

(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2004-501610

(P2004-501610A)

(43) 公表日 平成16年1月22日(2004.1.22)

(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 45/00</b>	A 6 1 K 45/00	2 G O 4 5
<b>A 6 1 K 51/00</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 6 3
<b>A 6 1 P 1/06</b>	A 6 1 P 1/16	4 B O 6 4
	審査請求 未請求 予備審査請求 有	4 B O 6 5
		(全 178 頁) 最終頁に続く

(21) 出願番号	特願2001-569007 (P2001-569007)	(71) 出願人	301005050
(86) (22) 出願日	平成13年3月22日 (2001. 3. 22)		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14年9月24日 (2002. 9. 24)		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/US2001/009220	(74) 代理人	100089266
(87) 国際公開番号	W02001/070807		弁理士 大島 陽一
(87) 国際公開日	平成13年9月27日 (2001. 9. 27)	(72) 発明者	ラル、ブリーティ
(31) 優先権主張番号	60/191, 631		アメリカ合衆国カリフォルニア州95056・サンタクララ・ビーオーボックス 5142
(32) 優先日	平成12年3月23日 (2000. 3. 23)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 Gタンパク質関連分子

## (57) 【要約】

本発明は、ヒトGタンパク質関連分子(GPAM)と、GPAMを同定及びコードするポリヌクレオチドとを提供する。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提供する。更に、本発明は、GPAMの異常発現に関連する疾患の診断・治療・予防方法を提供する。

## 【特許請求の範囲】

## 【請求項 1】

単離されたポリペプチドであって、

(a) SEQ ID NO: 1 及び SEQ ID NO: 2 (SEQ ID NO: 1 - 2) からなる群から選択されたアミノ酸を含むポリペプチドと、

(b) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列と少なくとも 80% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、

(c) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、

(d) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されることを特徴とする単離されたポリペプチド。 10

## 【請求項 2】

SEQ ID NO: 1 - 2 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: 3 - 4 からなる群から選択された請求項 4 の単離されたポリヌクレオチド。 20

## 【請求項 6】

請求項 3 のポリヌクレオチドに機能的に結合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。

## 【請求項 7】

請求項 6 の組換えポリヌクレオチドで形質転換された細胞。

## 【請求項 8】

請求項 6 の組換えポリヌクレオチドを含む遺伝子組換え生物。

## 【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結合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ステップと、

(b) そのように発現したポリペプチドを回収するステップとを含むことを特徴とする請求項 1 のポリペプチドの生産方法。

## 【請求項 10】

請求項 1 のポリペプチドに特異的に結合する単離された抗体。

## 【請求項 11】

単離されたポリヌクレオチドであって、 40

(a) SEQ ID NO: 3 - 4 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ポリヌクレオチドと、

(b) SEQ ID NO: 3 - 4 からなる群から選択されたポリヌクレオチド配列と少なくとも 80% の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、

(c) 前記 (a) のポリヌクレオチドに相補的なポリヌクレオチドと、

(d) 前記 (b) のポリヌクレオチドに相補的なポリヌクレオチドと、

(e) 前記 (a) 乃至 (d) の RNA 等価物とで構成される群から選択された単離されたポリヌクレオチド。

## 【請求項 12】

請求項 11 のポリヌクレオチドの少なくとも 60 個の連続するヌクレオチドを含む単離さ 50

れたポリヌクレオチド。

【請求項 13】

サンプルにおける請求項 11 に記載のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

(a) 前記サンプルをプローブでハイブリダイズするステップであって、前記プローブが、前記サンプル内の前記標的ポリヌクレオチドに相補的な配列を含む少なくとも 20 個の連続するヌクレオチドを含み、前記プローブと前記標的ポリヌクレオチドまたはその断片との間でハイブリダイゼーション複合体が形成される条件下で、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、該ステップと、

(b) 前記ハイブリダイゼーション複合体が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定するステップとを含むことを特徴とする標的ポリヌクレオチドの検出方法。

10

【請求項 14】

前記プローブが少なくとも 60 個の連続するヌクレオチドを含むことを特徴とする請求項 13 に記載の方法。

【請求項 15】

サンプルにおける請求項 11 のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

(a) ポリメラーゼ連鎖反応増幅法を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、

(b) 増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定するステップとを含むことを特徴とする標的ポリヌクレオチドの検出方法。

20

【請求項 16】

請求項 1 のポリペプチド及び医薬的に容認できる賦形剤を含む組成物。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を含むことを特徴とする請求項 16 の組成物。

【請求項 18】

機能的な GPAM (新規の G タンパク質関連分子) の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 16 の組成物を投与することを特徴とする治療方法。

30

【請求項 19】

請求項 1 のポリペプチドのアゴニストとして有効な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露するステップと、

(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 19 の方法によって同定されたアゴニスト化合物及び医薬的に容認できる賦形剤を含む組成物。

40

【請求項 21】

機能的な GPAM の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 20 の組成物を投与することを特徴とする治療方法。

【請求項 22】

請求項 1 のポリペプチドのアンタゴニストとして有効な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露するステップと、

(b) 前記サンプルにおける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とするスクリーニング方法。

50

## 【請求項 2 3】

請求項 2 2 の方法によって同定されたアンタゴニスト化合物及び医薬的に容認できる賦形剤を含む組成物。

## 【請求項 2 4】

機能的な G P A M の過剰な発現に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 2 3 の組成物を投与することを含むことを特徴とする治療方法。

## 【請求項 2 5】

請求項 1 のポリペプチドに特異的に結合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 のポリペプチドを好適な条件下で少なくとも 1 つの試験化合物と結合させるステップと、

10

( b ) 請求項 1 のポリペプチドと前記試験化合物との結合を検出して、請求項 1 のポリペプチドと特異的に結合する化合物を同定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 2 6】

請求項 1 のポリペプチドの活性を変化させ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 のポリペプチドを、そのポリペプチドの活性が許容される条件下で少なくとも 1 つの試験化合物と結合させるステップと、

( b ) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性を評価するステップと、

( c ) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性と、前記試験化合物の非存在下での請求項 1 のポリペプチドの活性とを比較するステップとを含み、

20

前記試験化合物の存在下での請求項 1 のポリペプチドの活性の変化が、請求項 1 のポリペプチドの活性を変化させる化合物の存在を示唆することを特徴とするスクリーニング方法。

## 【請求項 2 7】

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有効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、

( b ) 前記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、

30

( c ) 様々な量の前記化合物の存在下での前記標的ポリヌクレオチドの発現と、前記化合物の非存在下での前記標的ポリヌクレオチドの発現とを比較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 2 8】

試験化合物の毒性を評価する方法であって、

( a ) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、

( b ) 処理した前記生体サンプルの核酸を、請求項 1 1 のポリヌクレオチドの少なくとも 2 0 個の連続するヌクレオチドを含むプローブと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 1 1 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 ) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量と比較するステップとを含み、

前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差が試験化合物の毒性を示唆することを特徴とする試験化合物の毒性評価方法。

## 【請求項 2 9】

生体サンプルにおける G P A M の発現に関連する病態または疾患の診断テストであって、

50

( a ) 請求項 1 0 の抗体と前記ポリペプチドとの結合に好適な条件下で、前記生体サンプルを請求項 1 0 の抗体と結合させて、抗体 / ポリペプチド複合体を形成するステップと、  
 ( b ) 前記複合体を検出するステップとを含み、  
 前記複合体の存在が、前記生体サンプルにおける前記ポリヌクレオチドの存在と相関性を有することを特徴とする診断テスト。

【請求項 3 0】

前記抗体が、

- ( a ) キメラ抗体、
- ( b ) 一本鎖抗体、
- ( c ) F a b 断片、
- ( d ) F ( a b ' )<sub>2</sub> 断片、または
- ( e ) ヒト化抗体であることを特徴とする請求項 1 0 に記載の抗体。

10

【請求項 3 1】

請求項 1 0 の抗体及び許容される賦形剤を含む組成物。

【請求項 3 2】

患者において、G P A M の発現に関連する病態または疾患を診断する方法であって、請求項 3 1 の組成物を前記患者に有効量投与することを含むことを特徴とする方法。

【請求項 3 3】

前記抗体が標識されていることを特徴とする請求項 3 1 に記載の組成物。

【請求項 3 4】

患者において、G P A M の発現に関連する病態または疾患を診断する方法であって、請求項 3 3 の組成物を前記患者に有効量投与することを含むことを特徴とする方法。

20

【請求項 3 5】

請求項 1 0 の抗体の特異性を有するポリクローナル抗体を作製する方法であって、

- ( a ) 抗体反応が起こる条件下で、S E Q I D N O : 1 2 からなる群から選択されたアミノ酸配列またはその免疫原性断片を有するポリペプチドで動物を免疫化するステップと、
- ( b ) 前記動物から抗体を単離するステップと、
- ( c ) 前記単離された抗体を前記ポリペプチドでスクリーニングして、E Q I D N O : 1 2 からなる群から選択されたアミノ酸配列を有するポリペプチドに特異的に結合するポリクローナル抗体を同定するステップとを含むことを特徴とするポリクローナル抗体作製方法。

30

【請求項 3 6】

請求項 3 5 の方法により作製された抗体。

【請求項 3 7】

請求項 3 6 の抗体及び適当な担体を含む組成物。

【請求項 3 8】

請求項 1 0 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 ) 抗体反応が起こる条件下で、S E Q I D N O : 1 2 からなる群から選択されたアミノ酸配列またはその免疫原性断片を有するポリペプチドで動物を免疫化するステップと、
- ( b ) 前記動物から抗体を産生する細胞を単離するステップと、
- ( c ) 前記抗体産生細胞を不死化細胞と融合し、モノクローナル抗体産生ハイブリドーマ細胞を形成するステップと、
- ( d ) 前記ハイブリドーマ細胞を培養するステップと、
- ( e ) S E Q I D N O : 1 2 からなる群から選択されたアミノ酸配列を有するポリペプチドに特異的に結合するモノクローナル抗体を培養液から単離するステップとを含むことを特徴とするモノクローナル抗体作製方法。

40

【請求項 3 9】

請求項 3 8 の方法で作製したモノクローナル抗体。

50

## 【請求項 40】

請求項 39 の抗体及び適当な担体を含む組成物。

## 【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得られた請求項 10 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得られた請求項 10 の抗体。

## 【請求項 43】

サンプルにおける S E Q I D N O : 1 2 からなる群から選択されたアミノ酸配列を有するポリペプチドを検出する方法であって、

( a ) 請求項 10 の抗体と前記ポリペプチドとの特異的な結合が許容される条件下で、前記抗体を前記サンプルと共にインキュベートするステップと、 10

( b ) 特異的な結合を検出するステップとを含み、

前記特異的な結合が、前記サンプルに S E Q I D N O : 1 2 からなる群から選択されたアミノ酸配列を有するポリペプチドが存在することを示唆することを特徴とする方法。

## 【請求項 44】

S E Q I D N O : 1 2 からなる群から選択されたアミノ酸配列を有するポリペプチドをサンプルから精製する方法であって、

( a ) 請求項 10 の抗体と前記ポリペプチドとの特異的な結合が許容される条件下で、前記抗体を前記サンプルと共にインキュベートするステップと、

( b ) 前記サンプルから前記抗体を分離して、S E Q I D N O : 1 2 からなる群から選択されたアミノ酸配列を有する精製されたポリペプチドを得るステップとを含むことを特徴とする方法。 20

## 【請求項 45】

S E Q I D N O : 1 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 46】

S E Q I D N O : 2 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 47】

S E Q I D N O : 3 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

## 【請求項 48】

S E Q I D N O : 4 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。 30

## 【発明の詳細な説明】

## 【0001】

(技術分野)

