236 PCT

<220>

<221> misc\_feature

<223> Incyte ID No: 3837054CD1

<400> 5

```

Met Glu Asp Leu Phe Ser Pro Ser Ile Leu Pro Pro Ala Pro Asn
 1           5           10           15
Ile Ser Val Pro Ile Leu Leu Gly Trp Gly Leu Asn Leu Thr Leu
 20           25           30
Gly Gln Gly Ala Pro Ala Ser Gly Pro Pro Ser Arg Arg Val Arg
 35           40           45
Leu Val Phe Leu Gly Val Ile Leu Val Val Ala Val Ala Gly Asn
 50           55           60
Thr Thr Val Leu Cys Arg Leu Cys Gly Gly Gly Pro Trp Ala
 65           70           75
Gly Pro Lys Arg Arg Lys Met Asp Phe Leu Leu Val Gln Leu Ala
 80           85           90
Leu Ala Asp Leu Tyr Ala Cys Gly Gly Thr Ala Leu Ser Gln Leu
 95           100          105
Ala Trp Glu Leu Leu Gly Glu Pro Arg Ala Ala Thr Gly Asp Leu
 110          115          120
Ala Cys Arg Phe Leu Gln Leu Leu Gln Ala Ser Gly Arg Gly Ala
 125          130          135
Ser Ala His Leu Val Val Leu Ile Ala Leu Glu Arg Arg Arg Ala
 140          145          150
Val Arg Leu Pro His Gly Arg Pro Leu Pro Ala Arg Ala Leu Ala
 155          160          165
Ala Leu Gly Trp Leu Leu Ala Leu Leu Leu Ala Leu Pro Pro Ala
 170          175          180
Phe Val Val Arg Gly Asp Ser Pro Ser Pro Leu Pro Pro Pro Pro
 185          190          195
Pro Pro Thr Ser Leu Gln Pro Gly Ala Pro Pro Ala Ala Arg Ala
 200          205          210
Trp Pro Gly Glu Arg Arg Cys His Gly Ile Phe Ala Pro Leu Pro
 215          220          225
Arg Trp His Leu Gln Val Tyr Ala Phe Tyr Glu Ala Val Ala Gly
 230          235          240
Phe Val Ala Pro Val Thr Val Leu Gly Val Ala Cys Gly His Leu
 245          250          255
Leu Ser Val Trp Trp Arg His Arg Pro Gln Ala Pro Ala Ala Ala
 260          265          270
Ala Pro Trp Ser Ala Ser Pro Gly Arg Ala Pro Ala Pro Ser Ala
 275          280          285
Leu Pro Arg Ala Lys Val Gln Ser Leu Lys Met Ser Leu Leu Leu
 290          295          300
Ala Leu Leu Phe Val Gly Cys Glu Leu Pro Tyr Phe Ala Ala Arg
 305          310          315
Leu Ala Ala Ala Trp Ser Ser Gly Pro Ala Gly Asp Trp Glu Gly
 320          325          330
Glu Gly Leu Ser Ala Ala Leu Arg Val Val Ala Met Ala Asn Ser
 335          340          345

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

Ala Leu Asn Pro Phe Val Tyr Leu Phe Phe Gln Ala Gly Asp Cys  
 350 355 360  
 Arg Leu Arg Arg Gln Leu Arg Lys Arg Leu Gly Ser Leu Cys Cys  
 365 370 375  
 Ala Pro Gln Gly Gly Ala Glu Asp Glu Glu Gly Pro Arg Gly His  
 380 385 390  
 Gln Ala Leu Tyr Arg Gln Arg Trp Pro His Pro His Tyr His His  
 395 400 405  
 Ala Arg Arg Glu Pro Leu Asp Glu Gly Gly Leu Arg Pro Pro Pro  
 410 415 420  
 Pro Arg Pro Arg Pro Leu Pro Cys Ser Cys Glu Ser Ala Phe  
 425 430

&lt;210&gt; 6

&lt;211&gt; 339

&lt;212&gt; FRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 6157025CD1

&lt;400&gt; 6

Met Pro Gly His Asn Thr Ser Arg Asn Ser Ser Cys Asp Pro Ile  
 1 5 10 15  
 Val Thr Pro His Leu Ile Ser Leu Tyr Phe Ile Val Leu Ile Gly  
 20 25 30  
 Gly Leu Val Gly Val Ile Ser Ile Leu Phe Leu Leu Val Lys Met  
 35 40 45  
 Asn Thr Arg Ser Val Thr Thr Met Ala Val Ile Asn Leu Val Val  
 50 55 60  
 Val His Ser Val Phe Leu Leu Thr Val Pro Phe Arg Leu Thr Tyr  
 65 70 75  
 Leu Ile Lys Lys Thr Trp Met Phe Gly Leu Pro Phe Cys Lys Phe  
 80 85 90  
 Val Ser Ala Met Leu His Ile His Met Tyr Leu Thr Phe Leu Phe  
 95 100 105  
 Tyr Val Val Ile Leu Val Thr Arg Tyr Leu Ile Phe Phe Lys Cys  
 110 115 120  
 Lys Asp Lys Val Glu Phe Tyr Arg Lys Leu His Ala Val Ala Ala  
 125 130 135  
 Ser Ala Gly Met Trp Thr Leu Val Ile Val Ile Val Val Pro Leu  
 140 145 150  
 Val Val Ser Arg Tyr Gly Ile His Glu Glu Tyr Asn Glu Glu His  
 155 160 165  
 Cys Phe Lys Phe His Lys Glu Leu Ala Tyr Thr Tyr Val Lys Ile  
 170 175 180  
 Ile Asn Tyr Met Ile Val Ile Phe Val Ile Ala Val Ala Val Ile  
 185 190 195  
 Leu Leu Val Phe Gln Val Phe Ile Ile Met Leu Met Val Gln Lys  
 200 205 210  
 Leu Arg His Ser Leu Leu Ser His Gln Glu Phe Trp Ala Gln Leu

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

                215                220                225
Lys Asn Leu Phe Phe Ile Gly Val Ile Leu Val Cys Phe Leu Pro
                230                235                240
Tyr Gln Phe Phe Arg Ile Tyr Tyr Leu Asn Val Val Thr His Ser
                245                250                255
Asn Ala Cys Asn Ser Lys Val Ala Phe Tyr Asn Glu Ile Phe Leu
                260                265                270
Ser Val Thr Ala Ile Ser Cys Tyr Asp Leu Leu Leu Phe Val Phe
                275                280                285
Gly Gly Ser His Trp Phe Lys Gln Lys Ile Met Ala Tyr Gly Ile
                290                295                300
Val Phe Val Pro Leu Ala Thr Asn Tyr Ser Ile His Ile Cys Phe
                305                310                315
Leu Tyr Ile Gly Asn Lys Asn Gly Tyr Arg Gly Gly Lys Asn Gly
                320                325                330
Ile Ser Leu Leu Asp Gln Ser Met Pro
                335

```

&lt;210&gt; 7

&lt;211&gt; 549

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 55012817CD1

&lt;400&gt; 7

```

Met Ala Thr Pro Arg Gly Leu Gly Ala Leu Leu Leu Leu Leu
  1          5          10          15
Leu Pro Thr Ser Gly Gln Glu Lys Pro Thr Glu Gly Pro Arg Asn
  20          25          30
Thr Cys Leu Gly Ser Asn Asn Met Tyr Asp Ile Phe Asn Leu Asn
  35          40          45
Asp Lys Ala Leu Cys Phe Thr Lys Cys Arg Gln Ser Gly Ser Asp
  50          55          60
Ser Cys Asn Val Glu Asn Leu Gln Arg Tyr Trp Leu Asn Tyr Glu
  65          70          75
Ala His Leu Met Lys Glu Gly Leu Thr Gln Lys Val Asn Thr Pro
  80          85          90
Phe Leu Lys Ala Leu Val Gln Asn Leu Ser Thr Asn Thr Ala Glu
  95          100         105
Asp Phe Tyr Phe Ser Leu Glu Pro Ser Gln Val Pro Arg Gln Val
  110         115         120
Met Lys Asp Glu Asp Lys Pro Pro Asp Arg Val Arg Leu Pro Lys
  125         130         135
Ser Leu Phe Arg Ser Leu Pro Gly Asn Arg Ser Val Val Arg Leu
  140         145         150
Ala Val Thr Ile Leu Asp Ile Gly Pro Gly Thr Leu Phe Lys Gly
  155         160         165
Pro Arg Leu Gly Leu Gly Asp Gly Ser Gly Val Leu Asn Asn Arg
  170         175         180

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

Leu Val Gly Leu Ser Val Gly Gln Met His Val Thr Lys Leu Ala  
 185 190 195  
 Glu Pro Leu Glu Ile Val Phe Ser His Gln Arg Pro Pro Pro Asn  
 200 205 210  
 Met Thr Leu Thr Cys Val Phe Trp Asp Val Thr Lys Gly Thr Thr  
 215 220 225  
 Gly Asp Trp Ser Ser Glu Gly Cys Ser Thr Glu Val Arg Pro Glu  
 230 235 240  
 Gly Thr Val Cys Cys Cys Asp His Leu Thr Phe Phe Ala Leu Leu  
 245 250 255  
 Leu Arg Pro Thr Leu Asp Gln Ser Thr Val His Ile Leu Thr Arg  
 260 265 270  
 Ile Ser Gln Ala Gly Cys Gly Val Ser Met Ile Phe Leu Ala Phe  
 275 280 285  
 Thr Ile Ile Leu Tyr Ala Phe Leu Arg Leu Ser Arg Glu Arg Phe  
 290 295 300  
 Lys Ser Glu Asp Ala Pro Lys Ile His Val Ala Leu Gly Gly Ser  
 305 310 315  
 Leu Phe Leu Leu Asn Leu Ala Phe Leu Val Asn Val Gly Ser Gly  
 320 325 330  
 Ser Lys Gly Ser Asp Ala Ala Cys Trp Ala Arg Gly Ala Val Phe  
 335 340 345  
 His Tyr Phe Leu Leu Cys Ala Phe Thr Trp Met Gly Leu Glu Ala  
 350 355 360  
 Phe His Leu Tyr Leu Leu Ala Val Arg Val Phe Asn Thr Tyr Phe  
 365 370 375  
 Gly His Tyr Phe Leu Lys Leu Ser Leu Val Gly Trp Gly Leu Pro  
 380 385 390  
 Ala Leu Met Val Ile Gly Thr Gly Ser Ala Asn Ser Tyr Gly Leu  
 395 400 405  
 Tyr Thr Ile Arg Asp Arg Glu Asn Arg Thr Ser Leu Glu Leu Cys  
 410 415 420  
 Trp Phe Arg Glu Gly Thr Thr Met Tyr Ala Leu Tyr Ile Thr Val  
 425 430 435  
 His Gly Tyr Phe Leu Ile Thr Phe Leu Phe Gly Met Val Val Leu  
 440 445 450  
 Ala Leu Val Val Trp Lys Ile Phe Thr Leu Ser Arg Ala Thr Ala  
 455 460 465  
 Val Lys Glu Arg Gly Lys Asn Arg Lys Lys Val Leu Thr Leu Leu  
 470 475 480  
 Gly Leu Ser Ser Leu Val Gly Val Thr Trp Gly Leu Ala Ile Phe  
 485 490 495  
 Thr Pro Leu Gly Leu Ser Thr Val Tyr Ile Phe Ala Leu Phe Asn  
 500 505 510  
 Ser Leu Gln Gly Val Phe Ile Cys Cys Trp Phe Thr Ile Leu Tyr  
 515 520 525  
 Leu Pro Ser Gln Ser Thr Thr Val Ser Ser Ser Thr Ala Arg Leu  
 530 535 540  
 Asp Gln Ala His Ser Ala Ser Gln Glu  
 545

&lt;210&gt; 8

WO 02/26825

PCT/US01/30661

PI-0236 PCT

&lt;211&gt; 188

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7475061CD1

&lt;400&gt; 8

```

Met Pro Leu Val Gly Leu Gly Asp Tyr Val Pro Glu Pro Phe Gly
 1           5           10           15
Thr Ser Cys Thr Leu Asp Trp Trp Leu Ala Gln Ala Ser Val Gly
 20           25           30
Gly Gln Val Phe Ile Leu Asn Ile Leu Phe Cys Leu Leu Leu
 35           40           45
Pro Thr Ala Val Ile Val Phe Ser Tyr Val Lys Ile Ile Ala Lys
 50           55           60
Val Lys Ser Ser Ser Lys Glu Val Ala His Phe Asp Ser Arg Ile
 65           70           75
His Ser Ser His Val Leu Glu Met Lys Leu Thr Lys Val Ala Met
 80           85           90
Leu Ile Cys Ala Gly Phe Leu Ile Ala Trp Ile Pro Tyr Ala Val
 95           100          105
Val Ser Val Trp Ser Ala Phe Gly Arg Pro Asp Ser Ile Pro Ile
 110          115          120
Gln Leu Ser Val Val Pro Thr Leu Leu Ala Lys Ser Ala Ala Met
 125          130          135
Tyr Asn Pro Ile Ile Tyr Gln Val Ile Asp Tyr Lys Phe Ala Cys
 140          145          150
Cys Gln Thr Gly Gly Leu Lys Ala Thr Lys Lys Lys Ser Leu Glu
 155          160          165
Gly Phe Arg Leu His Thr Val Thr Thr Val Arg Lys Ser Ser Ala
 170          175          180
Val Leu Glu Ile His Glu Glu Val
 185

```

&lt;210&gt; 9

&lt;211&gt; 332

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7477374CD1

&lt;400&gt; 9

```

Met Val Asn Asn Phe Ser Gln Ala Glu Ala Val Glu Leu Cys Tyr
 1           5           10           15
Lys Asn Val Asn Glu Ser Cys Ile Lys Thr Pro Tyr Ser Pro Gly
 20           25           30
Pro Arg Ser Ile Leu Tyr Ala Val Leu Gly Phe Gly Ala Val Leu
 35           40           45