本発明は、G タンパク質関連分子の核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を利用した細胞増殖異常及び自己免疫/炎症性の疾患の診断・治療・予防に関する。本発明は更に、G タンパク質関連分子の核酸配列及びアミノ酸配列の発現における外来性化合物の影響についての評価に関する。

## 【0002】

(発明の背景)

シグナル伝達は、細胞が細胞外シグナルに応答する生化学的現象のプロセスである。細胞外シグナルは、生化学カスケードを介して伝達される。このカスケードは、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の細胞膜受容体への結合によって始まり、細胞内標的分子の活性化により終了する。シグナル伝達のプロセスにより、細胞増殖、分化、および遺伝子の転写を含む様々な細胞機能が調節される。 40

## 【0003】

G タンパク質は、直接或いはその他の分子に関連してグアニンヌクレオチドに結合する。グアニンヌクレオチド結合タンパク質 ( G T P 結合タンパク質 ) は、シグナル伝達経路のクリティカルなメディエーターである。ホルモン、成長因子、神経調節物質、またはその他のシグナル伝達分子などの細胞外リガンドが膜貫通受容体と結合し、そのシグナルが細胞内シグナル伝達タンパク質によってエフェクター分子に伝達される。これらのシグナル 50

伝達タンパク質の多くは、細胞内シグナル伝達経路を調節するGTP結合タンパク質である。GTP結合タンパク質は、代謝、成長、分化、細胞骨格構築、及び細胞内小胞体輸送や分泌を含む多様な他の調節機能に關与する。結合がGDPからGTPに変わり、GTPがGDPに加水分解され、そのエネルギーによりGTP結合タンパク質がその高次構造を変え、他の細胞成分と相互作用する。GTP結合タンパク質は、3つの異なったサブユニットから成るヘテロ三量体型GTP結合タンパク質と、単量体型低分子量(LMW)GTP結合タンパク質の構造的に異なった2つのクラスに分類される。

#### 【0004】

ヘテロ三量体型GTP結合タンパク質は、細胞膜の内面において三量体として結合している、 $\alpha$ 、 $\beta$ 、および $\gamma$ の3つのサブユニット(不活性な高次構造)からなる。GはGDP若しくはGTPと結合し、GTPアーゼ活性を有する。複合体は、Gの受容体への結合を増進させる。Gは、Gの折り畳みおよび活性に必要である(Neer, E. J.ら(1994) Nature 371: 297-300.)。各サブユニットの複数の相同体が哺乳動物組織で同定され、サブユニットの様々な組み合わせが特定の機能および組織特異性を有する(Spiegel, A. M. (1997) J. Inher. Metab. Dis. 20: 113-121)。Gタンパク質活性は、脂質類似体、アミノ酸及びその誘導体、ペプチド、サイトカイン、並びに光、食味、及び匂いなどの特定の刺激に応答する7回膜貫通型細胞表面受容体(Gタンパク質共役受容体)によって引き起こされる。このような刺激により受容体が活性化され、Gタンパク質に結合しているGDPがGTPに置換される。G-GTPが受容体および複合体から解離し、分離した成分のそれぞれが、下流のエフェクターに結合しこれらのエフェクターを調節する。シグナル伝達は、Gが結合しているGTPをGDPに加水分解し、複合体と再会合すると停止する(Neer, 前出)。

#### 【0005】

Gタンパク質またはトランスデュシンとしても知られるヘテロ三量体型Gタンパク質のサブユニットは、調節機能を有する多くのタンパク質に見られる7つのタンデムリピートであるWDリピート配列モチーフを含む。WDリピートタンパク質は、タンパク質-タンパク質相互作用に關与する約40個のアミノ酸からなる低度に保存されたりリピートの4個から8個のコピーを有する。トランスデュシンタンパク質の変異および変異体発現は、様々な疾患に關係する。ヒト血小板活性化因子アセチルヒドラーゼのサブユニットであるLIS1における変異は、ミラーディーカー脳回欠損症(Miller-Dieker lissencephaly)を引き起こす。RACK1が活性化されたプロテインキナーゼCと結合し、RbAp48が網膜芽細胞種タンパク質と結合する。CstTFは、in vitro哺乳動物mRNA前駆体のポリアデニル化に必要であり、切断刺激因子(cleavage-stimulating factor)のサブユニットと会合する。ヒト宿主細胞の感染のためのHIV補助受容体として機能する膜内在性糖タンパク質であるCD4は、小胞体におけるHIVによりコードされるVpuにより分解される。ヒトTrCP分子のWDリピートは、CD4-Vpu複合体の形成を仲介し、CD4タンパク質分解を誘導する(Neer, E. J.ら(1994) Nature 371: 297-300およびMargottin, F.ら(1998) Mol. Cell 11: 565-574)。

#### 【0006】

GPCRシグナル伝達カスケードの異常により、白血球およびリンパ球の異常な活性が起こり、リウマチ様関節炎、胆汁性肝硬変、溶血性貧血、紅斑性狼瘡、及び甲状腺炎などの様々な炎症性の疾患および自己免疫疾患に見られる組織の障害および破壊が引き起こされ得る。脳組織、甲状腺組織、副腎組織、及び性腺組織の増殖のサイクリックAMP刺激を含む異常な細胞増殖はGタンパク質により調節される。Gサブユニットにおける変異は、成長ホルモン分泌脳下垂体成長ホルモン分泌細胞癌、機能亢進甲状腺腺腫、卵巣腫瘍、及び副腎腫瘍に見られる(Meij, J. T. A. (1996) Mol. Cell Biochem. 157: 31-38; Ausseil, C.ら(1988) J. 50

Immunol. 140:215-220)。

【0007】

LMW GTP結合タンパク質は、細胞成長、細胞周期制御、タンパク質分泌、および細胞内小胞相互作用を調節する。LMW GTP結合タンパク質は1つのポリペプチドからなり、ヘテロ三量体型GTP結合タンパク質のサブユニットのようにGTPに結合して加水分解し、不活性状態から活性状態に、或いは活性状態から不活性状態に移行する。LMW GTP結合タンパク質は、受容体からの細胞外シグナルに応答し、様々な細胞機能に關与する分裂促進信号を伝達してタンパク質を活性化させる。GTPの結合および加水分解は、LMW GTP結合タンパク質の応答を調節し、このプロセス中のエネルギー源として働く(Bokoch, G. M. および Der, C. J. (1993) FASEB J. 7:750-759)。

【0008】

LMW GTP結合タンパク質スーパーファミリーの少なくとも6つのメンバーが同定され、現在はras、rho、arf、sar1、ran、及びrabの6つのサブファミリーに分類されている。活性化されたras遺伝子が初めにヒト癌で見出され、その後の研究によりこのras機能が、細胞が成長を続けるか或いは分化するかの決定において重要な役割を果たすことが分かった。LMW Gタンパク質スーパーファミリーの他のメンバーは、活性化された遺伝子の機能及び活性化を開始するGTP結合タンパク質の位置で変化するシグナル伝達において役割を果たす。Rho GTP結合タンパク質は、正常な細胞成長および分化に必要な成長因子受容体をアクチン重合に導くシグナル伝達経路を調節する。このタンパク質のrab、arf、およびsar1ファミリーは、タンパク質局在化、タンパク質プロセッシング、および分泌のために小胞の膜への移動及び膜からの移動を調節する。Ran GTP結合タンパク質は細胞の核に局在しており、核タンパク質移入、DNA合成の調節、及び細胞周期の進行において重要な役割を果たす(Hall, A. (1990) Science 249:635-640; Barbacid, M. (1987) Ann. Rev. Biochem. 56:779-827; および Sasaki, T. および Takai, Y. (1998) Biochem. Biophys. Res. Commun. 245:641-645)。

【0009】

LMW GTP結合タンパク質は、GTP結合活性型からGDP結合不活性型およびGDP結合不活性型からGTP結合活性型に転換するGTPアーゼである。この転換は、GDPの解離、GTPの会合、またはGTP加水分解の速度に影響を与えるタンパク質によって調節される。GDPの会合に影響を与えるタンパク質は、グアニンヌクレオチド解離インヒビター、およびグアニンヌクレオチド交換因子(GEF)によって代表される。Drosophila Son-of-Sevenlessタンパク質の哺乳動物相同体が最もその特徴を有している。GTP加水分解に影響を与えるタンパク質は、GTPアーゼ活性化タンパク質(GAP)によって代表される。GEFおよびGAPの両方の活性は、細胞外刺激に応答して制御され、RalBP1およびPOB1などのアクセサリタンパク質によって調節される。GDP結合型は、グアニンヌクレオチド放出因子により促進されるGDP/GTP交換反応によりGTP結合型に変換される。GTP結合型は、内在性GTPアーゼ活性によりGDP結合型に変換される。この変換は、GAPにより加速される(Ikeda, M.ら(1998) J. Biol. Chem. 273:814-821; Quilliam, L. A. (1995) Bioessays 17:395-404)。GTPに結合するが加水分解できない変異Rasファミリータンパク質は、永久に活性化した状態であり、LMW GTP結合タンパク質を活性化するGEFと同様に細胞増殖や癌を引き起こす(Drivias, G. T.ら(1990) Mol. Cell. Biol. 10:1793-1798; および Whitehead, I. P.ら(1998) Mol Cell Biol. 18:4689-4697)。

【0010】

新規のGタンパク質関連分子、およびそれらをコードするポリヌクレオチドの発見により

、新規の組成物を提供することで当分野の要望に応えることができる。この新規の組成物は、細胞増殖異常及び自己免疫/炎症性の疾患の診断・治療・予防において有用であり、また、Gタンパク質関連分子の核酸配列及びアミノ酸配列の発現における外来性化合物の影響についての評価にも有用である。

**【0011】**

(発明の要約)

本発明は、総称して「GPAM」、個別にはそれぞれ「GPAM-1」および「GPAM-2」と呼ぶGタンパク質関連分子である精製されたポリペプチドを提供する。本発明の一実施態様では、(a)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-2からなる群から選択されたアミノ酸と少なくとも80%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-2とからなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された単離されたポリペプチドを提供する。別法では、SEQ ID NO: 1-2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。

10

**【0012】**

更に本発明は、(a)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-2からなる群から選択されたアミノ酸と少なくとも80%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された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を提供する。別法では、このポリヌクレオチドは、SEQ ID NO: 1-2からなる群から選択されたポリペプチドをコードする。更なる別法では、このポリヌクレオチドは、SEQ ID NO: 3-4からなる群から選択される。

20

**【0013】**

更に、本発明は、(a)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-2からなる群から選択されたアミノ酸と少なくとも80%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドをコードするポリヌクレオチドと機能的に結合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を提供する。別法では、本発明は、この組換えポリヌクレオチドで形質転換された細胞を提供する。更なる別法では、本発明は、この組換えポリヌクレオチドを含む遺伝子組換え生物を提供する。

30

**【0014】**

また、本発明は、(a)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-2からなる群から選択されたアミノ酸と少なくとも80%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-2からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドの生産方法を提供する。この方法は、(a)このポリペプチドの発現に好適な条件下で、このポリペプチドをコードするポリヌクレオチドと機能的に結合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ステップと、(b)このように発現したポリペプチドを回収するステップとを含む。

40

**【0015】**

50

更に、本発明は (a) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸と少なくとも 80% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドに特異的に結合する単離された抗体を提供する。