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

Ala Ala Phe Gly Asn Leu Leu Val Met Ile Ala Ile Leu His Phe  
 50 55 60  
 Lys Gln Leu His Thr Pro Thr Asn Phe Leu Ile Ala Ser Leu Ala  
 65 70 75  
 Cys Ala Asp Phe Leu Val Gly Val Thr Val Met Pro Phe Ser Thr  
 80 85 90  
 Val Arg Ser Val Glu Ser Cys Trp Tyr Phe Gly Asp Ser Tyr Cys  
 95 100 105  
 Lys Phe His Thr Cys Phe Asp Thr Ser Phe Cys Phe Ala Ser Leu  
 110 115 120  
 Phe His Leu Cys Cys Ile Ser Val Asp Arg Tyr Met Leu Gly Tyr  
 125 130 135  
 Ala Trp Phe Phe Pro Gly Phe Phe Ser Val Thr Tyr Ser Phe Ser  
 140 145 150  
 Ile Phe Asn Thr Gly Ala Asn Glu Glu Gly Ile Glu Glu Leu Val  
 155 160 165  
 Val Ala Leu Thr Cys Val Gly Gly Cys Gln Ala Pro Leu Asn Gln  
 170 175 180  
 Asn Trp Val Leu Leu Cys Phe Leu Leu Phe Phe Ile Pro Asn Val  
 185 190 195  
 Ala Met Val Phe Ile Tyr Ser Lys Ile Phe Leu Val Ala Lys His  
 200 205 210  
 Gln Ala Arg Lys Ile Glu Ser Thr Ala Ser Gln Ala Gln Ser Ser  
 215 220 225  
 Ser Glu Ser Tyr Lys Glu Arg Val Ala Lys Arg Glu Arg Lys Ala  
 230 235 240  
 Ala Lys Thr Leu Gly Ile Ala Met Ala Ala Phe Leu Val Ser Trp  
 245 250 255  
 Leu Pro Tyr Leu Val Asp Ala Val Ile Asp Ala Tyr Met Asn Phe  
 260 265 270  
 Ile Thr Pro Pro Tyr Val Tyr Glu Ile Leu Val Trp Cys Val Tyr  
 275 280 285  
 Tyr Asn Ser Ala Met Asn Pro Leu Ile Tyr Ala Phe Phe Tyr Gln  
 290 295 300  
 Trp Phe Gly Lys Ala Ile Lys Leu Ile Val Ser Gly Lys Val Leu  
 305 310 315  
 Arg Thr Asp Ser Ser Thr Thr Asn Leu Phe Ser Glu Glu Val Glu  
 320 325 330  
 Thr Asp

<210> 10  
 <211> 948  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7479890CD1

<400> 10  
 Met Pro Ser Pro Pro Gly Leu Arg Ala Leu Trp Leu Cys Ala Ala

WO 02/26825

PCT/US01/30661

PI-0236 PCT

1	5	10	15
Leu Cys Ala Ser Arg	Arg Ala Gly Gly Ala	Pro Gln Pro Gly Pro	
	20	25	30
Gly Pro Thr Ala Cys	Pro Ala Pro Cys His	Cys Gln Glu Asp Gly	
	35	40	45
Ile Met Leu Ser Ala	Asp Cys Ser Glu Leu	Gly Leu Ser Ala Val	
	50	55	60
Pro Gly Asp Leu Asp	Pro Leu Thr Ala Tyr	Leu Leu Gly Cys Pro	
	65	70	75
Pro Pro Leu Gln Lys	Ala Gln Ala Val Gly	Gln Leu Gly Glu Tyr	
	80	85	90
Glu Lys Gln Phe Gly	Pro Arg Gln Val Lys	Leu Phe Pro Gln Ser	
	95	100	105
Leu Ser Lys Pro Glu	Leu Ala Cys Glu Val	Pro Ala Asn Leu Pro	
	110	115	120
His Tyr Cys Arg Arg	Leu Asp Ala Asn Leu	Ile Ser Leu Val Pro	
	125	130	135
Glu Arg Ser Phe Glu	Gly Leu Ser Ser Leu	Arg His Leu Trp Leu	
	140	145	150
Asp Asp Asn Ala Leu	Thr Glu Ile Pro Val	Arg Ala Leu Asn Asn	
	155	160	165
Leu Pro Ala Leu Gln	Ala Met Ala Leu Ala	Leu Asn Arg Ile Ser	
	170	175	180
His Ile Pro Asp Tyr	Ala Phe Gln Asn Leu	Thr Ser Leu Val Val	
	185	190	195
Leu His Leu His Asn	Asn Arg Ile Gln His	Leu Gly Thr His Ser	
	200	205	210
Phe Glu Gly Leu His	Asn Leu Glu Thr Leu	Asp Leu Asn Tyr Asn	
	215	220	225
Lys Leu Gln Glu Phe	Pro Val Ala Ile Arg	Thr Leu Gly Arg Leu	
	230	235	240
Gln Glu Leu Gly Phe	His Asn Asn Asn Ile	Lys Ala Ile Pro Glu	
	245	250	255
Lys Ala Phe Met Gly	Asn Pro Leu Leu Gln	Thr Ile His Phe Tyr	
	260	265	270
Asp Asn Thr Ile Gln	Phe Val Gly Arg Ser	Ala Phe Gln Tyr Leu	
	275	280	285
Pro Lys Leu His Thr	Leu Ser Leu Asn Gly	Ala Met Asp Ile Gln	
	290	295	300
Glu Phe Pro Gly Leu	Lys Gly Thr Thr Ser	Leu Glu Ile Leu Thr	
	305	310	315
Leu Thr Arg Ala Gly	Ile Arg Leu Leu Pro	Ser Gly Met Cys Gln	
	320	325	330
Gln Leu Pro Arg Leu	Arg Val Leu Glu Leu	Ser His Asn Gln Ile	
	335	340	345
Glu Glu Leu Pro Ser	Leu His Arg Cys Gln	Lys Leu Glu Glu Ile	
	350	355	360
Gly Leu Gln His Asn	Arg Ile Trp Glu Ile	Gly Ala Asp Thr Phe	
	365	370	375
Ser Gln Leu Ser Ser	Leu Gln Ala Leu Asp	Leu Ser Trp Asn Ala	
	380	385	390
Ile Arg Ser Ile His	Pro Glu Ala Phe Ser	Thr Leu His Ser Leu	

WO 02/26825

PCT/US01/30661

PI-0236 PCT

	395		400		405									
Val	Lys	Leu	Asp	Leu	Thr	Asp	Asn	Gln	Leu	Thr	Thr	Leu	Pro	Leu
	410								415					420
Ala	Gly	Leu	Gly	Gly	Leu	Met	His	Leu	Lys	Leu	Lys	Gly	Asn	Leu
	425								430					435
Ala	Leu	Ser	Gln	Ala	Phe	Ser	Lys	Asp	Ser	Phe	Pro	Lys	Leu	Arg
	440								445					450
Ile	Leu	Glu	Val	Pro	Tyr	Ala	Tyr	Gln	Cys	Cys	Pro	Tyr	Gly	Met
	455								460					465
Cys	Ala	Ser	Phe	Phe	Lys	Ala	Ser	Gly	Gln	Trp	Glu	Ala	Glu	Asp
	470								475					480
Leu	His	Leu	Asp	Asp	Glu	Glu	Ser	Ser	Lys	Arg	Pro	Leu	Gly	Leu
	485								490					495
Leu	Ala	Arg	Gln	Ala	Glu	Asn	His	Tyr	Asp	Gln	Asp	Leu	Asp	Glu
	500								505					510
Leu	Gln	Leu	Glu	Met	Glu	Asp	Ser	Lys	Pro	His	Pro	Ser	Val	Gln
	515								520					525
Cys	Ser	Pro	Thr	Pro	Gly	Pro	Phe	Lys	Pro	Cys	Glu	Tyr	Leu	Phe
	530								535					540
Glu	Ser	Trp	Gly	Ile	Arg	Leu	Ala	Val	Trp	Ala	Ile	Val	Leu	Leu
	545								550					555
Ser	Val	Leu	Cys	Asn	Gly	Leu	Val	Leu	Leu	Thr	Val	Phe	Ala	Gly
	560								565					570
Gly	Pro	Ala	Pro	Leu	Pro	Pro	Val	Lys	Phe	Val	Val	Gly	Ala	Ile
	575								580					585
Ala	Gly	Ala	Asn	Thr	Leu	Thr	Gly	Ile	Ser	Cys	Gly	Leu	Leu	Ala
	590								595					600
Ser	Val	Asp	Ala	Leu	Thr	Phe	Gly	Gln	Phe	Ser	Glu	Tyr	Gly	Ala
	605								610					615
Arg	Trp	Glu	Thr	Gly	Leu	Gly	Cys	Arg	Ala	Thr	Gly	Phe	Leu	Ala
	620								625					630
Val	Leu	Gly	Ser	Glu	Ala	Ser	Val	Leu	Leu	Leu	Thr	Leu	Ala	Ala
	635								640					645
Val	Gln	Cys	Ser	Val	Ser	Val	Ser	Cys	Val	Arg	Ala	Tyr	Gly	Lys
	650								655					660
Ser	Pro	Ser	Leu	Gly	Ser	Val	Arg	Ala	Gly	Val	Leu	Gly	Cys	Leu
	665								670					675
Ala	Leu	Ala	Gly	Leu	Ala	Ala	Ala	Leu	Pro	Leu	Ala	Ser	Val	Gly
	680								685					690
Glu	Tyr	Gly	Ala	Ser	Pro	Leu	Cys	Leu	Pro	Tyr	Ala	Pro	Pro	Glu
	695								700					705
Gly	Gln	Pro	Ala	Ala	Leu	Gly	Phe	Thr	Val	Ala	Leu	Val	Met	Met
	710								715					720
Asn	Ser	Phe	Cys	Phe	Leu	Val	Val	Ala	Gly	Ala	Tyr	Ile	Lys	Leu
	725								730					735
Tyr	Cys	Asp	Leu	Pro	Arg	Gly	Asp	Phe	Glu	Ala	Val	Trp	Asp	Cys
	740								745					750
Ala	Met	Val	Arg	His	Val	Ala	Trp	Leu	Ile	Phe	Ala	Asp	Gly	Leu
	755								760					765
Leu	Tyr	Cys	Pro	Val	Ala	Phe	Leu	Ser	Phe	Ala	Ser	Met	Leu	Gly
	770								775					780
Leu	Phe	Pro	Val	Thr	Pro	Glu	Ala	Val	Lys	Ser	Val	Leu	Leu	Val

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

785          790          795
Val Leu Pro Leu Pro Ala Cys Leu Asn Pro Leu Leu Tyr Leu Leu
800          805          810
Phe Asn Pro His Phe Arg Asp Asp Leu Arg Arg Leu Arg Pro Arg
815          820          825
Ala Gly Asp Ser Gly Pro Leu Ala Tyr Ala Ala Ala Gly Glu Leu
830          835          840
Glu Lys Ser Ser Cys Asp Ser Thr Gln Ala Leu Val Ala Phe Ser
845          850          855
Asp Val Asp Leu Ile Leu Glu Ala Ser Glu Ala Gly Arg Pro Pro
860          865          870
Gly Leu Glu Thr Tyr Gly Phe Pro Ser Val Thr Leu Ile Ser Cys
875          880          885
Gln Gln Pro Gly Ala Pro Arg Leu Glu Gly Ser His Cys Val Glu
890          895          900
Pro Glu Gly Asn His Phe Gly Asn Pro Gln Pro Ser Met Asp Gly
905          910          915
Glu Leu Leu Leu Arg Ala Glu Gly Ser Thr Pro Ala Gly Gly Gly
920          925          930
Leu Ser Gly Gly Gly Gly Phe Gln Pro Ser Gly Leu Ala Phe Ala
935          940          945
Ser His Val

```

```

<210> 11
<211> 315
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7482825CD1

```

```

<400> 11
Met Glu Ile Val Ser Thr Gly Asn Glu Thr Ile Thr Glu Phe Val
1          5          10          15
Leu Leu Gly Phe Tyr Asp Ile Pro Glu Leu His Phe Leu Phe Phe
20          25          30
Ile Val Phe Thr Ala Val Tyr Val Phe Ile Ile Ile Gly Asn Met
35          40          45
Leu Ile Ile Val Ala Val Val Ser Ser Gln Arg Leu His Lys Pro
50          55          60
Met Tyr Ile Phe Leu Ala Asn Leu Ser Phe Leu Asp Ile Leu Tyr
65          70          75
Thr Ser Ala Val Met Pro Lys Met Leu Glu Gly Phe Leu Gln Glu
80          85          90
Ala Thr Ile Ser Val Ala Gly Cys Leu Leu Gln Phe Phe Ile Phe
95          100          105
Gly Ser Leu Ala Thr Ala Glu Cys Leu Leu Leu Ala Val Met Ala
110          115          120
Tyr Asp Arg Tyr Leu Ala Ile Cys Tyr Pro Leu His Tyr Pro Leu
125          130          135

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

Leu Met Gly Pro Arg Arg Tyr Met Gly Leu Val Val Thr Thr Trp  
 140 145 150  
 Leu Ser Gly Phe Val Val Asp Gly Leu Val Val Ala Leu Val Ala  
 155 160 165  
 Gln Leu Arg Phe Cys Gly Pro Asn His Ile Asp Gln Phe Tyr Cys  
 170 175 180  
 Asp Phe Met Leu Phe Val Gly Leu Ala Cys Ser Asp Pro Arg Val  
 185 190 195  
 Ala Gln Val Thr Thr Leu Ile Leu Ser Val Phe Cys Leu Thr Ile  
 200 205 210  
 Pro Phe Gly Leu Ile Leu Thr Ser Tyr Ala Arg Ile Val Val Ala  
 215 220 225  
 Val Leu Arg Val Pro Ala Gly Ala Ser Arg Arg Arg Ala Phe Ser  
 230 235 240  
 Thr Cys Ser Ser His Leu Ala Val Val Thr Thr Phe Tyr Gly Thr  
 245 250 255  
 Leu Met Ile Phe Tyr Val Ala Pro Ser Ala Val His Ser Gln Leu  
 260 265 270  
 Leu Ser Lys Val Phe Ser Leu Leu Tyr Thr Val Val Thr Pro Leu  
 275 280 285  
 Phe Asn Pro Val Ile Tyr Thr Met Arg Asn Lys Glu Val His Gln  
 290 295 300  
 Ala Leu Arg Lys Ile Leu Cys Ile Lys Gln Thr Glu Thr Leu Asp  
 305 310 315

&lt;210&gt; 12

&lt;211&gt; 312

&lt;212&gt; PPT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7483087CD1

&lt;400&gt; 12

Met Lys Ala Gly Asn Phe Ser Asp Thr Pro Glu Phe Phe Leu Leu  
 1 5 10 15  
 Gly Leu Ser Gly Asp Pro Glu Leu Gln Pro Ile Leu Phe Met Leu  
 20 25 30  
 Phe Leu Ser Met Tyr Leu Ala Thr Met Leu Gly Asn Leu Leu Ile  
 35 40 45  
 Ile Leu Ala Val Asn Ser Asp Ser His Leu His Thr Pro Met Tyr  
 50 55 60  
 Phe Leu Leu Ser Ile Leu Ser Leu Val Asp Ile Cys Phe Thr Ser  
 65 70 75  
 Thr Thr Met Pro Lys Met Leu Val Asn Ile Gln Ala Gln Ala Gln  
 80 85 90  
 Ser Ile Asn Tyr Thr Gly Cys Leu Thr Gln Ile Cys Phe Val Leu  
 95 100 105  
 Val Phe Val Gly Leu Glu Asn Gly Ile Leu Val Met Met Ala Tyr  
 110 115 120