【0016】

更に、本発明は、(a) SEQ ID NO: 3 - 4 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ポリヌクレオチドと、(b) SEQ ID NO: 3 - 4 からなる群から  
10 選択されたポリヌクレオチド配列と少なくとも 80% の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c) 前記 (a) のポリヌクレオチドに相補的なポリヌクレオチドと、(d) 前記 (b) のポリヌクレオチドに相補的なポリヌクレオチドと、(e) 前記 (a) (d) の RNA 等価物とで構成され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。別法では、このポリヌクレオチドは、少なくとも 60 個の連続するヌクレオチドを含む。

【0017】

更に本発明は、(a) SEQ ID NO: 3 - 4 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ポリヌクレオチドと、(b) SEQ ID NO: 3 - 4 からなる群から  
20 選択されたポリヌクレオチド配列と少なくとも 80% の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c) 前記 (a) のポリヌクレオチドに相補的なポリヌクレオチドと、(d) 前記 (b) のポリヌクレオチドに相補的なポリヌクレオチドと、(e) 前記 (a) (d) の RNA 等価物とで構成される群から選択されたポリヌクレオチド配列を有するサンプルにおける標的ポリヌクレオチドを検出する方法を提供する。この方法は、(a) 前記サンプル内の標的ポリヌクレオチドと相補的な配列を含む少なくとも 20 個の連続するヌクレオチドを含むプローブと前記サンプルをハイブリダイズさせるステップであって、前記プローブと前記標的ポリヌクレオチドまたはその断片とでハイブリダイゼーション複合体が形成される条件下で、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、該ステップと、(b) 前記ハイブリダイゼーション複合体の存在するか否かを検出し、存在する場合には必要に応じてその収量を測定する  
30 ステップとを含む。別法では、前記プローブは、少なくとも 60 個の連続するヌクレオチドを含む。

【0018】

更に本発明は、(a) SEQ ID NO: 3 - 4 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ポリヌクレオチドと、(b) SEQ ID NO: 3 - 4 からなる群から  
40 選択されたポリヌクレオチド配列と少なくとも 80% の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c) 前記 (a) のポリヌクレオチドに相補的なポリヌクレオチドと、(d) 前記 (b) のポリヌクレオチドに相補的なポリヌクレオチドと、(e) 前記 (a) (d) の RNA 等価物とで構成される群から選択されたポリヌクレオチド配列を有するサンプルにおける標的ポリヌクレオチドを検出する方法を提供する。この方法は、(a) ポリメラーゼ連鎖反応増幅法を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、(b) 増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定する  
ステップとを含む。

【0019】

更に本発明は、(a) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 2 からなる群から選択されたア  
50 ミノ酸と少なくとも 80% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1 - 2 からなる群から選択さ

れたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された有効量のポリペプチド及び好適な医薬用賦形剤を含む組成物を提供する。一実施例では、SEQ ID NO: 1 - 2 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。本発明は更に、患者にこの組成物を投与することを含む、機能的G P A M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。

**【0020】**

更に本発明は、(a) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸と少なくとも80%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドのアゴニストとして有効な化合物をスクリーニングする方法を提供する。この方法は、(a)このポリペプチドを含むサンプルを化合物に曝露するステップと、(b)このサンプルのアゴニスト活性を検出するステップとを含む。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。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的G P A M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。

10

**【0021】**

更に、本発明は、(a) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸と少なくとも80%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドのアンタゴニストとして有効な化合物をスクリーニングする方法を提供する。この方法は、(a)このポリペプチドを含むサンプルを化合物に曝露するステップと、(b)このサンプルのアンタゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアンタ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。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的G P A Mの過剰な発現に関連した疾患やその症状の治療方法を提供する。

20

30

**【0022】**

更に本発明は、(a) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸と少なくとも80%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドに特異的に結合する化合物をスクリーニングする方法を提供する。この方法は、(a)このポリペプチドを好適な条件下で少なくとも1つの試験化合物と結合させるステップと、(b)このポリペプチドとこの試験化合物との結合を検出して、このポリペプチドと特異的に結合する化合物を同定するステップとを含む。

40

**【0023】**

更に本発明は、(a) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸と少なくとも80%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1 - 2 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される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。このスクリ

50

ーニング方法は、(a)このポリペプチドを、その活性が許容される条件下で少なくとも1つの化合物と結合させるステップと、(b)この試験化合物の存在下でのこのポリペプチドの活性を評価するステップと、(c)この試験化合物の存在下でのこのポリペプチドの活性と、この試験化合物の非存在下でのこのポリペプチドの活性とを比較するステップとを含み、この試験化合物の存在下でのこのポリペプチドの活性の変化が、このポリペプチドの活性を調節する化合物の存在を示唆する。

【0024】

更に本発明は、SEQ ID NO: 3-4からなる群から選択された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有効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、該スクリーニング方法を提供する。

10

【0025】

本発明はさらに、試験化合物の毒性を評価する方法を提供する。この方法は、(a)核酸を含む生体サンプルを前記試験化合物で処理するステップと、(b)処理した前記生体サンプルの核酸をプローブとハイブリダイズするステップと、(c)ハイブリダイゼーション複合体の収量を測定するステップと、(d)前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較するステップとを含み、前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆する。この方法における前記プローブが、(1)SEQ ID NO: 3-4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ポリヌクレオチドと、(2)SEQ ID NO: 3-4からなる群から選択されたポリヌクレオチド配列と少なくとも80%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(3)前記(1)に相補的な配列を有するポリヌクレオチドと、(4)前記(2)のポリヌクレオチドに相補的なポリヌクレオチドと、(5)前記(1)-(4)のRNA等価物とで構成される群から選択されたポリヌクレオチドの少なくとも20個の連続するヌクレオチドを含む。また、前記ハイブリダイゼーションは、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る。また、前記標的ポリヌクレオチドが、(1)SEQ ID NO: 3-4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ポリヌクレオチドと、(2)SEQ ID NO: 3-4からなる群から選択されたポリヌクレオチド配列と少なくとも80%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(3)前記(1)のポリヌクレオチドに相補的なポリヌクレオチドと、(4)前記(2)のポリヌクレオチドに相補的なポリヌクレオチドと、(5)前記(1)-(4)のRNA等価物とで構成される群から選択される。別法では、標的ポリヌクレオチドは、その標的ポリヌクレオチドの断片である。

20

30

【0026】

(本発明の記載について)

本発明のタンパク質及び核酸配列、方法について説明する前に、本発明は、ここに開示した特定の装置及び材料、方法に限定されず、その実施形態を変更できることを理解されたい。また、ここで用いられる用語は、特定の実施例のみを説明する目的で用いられたものであり、後述の請求の範囲によってのみ限定され、本発明の範囲を限定することを意図したものではないということも理解されたい。

40

【0027】

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」は複数の抗体は含まれ、当業者には周知の等価物なども含まれる。

【0028】

本明細書で用いた全ての科学技術用語は、別の方法で定義されていない限り、本発明の属

50

する技術分野の一般的な技術者が普通に解釈する意味と同じである。本明細書で記述したものと類似、或いは同等の全ての装置及び材料、方法は本発明の実施及びテストに使用できるが、好適な装置及び材料、方法をここに記す。本明細書に記載の全ての文献は、本発明に関連して使用する可能性のある文献に記載された細胞系、プロトコル、試薬、ベクターを記述し開示するために引用した。従来の特許を引用したからと言って、本発明の新規性が損なわれると解釈されるものではない。

【0029】

(定義)

用語「GPAM」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GPAMのアミノ酸配列を指す。

10

【0030】

用語「アゴニスト」は、GPAMの生物学的活性を強める、或いは模倣する分子を指す。このアゴニストは、GPAMに直接相互作用するか、或いはGPAMが関与する生物学的経路の成分と作用して、GPAMの活性を調節するタンパク質、核酸、糖質、小分子、任意の他の化合物や組成物を含み得る。

【0031】

用語「アレル変異配列」は、GPAMをコードする遺伝子の別の形を指す。アレル変異配列は、核酸配列における少なくとも1つの変異によって生じ、変異mRNA若しくは変異ポリペプチドになり、これらの構造や機能は変わる場合もあれば変わらない場合もある。ある遺伝子は、天然型のアレル変異配列が存在しないもの、1つ或いは多数存在するものがある。一般にアレル変異配列を生じる変異は、ヌクレオチドの自然な欠失、付加、或いは置換による。これらの各変異は、単独或いは他の変異と同時に起こり、所定の配列内で一回或いはそれ以上生じる。

20

【0032】

GPAM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、GPAMと同じポリペプチド或いはGPAM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。この定義には、GPAMをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置でのアレル変異配列との不相当或いは予期しないハイブリダイゼーション、並びにGPAM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な或いは検出困難な多形性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じGPAM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。意図的なアミノ酸置換は、生物学的或いは免疫学的にGPAMの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び/または両親媒性についての類似性に基づいて成され得る。例えば、負に荷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含まれ、正に荷電したアミノ酸にはリシン及びアルギニンが含まれ得る。類似の親水性の値をもち極性非荷電側鎖を有するアミノ酸には、アスパラギン、グルタミン、セリン、トレオニンが含まれ得る。類似の親水性の値をもち非荷電側鎖を有するアミノ酸には、ロイシン、イソロイシン、バリン、グリシン、アラニン、フェニルアラニン及びチロシンが含まれ得る。

30

40

【0033】

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、アミノ酸配列に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。

【0034】

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。一般に増幅は、この技術分野で周知のポリメラーゼ連鎖反応(PCR)技術によって行われる。

【0035】

50

用語「アンタゴニスト」は、G P A Mの生物学的活性を阻害或いは減弱する分子である。アンタゴニストは、G P A Mに直接相互作用するか、或いはG P A Mが関与する生物学的経路の成分と作用して、G P A Mの活性を調節する抗体、核酸、糖質、小分子、任意の他の化合物や組成物などのタンパク質を含み得る。

**【0036】**

用語「抗体」は、抗原決定基と結合可能なF a b及びF ( a b ' )<sub>2</sub>、及びそれらの断片、F v断片などの無傷の分子を指す。G P A Mポリペプチドと結合する抗体は、抗体を免疫する小ペプチドを含む無傷の分子またはその断片を用いて作製可能である。動物（例えば、マウス、ラット、若しくはウサギ）を免疫化するのに使用されるポリペプチド或いはオリゴペプチドは、R N Aの翻訳から、或いは化学的に合成可能であり、必要に応じて担体タンパク質と結合させることも可能である。ペプチドと化学的に結合した一般に用いられる担体は、ウシ血清アルブミン、チログロブリン、及びキーホールリンペットヘモニアン（K L H）を含む。次に、この結合したペプチドを用いて動物を免疫化する。

10

**【0037】**

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。タンパク質或いはタンパク質の断片が、宿主動物を免疫化するのに用いられるとき、このタンパク質の種々の領域は、抗原決定基（タンパク質上の特定の領域或いは三次元構造体）に特異的に結合する抗体の産生を誘発し得る。抗原決定基は、抗体と結合するために無傷の抗原（即ち、免疫応答を引き出すために用いられる免疫原）と競合し得る。

**【0038】**

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、D N Aと、R N Aと、ペプチド核酸（P N A）と、ホスホロチオネートやメチルホスホネート、ベンジルホスホネート（benzyl phosphonate）などの修飾された骨格（backbone linkage）を有するオリゴヌクレオチドと、2'-メトキシエチル糖または2'-メトキシエトキシ糖などの修飾された糖を有するオリゴヌクレオチドと、5-メチルシトシンまたは2'-deoxyuracil、7-deaza-2'-deoxyguanosineなどの修飾された塩基を有するオリゴヌクレオチドを含み得る。アンチセンス分子は、化学合成や転写を含む任意の方法で作ることができる。相補的アンチセンス分子は、一度細胞に導入されると、細胞によって作られた天然の核酸配列と塩基対となって二重鎖を形成し、転写や翻訳を阻害する。「負」または「マイナス」という表現はアンチセンス鎖であり、「正」または「プラス」という表現はセンス鎖である。

20

30

**【0039】**

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G P A M、合成のG P A M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。

**【0040】**

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'」が相補的な配列「3' T C A 5'」と対をなす。

40

**【0041】**

「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。この組成物は、乾燥した製剤或いは水溶液を含み得る。G P A M若しくはG P A M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、プローブは、塩（例えば、NaCl）及び界面活性剤（例えば、SDS：ドデシル硫酸ナトリウム）、その他の物質（例えば、デンハート液、乾燥ミルク、サケ精子D N Aなど）を含む水溶液に展開され得る。

50

## 【0042】

「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。伸長及び重複の両方によって構築されるコンセンサス配列もある。

## 【0043】

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。以下に、あるタンパク質の元のアミノ酸が別のアミノ酸に置換される保存的なアミノ酸置換を示す。

元の残基

保存的な置換

Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
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
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

一般に、保存されたアミノ酸置換の場合は、a)置換された領域のポリペプチドの骨格構造、例えば、シートやヘリックス高次構造、b)置換された部位の分子の電荷または疎水性、及び/または、c)側鎖の大半が維持される。

## 【0044】

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。

## 【0045】

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。ポリヌクレオチド配列の化学修飾には、例えば、アルキル基、アシル基、ヒドロキシル基、或いはアミノ基による水素の置換がある。誘導体ポリヌクレオチドは、自然分子(未修飾の分子)の生物学的或いは免疫学的機能の少なくとも1つを維持するポリペプチドをコードする。誘導体ポリペプチドとは、もとのポリペプチドの生物学的機能、或いは免疫学的機能の少なくとも1つを維持する、グリコシル化、ポリエチレングリコール化、或いは任意の同様のプロセスによって修飾されたポリペプチドのことである。

## 【0046】

「検出可能な標識」は、測定可能な信号を生成し得る、ポリヌクレオチドやポリペプチド

10

20

30

40

50

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。

【0047】

用語「断片」は、GPAMまたはGPAM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」の最大の長さは、親配列から1つのヌクレオチド/アミノ酸残基を差し引いた長さである。例えば、ある断片は、5~1000個の連続するヌクレオチド或いはアミノ酸残基を含む。プローブ、プライマー、抗原、治療用分子、またはその他の目的に用いる断片は、少なくとも5、10、15、16、20、25、30、40、50、60、75、100、150、250若しくは500個の連続するヌクレオチド或いはアミノ酸残基の長さである。断片は、優先的に分子の特定の領域から選択される場合もある。例えば、ポリペプチド断片は、所定の配列に示された最初の250若しくは500のアミノ酸 (或いは、ポリペプチドの最初の25%または50%) から選択された連続するアミノ酸の所定の長さを含み得る。これらの長さは一例であり、配列表及び表、図面を含む明細書に記載の任意の長さが、本発明の実施例に含まれ得る。

10

【0048】

SEQ ID NO: 3-4の断片は、例えば、この断片を得たゲノム内の他の配列とは異なる、SEQ ID NO: 3-4を明確に同定する固有のポリヌクレオチド配列の領域を含む。SEQ ID NO: 3-4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: 3-4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。ある断片と一致するSEQ ID NO: 3-4の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。

20

【0049】

SEQ ID NO: 1-2のある断片は、SEQ ID NO: 3-4のある断片によってコードされる。SEQ ID NO: 1-2のある断片は、SEQ ID NO: 1-2を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID NO: 1-2のある断片は、SEQ ID NO: 1-2を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。ある断片と一致するSEQ ID NO: 1-2の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に決定できる。

30

【0050】

「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)、それに続く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。

【0051】

「相同性」は、2つ以上のポリヌクレオチド配列間または2つ以上のポリペプチド配列間の配列類似性である。この配列類似性は配列同一性と言い換えることができる。

【0052】

ポリヌクレオチド配列についての用語「パーセントの同一性」または「%の同一性」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、標準化され再現できる方法で、2つの配列間のアラインメントを最適化するべく、配列にギャップを挿入して、より意味をもつ2つの配列間の比較を行うことができる。

40

【0053】

ポリヌクレオチド配列間の同一性のパーセントは、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。このプログラムはLASERGENEソフトウェアパッケージの一部であり、分子生物学分析プログラム一式 (DNASTAR, Madison WI) 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

50

る。ポリヌクレオチド配列の対のアライメントの場合、デフォルトパラメータは、`K t u p l e = 2`、`g a p p e n a l t y = 5`、`w i n d o w = 4`、「`d i a g o n a l s s a v e d`」= 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。同一性のパーセントは、アラインメントされたポリヌクレオチド配列間の「類似性のパーセント」としてCLUSTAL Vによって報告される。

【0054】

別法では、National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. 他 (1990) J. Mol. Biol. 215: 403 - 410) が提供する、広く用いられている無料の配列比較アルゴリズム一式が、NCBI (Bethesda, MD) を含む幾つかのソース及びインターネット (<http://WWW.ncbi.nlm.nih.gov/BLAST/>) で入手可能である。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、デフォルトを設定するギャップ及び他のパラメータと共に用いられる。例えば、2つのヌクレオチド配列を比較する場合、ある者は、デフォルトパラメータに設定された「BLAST 2 Sequences」ツール Version 2.0.12 (April - 21 - 2000) でblastnを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

【0055】

Matrix : B L O S U M 6 2  
 Reward for match : 1  
 Penalty for mismatch : - 2  
 Open Gap : 5 及び Extension Gap : 2 penalties  
 Gap x drop-off : 5 0  
 Expect : 1 0  
 Word Size : 1 1  
 Filter : on

同一性のパーセントは、例えば、特定の配列番号で決められた、所定の配列の全長に対して測定してもよいし、それより短い長さに対して、例えば、ある大きな所定の配列から得られた断片、例えば、連続する少なくとも、20または30、40、50、70、100、200のヌクレオチドの断片の長さに対して測定してもよい。このような長さは単なる例であり、配列表及び表、図面を含む明細書に記載の配列の任意の長さの断片を用いて、同一性のパーセントが測定される長さを示すことができる。

【0056】

高い同一性を示さない核酸配列でも、遺伝子コードの縮重によって類似のアミノ酸配列をコードし得る。縮重を利用して核酸配列を変え、それぞれが実質的に同じタンパク質をコードする様々な核酸配列を作製できることを理解されたい。

【0057】

ポリペプチド配列に用いられる用語「パーセントの同一性」または「%の同一性」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周知である。アラインメント方法の中には、保存的なアミノ酸置換を考慮したものもある。詳細に上述したこのような保存的な置換は、一般に、置換部位の電荷や疎水性が保存され、ポ

リペプチドの構造（従って機能も）が保存される。

【0058】

ポリペプチド配列間の同一性のパーセントは、MEGALIGNバージョン3.12e配列アラインメントプログラム（上記）に組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。CLUSTAL Vを用いる対方式のポリペプチド配列のアラインメントの場合、デフォルトパラメータは、Ktuple = 1、gap penalty = 3、window = 5、及び「diagonals saved」= 5と設定する。PAM250マトリクスが、デフォルトの残基重み付け表として選択される。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、「類似性のパーセント」としてCLUSTAL Vによ

10

【0059】

別法では、NCBI BLASTソフトウェア式が用いられる。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。そのようなデフォルトパラメータは、例えば、以下のようにする。

【0060】

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

同一性のパーセントは、例えば、特定の配列番号で決められた、所定のポリペプチド配列の全長に対して測定してもよいし、それより短い長さに対して、例えば、ある大きな所定のポリペプチド配列から得られた断片、例えば、連続する少なくとも15、20または30、40、50、70、150の残基の断片の長さに対して測定してもよい。このような長さは単なる例であり、配列表及び表、図面を含む明細書に記載の配列の任意の長さの断片を用いて、同一性のパーセントが測定される長さを示すことができる。

30

【0061】

「ヒト人工染色体（HAC）」は、約6kb（キロベース）～10MbのサイズのDNA配列を含み得る、安定した有糸分裂染色体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。

【0062】

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。

【0063】

「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。特異的なハイブリダイゼーションとは、2つの核酸配列が高い相同性を有することを意味する。アニーリングが許容される条件下で、特異的なハイブリダイゼーション複合体が形成され、洗浄過程の後もハイブリダイズしたままである。洗浄過程は、ハイブリダイゼーションプロセスの厳密性即ちストリンジェンシー（stringency）の決定において特に重要であり、よりストリンジェントな条件では、非特異的な結合、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。核酸配列間のアニーリングが許容される条件は、当業者によって日常的に決定され、ハイブリダイゼーションの間は一定であるが、洗浄過程は、目的のストリンジェンシーにするためにその最中に条件の変更が可能であり、ハイブリダイゼーション特異性が得られる。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100µg/m

40

50

1のせん断して変性したサケ精子DNAが含まれる。

【0064】

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、洗浄過程を行う際の温度によっても左右される。この洗浄温度は通常、所定のイオン強度とpHにおける特定の配列の熱融点( $T_m$ )より約5~20 低く選択される。この $T_m$ は、(所定のイオン強度とpHの下)標的の配列の50%が完全に一致するプローブとハイブリダイズする温度である。 $T_m$ を計算する式及び核酸のハイブリダイゼーションの条件は、周知であり、Sambrook, J. 他による, 1989, Molecular Cloning: A Laboratory Manual, 第2版の1-3巻, Cold Spring Harbor Press, Plainview NY; 特に2巻の9章に記載されている。 10

【0065】

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0.2x 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の濃度は、約0.1%のSDSが存在の下、約0.1~2x SSCの範囲である。通常は、ブロッキング試薬を用いて非特異的な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試薬には、例えば、約100~200  $\mu\text{g}/\text{ml}$ の切断され変性したサケ精子DNAが含まれる。約35~50% v/vの濃度のホルムアミドなどの有機溶剤が、例えば、RNAとDNAのハイブリダイゼーションなどの特定の場合に用いることができる。これらの洗浄条件の有用な改変は、当業者には周知である。特に高いストリンジェントな条件でのハイブリダイゼーションは、ヌクレオチド間の進化における類似性を示唆し得る。このような類似性は、それらのヌクレオチド及びコードされたポリペプチドが類似の役割を果たしていることを強く示唆する。 20

【0066】

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液中(例えば、C<sub>0</sub>tまたはR<sub>0</sub>t分析)で形成されるか、或いは溶液中の1つの核酸配列と固体の支持物(例えば、紙、膜、フィルタ、チップ、ピン、或いはスライドガラス、または細胞及びその核酸を固定する任意の適当な基板)に固定されたもう一つの核酸配列とで形成され得る。 30

【0067】

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。

【0068】

「免疫応答」は、炎症性疾患及び外傷、免疫異常、感染症、遺伝病などに関連する症状を指す。これらの症状は、細胞系及び全身防衛系に影響を及ぼすサイトカイン及びケモカイン、別の情報伝達分子などの様々な因子の発現という特徴をもつ。

【0069】

用語「免疫原性断片」は、例えば哺乳動物などの生きている動物に導入すると、免疫反応を引き起こすGPAMのポリペプチド断片またはオリゴペプチド断片を指す。用語「免疫原性断片」はまた、本明細書で開示するまたは当分野で周知のあらゆる抗体生産方法に有用なGPAMのポリペプチド断片またはオリゴペプチド断片を含む。 40