WO 02/26825

PCT/US01/30661

PI-0236 PCT

Asp Arg Phe Val Ala Ile Cys His Pro Leu Arg Tyr Asn Val Ile  
 125 130 135  
 Met Asn Pro Lys Leu Cys Gly Leu Leu Leu Leu Ser Phe Ile  
 140 145 150  
 Val Ser Val Leu Asp Ala Leu Leu His Thr Leu Met Val Leu Gln  
 155 160 165  
 Leu Thr Phe Cys Ile Asp Leu Glu Ile Pro His Phe Phe Cys Glu  
 170 175 180  
 Leu Ala His Ile Leu Lys Leu Ala Cys Ser Asp Val Leu Ile Asn  
 185 190 195  
 Asn Ile Leu Val Tyr Leu Val Thr Ser Leu Leu Gly Val Val Pro  
 200 205 210  
 Leu Ser Gly Ile Ile Phe Ser Tyr Thr Arg Ile Val Ser Ser Val  
 215 220 225  
 Met Lys Ile Pro Ser Ala Gly Gly Lys Tyr Lys Ala Phe Ser Ile  
 230 235 240  
 Cys Gly Ser His Leu Ile Val Val Ser Leu Phe Tyr Gly Thr Gly  
 245 250 255  
 Phe Gly Val Tyr Leu Ser Ser Gly Ala Thr His Ser Ser Arg Lys  
 260 265 270  
 Gly Ala Ile Ala Ser Val Met Tyr Thr Val Val Thr Pro Met Leu  
 275 280 285  
 Asn Pro Leu Ile Tyr Ser Leu Arg Asn Lys Asp Met Leu Lys Ala  
 290 295 300  
 Leu Arg Lys Leu Ile Ser Arg Ile Pro Ser Phe His  
 305 310

&lt;210&gt; 13

&lt;211&gt; 309

&lt;212&gt; FRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7483134CD1

&lt;400&gt; 13

Met Gly Ala Lys Asn Asn Val Thr Glu Phe Val Leu Phe Gly Leu  
 1 5 10 15  
 Phe Glu Ser Arg Glu Met Gln His Thr Cys Phe Val Val Phe Phe  
 20 25 30  
 Leu Phe His Val Leu Thr Val Leu Gly Asn Leu Leu Val Ile Ile  
 35 40 45  
 Thr Ile Asn Ala Arg Lys Thr Leu Lys Ser Pro Met Tyr Phe Phe  
 50 55 60  
 Leu Ser Gln Leu Ser Phe Ala Asp Ile Cys Tyr Pro Ser Thr Thr  
 65 70 75  
 Ile Pro Lys Met Ile Ala Asp Thr Phe Val Glu His Lys Ile Ile  
 80 85 90  
 Ser Phe Asn Gly Cys Met Thr Gln Leu Phe Ser Ala His Phe Phe  
 95 100 105  
 Gly Gly Thr Glu Ile Phe Leu Leu Thr Ala Met Ala Tyr Asp Arg

WO 02/26825

PCT/US01/30661

FI-0236 PCT

```

110          115          120
Tyr Val Ala Ile Cys Arg Pro Leu His Tyr Thr Ala Ile Met Asp
125          130          135
Cys Arg Lys Cys Gly Leu Leu Ala Gly Ala Ser Trp Leu Ala Gly
140          145          150
Phe Leu His Ser Ile Leu Gln Thr Leu Leu Thr Val Gln Leu Pro
155          160          165
Phe Cys Gly Pro Asn Glu Ile Asp Asn Phe Phe Cys Asp Val His
170          175          180
Pro Leu Leu Lys Leu Ala Cys Ala Asp Thr Tyr Met Val Gly Leu
185          190          195
Ile Val Val Ala Asn Ser Gly Met Ile Ser Leu Ala Ser Phe Phe
200          205          210
Ile Leu Ile Ile Ser Tyr Val Ile Ile Leu Leu Asn Leu Arg Ser
215          220          225
Gln Ser Ser Glu Asp Arg Arg Lys Ala Val Ser Thr Cys Gly Ser
230          235          240
His Val Ile Thr Val Leu Leu Val Leu Met Pro Pro Met Phe Met
245          250          255
Tyr Ile Arg Pro Ser Thr Thr Leu Ala Ala Asp Lys Leu Ile Ile
260          265          270
Leu Phe Asn Ile Val Met Pro Pro Leu Leu Asn Pro Leu Ile Tyr
275          280          285
Thr Leu Arg Asn Asn Asp Val Lys Asn Ala Met Arg Lys Leu Phe
290          295          300
Arg Val Lys Arg Ser Leu Gly Glu Lys
305

```

<210> 14  
 <211> 309  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7478550CD1

```

<400> 14
Met Met Thr Asn Arg Asn Gln Val Val Leu Gly Arg Met Arg His
1          5          10          15
Gln Cys Leu Pro Gln Thr Glu Arg Ala His Thr Lys His Asp Leu
20          25          30
Ser Leu Gln Ala Gln Leu Gln Gln Lys Val Phe Met Glu Lys Trp
35          40          45
Asn His Thr Ser Asn Asp Phe Ile Leu Leu Gly Leu Leu Pro Pro
50          55          60
Asn Gln Thr Gly Ile Phe Leu Leu Cys Leu Ile Ile Leu Ile Phe
65          70          75
Phe Leu Ala Ser Val Gly Asn Ser Ala Met Ile His Leu Ile His
80          85          90
Val Asp Pro Arg Leu His Thr Pro Met Tyr Phe Leu Leu Ser Gln
95          100         105

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

Leu Ser Leu Met Asp Leu Met Tyr Ile Ser Thr Thr Val Pro Lys  
 110 115 120  
 Met Ala Tyr Asn Phe Leu Ser Gly Gln Lys Gly Ile Ser Phe Leu  
 125 130 135  
 Gly Cys Gly Val Gln Ser Phe Phe Phe Leu Thr Met Ala Cys Ser  
 140 145 150  
 Glu Gly Leu Leu Leu Thr Ser Met Ala Tyr Asp Arg Tyr Leu Ala  
 155 160 165  
 Ile Cys His Ser Leu Tyr Tyr Pro Ile Arg Met Ser Lys Met Met  
 170 175 180  
 Cys Val Lys Met Ile Gly Gly Ser Trp Thr Leu Gly Ser Ile Asn  
 185 190 195  
 Ser Leu Ala His Thr Val Phe Ala Leu His Ile Pro Tyr Cys Arg  
 200 205 210  
 Ser Arg Ala Ile Asp His Phe Phe Cys Asp Val Pro Ala Met Leu  
 215 220 225  
 Leu Leu Ala Cys Thr Asp Thr Trp Val Tyr Glu Tyr Met Val Phe  
 230 235 240  
 Val Ser Thr Ser Leu Phe Leu Leu Phe Pro Phe Ile Gly Ile Thr  
 245 250 255  
 Ser Ser Cys Gly Arg Val Leu Phe Ala Val Tyr His Met His Ser  
 260 265 270  
 Lys Glu Gly Arg Lys Lys Ala Phe Thr Thr Ile Ser Thr His Leu  
 275 280 285  
 Thr Val Val Ile Phe Tyr Tyr Ala Pro Phe Val Tyr Thr Tyr Leu  
 290 295 300  
 Arg Pro Thr Glu Ser Pro Leu Thr Ser  
 305

&lt;210&gt; 15

&lt;211&gt; 315

&lt;212&gt; FRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7483142CD1

&lt;400&gt; 15

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

WO 02/26825

PCT/US01/30661

PI-0236 PCT

95 100 105  
 Phe Phe Val Gly Met Gly Leu Ser Glu Cys Leu Leu Thr Ala  
 110 115 120  
 Met Ala Tyr Asp Arg Tyr Ala Ala Ile Ser Ser Pro Leu Leu Tyr  
 125 130 135  
 Pro Thr Ile Met Thr Gln Gly Leu Cys Thr Arg Met Val Val Gly  
 140 145 150  
 Ala Tyr Val Gly Gly Phe Leu Ser Ser Leu Ile Gln Ala Ser Ser  
 155 160 165  
 Ile Phe Arg Leu His Phe Cys Gly Pro Asn Ile Ile Asn His Phe  
 170 175 180  
 Phe Cys Asp Leu Pro Pro Val Leu Ala Leu Ser Cys Ser Asp Thr  
 185 190 195  
 Phe Leu Ser Gln Val Val Asn Phe Leu Val Val Thr Val Gly  
 200 205 210  
 Gly Thr Ser Phe Leu Gln Leu Leu Ile Ser Tyr Gly Tyr Ile Val  
 215 220 225  
 Ser Ala Val Leu Lys Ile Pro Ser Ala Glu Gly Arg Trp Lys Ala  
 230 235 240  
 Cys Asn Thr Cys Ala Ser His Leu Met Val Val Thr Leu Leu Phe  
 245 250 255  
 Gly Thr Ala Leu Phe Val Tyr Leu Arg Pro Ser Ser Ser Tyr Leu  
 260 265 270  
 Leu Gly Arg Asp Lys Val Val Ser Val Phe Tyr Ser Leu Val Ile  
 275 280 285  
 Pro Met Leu Asn Pro Leu Ile Tyr Ser Leu Arg Asn Lys Glu Ile  
 290 295 300  
 Lys Asp Ala Leu Trp Lys Val Leu Glu Arg Lys Lys Val Phe Ser  
 305 310 315

&lt;210&gt; 16

&lt;211&gt; 307

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7483151CD1

&lt;400&gt; 16

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

WO 02/26825

PCT/US01/30661

PI-0236 PCT

	80		85		90									
Met	Leu	Ser	Met	Ala	Lys	Cys	Ala	Thr	Gln	Phe	Tyr	Phe	Phe	Gly
	95								100					105
Ser	Asn	Ala	Thr	Thr	Glu	Cys	Phe	Leu	Leu	Val	Val	Met	Ala	Tyr
	110								115					120
Asp	Arg	Tyr	Val	Ala	Ile	Cys	Asn	Pro	Leu	Leu	Tyr	Pro	Val	Val
	125								130					135
Met	Ser	Asn	Ser	Leu	Cys	Thr	Gln	Phe	Ile	Gly	Ile	Ser	Tyr	Phe
	140								145					150
Ile	Gly	Phe	Leu	His	Ser	Ala	Ile	His	Val	Gly	Leu	Leu	Phe	Arg
	155								160					165
Leu	Thr	Phe	Cys	Arg	Ser	Asn	Ile	Ile	His	Tyr	Phe	Tyr	Cys	Glu
	170								175					180
Ile	Leu	Gln	Leu	Phe	Lys	Ile	Ser	Cys	Thr	Asn	Pro	Thr	Val	Asn
	185								190					195
Ile	Leu	Leu	Ile	Phe	Ile	Phe	Ser	Ala	Phe	Ile	Gln	Val	Phe	Thr
	200								205					210
Phe	Met	Thr	Leu	Ile	Val	Ser	Tyr	Ser	Tyr	Ile	Leu	Ser	Ala	Ile
	215								220					225
Leu	Lys	Lys	Lys	Ser	Glu	Lys	Gly	Arg	Ser	Lys	Ala	Phe	Ser	Thr
	230								235					240
Cys	Ser	Ala	His	Leu	Leu	Ser	Val	Ser	Leu	Phe	Tyr	Gly	Thr	Leu
	245								250					255
Phe	Phe	Met	Tyr	Val	Ser	Ser	Arg	Ser	Gly	Ser	Ala	Ala	Asp	Gln
	260								265					270
Ala	Lys	Met	Tyr	Ser	Leu	Phe	Tyr	Thr	Ile	Ile	Ile	Pro	Leu	Leu
	275								280					285
Asn	Pro	Phe	Ile	Tyr	Ser	Leu	Arg	Asn	Lys	Glu	Val	Ile	Asp	Ala
	290								295					300
Leu	Arg	Arg	Ile	Met	Lys	Lys								305

<210> 17

<211> 2422

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2536292CB1

<400> 17

```

gacaagggtc tcactttggt gccacgggtg gtctcgaact cccaggtcgg agcggctctc 60
ctgttttggc ctcccaaagc attgggatta caggcatgca ccacaattcc tgaactacag 120
ttttctctct atgggcctttg tacatttctt gccctcctgt atttttaagg aaatgtaaga 180
gtgtgaatca ggtactgatg gttttgactt cattatgatc agcatctttg gcttccatca 240
gagtagacac tgaacgtttt catccaaccg gatgatgatg gcattagaac tacataagag 300
atgaaatcct ttctccctgg gacatgtatc cttttgtggt ctgcttttaa tttgatgttt 360
ttcagtcctt tcaggctgaa atataatatt tgtatcctc tgagagcctg taacaccatg 420
ctttccagca atacaattat gaaaattttt ttctctctc atattgacat tgggatttgg 480
aggaacttac tgctactcct gatgcctatt tacacctttt tgatagtccc ccagcagaag 540
aagcccatgg gcttgctttt ccttcacttg totgtgtcta atacgatgac acttctccgc 600

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

aaagtatttc cattggcagt aaaatcttcc aacctataaa atctttttaa ttatactgga 660
tgtaggggaat ttgaattttt atabagagta tcttggggac ttccccatg tacpacatag 720
ctctcaagga tgggtgaggg cctccgtggg agccccagca aatcccaggt catcaacagt 780
ctatctcaca tcaagcttgt gccatttgtt gacactatca aatatggcag tgcaccaag 840
aatctctcca taaaaatgtg ttagcttaca cccacatgg gcaacacaat tgcgtctct 900
catacagagtg ttatcacatt ccaggactta atttttctgg ttctcatgag ttgagccagt 960
ggctacttgg tgattttcct tcacagacac cagaaaaata atcccatct tcacaacaac 1020
agctgttctt ctatagcctc ccattgagcc ggaaccatca cgaactgcct gctgcttatg 1080
atltgcttgc ttgtatttaa tggtagcaac tctgtccaca gcatttacct aagtaaggtg 1140
aagaaaaagg atcagttgtg gaccatctca gatttgattt cctcatgta cccattttt 1200
tggctcattt ttgctcattg gttagaagag tcttctctg aattcaaat ctatgaagta 1260
gagaagtgct tttatcccca tagatagtgc aaagtaact ttcaccaata taacttttt 1320
ccaaaaaac atctataaa gagtgggtat tctgaagaaa tgatgggatt taggtctgaa 1380
aaatgagata tttctcagtt cactgcatga aatataatct aatgacctt accttcagta 1440
aagacaatat tgcataagctt agtatttaat gttgtatag atagatata tcatttaata 1500
gacaagtaga tagataatct ttcattgtag atgtttataat ggactcttg cactagtcac 1560
tcaatatcca tatttttctt actttaattt ggttccagca tgaaaaatca tttcaaaag 1620
agatcagggg ctcccaggaag cagagcgtaa ctcccactt ctaactgtg ggttccacat 1680
atgacatcc taccagaata tcagtatagc aaaggaggaa aagagtaac ttcgcttgg 1740
agaaacctaa caaacactgt ctacgcccag tgatcaaggc cactgtgagc agtgatcagt 1800
caagtccgaaa gtgcataatca tatgacagga tgagaatggc attttacctc gacatctcc 1860
ctcccaaaa taataatcct actacatcag acaaacctca gctgagaaac agtttccaaa 1920
aacacctgac tagtaccctc taaaacctc aaggtcatca agacaatgt agtctcgaga 1980
aatgtacaca gccgggaaga acctaaggac ttggttctct ttggtatctt gatcggaa 2040
cctggagcag aaaaggacat taagtaaaaa ccaaggaaat agaatcaag tataactta 2100
ataatatttc aatattaatt catatattgt gacaatgta acatgtaat gtaagacatt 2160
aaaatggggg aaattgggta tgatgtggag gagaactctg tttattctat tgcactct 2220
totgtaaatg taacctatt tgaacatta aaaatttat ttttcaaaa caaaaaaaa 2280
aaaacaaaac aacaacaaca caaaaacaca gaggggcygc gcgacaaaa atatcaaac 2340
gcccaccgcy gggggggcgc cccaccata aagagataaa cacacacgga gytaaaaaac 2400
gggggaaagc ggtccctctc gc 2422

```

&lt;210&gt; 18

&lt;211&gt; 1912

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 747708CB1

&lt;400&gt; 18

```

aggaaaacag aaagtcagct gtgtggaaga caatgactca tatacttttg ctgtaactact 60
tgggttttct ttgcccaca gagtctctga ggacattgta tcaggtgtat gcgatgggt 120
ctgtacagac tactcccagt gtactcaacc ttgccctcca gacactcagg gaatatggg 180
gttttcatgc aggcaaaaga catggcacaa gatcactgac acctgcggga ctcttaatgc 240
cctcaacatc tttgaggaag attcagcttt ggttcagcca ttgtaagaca atataaaaat 300
aagtgtatat actggaagat ctgagccat aacagatag ttgtacaaa agtgcaccac 360
agatctgtct tgtgaaata gaaacattca gcagctctcc tggataccag gaaacattgc 420
cgtaattgtg cagctcttac acaacatctc aacagcaata tggccaggtg ttgatgagc 480
aaagatgcag agttacagca ccatagccaa cccattctt aacagcaaaa gcatctccaa 540
ctggactttc attcctgaca gaaacagcag ctatatctgt ctacattcag tcaactcct 600

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

tgcagaagg ctatccatag ataacatccc tgttgacata tcagatgtct tealtcatac 660
tatgggcaacc accatbatctg gagataacat tggaaaaaat ttcacttttt ctatgagaat 720
taatgacacc agcaaatgag tcactgggag agtgttgatc agcagagatg aacttcggaa 780
ggtgccttcc cctctccagg tcatcagcat tgcatttcca actattgggg ctattttgga 840
agccagctct ttggaaaaat ttactgtaaa tgggcttgtc ctgtctgcca tttgcccaca 900
ggaacttaaa agaactcac tgatttttga aaagatcagc aagtcagagg agaggaggac 960
acagttgttt ggcctggcac ctgttgagaa cagatgggac cagcaggcct gcaaaatgat 1020
tcaagaaaac tcccagcaag ctgtttgcaa atgtaggcca agcaaatgtt ttacctcttt 1080
ctcaattctt atgtcacctc acatctttaga gactctgatt ctgacttaca tccatctgt 1140
aggcctgggc atttctattt gcagcctgat cctttgcttg tccattgagg tccctgtctg 1200
gagcccaatg acaaaagcac agatcaccta tttacgcoat gtgtgcatg tcaacatgc 1260
agccactttg ctgatggcag atgtgtggtt cattgtggct tcocttctta gtggcccact 1320
aacaacccac aagggatgtg tggcagccac attttttggc catttctttt accttctgt 1380
atttttctgg atgcttgcca aggcactcct tatcctctat ggaatcatga ttgttttcca 1440
taccttgccc aagtcagtc tgggtggcac tctgttttca gtgggctatg gatgcccctt 1500
ggccattgct gccatcactg ttgtggccac tgaacctggc aaaggctatc tacgacctga 1560
gatctgctgg ctcaactggg acatgaccaa agccctcctg gcctctgta tcccagcttt 1620
ggccactgct gtatgaacc tcatcacagt cacactgggt atgtcaaga cccagcagc 1680
tgccattggc aattccatgt tccaggaagt gagagccatc gtgagaatca gcaagaacat 1740
cgccatctcc acaccacttc tgggactgac ctgggggattt gggatagcca ctgtcatcga 1800
tgacagatcc ctggcctccc acattatctt ctccctgctc aatgcatcc aggtaagtcc 1860
agatgctctt gaccaagtgc aaagtgagag aattcatgaa gatgtctctg ga 1912

```

<210> 19

<211> 1326

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7474823CB1

<400> 19

```

atgyttctag gtaaaaaatgt ctccatgtct ggcccaggc cagcatcatg gcagtccac 60
cctcaagggc ttgagcttgt ctttggaaaa tggccctgca gatgtagta ttcagcagtc 120
ctggatgatc cctccatcag cagttctgga gctggggaca tcccagatca agattctggc 180
caatattggt tcoctgatgag ggcctgtctt ctggcttcca gacggctgcc gtctcctgc 240
gtcctcaaga ggcctttctc tgagtgcgca cagagagaga gaactaattt ggttctcatg 300
aaaaaatggg aattcctgga agtacctgat acatttgaag taactcaaca aagtgtgatc 360
tccattcctt tgtacatccc tcacacgctg ttcgaaatgg attttggaaa ggaatctctg 420
gtattttggc tcaactactga ctactctgta tgtacagcat ctgtatataa cattgtcctc 480
atcagctatg atcgatacct gtcagtctca aatgctgtgt ctatagaac tcaacatac 540
ggggtcttga agattgttac tctgatgggt gccgtttggy tcttgccctt cttagtgaat 600
gggccaatga tctagtttc agagtcttgg aaggatgaa gtagtgaatg tgaacctgga 660
ttttttctgg aatggtacat ccttgcacat acatcattct tggaaatctg gatcccagtc 720
atcttagtgc cttatttcaa catgaaatatt tatttggagc ttgtggaagc tgatcatctc 780
aggcttgggc atcccagggt atgggcccag ctggtgctca gactgccaca tggggttgag 840
ggacagcctt ggcgctgca gctggtccct cggatggggt acattgaag agggagcttg 900
ttgtgcaact ctgctggaga gatgtcaact cagctagaaa gtgtcaaat cctgtcaaca 960
ggcagtgaaa atgacacctt gccagttcca agcctagcct caagaagcct gtgcccactc 1020
gttctatctc ttggcagttt tcccagctgc cagagctgcc ttagtacca gatgtcaaca 1080
tgtgatcacag agccagagag gaagtcattt ttgtccatga tgcaaggaac ccagcatttt 1140

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

gacaaccag atggaatgtg gagctccat ggaagaaeg tgtcaatctgg aggcctgcat 1200
aaccactgca tctccagat gggcactgga agcctgggg cttcccaccc agaaggtcca 1260
cgtggtggac aagggaagt gacaaccaga gccacgactc aaaaagaggt ggctgttca 1320
ggctga 1326

```

<210> 20

<211> 3058

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 644692CB1

<400> 20

```

aggaaattga aagcagagta tgcacctttt attaggagat tcaaaactgca tcctactgga 60
ttagcctcaa aagtcctaaa atacaaagac atccatctga cagatcactg aggggaggac 120
ttgtttttct gttttagaat agtttccgat taaacttttt agctcaagaa gaaaagaagc 180
tagttatttc tcaccaggga gtggatttgt ggtttggett caccatggct tctgcccctg 240
cctggaacct taggtgctg gtggctgctg tgtgtggact actgactggc atcaattttg 300
gactgggcat ctgggggatt gtgacagga tccaagagg aaaaactact tctctatcaa 360
gcaccctcac agagtctctc agaatgggt gaacctggga aaatgcaga tgtatttcta 420
cagaagagtg gaaaggactg agatgtacaa ttgctaatt ttgtgaaat agtacctata 480
tgggttttac ttttgccaga atccagtggt gcagatattg accatccttg caaacatgtg 540
gcaaggatca tccaatgctg gccaatccaa tggcagtcg gttgtgcat ctctctat 600
atggagagat agaattacaa aaagtgcaga taggaattgt caatgaaat ctggaacccc 660
tggaaaagca ggtaaaggat gtcacagcac cacttaataa catttcttct gaagtcoga 720
ttttacatct tgatgccaat aaattactg ctgagacat cactagtgtc acggagatgg 780
ttggacagat attcaacct tccagaaatg cttcaactga ggaagaaa gttgccatag 840
taacagtgag tcaactctca gatgccagtg aagatgcttt tcaaaagatt gctgctactg 900
cfaatgatga tgcctttaca acgcttattg agcaaatgga gacttattcc ttgtctctg 960
gtaacatcat agtgggtgaa cctaacatag caatacagtc agcaaatctc tcttcagaaa 1020
atgcygtggg gcttcaaat gttcgtctct ctgtgcagaa agggagctagc agttctctag 1080
tttctagtcc aacatttata catacaaatg tggatggcct taaccagat gcacagactg 1140
agcttcaggt ctgtctaat atgacgaaaa attacaccaa gacatggggc tttgtagttt 1200
atcaaaatga caagcttttc caatcaaaaa cttttacagc taaatcggat tttagctaaa 1260
aaattatctc aagcaaaact gatgaaaaat agcaagatca gagtgtctct gttgacatgg 1320
tcttttagtc aaagtacaa caaaaagaat ttcaactcta tctctatgcc tgtgtctatt 1380
ggaattgttc agcgaaggc tgggacacat atggctgtca aaaagacaag ggcactgatg 1440
gattcctgct ctgcccgtgc aaccatacta ctaattttgc tgtattaatg accttcaaaa 1500
aggattatca atatcccaaa tcaactgaca tattatccaa cgttggatgt gcactgtctg 1560
ttactgtctc gctctcaca gttatatttc agattgtcac caggaaagtc agaaaaactc 1620
cagttaacct ggttttggtc aatctgtgca tatcaatggt gattttcaac ctctctttg 1680
tgtttggaat tgaaaactcc aataagaact tgcagacaag tgatggtgac atcaataata 1740
ttgacttga caataatgac ataccagga cagacacatc taacatccc atcccatgt 1800
gcactcgat tgcgcctta ctgcaactatt tctgttagt gacatttacc tggaaagcac 1860
tcagcgtgc acagctctat taccctctaa taaggaccat gaagctctct cctggcatt 1920
tcattctttt catctcatta attggatggg gagtcccagc tatagtatgt gctataacag 1980
tggagttat ttattctcag aatggaaata atccacagtg ggaattagac taccggcaag 2040
agaaaatctg ctggttggca attccagaac ccaatgggtt tataaaaagt ccgctgtgtg 2100
ggtcaatcat cgtacctgta accattatcc tcatcagcaa tgtgttatg tttattacaa 2160
tctcgatcaa agtgcgtgg aagaataacc agaactgac aagcacaanaa aaagtctcat 2220

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

ccatgaagaa gattgttagc acattatctg ttgcagttgt ttttggatt acctggatc 2280
tagcatacct gatgctagtt aatgatgata gcacaggat cgtcttcagc tacatattct 2340
gccttttcaa cactcacacag ggattgcaaa tttttatcct gtacactgtt agaacaaaag 2400
tctccacag tgaagcttcc aaagtgttga ttgtctatc gtctatgtgg agaaggaagt 2460
cattgccttc agtgacgctg ccgaggctgc gtgtaaaagt gtataatttc ctcaagtcac 2520
tgccaacctt acatgaacgc tttaggctac tggaaacctc tccagatctc gaggaaatca 2580
cactctctga aagtacaat gcaaaggaaa gcactctagc agtaaaactt acctgtttgt 2640
gtctttttaa tcacctggtt tgagttttat ctgtttctct cctttatttc ccagtctctc 2700
cagaaagtct tctcaatgt attttgctca ggattaagaa ttagataaaa cctgtttgtt 2760
attattatct ggcataatgg acttggtagt tttctattt ttcaatagat ttgtacttga 2820
ataaggtgaa gaatttcaca caacatacaa gactaccatt gttccttata tctgtaatac 2880
tttgtgacac accttgacaa aaatgtagaa cctatacaaa atctctttac aagttactat 2940
aaaggacaca aagagaaaac ttacacctcc agaacaaaat gactcctgat gaacagtgtg 3000
tgggatttg atgtatgta ttaactttg gacctctgaa ctttttaaaa aaaaaaaaa 3058

```

<210> 21

<211> 1993

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3837054CB1

<400> 21

```

ctaactttgg gaactcgtat agaccacgag tcgctcccg cygtgcctcy cctccacttt 60
ggtttccgc gtcctgcgcy ctctcttcgg tgcctcctct tccctccgga caaggatgga 120
ggatctcttt agccctcaa ttctgcgccc ggcgcccaac atttccgtgc ceatcttct 180
gggtgggggt ctcaacctga ccttggggca aggagccctc gcctctgggc cgcaccagcc 240
ccgctccgcy ctggtgttcc tgggggtcat cctggtgggt gcggtggcag gcaacaccac 300
agtctctgcy cgcctgtgcy ggcgagcggc gccctggggc gcccacaagc gtcgcaagat 360
ggacttctcy ctggtgcaag tggccctggc ggaacctgac gcctgcgggg gcaacgcgct 420
gtcacagcty gcctgggaac tgctgggca gcccgcgcy gccacggggc acctggcgty 480
ccgcttctcy cagctgctgc aggcctccg ggcgggagcc tggcccaacc tctgtgtctc 540
catgcctcct gagcgcggcy gcgcggtgcy tcttcgcaac ggcggccgcy tgcctgcgcy 600
tgccctcgc gccctgggct ggctgctgcy actgtgctg gcctgcctcc cggccttctc 660
ggctgcggcy gactccctcc cgcctgctgc gccgcgccc cgcaccaact cctgcagcc 720
aggcgcgccc ccggccgccc gcgcctgccc ggggagcgt cgcctgcagc gcatctctgc 780
gccctcgcgc cgcctgcaac tgcaggtcta cgccttctac gaggccctcy cggctctctc 840
cgcgcctggt acggtcctgg gcctgcttgy cgcaccacta ctctccgtct ggtggcgga 900
ccggccgca gcccgcggcy ctgcagcctc ctggtcggcy agcccagtc gaccctctgc 960
gccacgcgcy ctgcccgcgc caaaggtgca gagcctgaag atgagcctgc tctggcgct 1020
gctgtctgcy ggctgcgagc tgcctactt tgcccgcgcy ctgagcgcc cgtgtgctc 1080
cgggcccgcy ggaactggg agggagaggg cctgtcggcy gcctgcgcy tgggtggcat 1140
ggccaacagc gctctcaatc ccttctctca cctctctctc cagcggggcy actgcggct 1200
cggcgagcag ctgcggaagc ggctgggtc tctgtctgcy gcgcgagc gaggcggga 1260
ggagagagag ggcggcgcy gccaccagc gctctaccgc caacgtgcy cccacctca 1320
ttatcccat gctcggcgcy aaccgctgga cgaaggcgc ttgcgcccac cccctccgcy 1380
ccccagacc ctgccttctc cctgcgaagc tgcctctag gtgtgtgtg gtcagagac 1440
ggctactctg cgtcaaggcy caacctccag gaaactcag gcctgcagc gctctccag 1500
atcacaagcy gcagagagct ctgtgagaga gtgacctga agtctccc tctctccat 1560
ctctactcc ctctctctgt ttacattcc ctatgctct ccagttctc tctctccta 1620