【0070】

用語「マイクロアレイ」は、基質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。

【0071】

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。

【0072】

用語「調節」は、GPAMの活性の変化を指す。例えば、調節によって、GPAMのタン 50

パク質活性、或いは結合特性、またはその他の生物学的特性、機能的特性或いは免疫学的特性の変化が起こる。

【0073】

用語「核酸」及び「核酸配列」は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチド、或いはそれらの断片を指し、一本鎖若しくは二本鎖であって、センス鎖或いはアンチセンス鎖であるゲノム起源若しくは合成起源のDNA或いはRNA、ペプチド核酸(PNA)、任意のDNA様物質、及びRNA様物質である。

【0074】

「機能的に結合した」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与える場合、そのプロモーターはそのコード配列に機能的に結合している。一般に、機能的に結合したDNA配列は、同じ読み枠内で2つのタンパク質をコードする領域が結合する必要がある場合は、非常に近接或いは連続する。

10

【0075】

「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骨格に結合した、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。この末端のリシンにより、この組成物が溶解性となる。PNAは、相補的な一本鎖DNAやRNAに優先的に結合して転写物の伸長を止め、ポリエチレングリコール化して細胞における寿命を延ばし得る。

【0076】

GPAM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を含まれ得る。これらのプロセスは、合成或いは生化学的に生じ得る。生化学的修飾は、GPAMの酵素環境に依存し、細胞の種類によって異なり得る。

20

【0077】

「プローブ」とは、同一配列或いはアレル核酸配列、関連する核酸配列の検出に用いる、GPAM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、検出可能な標識またはレポーター分子が結合され単離されたオリゴヌクレオチドやポリヌクレオチドである。典型的な標識には、放射性アイソトープ及びリガンド、化学発光試薬、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。プライマーがポリヌクレオチドにアニーリングした後、あるDNAポリメラーゼ酵素によって、標的のDNA一本鎖に沿って伸長される。プライマーの組は、例えば、PCR法における核酸配列の増幅(及び同定)に用いることができる。

30

【0078】

本発明に用いられるプローブ及びプライマーは、既知の配列の少なくとも15の連続するヌクレオチドを含む。特異性を高めるために、より長いプローブ及びプライマーが用いることも可能である。例えば、開示した核酸配列の連続する少なくとも20または25、30、40、50、60、70、80、90、100、150のヌクレオチドを含む。プローブ及びプライマーは、上記した例より相当長いものも用いることができ、本明細書の表及び図面、配列表に示された任意の長さのヌクレオチドも用いることができることを理解されたい。

40

【0079】

プローブ及びプライマーの準備及び使用方法については、例えば、Sambrook, J. 他による、1989年、名称「Molecular Cloning: A Laboratory Manual」、第2版の1-3巻(Cold Spring Harbor Press, Plainview NY)、またはAusubel, F.M. 他による、1987年、名称「Current Protocols in Molecular Biology」(Greene Publi. Assoc. & Wiley-Intersciences, New York NY)、並びにInnis他による、199

50

0年、名称「PCR Protocols, A Guide to Methods and Applications」(Academic Press, San Diego CA.)を参照されたい。PCR用のプライマーの組は、例えば、Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA)などのそのような目的のためのコンピュータプログラムを用いて、ある既知の配列から引き出すことができる。

【0080】

プライマーとして用いるオリゴヌクレオチドは、当分野で周知のプライマー選択用のコンピュータプログラムで選択される。例えば、OLIGO 4.06ソフトウェアは、それぞれが最大100ヌクレオチドまでのPCR用のプライマーの対の選択、及び32,000塩基までの入力ポリヌクレオチド配列から最大5,000ヌクレオチドまでの大きなポリヌクレオチド及びオリゴヌクレオチドの分析に有用である。類似のプライマー選択用プログラムには、能力を拡大する追加の機能が含まれている。例えば、PrimOUプライマー選択プログラム(Genome Center at University of Texas South West Medical Center, Dallas TXより入手可能)は、メガベース配列から特定のプライマーを選択できるため、ゲノムワイドスコープ(genome-wide scope)におけるプライマーの設計に有用である。Primer3プライマー選択プログラム(Whitehead Institute/MIT Center for Genome Research, Cambridge MA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(mispriming library)」を入力できる。また、Primer3は、特にマイクロアレイのオリゴヌクレオチドの選択に有用である(後の方の2つのプライマー選択プログラムのソースコードは、それぞれのソースから得ることができ、ユーザーのニーズを満たすように変更することもできる)。PrimerGenプログラム(UK Human Genome Mapping Project Resource Centre, Cambridge UKより入手可能)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するため、アラインメントされた核酸配列の最も保存された領域或いは最も保存されていない領域のどちらかとハイブリダイズするプライマーを選択することができる。従って、このプログラムは、固有及び保存されたオリゴヌクレオチドやポリヌクレオチドの断片の同定に有用である。上記した任意の選択方法で同定されたオリゴヌクレオチドやポリヌクレオチドの断片は、例えば、PCR法やシーケンシングプライマー、マイクロアレイエレメント、或いはサンプルの核酸の完全或いは部分的に相補的なポリヌクレオチドを同定する特定のプローブなどの、ハイブリダイゼーション技術に有用である。オリゴヌクレオチドの選択方法は、上記した方法に制限されるものではない。

【0081】

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。この人工の組み合わせは、化学合成によって作られる場合も多いが、前出のSambrookに記載されたような遺伝子工学の技術を用いて核酸の離れたセグメントを人工的に操作する方がより一般的である。この「組換え核酸」には、単に核酸の一部の追加または置換、欠失によって変更された核酸も含む。組換え核酸は、あるプロモーター配列に機能的に結合した核酸配列を含む場合もある。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のに用いられるベクターの一部であり得る。

【0082】

別法では、このような組換え核酸は、この組換え核酸を発現する哺乳動物のワクチン接種に用いると、その哺乳動物の防衛的な免疫応答を誘発する、ワクシニアウイルスに基づいたウイルスベクターの一部であり得る。

【0083】

「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

10

20

30

40

50

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(UTR)を含む。調節エレメントは、転写や翻訳、またはRNAの安定性を調節する宿主またはウイルスタンパク質と相互作用する。

【0084】

「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学的部分である。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、インヒビター、磁気粒子及びその他の当分野で既知の成分が含まれる。

【0085】

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、糖鎖の背骨がデオキシリボースではなくリボースからなる。

10

【0086】

用語「サンプル」は、その最も広い意味で用いられている。GPAM、GPAMをコードする核酸、またはその断片を含むと推定されるサンプルは、体液と、細胞からの抽出物や細胞から単離された染色体や細胞内小器官、膜と、細胞と、溶液中に存在するまたは基板に固定されたゲノムDNA、RNA、cDNAと、組織と、組織プリント等を含み得る。

【0087】

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、結合する分子によって認識される、例えば、抗原決定基つまりエピトープなどのタンパク質の特定の構造の存在によって左右される。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

【0088】

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。

30

【0089】

「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。

【0090】

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、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。この基板には、壁または塹壕、ピン、チャンネル、細孔などの様々な表面形態があり、そこにポリヌクレオチドやポリペプチドが結合する。

【0091】

「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。

40

【0092】

「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、当分野で周知の種々の方法により、自然或いは人工の条件下で起こり、原核宿主細胞若しくは真核宿主細胞の中に外来核酸配列を挿入する任意の周知の方法によって行うことができる。この形質転換の方法は、形質転換される宿主細胞のタイプによって選択される。この方法には、バクテリオファージまたはウイルス感染、電気穿孔法(エレクトロポレーション)、リポフェクション、及び微粒照射が含まれるが、これらに限定されるものではない。「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

50

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。

#### 【0093】

本明細書における「遺伝子組換え生物」とは、当分野で周知の遺伝子組換え技術などを用いて、人間が生物の1つ以上の細胞に異種の核酸を導入した任意の生物であり、動物及び植物を含むが、それらに限定されるものではない。微量注入や組換えウイルスに感染させるなどの慎重な遺伝子操作によって、細胞の前駆体に直接或いは間接的に異種核酸を細胞に導入する。「遺伝子操作」とは、典型的な交雑育種や*in vitro*での受精ではなく、組換えDNA分子を導入することである。本発明に従った遺伝子組換え生物には、細菌及びラン藻類、菌類、植物、動物が含まれる。本発明の単離されたDNAは、当分野で周知の、例えば、感染、形質移入、形質転換、トランス接合(*trans conjugation*)などの方法によって、宿主に導入することができる。本発明のDNAをそのような生物に導入する技術は周知であり、前出のSambrook他(1989)に記載されている。

10

#### 【0094】

特定の核酸配列の「変異配列」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツールVersion 2.0.9 (May-07-1999)を用いるblastnによって、ある核酸配列のある長さに対する該特定の核酸配列の同一性が、少なくとも40%と決定された核酸配列のことである。このような核酸の対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、85%、90%、95%、98%、或いはそれ以上の同一性を示し得る。ある変異配列は、例えば、「アレル」変異配列(上述)または「スプライス」変異配列、「種」変異配列、「多型」変異配列と表すことができる。スプライス変異配列は基準分子と同一性が極めて高い可能性があるが、mRNAプロセッシング中のエキソンの択一的スプライシングによってポリヌクレオチドの数が多くなったり、少なくなったりする。対応するポリペプチドは、基準分子に存在する追加の機能ドメインを有したり、基準分子に存在するドメインが欠落したりし得る。種変異配列は、種によって異なるポリヌクレオチド配列である。得られるポリペプチドは、互いに高いアミノ酸同一性を有する。多型変異配列は、所定の種と種における特定の遺伝子のポリヌクレオチド配列が異なる。多型変異配列はまた、ポリヌクレオチド配列の1つのヌクレオチドが異なる「1ヌクレオチド多型」(SNP)も含み得る。SNPの存在は、例えば、或る集団、病態、病態の性向を示唆し得る。

20

30

#### 【0095】

特定のポリペプチド配列の「変異体」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツールVersion 2.0.9 (May-07-1999)を用いるblastpによって、あるポリペプチド配列のある長さに対する該特定のポリペプチド配列の同一性が、少なくとも40%と決定されたポリペプチド配列のことである。このようなポリペプチドの対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、85%、90%、95%、98%、或いはそれ以上の同一性を示し得る。

40

#### 【0096】

(発明)

本発明は、新規のヒトGタンパク質関連分子(GPAM)及びGPAMをコードするポリヌクレオチドの発見に基づき、これらの組成物を利用した細胞増殖異常及び自己免疫/炎症性の疾患の診断、治療、及び予防に関する。

#### 【0097】

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識別番号を示す。各ポリヌクレオチドおよびそれに対応するポリペプチドは、1つのインサイトプロジェクト識別番号(Incyte Project ID)に相関する。各ポリペプチド配列は、記載されているようにポリペプチド配列識別番号(Polypeptide SEQ ID

50

NO : ) およびインサイトポリペプチド配列番号 ( Incyte Polypeptide ID ) の両方によって示されている。各ポリヌクレオチド配列は、記載されているようにポリヌクレオチド配列識別番号 ( Polypeptide SEQ ID NO : ) およびインサイトポリヌクレオチドコンセンサス配列番号 ( Incyte Polypeptide ID ) の両方によって示されている。

【0098】

表2は、GenBankタンパク質 ( genept ) データベースにおいてBLAST解析により同定された本発明のポリペプチドに相同性を有する配列を示す。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号 ( Polypeptide SEQ ID NO : ) およびそれに対応するインサイトポリペプチド配列番号 ( Incyte Polypeptide 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。これらを引用することを以って本明細書の一部とする。