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

cagttcctct catabctccc catttgaga cagtggagcca ctggaaagt gtaaaaaaa 1680
aaacagttat ttttgagtt ttcttccacg catttatagt gctctggata atgocattta 1740
tttttgctga ttaaccaact ttcagttatt gctgtgttat catctgtatt tacttatttt 1800
gaatcgtgct taaatcaaat gtaccttcag cacctgcaag tttgoccttt ctttccagga 1860
ggaaaatccc cacgttgctc tccttgggga gctcgagaaat tataccagtg ctgtccagaaa 1920
tgtaatcatg ctgtcatttc agagccacag agtatttata aaataaaaaa ctttcccacg 1980
gaaaaaaaaa aaa                                     1993

```

&lt;210&gt; 22

&lt;211&gt; 1499

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 6157025CE1

&lt;400&gt; 22

```

aaatgcatgc agagcatgga aatgcccoca gctgcctgc tgttgaaca gaatcctatt 60
tgggaaggcag acatgtggcc catctctgta gccatcactg agaaatctgg attttcaagg 120
gcctttctct ctgttgccca ggctggagtt tagcgactca atcatggctc actgactgca 180
goatcgacct ccggggctca agtactcctt tcactctcag cctctcagta gctgagacta 240
caggttttcc tgggatcagc tgcaactcct agcaaaagta tattggagaa tcaactgaga 300
agtaactgga gacatttcaa tcattcttag gtgtaaaaga agaccagatc ccaggaaaaa 360
attacggtaga ctcccaagt atgcctggcc acaataacct caggaattcc tcttgcgatc 420
ctatagtgac accccactta atcagcctct acttcatagt gcttattggc gggctgggtg 480
gtgtcatttc cattcttttc ctctcgtgta aaatgaacac ccggtcagtg accaccatgy 540
cggctattaa cttgggtggtg gtcccacagc tttttctgct gacagtgcca tttcgcttga 600
cctaccctcat caagaagact tggatgtttg ggctgcccct ctgcaaatat gtgagtgcca 660
tgctgcaaat ccacatgtac ctcaagctcc tattctatgt ggtgatcctg gtcaccagat 720
acctcatctt ctccaagtgc aaagacaaag tggaaattca cagaaaactg catgctgtgg 780
ctgccaagtgc tggcatgtgg acgctgggta ttgtcattgt ggtaccctcg gttgtctccc 840
ggataggaaat ccatgaggaa tacaactata tgatagtcat ttttgcata gccgttctg 900
cttaacacata tgtgaaaatc atcaactata tgatagtcat ttttgcata gccgttctg 960
tgattctggt ggtcttcacg gctctcatca ttatgttgat ggtgcagaag ctaccaccat 1020
ctttactatc ccaccaggag ttctgggctc agctgaaaaa cctatttttt atagggttca 1080
tcctgttttg ttctctccc taccagttct ttatgactca ttacttgaat gttgtgacgc 1140
attccaatgc ctgtaacagc aaggttgcat ttataacga aatctctctg agtgaacag 1200
caattagctg ctatgatttg ctctctcttg tcttgggggg aagccattgg ttaagcaaa 1260
agataabggc ttatggaatt gtgtttgtgc cgttagccac aaactacagt attcatatt 1320
gcttctctta tattgggaat aaaaatgggt atagggggag taagaatggt atttcattc 1380
ttgatcaaaag catgccttga tgtaaccaaa acaaaaaggc tataaatgca agagcctca 1440
ttgtagtcct attgggatcc tccatctcga gtgatggcgt acaagaccgg tgtgtcgc 1499

```

&lt;210&gt; 23

&lt;211&gt; 2455

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 55012817CE1

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

<400> 23
cgttgctgtc gagagctcct ggcgtgggca aggcctggcca aggatggcga cggccagggg 60
cctggggggcc ctgctcctgc tcctcctgct cccgacctca ggtcaggaaa agcccaccga 120
agggccaaga aacacctgoc tggggagcaa caacatgtac gacatcttca acttgaatga 180
caaggctttg tgcttcacca agtgcaggca gtcgggcagc gactcctgca atgtggaata 240
cttgcaagaga taactggctaa actacagggc ccactctgat aaggagggtt tgacgcagaa 300
ggtgaacacg cctttcctga aggtcttggc ccagaacctc agccaacaac ctgcagaaga 360
cttctatttc tctctggagc cctctcaggt tccgaggcag gtagtgaagg acgaggacaa 420
gccccttgac agagtggcgc ttcccaagag cctttttcga tcctctccag gcaacaggto 480
tgtggtcgcg ttggccgtca ceattctgga cattggtcca gggactctct tcaagggtcc 540
ccggtctggc ctgggagatg gcagcggcgt gttgaacaat cgcctggtgg gtttgagtgt 600
gggacaaatg catgtcacca agctggctga gctctggag atcgtcttct ctaccacgcy 660
accgccccct aacatgacct teacctgtgt attctgggat gtgactaaag ggaacctggy 720
agactgtgtc tctgagggtc gctccacgga ggtcagacct gaggggaccg tgtgctgcty 780
tgaccaactg acctttttcg cctctgctct gagaccacc ttggaccagt ccaagtgca 840
tatcctcaca cgcactctcc aggcgggctg tggggctctc atgatcttc tggcctcac 900
cattattctt tatgccttc tgaggctttc ccgggagagg tcaagtcag aagatgccc 960
aaagatccac gtggccctgg gtggcagcct gttctcctg aatctggcct tcttggcca 1020
tgtggggagt ggctcaaagg ggtctgatgc tgcctctgg gccccggggg ctgtcttcca 1080
ctacttctct ctctgtgctc tcacctggat gggccttgaa cctctccacc tctactctg 1140
cgctgtcagg gtcttcaaca cctactctgg gcactacttc ctgaagctga gcttgggtgg 1200
ctggggctcg ccgcccctga tggctatcgg cactgggagt gccaacagct acggcctcta 1260
caccatcctg gataggggaga accgaccttc tctggagcta tctggttcc gtraagggac 1320
aacatgtac gccctctata tcaccgtcca cggctacttc ctcatcact tctctcttg 1380
catggtggtc ctggccctgg tggctctgaa gatctcacc ctgtcccgt ctacagcgtt 1440
caaggagcgg gggaaagaacc ggaagaaggt gctcacccty ctggccctct cgaagcctgt 1500
gggtgtgaca tgggggttgg ccactctcac ccggttggc ctctccaccg tctacatct 1560
tgcaactttc aactccttgc aaggtgtctt catctgtgc tggttacca tctttactt 1620
ccccagtcag agcaccacag tctcctcttc taectgaaga ttggaccagg cccactccg 1680
atctcaagaa taggaaggca cggccctgca atatggactc agctctggct ctctgtgtga 1740
ccttgggagc ctccgtgctc ctctctgtac tccctcagtt tccctctctg tacaatgtg 1800
ctggggaggg agaggatggg accaggttgg accacgtgac atcagaggtc ccatccagat 1860
ccaactatag gtccaagagt ccaagtaagc aggtttgcaa ggctctaaag tccctatagt 1920
cctgagacct cctgccaaga aagagtgaac gtcacctcca tgcctgccc tcaattgcaa 1980
gccctcactc accttctggt ctcaagcaag gaggagatc tgttctggc atagccctgg 2040
aaggagcccc cagcctctcc cctcctctcc ctgtcactg gccctccaca actcccttc 2100
tggctgctg taaccttgag gggcattcag gaggccagc ttcctcagg cactgggggt 2160
ttgttttggg ggttgggagt tgatctccc acccagctg cccctggtct ctgccatcc 2220
aatcagagcc caccctcctg gaagagacct ccgtgttcc agtctgtgca gccctgcag 2280
tgtccagggc cactgcattt caaagaacca ctgagtggt gagtaacct gggcaaaccc 2340
ccccctctg actctgactg ccacgtgggt gggccgacct ctgacctgt gtcactgtag 2400
aggtagaaga caaacaatct ggggctcagc acacctgggg gtgctccac tcatt 2455

```

<210> 24

<211> 2056

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7475061CB1

WO 02/26825

PCT/US01/30661

PI-0236 PCT

<220>  
<221> unsure  
<222> 1863  
<223> a, t, c, g, or other

```

<400> 24
gatcacgagc caccaagccc ottccacaag ttcttaggaa gccagetaag actagcaaca 60
caaaagctgt tagcaggaag ggcgttttaa agacacaagc catgttttgt gcttctctac 120
ctcccagaag gattttccct ggatgatcca ctaacacacc ctttccgaga tgttttctca 180
agggcctcca ttgtcttcag catcctctca tctctttcct gtccctttc ceagctgttg 240
tcactctctc ctcccgaagc tcccagatgc ctgagaaaaa ccttccaatt gtgtctctcc 300
ggagaaaaggc ccaagctctg gaactttttt ccttttaact cactctccct caggctctcag 360
ttatctgccc aggtctccca gaaatcaatg gaataaatct ttggaggttg aatttgacat 420
atcgagggga ggtatctggg tgtaggcagc attaatcccg ggtaatgctt tcttcccgcc 480
tccaggggtt tggctgaaaa gaaagcagc ctacatctgc ctggcagcca tctgggccta 540
tgtctcttc tggaccacca tgccttgggt aggtctgggg gactacgtac ctgagccctt 600
cggaacctcg tgcacctggc actggtggct ggcccaggcc tctgtagggg gccaggtttt 660
catctgaac atcctctctc tctgctctct cctcccacgc gctgtgacg tgttctctca 720
cgtaaagatc attgccaagg ttaagtctct tcccagaaga gtatgctcatt tgcagactgc 780
gatccatagc agccatgtgc tggaaatgaa actgacaaag gtatgctatg tgettgtgc 840
tggattctct attgctggga tctcttatgc agtgggtgtc gtgtgtgtag cttttggaag 900
gccagactcc attcccatcc agctctctgt ggtgccaac ctaactgcaa aatctgagc 960
gatgtacaat cccatcattt accaatgat tgaattacaa ttgctctgtt gccaaactgg 1020
tggtttgaaa gcaaccaaga agaagtctct ggaaggcttc aggtgcaca ccgtaaccac 1080
agtcaggagc tcttctgctg tctgtgaaat tcatgaagag gtatgaagat ggtacagca 1140
tcaatfagga cactcgtatt caattatttg cotcttact gctgtaaca ttgattgtg 1200
gccacacttt tgtcttata ttatatattt atattttgta cagtttccct caagccaggg 1260
gttgcctggg aattccaaty gctaaatgga aactagatta caggatacta atttaaggaa 1320
tattatcagc atgagcgtcc ctggaaacttt cttttctcgt actttaaac ctattgtcat 1380
gccaaaacga attgaaatgt agatagtgag aaaaatgttg agaagtttta tattaggggg 1440
ttattlaatt ttattacatc attgtcatta ttgtttctta atgtcgggtt agggaggcag 1500
tttctggggc tcaattcaggt ggacaatcat gtatgttggg ccatggctct ttggtagta 1560
ctttggttct cacagccatt ttctatgctg gaggaggttt gtactctgat ggcagtgatt 1620
ccacaaggga gtccctctggg gaaatcagtt ttgctttaa tcaactaat gaaaactctc 1680
ctgtatctcc aaatttgaag acaacagtaa tttccgctga tgaaaaaaa aaaaaggggc 1740
ggccgctgac ttagtgagcc tctctgacc cggaaataaa tcccggacc gtaccctggg 1800
agaggcgtcc ttcccagac acattgggag gctcaaggtg gccccaagtt ctaatagtg 1860
cncataatc cgtttgtgtt gacacatcag gtctggatat tgaccggcta aggagcgaac 1920
aaataaaaca acaacaagaa aaataagaa aacgtactca cacggaacac aaaccacaa 1980
caaaaagaca acaaaaatca gaataaaaa aaacaacagc acatacaaca caaaccaaca 2040
aacaataaca aacaaa 2056

```

<210> 25  
<211> 999  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 7477374CB1

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

<400> 25
atggggaaca atttctocca agctgaggct gtggagctgt gttacaagaa cgtgaacgaa 60
tcctgcatta aaactcotta ctogccaggt cctcgatcta tcctctacgc cgtccttggg 120
tttggggctg tgcctggcagc gtttggaaac ttactggtea tgattgctat ccttcaettc 180
aaacaactgc acacacctac aaactttctg attgctgctc tggcctgtgc tgaactcttg 240
gtgggagtea ctgtgatgcc ctccagcaca gtgaggtctg tggagagctg ttggtacttt 300
ggggacagtt actgtaaat ccatacatgt ttgacacat ccttctgttt tgcttcttta 360
tttcattat gctgtatctc tgttgataga taatgtctgg gatatgcatg gttctttcct 420
ggtttctttt ctgtcacata cagctttctg atcttaaca cgggagccaa cgaagaagg 480
attgaggaat tagtagttgc tctaacctgt gtaggagctc gccaggctcc actgaatcaa 540
aacgggtcc tactttgttt tcttctattc ttataccca atgtcgccat ggtgtttata 600
tacagtaaga tatttttggg ggccaagcat caggctagga agatagaaga tacagccagc 660
caagctcagt cctcctcaga gagttaacaag gaaagagtag caaaaagaga gagaaggct 720
gccaaacct tgggaattgc tatggcagca ttcttctgtc cttggctacc atacctcgtt 780
gatgcaatga ttgatgctta tatgaatttt ataactctc cttatgttta tgagatttta 840
gtttgggtg tttattataa ttcagctatg aacctctga tttatgcttt cttttacc 900
tggtttggga aggcaataaa acttattgta agcggcaagg tcttaaggac tgattcgtca 960
acaactaatt tatttctga agaagtagag acagattaa 999

```

<210> 26

<211> 3429

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Inyte ID No: 7479890CB1

```

<400> 26
atgccagacc cgcggggct ccggggccta tggctttgcy ccgcgctgtg cgttcccgg 60
agggccggcg gcgccccca gcccgcccg gggcccacg cctgccggc ccctgcac 120
tgccaggagg acggcatcat gctgtctgcc gactgctctg agctggggct gtcggcgtt 180
ccgggggacc tggaccctct gacggcttac ctgttgggat gtcctctcc actccaaag 240
gcaacagctg tgggccaagt gggggagtat gagaagcagt ttggcctcg acaggttaag 300
ctgtttcccc agagcctaag caagcccga cttggcctgt aggtccctgc taatctgcc 360
cattactgca ggcgcctaga tgccaaactc atctccctgg tccgggagag gagctttgag 420
gggctgtcct cctcctcgca cctctggctg gacgacaatg cactcacgga gatcctgtc 480
agggccctca acaactctcc tgcctgcag gccatggccc tggccctca ccgcatcag 540
cacatccccg actacgctt ccagaatctc accagccttg tgggtctgca ttgcaaac 600
aacgcacacc agcatctggg gaccacagc ttcgaggggc tgcaaatct ggagacacta 660
gacctgaatt ataacaagct gcaggagttc cctgtggcca tccggacctt gggcagactg 720
caggaactgg ggttccataa caacaacatc aaggccatcc cagaaaagc cttcatggg 780
aacctctgct tacagcagat acacttttat gataacaaa tccagtttgt gggagatcg 840
gcattccagt acctgcctaa actccacaca ctatctctga atgggtccat ggacatccag 900
gagtttccag gtctcaaaag caaccacag ctggagatcc tgacctgac ccgcagggc 960
atccgctgc tccatcggg gatgtgccc caactgccc ggtccagat cctggaaact 1020
tctcaaatc aaattgggga gctggccagc ctgcaacagt gtcagaatt ggggaaatc 1080
ggcctccaac acaaccgat ctgggaaatt ggagctgaca ccttcagcca gctgagctc 1140
ctgcaagccc tgatcttag ctggaaagcc atccggtcca tccaccocga ggccttctc 1200
acctgcact cctggtcaa gctggaactg acagacaacc agctgaccac actgcccctg 1260
gctggacttg gggccttgat gcatctgag ctcaagggga acctgtctct ctcccggcc 1320
ttctccaagg acagtttccc aaaaactgag atcctggagg tgcttatgc ctaccagtc 1380

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

tgtcccatg ggatgtgtg cagcttctc aaggcctctg ggcagtggga ggctgaagac 1440
cttcaacctg atgatgagga gtcttcaaaa aggccctctg gcctccttgc cagacaagca 1500
gagaaccaact atgaccagga cctggatgag ctccagctgg agatggagga ctcaaaagca 1560
caccccagtg tccagtgtag ccctactcca gccccttcca agccctgtga gtacctcttt 1620
gaaactctgg gcatccgctt ggcctgtgtg gccatctgtt tgctctcctg gctctgcaat 1680
ggaactgtgc tgcctgacct gttcctctgg ggcctgtccc cctctccccg ggtcaagttt 1740
gtgttaggtg cgattgcagg cgccaacacc ttgactggca tttcctgtgg ccttctagcc 1800
tcagtgcgat cctgacctt ttgtcagttc tctgagtacg gagcccgctg ggagacgggg 1860
ctaggtctgc gggccactgg ctctcctgca gtaactgggt cggaggcctc ggtgtgctg 1920
ctcaactctg ccgcagtgca gtgcagctgc tccgtctcct gtgtccgggc ctatgggaag 1980
tccccctccc tggcagcgtt tcgagcaggg gtccctaggt gcctggcact ggcagggctg 2040
gcgcgcgcac tccccctggc ctcaagtggg gaatacgggg cctccccact ctgctgccc 2100
tacgcgccac ctgaggggta gccagcagcc ctgggcttca ccgtggccct ggtgatgatg 2160
aactcctctt gtttccctgt cgtggccggt gccacatcca aactgtactg tgacctgccc 2220
cggggcgact ttgagggcct gtgggaactg gccatggtga ggcacgtggc ctgctctcct 2280
ttcgcagcag ggtcctctca ctgtcccctg gccctcctca gctttgctcc catgtgggc 2340
ctcttccctg tcacgccgca ggcctcaag tctgtcctgc tgggtggtgt gccctgctc 2400
gcctgcctca acccaactgt gtaactgctc ttcaaccccc acttccggga tgaccttgg 2460
cggcttctgg ccgcgcagag ggaactcaggg cccctagcct atgtctgggc cggggagctg 2520
gagaagagct cctgtgatcc taaccagccc ctggtagcct tctctgatgt ggtatctcatt 2580
ctgaaagctt ctgaaactgg gcccgcctct gggctggaga cctatggctt ccctcagtg 2640
accctcctct cctgtcagca gccagggccc cccagctggg agggcagcca ttgtgtagag 2700
ccagagggga accactttgg gaacccccaa cctccatgag atggagaact gctgtgagg 2760
gcagagggat ctacgccagc aggtggaggc ttgtcagggg gtggggctt tcagcctct 2820
ggctggcctt ttgctcaaca cgtgtaaata tccctcccca tcttctctt cccctctctt 2880
cccttctctc tctccccctc ggtgaatgat ggtgtctctt aaaaacaata caacaaaac 2940
tcagcagtgat gatctatagc aggatggccc agtacctggc tccaactgatc acctctctcc 3000
tgtgaccctc accaaagggg gccctctggc ctggctttcc cttggccttc ctccagctca 3060
ccttgatact gggcctcttc ctgtctatgt ctgaaactgt ggaccagaga cctggacttt 3120
tgtctgctta agggaaatga gggaaagtaa gacagtgaag ggggtggagg ttgatcaggg 3180
cacagtggac agggagacct cacagagaaa ggcctggaag gtgatttccc gtgtgactca 3240
tggataggat acaaaaatgt ttccatgtac cattaactct gacatatgcc atgcataaag 3300
acttccattt aaaaatgct ttggaagaga ttacacatga tgtcttttcc ttgagatctc 3360
acagtgcctg ttagtgtaat aaagagataa gtccctacagt agtaaaatct attgaacttt 3420
gtttcaaaa 3429

```

<210> 27

<211> 948

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7482825CE1

<400> 27

```

atggaattg tctccacagg aaacgaaact attactgaat ttgtctctct tggettctat 60
gacatccctg aactgcattt cttgtttttt attgtattca ctgctgteta tgtcttctac 120
atcatagggg atagtctgat tattgttagca gtggttagct cccagaggct ccaaaaacc 180
atgtatatatt tcttggcgaa tctgtctctc ctggatattc tctacacctc cgcagtgatg 240
ccaaaatgc tggagggctt cctgcaagaa gcaactatct ctgtggctgg ttgcttctc 300
cagttcttta tottccgctc tctagccaca gctgaatgct tactgctggc tgtctatgca 360

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

tatgacogct aactggcaat ttgtaccoca ctccactacc cactcctgat ggggccaga 420
cggtaacatgg ggetgggtgt cacaacctgg ctctctggat ttgtggtaga tggactggtt 480
gtggccctgg tggcccagct gaggttctgt gyccccaacc acattgacca gttttactgt 540
gacattatgc ttttcgtggg cctggcctgc toggatccca gattggctca ggtgacaact 600
ctcaattctgt ctgtgtctgt cctcactatt ccttttggac tgattctgac atcttatgcc 660
agaattgtgg tggcagtgct gagagttcct gctggggcaa gcaggagaag ggtttctccc 720
acatgotcct cccacctagc tgtagtgacc acattctatg gaacgctcat gatcttttat 780
gttgcaacct ctgctgtcca ttcccagctc ctctccaagg tcttctccct gctctacact 840
gtggtcaacc ctctcttcaa tcctgtgabc tataccatga ggaacaagga ggtgcatcag 900
gcacttcgga agattctctg tatcaaacaa actgaaacac ttgattga 948

```

<210> 28

<211> 939

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7483087CB1

<400> 28

```

atgaaagcag gaaactctc agacactoca gaattcttcc tcttgggatt gtcaggggat 60
cgggagctgc agcccactct cttcatgctg ttcotgtcca tgtacctgac cacaatgctg 120
gggaacctgc tcaatcactc ggccgtcaac tctgactccc acctccacac ccccatgtac 180
ttcotcctct ctatcctgtc ctgggtgac atctgtttca cctccaccac gatcccaga 240
atgtcgttga acatccaggc acaggctcaa tccatcaatt acacaggctg cctcccacaa 300
atctgctttg tctcgttttt tgttggattg gaaaatggaa tctcgttcat gatggcctat 360
gatcagattg tggccactct tcacccactg aggtacaatg tcatcatgaa ccccaaacctc 420
tgtgggctgc tgcttctgct gtccttcctc gttagtgtcc tggatgctct gctgcaacgc 480
ttgatgtgac tacagctgac ctctgcata gaactggaaa ttcccactt tttctgtgaa 540
ctagctcata ttctcaagct gcctgttct gatgtctcca tcaataacat cctgggtgat 600
ttggtgacca gectgttagg tgttcttct ctctctggga tcattttctc ttacacacga 660
attgtctct ctgctcatgaa aattccatca gctgggtgaa agtataaagc tttttccatc 720
tgcgggtcac atttaactgt tgtttccttg ttttatggaa cagggtttgg ggtgtacctt 780
agttctgggg ctacccaactc ctccaggaag ggtgcaatag catcagtgat gtataccgtg 840
gtcaacccca tgctgaaccc actcaattac agcctgagaa acaaggacat gttgaaggct 900
ttgaggaaac taatatctag gataccatct ttccattga 939

```

<210> 29

<211> 930

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7483134CB1

<400> 29

```

atgggtgcca agaacaatgt gactgagttt gttttatttg gcctttttga gagcagagag 60
atgcagcata catgctttgt ggtattcttc ctcttctcat tgetcaactg cctggggaac 120
cttctgttca tcaatccact caatgctaga aagaccctga agtctcccat gtattctctc 180
ctgagccagt tgtcttttgc tgacatagtg tatccatcca ctaccatacc caegatgatt 240

```

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

gctgacact ttgtggaga taagatcgc tcttcaatg gctgcatgac ccagctcttt 300
tcggccact ttttggtyg cactgagac ttctctcta cagccatggc ctatgaocgc 360
tatgtggca tctgtaggcc cctgcactac acagccatca tggattgocg gaagtgttgc 420
ctgctagcgg gggctcctg gttagctggc ttctctgatt ccattcctga gacctctctc 480
acggttcagc tgcctttttg tgggccaat gagatagaca acttctctg tgatgttcat 540
ccctgtctca agttggcctg tgcagacacc tacatggtag gttctctctg ggtggccaac 600
agcggatga tttctttagc atcctttttt atccttatca ttctctatgt tatcatctta 660
ctgaacctaa gaagccagtc atctgaggac cggcgttaag ctgtctccac atgtggctca 720
cacgtaatca ctgtcctttt ggtttctatg ccccccatgt tcatgtacat tctgtccctc 780
accaccctgg ctgctgacaa acttatctac ctctttaaca ttgtgatgac accttctgtg 840
aaccttttga tctatacact aaggacaac gatgtgaaa atgcatgag gaagctgttt 900
agggtaaga gtagcttagg ggagaagta 930

```

<210> 30

<211> 1161

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7478550CB1

<400> 30

```

gaggacggc tgctgcgaga ccagccggc cgtctttggc agtagtgggc gtccttggc 60
gtccaggagg gccctctccc cgcgaccgcc gaccacgatg agagcgtgaa gccctctcg 120
aaaggacagg ctctgctcta cacactggtg tctctcggg aaggcagct gggcagcct 180
tccagaccga gcaaaagtgg aaaaatcagt aacccaatga tgactaacg aaataagta 240
gtgtggggca ggtgagaca tcaatgctg cctcagacag aacgagaca cacaaacat 300
gacctctcac tccaagccca gttacagcag aaagttttca tggagaaatg gaatcacact 360
tcaaatgatt tcaatttggc ggttctgctt ccccccaatc aaactggaat atttctctg 420
tgccttatca tctcatatt ctctctggcc tgggtgggta actcggccat gattcacctc 480
atccacgtgg atctctgtct ccaacacccg atgtacttcc ttctagcca gctctcctt 540
atggacctga tgtacatctc caccacgctc cccaagatgg cgtacaactt cctgtccggc 600
cagaagggca tctctctctt gggatgtggt gtgcaagct tcttctctc gaccatggcg 660
tgttttgaag gcttactcct gacctccatg gctcaagacc gttatttggc catctgccac 720
tctctctatt atctatccg catgagtata atgatgtgtg tgaagatgat tggagctct 780
tggacactgg ggtccatcaa ctctctggca cacacagtct ttgacctca tattccctac 840
tgcaggtcta gggctattga ccattctctc tgcgatgtcc cagccatgtt gcttcttgc 900
tgtacagata cttgggtcta tgaatatatg gttttgttaa gtacaagcct ctttctctc 960
ttcccttcca ttggcatcac ttctctctgt ggccgagtc tatttctgt ctatcatatg 1020
cactcaaaag aggggagaaa aaaggccttc accaccattt caacacattt aactgtagt 1080
atcttttact atgcaccttt tgtctacacc tatctctggc ccacggaaac tccgtcacc 1140
agctgaagac aagatcttgc c 1161

```

<210> 31

<211> 948

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7483142CB1

WO 02/26825

PCT/US01/30661

PI-0236 PCT

```

<400> 31
atgtccataa ccaagcctg gaacagctca teagtacca tgttcaccc cctgggattc 60
acagaccacc cagaactcca ggcctctctc ttgtgacct tcctgggcat ctatcttaac 120
acctggcctt ggaacctggc cctcattttt ctgacagag gtgacacca cctgcacaca 180
cccattgact tcttctcaag caacttatct ttcattgaca tctgctactc tctgctctg 240
gtccccataa tgotcaotga cttctctctg gagcagaaga ccatatcatt tgtggctgt 300
gotgctcaat ttttttctt tctcggctag ggtctgtctg agtgcctcct cctgactgct 360
atggcaacag acagatagtc agccatctcc agcccccttc tctacccccc tatcatgacc 420
caggcctctc gtaacacgat ggtggttggg gcatatggtg gtggcttctc gactccctg 480
atccaggcca gctccatatt taggcttccac ttttgctgac ccaacatcat caaccacttc 540
tctgctgacc tcccaccagt cctgctctctc tctgctctg acaccttctc cagtcaagt 600
gtgaatttcc tctggtggtt cactgtcggg ggaacatcgt tctctcaact ccttatctcc 660
tatggttaca tagtctctgc ggtcctgaag atccctcag cagagggccg atggaaaagc 720
tgcaacacgt gtgcctcgcg tctgatggtg gtgactctgc tgtttgggac agcccccttc 780
gtgtacttgc gacccagctc cagctacttg ctaggcaggg acaaggtggt gctgttttcc 840
tattcattgg tgatccccat gctgaaccct ctcatttaca gtttgaggaa caaagagatc 900
aaggatgccc tgtggaaggt gttggaaggt aagaaggtgt tttcttag 948

```

&lt;210&gt; 32

&lt;211&gt; 924

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7483151CB1

```

<400> 32
atggcagaag aaaataagat tctggtgact cactttgtcc tcacaggact cacagatcat 60
ccaaggctgc aggcgcctct gttcctggtg ttcttggta totacctcat caccctggtg 120
ggcaaccttg gctgatggc tctcatctgg aaggacccc accctcacac ccccatatac 180
ttattctctg gcagtttagc ctttgagat gcatgcactc catcctctgt aactcttaag 240
atgctttcaa tttttttatc aaagaatcat atgctatcca tggctaagtg tgcccaccag 300
ttttactttt ttggttccaa tgcaaccaca gaatgcttcc tgcctgtagt gatggcctat 360
gaccgctatg tagccataty caatcccttg ctttatccag tggtagatgc caatagcctc 420
tgtactcagt ttataggtat ttcatatttt attggttttc tgcattcagc gattcatgtg 480
ggtttgttat ttatagtaaac ttctctgagg tccaatatta tacattattt ctactgtgae 540
attttacagc tgttcaaaat ttcttgcaac aatcctacag ttaataatct tctgattttc 600
atcctttcag catttataca agtcttcaat tttatgactc ttatgctc ttactcttat 660
atctctctct ccatcctgaa aaagaagtct gagaagggta gaagcaaac cttctctact 720
tgcaagtccc atctgctctc tgtctctttg ttctacggca cctctctct catgtatgtg 780
agttctaggt ctggtcagc tgcagatcag gccaaaagt attctttatt ttacacaata 840
ataattcctt tactaaatcc ttttatttac agcctaagga acaagaggtt tatagatgcc 900
ctgagaagaa tcatgaagaa ataa 948