10

【0099】

表3は、本発明の各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インサイトポリペプチド配列番号 ( Incyte Polypeptide 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、潜在的なリン酸化部位および潜在的なグリコシル化部位を示す。列6は、シグネチャ ( signature ) 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

【0100】

表2および表3は共に、本発明の各ポリペプチドの特性を一覧にしたものであって、これらの特性が請求するポリペプチドがGタンパク質関連分子であることを立証する。例えば、SEQ ID NO : 1は残基M176から残基I914において、マウスオンコジーン etc 2 ( GenBank ID g293332 ) と93%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を参照 ) 。BLASTの確率スコアは0.0である。これは、探しているポリペプチド配列アラインメントが偶然の一致により得られる確率を示す。SEQ ID NO : 1はまた、RhoGEF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を参照 ) 。MOTIFS、更にBLAST分析から得られたデータによって、SEQ ID NO : 1がグアニンヌクレオチド交換因子であることが裏付けられた。SEQ ID NO : 2も同様の方法で解析してアノテーションを付けた。SEQ ID NO : 1 - 2の解析のためのアルゴリズムおよびパラメータを表7に記載する。

30

40

【0101】

表4に示されているように、本発明の完全長ポリヌクレオチド配列は、cDNA配列、またはゲノムDNA由来のコード ( エキソン ) 配列、或いはこれらの2種類の配列のあらゆる組み合わせを用いて組み立てた。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 ( Polynucleotide SEQ ID NO : ) およびそれに対応するインサイトポリヌクレオチドコンセンサス配列番号 ( Incyte Polynucleotide ID ) 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 : 3 - 4を同定するため、或いはSEQ ID NO : 3 - 4と関連するポリヌクレオチド配列とを区別するた

50

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から推定されるコード配列(エキソン)、および/またはcDNAおよびゲノムDNAの両方からなる群に対応する識別番号を示す。これらの配列を用いて本発明の完全長ポリヌクレオチド配列を組み立てた。表4の列6および列7はそれぞれ、それぞれの完全長配列に対する列5のcDNA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。

#### 【0102】

表4の列5に示されている識別番号は、具体的には、例えばインサイトcDNAおよびそれらに対応するcDNAライブラリの識別番号を示す。例えば、1272429F1はインサイトcDNA配列の識別番号であり、TESTTUT02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インサイトcDNAは、プールされているcDNAライブラリ(例えば、SCIA00029V1)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の識別番号の場合もある。または、列5の識別番号は、Genscan分析によって推定されるゲノムDNAのコード領域の場合もある。このGenscan推定コード配列は、配列を組み立てる前に編集する場合がある(実施例4を参照)。または、列5の識別番号は、“exon-stitching”アルゴリズムによってcDNAおよびGenscan推定エキソンの両方からなる群の場合もある(実施例5を参照)。または、列5の識別番号は、“exon-stretching”アルゴリズムによってcDNAおよびGenscan推定エキソンの両方からなる群の場合もある(実施例5を参照)。場合によっては、列5に示されている配列の範囲と重複するインサイトcDNAの範囲が得られ、最終的なコンセンサス配列が決定されるが、それに相当するインサイトcDNAの識別番号は示されていない。

#### 【0103】

表5は、インサイトcDNA配列を用いて組み立てられたこれらの完全長ポリヌクレオチド配列が由来する代表的なcDNAライブラリを示す。代表的なcDNAライブラリとは、上記ポリヌクレオチド配列の組み立ておよび決定に用いられたインサイトcDNA配列を最も多く含むインサイトcDNAライブラリのことである。表5に示されているcDNAライブラリを作製するために用いた組織およびベクターが表6に示されている。

#### 【0104】

本発明はまた、GPAMの変異体も含む。好適なGPAMの変異体は、GPAMの機能的或いは構造的特徴の少なくともどちらか一方を有し、かつGPAMアミノ酸配列に対して少なくとも約80%のアミノ酸配列同一性、或いは少なくとも約90%のアミノ酸配列同一性、更には少なくとも約95%のアミノ酸配列同一性を有する。

#### 【0105】

本発明はまた、GPAMをコードするポリヌクレオチドを提供する。特定の実施例において、本発明は、GPAMをコードするSEQ ID NO: 3-4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を提供する。配列表に示したSEQ ID NO: 3-4のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖の背骨がデオキシリボースではなくリボースからなる等価RNA配列を含む。

#### 【0106】

本発明はまた、GPAMをコードするポリヌクレオチド配列の変異配列を含む。詳細には、このようなポリヌクレオチド配列の変異配列は、GPAMをコードするポリヌクレオチド配列と少なくとも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。本発明の特定の実施形態は、SEQ ID NO: 3-4からなる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SEQ ID NO: 3-4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、GPAM

10

20

30

40

50

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。

【0107】

遺伝暗号の縮重により作り出され得るG P A Mをコードする種々のポリヌクレオチド配列には、既知の自然発生する任意の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、可能なコドン選択に基づいた組み合わせの選択によって作り出され得る可能なポリヌクレオチド配列の変異の全てが含まれ得る。これらの組み合わせは、天然のG P A Mのポリヌクレオチド配列に適用される標準的なトリプレット遺伝暗号を基に作られ、全ての変異が明確に開示されていると考慮する。

【0108】

G P A Mをコードするヌクレオチド配列及びその変異配列は一般に、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 P A Mのヌクレオチド配列とハイブリダイズ可能であるが、非天然のコドンを含めるなどの実質的に異なった使い方のコドンを有するG P A M或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。特定のコドンが宿主によって利用される頻度に基づいてコドンを選択して、ペプチドの発現が特定の真核細胞または原核宿主に発生する割合を高め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いで、G P A M及びその誘導体をコードするヌクレオチド配列を実質的に変更する別の理由は、天然の配列から作られる転写物より例えば長い半減期など好ましい特性を備えるRNA転写物を作ることにある。

【0109】

本発明はまた、G P A M及びその誘導体をコードするDNA配列またはそれらの断片を完全に合成化学によって作り出すことも含む。作製後にこの合成配列を、当分野で良く知られた試薬を用いて、種々の入手可能な発現ベクター及び細胞系の何れの中にも挿入可能である。更に、合成化学を用いて、G P A Mまたはその任意の断片をコードする配列の中に突然変異を導入することも可能である。

【0110】

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、SEQ ID NO: 3 - 4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、Wahl, G. M. 及びS. L. Berger (1987) *Methods Enzymol.* 152: 399 - 407; and Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507 - 511. 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。

【0111】

当分野で周知のDNAのシーケンシング方法を用いて、本発明の何れの実施例も実行可能である。この方法には、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech Piscataway NJ)、或いはELONGASE増幅システム (Life Technologies, Gaithersburg MD) にみられるような校正エキソヌクレアーゼとポリメラーゼとの組み合わせなどの酵素が用いられる。好ましくは、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200 Thermal Cycler 200 (MJ Research, Watertown MA) 及びABI CATALYST 800 (PE Biosystems) など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 (PE Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野で周知の他の方法を用いてシーケンシングを行う。得られた配列を当分野で周知の様々なアルゴリズムを用いて分析する(例えば、Ausubel, F. M. (1997) Short

10

20

30

40

50

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. を参照)。

【0112】

当分野で周知のPCR法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、GPAMをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば制限部位PCR法を利用する1つの方法では、一般的なプライマー及び入れ子プライマー(nested primer)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(例えば、Sarkar, G. (1993) PCR Methods Applic 2: 318 - 322を参照)。逆PCR法を用いる別法では、広範な方向に伸長して環状化した鋳型から未知の配列を増幅するプライマーを用いる。この鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限断片に由来する(例えば、Triglia, T.ら(1988) Nucleic Acids Res 16: 8186を参照)。キャプチャPCR法を用いる第3の方法は、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断片のPCR増幅を含む(例えば、Lagerstrom, M.他(1991) PCR Methods Applic 1: 111 - 119を参照)。この方法では、多数の制限酵素による消化及びライゲーションを用いて、PCRを行う前に未知の配列の領域の中に組換え二本鎖配列を挿入することが可能である。また、当分野で周知の別の方法を用いて未知の配列を得ることも可能である。(例えば、Parker, J. D. 他(1991) Nucleic Acids Res. 19: 3055 - 3060を参照)。更に、PCR、ネスト化プライマー、PROMOTERFINDERライブラリ(Clontech, Palo Alto CA)を用いれば、ゲノムDNA内の歩行が可能である。この方法で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/エキソン接合部を探すのに有用である。全てのPCR法をベースにした方法では、プライマーは、市販のOLIGO 4.06 Primer Analysis software(National Biosciences, Plymouth MN)或いは別の好適なプログラムなどを用いて、長さが22~30ヌクレオチド、GC含量が50%以上、約68~72の温度で鋳型に対してアニーリングするよう設計される。

【0113】

完全長のcDNAをスクリーニングする場合は、大きなcDNAを含むようにサイズが選択されたライブラリを用いるのが好ましい。更に、オリゴd(T)ライブラリが完全長のcDNAを産生できない場合は、遺伝子の5'領域を有する配列を含むものが多いランダムに初回抗原刺激を受けたライブラリが有用である。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。

【0114】

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産物のヌクレオチド配列のサイズの分析、または確認が可能である。詳しくは、キャピラリーシーケンシングには、電気泳動による分離のための流動性ポリマー、及び4つの異なるヌクレオチドに特異的なレーザーで活性化される蛍光色素、放出された波長の検出に利用するCCDカメラを使用することが可能である。出力/光強度は、適切なソフトウェア(例えば、GENOTYPER及びSEQUENCE NAVIGATOR、PE Biosystems)を用いて電気信号に変換され、サンプルのローディングからコンピュータ分析までのプロセス及び電子データ表示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場合もあるDNAの小片のシーケンシングに特に適している。

【0115】

本発明の別の実施例では、GPAMをコードするポリヌクレオチド配列またはその断片を組換えDNA分子にクローニングして、適切な宿主細胞内にGPAM、その断片または機

10

20

30

40

50

能的等価物を発現させることが可能である。遺伝暗号固有の縮重により、実質的に同じ或いは機能的に等価のアミノ酸配列をコードする別のDNA配列が作られ得り、これらの配列をGPAMのクローン化及び発現に利用可能である。

【0116】

種々の目的でGPAMをコードする配列を変えるために、当分野で一般的に知られている方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び/または発現の調節が含まれるが、これらに限定されるものではない。ランダムな断片によるDNAの混合や遺伝子断片と合成オリゴヌクレオチドのPCR再組み立てを用いて、ヌクレオチド配列の組換えが可能である。例えば、オリゴヌクレオチド媒介性定方向突然変異誘発を利用して、新しい制限部位を生成する突然変異の導入、グリコシル化パターンの変更、コドン選択の変更、スプライスバリアントの作製等が可能である。

10

【0117】

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シャフリング技術を用いてシャフリングして、GPAM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GPAMの生物学的特性を変更或いは改良することができる。DNAシャフリングは、PCR法による遺伝子断片の組換えで遺伝子変異体のライブラリを作製するプロセスである。次に、このライブラリを、目的の特性を有する遺伝子変異体を同定するために選択或いはスクリーニングする。これらの好ましい変異体をプールし、DNAシャフリング及び選択/スクリーニングを繰り返す。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、ランダムな位置に変異がある1つの遺伝子の断片を、目的の特性が最適化するまで、組換え及びスクリーニング、シャフリングを実施することもできる。別法では、所定の遺伝子の断片を、同じ或いは異なった種の同じ遺伝子ファミリーの相同な遺伝子の断片で組換え、それによってプロトコルに従った調節可能な方法で、多数の天然遺伝子の遺伝子多様性を最大にすることができる。