```

## 【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  
4 April 2002 (04.04.2002)

PCT

(10) International Publication Number  
WO 02/026825 A3(51) International Patent Classification: C12N 15/12,  
C07K 14/705, 16/28, 14/72

(21) International Application Number: PCT/US01/30661

(22) International Filing Date:  
28 September 2001 (28.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/236,546 29 September 2000 (29.09.2000) US  
60/240,589 13 October 2000 (13.10.2000) US  
60/242,322 20 October 2000 (20.10.2000) US  
60/242,223 20 October 2000 (20.10.2000) US  
60/245,855 3 November 2000 (03.11.2000) US  
60/245,900 3 November 2000 (03.11.2000) US  
60/247,587 9 November 2000 (09.11.2000) US  
60/249,343 15 November 2000 (15.11.2000) US(71) Applicant (for all designated States except US): INCYTE  
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo  
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BAUGHN, Mariah,  
R. [US/US]; 14244 Santiago Road, San Leandro, CA  
94577 (US). GRAUL, Richard, C. [US/US]; 682-29th Ave-  
nue, San Francisco, CA 94121 (US). WALIA, Narinder,  
K. [US/US]; 890 Davis Street #205, San Leandro, CA  
94577 (US). GANDHI, Ameena, R. [US/US]; 837 Roble  
Avenue, #1, Menlo Park, CA 94025 (US). HAFALIA,  
April, J., A. [US/US]; 2227 Calle de Primavera, Santa  
Clara, CA 95054 (US). RAMKUMAR, Jaylaxmi  
[IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US).  
TRIBOULEY, Catherine, M. [FR/US]; 1121 Tennessee  
Street, #5, San Francisco, CA 94107 (US). THORNTON,  
Michael [US/US]; 9 Medway Road, Woodside, CA 94062  
(US). KALLICK, Deborah, A. [US/US]; 900 Olive  
Street, Menlo Park, CA 94025 (US). YAO, Monique,  
G. [US/US]; 111 Frederick Court, Mountain View, CA  
94043 (US). ELLIOTT, Vicki, S. [US/US]; 3770 Polton  
Place Way, San Jose, CA 95121 (US). BURFORD, Neil  
[GB/US]; 105 Wildwood Circle, Durham, CT 06422(US). KHAN, Farrah, A. [IN/US]; 3617 Central Road  
#102, Glenview, IL 60025 (US). YUE, Henry [US/US];  
826 Lois Avenue, Sunnyvale, CA 94087 (US). LU, Yan  
[CN/US]; 3885 Corinna Way, Palo Alto, CA 94303 (US).  
ARVIZU, Chandra [US/US]; 490 Sherwood Way #1,  
Menlo Park, CA 94025 (US). ROOPA, Reddy [IN/US];  
1233 W. McKinley Avenue, #3, Sunnyvale, CA 94086  
(US). NGUYEN, Dannel, B. [US/US]; 1403 Ridgewood  
Drive, San Jose, CA 95118 (US). LEE, Ernestine, A.  
[US/US]; 624 Kains Street, Albany, CA 94706 (US).  
LU, Dyung, Aina, M. [US/US]; 233 Coy Drive, San  
Jose, CA 95123 (US). ISON, Craig, H. [US/US]; 1242  
Weathersfield Way, San Jose, CA 95118 (US). WALSH,  
Roderick, T. [GB/US]; 8 Boundary Court, St. Lawrence  
Road, Canterbury, Kent CT1 3FZ (GB). POLICKY,  
Jennifer, L. [US/US]; 1511 Jarvis Court, San Jose, CA  
95118 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,  
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EG, ES, FI, GB, GD, GH, GI,  
GM, GR, GU, HT, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,  
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,  
ZW.(84) Designated States (regional): ARIPO patent (GI, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,  
TG).Published:  
with international search report(88) Date of publication of the international search report:  
24 April 2002For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: G-PROTEIN COUPLED RECEPTORS

(57) Abstract: The invention provides human G-protein coupled receptors (GCR) and polynucleotides which identify and en-  
code GCR. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also  
provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of GCR.

WO 02/026825 A3

## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/30661
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/705 C07K16/28 C07K14/72		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, BIOSIS, MEDLINE, SEQUENCE SEARCH		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] 8 April 1998 (1998-04-08) ADAMS, M.D. ET AL.: "Use of random BAC End Sequence Database for Sequence-Ready Map Building (1998)" Database accession no. B99531 XP002202568 the whole document	3,13
P,X	WO 01 18049 A (HUMAN GENOME SCIENCES INC ;NI JIAN (US); RUBEN STEVEN M (US); SHI) 15 March 2001 (2001-03-15)  the whole document	1,3,6,7, 9,12-20, 22,23, 25-42, 44,45
	---	-/--
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "&" document member of the same patent family		
Date of the actual completion of the international search 19 June 2002		Date of mailing of the international search report 27 11 2002
Name and mailing address of the ISA European Patent Office, P.B. 5816 Patentlaan 2 NL - 2200 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3015		Authorized officer Schulz, R

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 01/30661

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	STADEL J M ET AL: "Orphan G protein-coupled receptors: a neglected opportunity for pioneer drug discovery" TRENDS IN PHARMACOLOGICAL SCIENCES, ELSEVIER TRENDS JOURNAL, CAMBRIDGE, GB, vol. 18, no. 11, 1 November 1997 (1997-11-01), pages 430-437, XP004099345 ISSN: 0165-6147 the whole document -----	

4

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 01/30661
<b>Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)</b>	
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
<p>1. <input checked="" type="checkbox"/> Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210</p> <p>2. <input checked="" type="checkbox"/> Claims Nos.: 21, 24, 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</p> <p>3. <input type="checkbox"/> Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).</p>	
<b>Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)</b>	
This International Searching Authority found multiple inventions in this international application, as follows:	
see additional sheet	
<p>1. <input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.</p> <p>2. <input type="checkbox"/> As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.</p> <p>3. <input type="checkbox"/> As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:</p> <p>4. <input checked="" type="checkbox"/> 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 claims Nos.: 1-61 partially and 62, 78 completely</p>	
<p><b>Remark on Protest</b></p> <p><input type="checkbox"/> The additional search fees were accompanied by the applicant's protest.</p> <p><input type="checkbox"/> No protest accompanied the payment of additional search fees.</p>	

International Application No. PCT/US 01/30661

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 33 and 35 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition and although claims 19, 22, 25 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 21, 24,

Claims 21 and 24 refer to compositions comprising agonists or antagonists respectively identified in screens designed to find either of these compounds.

No specific compounds are defined or characterised in the application. In consequence, the scope of these claims is ambiguous and vague and its subject-matter is not sufficiently disclosed and supported (Art. 6 PCT).

No search can be carried out for such purely speculative claims 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 01/30661

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 - 61 (partially); 62, 78 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 1 and fragments thereof; isolated polynucleotides (SEQ ID NO: 17) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide; a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

2. Claims: 1 - 61 (partially); 63, 79 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 2 and fragments thereof; isolated polynucleotides (SEQ ID NO: 18) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide; a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

International Application No. PCT/US 01/30661

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## 3. Claims: 1 - 61 (partially); 64, 80 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 3 and fragments thereof; isolated polynucleotides (SEQ ID NO: 19) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

## 4. Claims: 1 - 61 (partially); 65, 81 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 4 and fragments thereof; isolated polynucleotides (SEQ ID NO: 20) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

## 5. Claims: 1 - 61 (partially); 66, 82 (completely)

International Application No. PCT/US 01/30661

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 5 and fragments thereof; isolated polynucleotides (SEQ ID NO: 21) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide; a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

6. Claims: 1 - 61 (partially); 67, 83 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 6 and fragments thereof; isolated polynucleotides (SEQ ID NO: 22) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide; a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

7. Claims: 1 - 61 (partially); 68, 84 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 7 and fragments thereof; isolated polynucleotides (SEQ ID NO: 23) encoding said polypeptide; cells transformed with

International Application No. PCT/US 01/30661

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

8. Claims: 1 - 61 (partially); 69, 85 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 8 and fragments thereof; isolated polynucleotides (SEQ ID NO: 24) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

9. Claims: 1 - 61 (partially); 70, 86 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 9 and fragments thereof; isolated polynucleotides (SEQ ID NO: 25) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide;

International Application No. PCT/US 01/30661

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide; a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

10. Claims: 1 - 61 (partially); 71, 87 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 10 and fragments thereof; isolated polynucleotides (SEQ ID NO: 26) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide; a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

11. Claims: 1 - 61 (partially); 72, 88 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 11 and fragments thereof; isolated polynucleotides (SEQ ID NO: 27) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a

International Application No. PCT/US 01/30661

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

## 12. Claims: 1 - 61 (partially); 73, 89 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 12 and fragments thereof; isolated polynucleotides (SEQ ID NO: 28) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

## 13. Claims: 1 - 61 (partially); 74, 90 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 13 and fragments thereof; isolated polynucleotides (SEQ ID NO: 29) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist

International Application No. PCT/US 01/30661

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

or modulators 6) of assessing toxicity of a test compound;  
7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

## 14. Claims: 1 - 61 (partially); 75, 91 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 14 and fragments thereof; isolated polynucleotides (SEQ ID NO: 30) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody;  
methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

## 15. Claims: 1 - 61 (partially); 76, 92 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 15 and fragments thereof; isolated polynucleotides (SEQ ID NO: 31) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody;  
methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests

International Application No. PCT/US 01/30661

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

16. Claims: 1 - 61 (partially); 77, 93 (completely)

Isolated polypeptide comprising amino acid sequence SEQ ID NO: 16 and fragments thereof; isolated polynucleotides (SEQ ID NO: 32) encoding said polypeptide; cells transformed with said recombinant polynucleotide; transgenic organism comprising said recombinant polynucleotide; an isolated antibody that specifically binds to said polypeptide; compositions comprising said polypeptide or agonists or antagonists of said polypeptide or said antibody; methods: 1) of producing said polypeptide, 2) of preparing a said antibodies, 3) of detecting or purifying said polypeptide 4) of detecting a target polynucleotide, 5) of screening for a compound effective as an agonist, antagonist or modulators 6) of assessing toxicity of a test compound; 7) a method of diagnosing a condition or disease associated with the expression of said polypeptide; diagnostic tests comprising said antibody; use of said polypeptide or any of its agonists or antagonists for the preparation of medicament to treat a disease associated with the expression of said polypeptide;  
a microarray wherein at least one element is said polynucleotide and a method of generating a transcript image of a sample which contains said polynucleotide.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International Application No  
PCT/US 01/30661

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0118049	A	15-03-2001	
		AU 7354600 A	10-04-2001
		EP 1214342 A1	19-06-2002
		WO 0118049 A1	15-03-2001
		US 2002165384 A1	07-11-2002
		US 2002120110 A1	29-08-2002

## フロントページの続き

(51)Int.Cl. <sup>7</sup>	F I	テーマコード(参考)
A 6 1 P 1/08	A 6 1 P 1/10	4 B 0 6 5
A 6 1 P 1/10	A 6 1 P 1/12	4 C 0 8 4
A 6 1 P 1/12	A 6 1 P 1/14	4 H 0 4 5
A 6 1 P 1/14	A 6 1 P 1/16	
A 6 1 P 1/16	A 6 1 P 1/18	
A 6 1 P 1/18	A 6 1 P 3/00	
A 6 1 P 3/00	A 6 1 P 3/04	
A 6 1 P 3/04	A 6 1 P 3/10	
A 6 1 P 3/10	A 6 1 P 7/00	
A 6 1 P 7/00	A 6 1 P 7/02	
A 6 1 P 7/02	A 6 1 P 7/06	
A 6 1 P 7/06	A 6 1 P 9/00	
A 6 1 P 9/00	A 6 1 P 9/04	
A 6 1 P 9/04	A 6 1 P 9/08	
A 6 1 P 9/08	A 6 1 P 9/10	
A 6 1 P 9/10	A 6 1 P 9/10	1 0 1
A 6 1 P 9/14	A 6 1 P 9/14	
A 6 1 P 11/06	A 6 1 P 11/06	
A 6 1 P 11/16	A 6 1 P 11/16	
A 6 1 P 17/00	A 6 1 P 17/00	
A 6 1 P 17/02	A 6 1 P 17/02	
A 6 1 P 17/06	A 6 1 P 17/06	
A 6 1 P 17/12	A 6 1 P 17/12	
A 6 1 P 19/02	A 6 1 P 19/02	
A 6 1 P 19/04	A 6 1 P 19/04	
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 21/00	A 6 1 P 21/00	
A 6 1 P 21/02	A 6 1 P 21/02	
A 6 1 P 21/04	A 6 1 P 21/04	
A 6 1 P 25/00	A 6 1 P 25/00	
A 6 1 P 25/02	A 6 1 P 25/02	
A 6 1 P 25/04	A 6 1 P 25/02	1 0 3
A 6 1 P 25/08	A 6 1 P 25/04	
A 6 1 P 25/14	A 6 1 P 25/08	
A 6 1 P 25/16	A 6 1 P 25/14	
A 6 1 P 25/18	A 6 1 P 25/16	
A 6 1 P 25/20	A 6 1 P 25/18	
A 6 1 P 25/22	A 6 1 P 25/20	
A 6 1 P 25/28	A 6 1 P 25/22	
A 6 1 P 27/02	A 6 1 P 25/28	
A 6 1 P 29/00	A 6 1 P 27/02	
A 6 1 P 31/04	A 6 1 P 29/00	
A 6 1 P 31/10	A 6 1 P 31/04	
A 6 1 P 31/12	A 6 1 P 31/10	
A 6 1 P 31/18	A 6 1 P 31/12	
A 6 1 P 31/20	A 6 1 P 31/18	

A 6 1 P 31/22	A 6 1 P 31/20	
A 6 1 P 33/00	A 6 1 P 31/22	
A 6 1 P 33/02	A 6 1 P 33/00	
A 6 1 P 33/10	A 6 1 P 33/02	
A 6 1 P 35/00	A 6 1 P 33/10	
A 6 1 P 35/02	A 6 1 P 35/00	
A 6 1 P 37/00	A 6 1 P 35/02	
A 6 1 P 37/02	A 6 1 P 37/00	
A 6 1 P 37/08	A 6 1 P 37/02	
A 6 1 P 43/00	A 6 1 P 37/08	
C 0 7 K 14/705	A 6 1 P 43/00	1 0 5
C 0 7 K 16/18	C 0 7 K 14/705	
C 1 2 M 1/00	C 0 7 K 16/18	
C 1 2 N 1/15	C 1 2 M 1/00	A
C 1 2 N 1/19	C 1 2 N 1/15	
C 1 2 N 1/21	C 1 2 N 1/19	
C 1 2 N 5/10	C 1 2 N 1/21	
C 1 2 P 21/02	C 1 2 P 21/02	C
C 1 2 P 21/08	C 1 2 P 21/08	
C 1 2 Q 1/02	C 1 2 Q 1/02	
C 1 2 Q 1/68	C 1 2 Q 1/68	A
G 0 1 N 33/15	C 1 2 Q 1/68	Z
G 0 1 N 33/50	G 0 1 N 33/15	Z
G 0 1 N 33/53	G 0 1 N 33/50	Z
G 0 1 N 33/566	G 0 1 N 33/53	D
G 0 1 N 37/00	G 0 1 N 33/566	M
	G 0 1 N 37/00	1 0 2
	C 1 2 N 5/00	A
	C 1 2 N 15/00	F
	A 6 1 K 37/02	