20

30

【0118】

別の実施例によれば、GPAMをコードする配列は、当分野で周知の化学的方法を用いて、全体或いは一部が合成可能である(例えば、Caruthers, M.H.ら(1980) Nucl. Acids Res. Symp. Ser. 7:215-223; 及びHorn, T. 他(1980) Nucl. Acids Res. Symp. Ser. 225-232を参照)。別法として、化学的方法を用いてGPAM自体またはその断片を合成することが可能である。例えば、ペプチド合成は種々の固相技術を用いて実行可能である(例えば、Creighton, T. (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J.Y.ら(1995) Science 269:202-204を参照)。また、合成の自動化は例えばABI 431Aペプチドシンセサイザー(PE Biosystems)を用いて達成し得る。更にGPAMのアミノ酸配列または任意のその一部は、直接的な合成の際の変更、及び/または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

40

【0119】

このペプチドは、分離用高速液体クロマトグラフィー(例えば、Chiez, R.M. 及びF.Z. Regnier (1990) Methods Enzymol. 182:392-421を参照)を用いて実質的に精製可能である。合成されたペプチドの組成

50

は、アミノ酸分析或いはシーケンシングにより確認することができる（例えば、C r e i g h t o n、前出、p p 2 8 - 5 3を参照）。

【0120】

生物学的に活性なG P A Mを発現させるために、G P A Mをコードするヌクレオチド配列またはその誘導体を好適な発現ベクターに挿入する。この発現ベクターは、好適な宿主に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エレメントを含む。これらのエレメントには、ベクター及びG P A M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のプロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。このようなエレメントは、その長さ及び特異性が様々である。特定の開始シグナルによって、G P A Mをコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、A T G開始コドン及びコザック配列などの近傍の配列が含まれる。G P A M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A T G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なければならない。外来性の翻訳要素及び開始コドンは、自然及び合成の様々なものから得ることが可能である。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、S c h a r f , D . 他（1994）Results Probl. Cell Differ. 201-18-162.を参照）。

10

20

【0121】

当業者に周知の方法を用いて、G P A Mをコードする配列、好適な転写及び翻訳調節エレメントを含む発現ベクターを作製することが可能である。これらの方法には、i n v i t r o組換えDNA技術、合成技術、及びi n v i v o遺伝子組換え技術が含まれる。（例えば、S a m b r o o k , 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；及びA u s u b e l , F . M . 他（1995）Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9章及び13章1-4章を参照）。

30

【0122】

種々の発現ベクター/宿主系を利用して、G P A Mをコードする配列の保持及び発現が可能である。これらには、限定するものではないが、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れた細菌などの微生物や、酵母菌発現ベクターで形質転換された酵母菌や、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、C a M V ; タバコモザイクウイルス、T M V）または細菌発現ベクター（例えば、T i またはp B R 3 2 2 プラスミド）で形質転換された植物細胞系や、動物細胞系などが含まれる（例えば、前出のS a m b r o o k、前出のA u s u b e l、V a n H e e k e , G . およびS . M . S c h u s t e r（1989）J. Biol. Chem. 264:5503-5509、E n g e l h a r d , E . K . 他（1994）Proc. Natl. Acad. Sci. USA 91:3224-3227、S a n d i g , V . 他（1996）Hum. Gene Ther. 7:1937-1945、T a k a m a t s u , N .（1987）EMBOJ. 6:307-311；The McGraw Hill Yearbook of Science and Technology（1992）McGraw Hill, New York NY, pp. 191-196、L o g a n , J . a n d T . S h e n k（1984）Proc. Natl. Acad. Sci. USA 81:3655-3659、H a r r i n g t o n , J . J . 他（1997）Nat. Genet. 15:345-355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

40

50

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. 他 (1998) *Cancer Gen. Ther.* 5 (6): 350-356、Yu, M. 他 (1993) *Proc. Natl. Acad. Sci. USA* 90 (13): 6340-6344、Buller, R. M. 他 (1985) *Nature* 317 (6040): 813-815; McGregor, D. P. 他 (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。

#### 【0123】

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、GPAM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、GPAM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、増殖には、PBLUESCRIPT (Stratagene, La Jolla CA) または pSPORT1 プラスミド (GIBCO BRL) などの多機能の大腸菌ベクターを用いることができる。ベクターの多数のクローニング部位にGPAMをコードする配列をライゲーションすると 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の同定のための比色スクリーニング法が可能となる。更に、これらのベクターを用いて、クローニングされた配列の *in vitro* での転写、ジデオキシンスクリーニング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失を作り出すことが可能である (例えば、Van Heeke, G. 及び S. M. Schuster (1989) *J. Biol. Chem.* 264: 5503-5509. を参照)。例えば、抗体の産生のためなどに多量のGPAMが必要な場合は、GPAM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、強力に発現を誘発する T5 または T7 バクテリオファージプロモーターを含むベクターを使用できる。

#### 【0124】

GPAMの発現に酵母の発現系の使用が可能である。因子やアルコールオキシダーゼやPGHプロモーターなどの構造型或いは誘導型のプロモーターを含む多種のベクターが、酵母菌サッカロミセス・セレビジエまたは *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* 121-181-184. を参照)。

#### 【0125】

植物系もGPAMの発現に使用可能である。GPAMをコードする配列の転写は、例えば、CaMV由来の35S及び19Sプロモーターなどのウイルスプロモーターが単独で、或いはTMV (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) 由来のオメガリーダー配列と組み合わせて促進される。これらの作製物は、直接のDNA形質転換或いは病原体を介したトランスフェクションによって、植物細胞の中に導入可能である。(例えば、The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill NY, pp. 191-196を参照)。

#### 【0126】

哺乳動物細胞では、多種のウイルスベースの発現系が利用され得る。アデノウイルスが発現ベクターとして用いられる場合、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物/翻訳複合体にGPAMをコードする配列を結合し得る。ウイルスのゲ

10

20

30

40

50

ノムの非必須のE1またはE3領域への挿入により、感染した宿主細胞にGPAMを発現する生ウイルスを得ることが可能である(Logan, J. 及びShenk, T. (1984) Proc. Natl. Acad. Sci. 81:3655-3659を参照)。さらに、ラウス肉腫ウイルス(RSV)エンハンサーなど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ることが可能である。タンパク質を高レベルで発現させるために、SV40またはEBVを基にしたベクターを用いることが可能である。

#### 【0127】

ヒト人工染色体(HAC)を用いて、プラスミドで発現しそれに含まれているものより大きなDNAの断片を供給可能である。治療のために約6kb~10MbのHACsを作製し、従来の輸送方法(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)。

10

#### 【0128】

哺乳動物系の組換えタンパク質の長期にわたる産生のためには、株化細胞におけるGPAMの安定した発現が望ましい。例えば、発現ベクターを用いて、GPAMをコードする配列を株化細胞に形質転換することが可能である。このような発現ベクターは、ウイルス起源の複製及び/または内在性の発現要素や、同じ或いは別のベクターの上の選択マーカー遺伝子を含む。ベクターの導入の後、細胞を選択培地に移す前に、強化培地で約1~2日の間増殖させる。選択マーカーの目的は選択的な媒介物に対する抵抗性を与えるとともに、その存在により導入された配列を確実に発現する細胞の増殖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、その細胞型に好適な組織培養技術を用いて増殖可能である。

20

#### 【0129】

任意の数の選択系を用いて、形質転換された細胞系を回収することが可能である。選択系には、以下のものに限定はしないが、単純ヘルペスウイルスチミジンキナーゼ遺伝子及びアデニンホスホリボシルトランスフェラーゼ遺伝子が含まれ、それぞれtk<sup>r</sup>またはap<sup>r</sup>細胞において使用される。(例えば、Wigler, M. 他(1977) Cell 11:223-232; 及びLowy, I. 他(1980) Cell 22:817-823を参照)。また代謝拮抗物質、抗生物質或いは除草剤への耐性を選択のベースとして用いることができる。例えばdhfrはメトトレキセ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或いはpatはクロルスルフロン(chlorosulfuron)、ホスフィノトリシンアセチルトランスフェラーゼ(phosphinotricin acetyltransferase)に対する耐性を与える(例えば、Wigler, M. 他(1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. 他(1981) J. Mol. Biol. 150:1-14を参照)。さらに選択に利用できる遺伝子、例えば、代謝のために細胞が必要なものを変えるtrpB及びhisDが文献に記載されている(例えば、Hartman, S. C. 及びR. C. Mulligan(1988) Proc. Natl. Acad. Sci. 85:8047-51を参照)。アノトシアニン、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GUS、ルシフェラーゼ及びその基質ルシフェリンなどの可視マーカーが用いられる。緑色蛍光タンパク質(GFP)(Clontech, Palo Alto, CA)も使用できる。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(例えば、Rhodes, C. A. 他(1995) Methods Mol. Biol. 55:121-131を参照)。

30

40

#### 【0130】

マーカー遺伝子の発現の存在/非存在によって目的の遺伝子の存在が示されても、その遺

50



細胞膜及び真核細胞膜を透過するG P A Mの分泌を誘導するシグナル配列を含むように設計できることは、当業者には理解されよう。

【0135】

更に、挿入した配列の発現調節能力または発現したタンパク質を所望の形にプロセシングする能力によって宿主細胞株が選択される。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 (lipidation)、及びアシル化が含まれるが、これらに限定されるものではない。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特定の細胞装置及び特徴のある機構をもつ種類の宿主細胞 (例えば、CHO、HeLa、MDCK、MEK293、WI38) が American Type Culture Collection (ATCC; Bethesda, MD) より入手可能であり、外来のタンパク質の正しい修飾及びプロセシングを確実にするために選択される。

10

【0136】

本発明の別の実施例では、G P A Mをコードする自然或いは変更された、または組換えの核酸配列を上記した任意の宿主系の融合タンパク質の翻訳となる異種配列に結合させる。例えば、市販の抗体によって認識できる異種部分を含むキメラG P A Mタンパク質が、G P A Mの活性のインヒビター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が、市販の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。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-mc、赤血球凝集素 (HA) が含まれるが、これらに限定されるものではない。GST及びMBP、Trx、CBP、6-Hisによって、固定されたグルタチオン、マルトース、フェニルアルシン酸化物 (phenylarsine oxide)、カルモジュリン、金属キレート樹脂のそれぞれで同族の融合タンパク質の精製が可能となる。FLAG、c-mc、及び赤血球凝集素 (HA) によって、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体及びポリクローナル抗体を用いた融合タンパク質の免疫親和性の精製ができる。また、G P A Mをコードする配列と異種タンパク質配列との間にあるタンパク質分解切断部位を融合タンパク質が含むように遺伝子操作すると、G P A Mが精製の後に異種部分から切断され得る。融合タンパク質の発現と精製の方法は、Ausubel (1995、前出ch 10) に記載されている。市販されている様々なキットを用いて、融合タンパク質の発現及び精製を促進できる。

20

30

【0137】

本発明の別の実施例では、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 *in vitro* で放射能標識したG P A Mの合成が可能である。これらの系は、T7またはT3、SP6プロモーターと機能的に結合したタンパク質をコードする配列の転写と翻訳をつなげる。翻訳は、例えば、<sup>35</sup>Sメチオニンである放射能標識されたアミノ酸前駆体の存在の下で起こる。

【0138】

本発明のG P A Mまたはその断片を用いて、G P A M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、G P A M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。

40

【0139】

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G P A M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、G P A M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