(31)優先権主張番号 60/242,223  
(32)優先日 平成12年10月20日(2000.10.20)  
(33)優先権主張国 米国(US)

(31)優先権主張番号 60/245,855  
(32)優先日 平成12年11月3日(2000.11.3)  
(33)優先権主張国 米国(US)

(31)優先権主張番号 60/245,900  
(32)優先日 平成12年11月3日(2000.11.3)  
(33)優先権主張国 米国(US)

(31)優先権主張番号 60/247,587  
(32)優先日 平成12年11月9日(2000.11.9)  
(33)優先権主張国 米国(US)

(31)優先権主張番号 60/249,343  
(32)優先日 平成12年11月15日(2000.11.15)  
(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,GQ,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, EC, 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 グラール、リチャード・シー  
アメリカ合衆国カリフォルニア州94121・サンフランシスコ・トゥエンティナインズアベニュー 682
- (72)発明者 ワライヤ、ナリンダー・ケイ  
アメリカ合衆国カリフォルニア州94577・サンレアンドロ・#205・デービスストリート 890
- (72)発明者 ガンディー、アミーナ・アール  
アメリカ合衆国カリフォルニア州94025・メンロパーク・#1・ローブルアベニュー 837
- (72)発明者 ハファリア、エープリル・ジェイ・エイ  
アメリカ合衆国カリフォルニア州95054・サンタクララ・コーレデプリマベラ 2227
- (72)発明者 ランクマール、ジャヤラクシミ  
アメリカ合衆国カリフォルニア州94555・フレモント・メイパードサークル 34359
- (72)発明者 トリボレー、キャサリン・エム  
アメリカ合衆国カリフォルニア州94107・サンフランシスコ・#5・テネシーストリート 1121
- (72)発明者 ソントン、マイケル  
アメリカ合衆国カリフォルニア州94062・ウッドサイド・メッドウェイロード 9
- (72)発明者 カリック、デボラー・エイ  
アメリカ合衆国カリフォルニア州94025・メンロパーク・オリーブストリート 900
- (72)発明者 ヤオ、モニック・ジー  
アメリカ合衆国カリフォルニア州94043・マウンテンビュー・フレデリックコート 111
- (72)発明者 エリオット、ビッキー・エス  
アメリカ合衆国カリフォルニア州95121・サンノゼ・ポルトンプレイスウェイ 3770
- (72)発明者 パーフォード、ニール  
アメリカ合衆国コネチカット州06422・ダラム・ワイルドウッドサークル 105
- (72)発明者 カーン、ファラ・エイ  
アメリカ合衆国イリノイ州60025・グレンビュー・#102・セントラルロード 3617
- (72)発明者 ユエ、ヘンリー  
アメリカ合衆国カリフォルニア州94087・サニーベイル・ルイスアベニュー 826
- (72)発明者 リュ、ヤン  
アメリカ合衆国カリフォルニア州94303・パロアルト・コリーナウェイ 3885
- (72)発明者 アービズ、チャンドラ  
アメリカ合衆国カリフォルニア州94025・メンロパーク・#1・シャーウッドウェイ 490
- (72)発明者 ルーバ、レディ  
アメリカ合衆国カリフォルニア州94086・サニーベイル・#3・ウェストマッキンレーアベニュー 1233
- (72)発明者 ニュエン、ダニエル・ビー  
アメリカ合衆国カリフォルニア州95118・サンノゼ・リッジウッドドライブ 1403
- (72)発明者 リー、アーンステーション・エイ  
アメリカ合衆国カリフォルニア州94706・アルバニー・ケインズストリート 624
- (72)発明者 リュ、デュング・アイナ・エム  
アメリカ合衆国カリフォルニア州95123・サンノゼ・コイドライブ 233
- (72)発明者 アイゾン、クレイグ・エイチ  
アメリカ合衆国カリフォルニア州95118・サンノゼ・ウェザーズフィールドウェイ 1242
- (72)発明者 ウォルシュ、ロドリック・ティー  
イギリス国ケント州・シーティー1 3イージー・カンタベリー・セントローレンスロード・バ

## ウンダリーコート 8

(72)発明者 ポリッキー、ジェニファー・エル

アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ ・ ジャービスコート 1 5 1 1

F ターム(参考) 2G045 AA34 AA35 BA14 BB01 BB07 BB14 BB20 BB46 BB48 BB50  
BB51 CB01 DA13 DA36 FA16 FA29 FB02 FB03 FB05 FB08  
FB12 GC15  
4B024 AA01 AA11 BA63 CA04 CA09 CA11 DA02 DA06 EA02 EA03  
HA11 HA14 HA17  
4B029 AA07 FA12  
4B063 QA08 QA13 QA17 QA18 QQ05 QQ42 QQ52 QR32 QR35 QR55  
QR62 QR74 QR84 QS25 QS33 QS34 QX02 QX07  
4B064 AG27 CA10 CA20 CC24 DA01 DA13  
4B065 AA26X AA93Y AB01 CA24 CA44 CA46  
4C084 AA02 AA07 NA14 ZA01 ZA02 ZA05 ZA06 ZA08 ZA12 ZA15  
ZA16 ZA18 ZA20 ZA22 ZA23 ZA24 ZA33 ZA36 ZA39 ZA40  
ZA41 ZA42 ZA45 ZA51 ZA55 ZA59 ZA60 ZA66 ZA68 ZA69  
ZA70 ZA71 ZA72 ZA73 ZA75 ZA81 ZA91 ZA94 ZA96 ZB05  
ZB07 ZB11 ZB13 ZB15 ZB21 ZB26 ZB27 ZB33 ZB35 ZB38  
ZB39 ZC21 ZC35 ZC55  
4H045 AA10 AA11 AA20 AA30 BA10 CA40 DA50 DA75 DA76 EA20  
EA50 FA74

专利名称(译)	G蛋白偶联受体		
公开(公告)号	<a href="#">JP2004528007A</a>	公开(公告)日	2004-09-16
申请号	JP2002531207	申请日	2001-09-28
[标]申请(专利权)人(译)	洞察Genomics公司		
申请(专利权)人(译)	洞察基因组公司		
[标]发明人	ボーグンマライアアール グラールリチャードシー ワライヤナリンダーケイ ガンディーアミーナアール ハファリアエープリルジェイエイ ランクマールジャヤラクシミ トリボレーキャサリーンエム ソーントンマイケル カリックデボラーエイ ヤオモニークジー エリオットビッキーエス バーフォードニール カーンファラエイ ユエヘンリー リュヤン アービズチャンドラ ルーパレディ ニュエンダニエルビー リーアーンステイーンエイ リュデユングアイナエム アイソクレイグエイチ ウォルシュロドリックティー ポリッキージェニファーエル		
发明人	ボーグン、マライア・アール グラール、リチャード・シー ワライヤ、ナリンダー・ケイ ガンディー、アミーナ・アール ハファリア、エープリル・ジェイ・エイ ランクマール、ジャヤラクシミ トリボレー、キャサリーン・エム ソーントン、マイケル カリック、デボラー・エイ ヤオ、モニーク・ジー エリオット、ビッキー・エス バーフォード、ニール カーン、ファラ・エイ ユエ、ヘンリー リュ、ヤン アービズ、チャンドラ ルーパ、レディ ニュエン、ダニエル・ビー リー、アーンステイーン・エイ リュ、デユング・アイナ・エム アイソン、クレイグ・エイチ		

ウォルシュ、ロドリック・ティー  
ポリッキー、ジェニファー・エル

IPC分类号	G01N33/50 A61K38/00 A61P1/00 A61P1/04 A61P1/06 A61P1/08 A61P1/10 A61P1/12 A61P1/14 A61P1/16 A61P1/18 A61P3/00 A61P3/04 A61P3/10 A61P7/00 A61P7/02 A61P7/06 A61P9/00 A61P9/04 A61P9/08 A61P9/10 A61P9/14 A61P11/06 A61P11/16 A61P17/00 A61P17/02 A61P17/06 A61P17/12 A61P19/02 A61P19/04 A61P19/08 A61P19/10 A61P21/00 A61P21/02 A61P21/04 A61P25/00 A61P25/02 A61P25/04 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/28 A61P27/02 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P31/20 A61P31/22 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/08 A61P43/00 C07K14/705 C07K16/18 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N5/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 G01N37/00
CPC分类号	A61P1/00 A61P1/04 A61P1/06 A61P1/08 A61P1/10 A61P1/12 A61P1/14 A61P1/16 A61P1/18 A61P3/00 A61P3/04 A61P11/06 A61P11/16 A61P17/00 A61P17/02 A61P17/06 A61P17/12 A61P19/02 A61P19/04 A61P19/08 A61P19/10 A61P21/00 A61P21/02 A61P21/04 A61P25/00 A61P25/02 A61P25/04 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/28 A61P27/02 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P31/20 A61P31/22 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 C07K14/705
FI分类号	C12N15/00.ZNA.A A61P1/00 A61P1/04 A61P1/06 A61P1/08 A61P1/10 A61P1/12 A61P1/14 A61P1/16 A61P1/18 A61P3/00 A61P3/04 A61P3/10 A61P7/00 A61P7/02 A61P7/06 A61P9/00 A61P9/04 A61P9/08 A61P9/10 A61P9/10.101 A61P9/14 A61P11/06 A61P11/16 A61P17/00 A61P17/02 A61P17/06 A61P17/12 A61P19/02 A61P19/04 A61P19/08 A61P19/10 A61P21/00 A61P21/02 A61P21/04 A61P25/00 A61P25/02 A61P25/02.103 A61P25/04 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/28 A61P27/02 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P31/20 A61P31/22 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/08 A61P43/00.105 C07K14/705 C07K16/18 C12M1/00.A C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/02 C12Q1/68.A C12Q1/68.Z G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 G01N37/00.102 C12N5/00.A C12N15/00.F A61K37/02
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/BA14 2G045/BB01 2G045/BB07 2G045/BB14 2G045/BB20 2G045/BB46 2G045/BB48 2G045/BB50 2G045/BB51 2G045/CB01 2G045/DA13 2G045/DA36 2G045/FA16 2G045/FA29 2G045/FB02 2G045/FB03 2G045/FB05 2G045/FB08 2G045/FB12 2G045/GC15 4B024/AA01 4B024/AA11 4B024/BA63 4B024/CA04 4B024/CA09 4B024/CA11 4B024/DA02 4B024/DA06 4B024/EA02 4B024/EA03 4B024/HA11 4B024/HA14 4B024/HA17 4B029/AA07 4B029/FA12 4B063/QA08 4B063/QA13 4B063/QA17 4B063/QA18 4B063/QQ05 4B063/QQ42 4B063/QQ52 4B063/QR32 4B063/QR35 4B063/QR55 4B063/QR62 4B063/QR74 4B063/QR84 4B063/QS25 4B063/QS33 4B063/QS34 4B063/QX02 4B063/QX07 4B064/AG27 4B064/CA10 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA26X 4B065/AA93Y 4B065/AB01 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA07 4C084/NA14 4C084/ZA01 4C084/ZA02 4C084/ZA05 4C084/ZA06 4C084/ZA08 4C084/ZA12 4C084/ZA15 4C084/ZA16 4C084/ZA18 4C084/ZA20 4C084/ZA22 4C084/ZA23 4C084/ZA24 4C084/ZA33 4C084/ZA36 4C084/ZA39 4C084/ZA40 4C084/ZA41 4C084/ZA42 4C084/ZA45 4C084/ZA51 4C084/ZA55 4C084/ZA59 4C084/ZA60 4C084/ZA66 4C084/ZA68 4C084/ZA69 4C084/ZA70 4C084/ZA71 4C084/ZA72 4C084/ZA73 4C084/ZA75 4C084/ZA81 4C084/ZA91 4C084/ZA94 4C084/ZA96 4C084/ZB05 4C084/ZB07 4C084/ZB11 4C084/ZB13 4C084/ZB15 4C084/ZB21 4C084/ZB26 4C084/ZB27 4C084/ZB33 4C084/ZB35 4C084/ZB38 4C084/ZB39 4C084/ZC21 4C084/ZC35 4C084/ZC55 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA50 4H045/DA75 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA74
優先権	60/236546 2000-09-29 US 60/240589 2000-10-13 US 60/242322 2000-10-20 US 60/242223 2000-10-20 US 60/245855 2000-11-03 US 60/245900 2000-11-03 US 60/247587 2000-11-09 US 60/249343 2000-11-15 US

摘要(译)

本发明提供了鉴定和编码GCREC的人G蛋白偶联受体 ( GCREC ) 和多核苷酸。 本发明还提供表达载体，宿主细胞，抗体，激动剂和拮抗剂。 本发明还提供了用于诊断，治疗或预防与GCREC异常表达有关的疾病的方法。

						(43) 公表日 平成16年9月16日 (2004.9.16)	
						(P2004-528007A)	
(5) Int.Cl. <sup>7</sup>		F 1					テーマコード (参考)
C 1 2 N	15/00	C 1 2 N	15/00	Z N A A			2 G 0 4 5
A 6 1 K	38/00	A 6 1 P	1/00				4 B 0 2 4
A 6 1 P	1/00	A 6 1 P	1/04				4 B 0 2 9
A 6 1 P	1/04	A 6 1 P	1/06				4 B 0 6 3
A 6 1 P	1/06	A 6 1 P	1/08				4 B 0 6 4
						審査請求 未請求 予備審査請求 有 (全 283 頁) 最終頁に続く	
(21) 出願番号	特願2002-531207 (P2002-531207)	(71) 出願人	301005050				
(86) (22) 出願日	平成13年9月28日 (2001.9.28)		インサイト・ゲノミクス・インコーポレイテッド				
(85) 翻訳文提出日	平成15年3月28日 (2003.3.28)		アメリカ合衆国カリフォルニア州94304・パロアルト・ボータードライブ 3160				
(86) 国際出願番号	PCT/US2001/030661						
(87) 国際公開番号	W02002/026825	(74) 代理人	100089266				
(87) 国際公開日	平成14年4月4日 (2002.4.4)		弁理士 大島 陽一				
(31) 優先権主張番号	60/236,546	(72) 発明者	ボーグン、マライア・アール				
(32) 優先日	平成12年9月28日 (2000.9.28)		アメリカ合衆国カリフォルニア州94577・サンレアンドロ・サンティアゴロード 14244				
(33) 優先権主張国	米国 (US)						
(31) 優先権主張番号	60/240,589						
(32) 優先日	平成12年10月13日 (2000.10.13)						
(33) 優先権主張国	米国 (US)						
(31) 優先権主張番号	60/242,322						
(32) 優先日	平成12年10月20日 (2000.10.20)						
(33) 優先権主張国	米国 (US)						

最終頁に続く