50

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G P A M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、大腸菌からの細胞が含まれる。G P A Mを発現する細胞またはG P A Mを含有する細胞膜断片を試験化合物と接触させて、G P A M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。

#### 【0140】

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、フルオロフォア、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G P A Mと結合させるステップと、G P A M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、細胞遊離剤、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。

10

#### 【0141】

本発明のG P A Mまたはその断片を用いて、G P A M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、G P A Mが少なくとも1つの試験化合物と結合する、G P A M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G P A Mの活性が試験化合物非存在下でのG P A Mの活性と比較する。試験化合物の存在下でのG P A Mの活性の変化は、G P A M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G P A Mの活性に適した条件下でG P A M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、G P A M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。

20

#### 【0142】

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、G P A M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*: *Capecchi, 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ロックアウトする（*Marth, 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、潜在的な治療薬や毒性薬剤で検査することができる。

30

40

#### 【0143】

G P A M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

50

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)。

【0144】

G P A M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、G P A M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、潜在的な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G P A Mを乳汁内に分泌するなどG P A M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)。

10

【0145】

(治療)

G P A Mのある領域とGタンパク質関連分子のある領域との間に、例えば配列及びモチーフの文脈における化学的及び構造的類似性が存在する。従って、G P A Mは、細胞増殖異常及び自己免疫/炎症性の疾患においてある役割を果たすと考えられる。G P A Mの発現若しくは活性の増大に関連する疾患の治療においては、G P A Mの発現または活性を低下させることが望ましい。また、G P A M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、G P A Mの発現または活性を増大させることが望ましい。

20

【0146】

従って、一実施例において、G P A M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G P A M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、このような疾患には細胞増殖異常及び自己免疫/炎症性の疾患が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、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、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、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、子宮の癌などが含まれ、自己免疫/炎症性の疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群(AIDS)及び副腎機能不全、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、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ(APECED)、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的リンパ球減少症、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、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、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、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌の合併症、血液透析及び体外循環による炎症、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、蠕虫感染症、外傷による炎症が含まれる。

30

40

【0147】

別の実施例では、限定するものではないが上に列記した疾患を含むG P A M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、G P A M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することも可能である。

【0148】

更に別の実施例では、限定するものではないが上に列記した疾患を含むG P A M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、実質的に精製されたG P A M

50

を含む組成物を好適な医薬用担体と共に患者に投与することも可能である。

【0149】

更に別の実施例では、限定するものではないが上に列記した疾患を含むG P A M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、G P A Mの活性を調節するアゴニストを患者に投与することも可能である。

【0150】

更なる実施例では、G P A Mの発現または活性の増大に関連した疾患の治療または予防のために、患者にG P A Mのアンタゴニスト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には、上記した細胞増殖異常及び自己免疫/炎症性の疾患が含まれる。一実施態様では、G P A Mと特異的に結合する抗体が直接アンタゴニストとして、或いはG P A Mを発現する細胞または組織に薬剤を運ぶターゲッティング或いは運搬機構として間接的に用いられ得る。

10

【0151】

別の実施例では、限定するものではないが上に列記した疾患を含むG P A Mの発現または活性の増大に関連した疾患の治療または予防のために、G P A Mをコードするポリヌクレオチドの相補配列を発現するベクターを患者に投与することも可能である。

【0152】

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、相補的な配列、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。当業者は、従来の医薬原理にしたがって併用療法で用いる好適な治療薬を選択可能である。治療薬との組み合わせにより、上に列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。この方法を用いて少ない量の各薬剤で医薬効果をあげることが可能であり、広範囲な副作用の可能性を低減し得る。

20

【0153】

G P A Mのアンタゴニストは、当分野で一般的な方法を用いて製造することが可能である。詳しくは、精製されたG P A Mを用いて抗体を作ったり、治療薬のライブラリをスクリーニングしてG P A Mと特異的に結合するものを同定が可能である。G P A Mの抗体も、当分野で一般的な方法を用いて製造することが可能である。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、F a bフラグメント、及びF a b発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。治療用には、中和抗体(即ち、二量体の形成を阻害するもの)が特に好ましい。

30

【0154】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒト及びその他のものを含む種々の宿主が、G P A Mまたは任意の断片、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。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。このようなアジュバントにはフロイントアジュバント、水酸化アルミニウムなどのミネラルゲルアジュバント、リゾレシチン、ブルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、キーホールリンペットヘモシニアン、及びジニトロフェノールなどの界面活性剤が含まれるが、これらに限定されるものではない。ヒトに用いられるアジュバントの中では、B C G ( b a c i l l i C a G P A M t t e - G u e r i n ) 及び 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

【0155】

G P A Mに対する抗体を誘発するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約5個のアミノ酸からなり、一般的には約10個以上のアミノ酸からなるものが好ましい。これらのオリゴペプチド或いはペプチド、またはそれらの断片は、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。G P A Mアミノ酸の短いストレッチは、K L Hなどの別のタンパク質の配列と融合し、キメラ分子に対する抗体が産生され得る。

【0156】

50

G P A Mに対するモノクローナル抗体は、培地内の連続した細胞株によって、抗体分子を産生する任意の技術を用いて作製することが可能である。これらの技術には、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E B V - ハイブリドーマ技術が含まれるが、これらに限定されるものではない(例えば、K o h l e r , G . ら . ( 1 9 7 5 ) Nature 256 : 495 - 497 ; K o z b o r , D . ら . ( 1 9 8 5 ) . J . Immunol . Methods 81 - 8 - 42 ; C o t e , R . J . ら . ( 1 9 8 3 ) Proc . Natl . Acad . Sci . 80 : 2026 - 2030 ; C o l e , S . P . ら . ( 1 9 8 4 ) Mol . Cell Biol . 62 : 109 - 120を参照)。

【0157】

更に、「キメラ抗体」作製のために発達した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、M o r r i s o n , S . L . 他 . ( 1 9 8 4 ) Proc . Natl . Acad . Sci . 81 - 4851 - 4855 ; N e u b e r g e r , M . S . 他 . ( 1 9 8 4 ) Nature 312 : 604 - 608 ; T a k e d a , S . ら . ( 1 9 8 5 ) Nature 314 : 452 , 454を参照)。別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、G P A M特異性一本鎖抗体を生成する。関連する特異性を備えるが別のイデオタイプの組成の抗体は、ランダムな組み合わせの免疫グロブリンライブラリから鎖混合によって生成することもできる(例えば、B u r t o n D . R . ( 1 9 9 1 ) Proc . Natl . Acad . Sci . 88 : 11120 - 3を参照)。

【0158】

抗体は、リンパ球集団の中の i n v i v o 産生を誘発することによって、また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示されているような、高度に特異的な結合試薬のパネルをスクリーニングすることによって、得ることもできる(例えば、O r l a n d i , R . 他 . ( 1 9 8 9 ) Proc . Natl . Acad . Sci . 86 : 3833 - 3837 ; W i n t e r , G . 他 . ( 1 9 9 1 ) Nature 349 : 293 - 299を参照)。

【0159】

G P A Mに対する特異的な結合部位を含む抗体も得ることができる。例えば、限定するものではないが、抗体分子のペプシン消化によって生成されるF ( a b ' )<sub>2</sub>断片、及びF ( a b ' )<sub>2</sub>断片のジスルフィド架橋を減少させて生成されるF a b断片が含まれる。別法では、所望の特異性を有するモノクローナルF a b断片の迅速且つ容易な同定が可能となるように、F a b発現ライブラリを作製することもできる(例えば、H u s e , W . D . ら . ( 1 9 8 9 ) Science 254 : 1275 - 1281を参照)。

【0160】

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。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競合的な結合アッセイ、または免疫放射線アッセイの様々なプロトコルが、当分野では周知である。通常このようなイムノアッセイでは、G P A Mとその特異的抗体との間で形成された複合体の測定が含まれる。二つの非干渉性G P A Mエピトープに対して反応性のモノクローナル抗体を用いる、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、競合的結合アッセイも利用することができる(P o u n d、前出)。

【0161】

ラジオイムノアッセイ技術と共にS c a t c h a r d分析などの様々な方法を用いて、G P A Mに対する抗体の親和性を評価する。親和性を結合定数K aで表すが、このK aは、平衡状態の下でG P A M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て得られる値である。多数のG P A Mエピトープに対して親和性が不均一なポリクローナル抗体医薬のK aは、G P A Mに対する抗体の平均親和性または結合活性を表す。特定のG P A Mエピトープに単一特異的なモノクローナル抗体医薬のK aは、親和性の真の測定値

10

20

30

40

50

を表す。Ka値が $10^9 \sim 10^{12}$  L/molの高親和性抗体医薬は、GPAM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。Ka値が $10^6 \sim 10^7$  L/molの低親和性抗体医薬は、GPAM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 (immunopurification) 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 Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0162】

ある下流での適用におけるこのような医薬品の品質及び適性を調べるために、ポリクローナル抗体医薬の抗体価及び結合活性を更に評価する。例えば、少なくとも $1 \sim 2$  mg/mlの特異的な抗体、好ましくは $5 \sim 10$  mg/mlの特異的な抗体を含むポリクローナル抗体医薬は一般に、GPAM抗体複合体を沈殿させなければならない処理に用いられる。様々な適用例における抗体の特異性及び抗体価、結合活性、抗体の品質や使用法の指針は一般に入手可能である。(例えば、Catty, 前出, 及びColigan 他、前出を参照)。

【0163】

本発明の別の実施例では、GPAMをコードするポリヌクレオチド、またはその任意の断片や相補配列が、治療目的で使用することができる。ある実施態様では、GPAM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、GPAM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。

【0164】

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送達することができる(例えば、Slater, J. E. 他 (1998) *J. Allergy Clin. Immunol.* 102(3): 469 - 475; and Scanlon, K. J. 他 (1995) 9(13): 1288 - 1296. を参照)。また、アンチセンス配列は、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(例えば、Miller, A. D. (1990) *Blood* 76: 271; Ausubel, 前出; Ucker, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3): 323 - 347を参照)。その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロープ、及び当分野で周知のその他の系が含まれる(Rossi, J. J. (1995) *Br. Med. Bull.* 51(1): 217 - 225; Boado, R. J. 他 (1998) *J. Pharm. Sci.* 87(11): 1308 - 1315; and Morris, M. C. 他 (1997) *Nucleic Acids Res.* 25(14): 2730 - 2736. を参照)。

【0165】

本発明の別の実施例では、GPAM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の遺伝子治療に用いることが可能である。遺伝子治療は、(i) 遺伝子欠損症(例えば、X染色体連鎖遺伝 (Cavazzana-Calvo, M. 他 (2000) *Science* 288: 669 - 672) によって特徴づけられる重度の複合型免疫欠損 (SCID) - X1)、遺伝性アデノシン - デアミナーゼ (ADA) 欠損症 (Blaese, R. M. 他 (1995) *Science* 270: 475 - 480; Bordignon, C. 他 (1995) *Science* 270: 470 - 475) に関連す

10

20

30

40

50

る重度の複合型免疫欠損、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因子欠損による血友病 (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and Somia, N. (1997) *Nature* 389:239-242) を治療したり、(ii) 条件的致死性遺伝子産物 (例えば、細胞増殖の制御不能による癌の場合) を発現させたり、及び (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) や、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、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て行うことができる。GPAMの発現若しくは調節に必要な遺伝子の欠損が疾患を引き起こす場合、導入した細胞の好適な集団からGPAMを発現させて、遺伝子欠損によって起こる症状の発現を緩和することが可能である。

10

20

30

40

50

## 【0166】

本発明の更なる実施例では、GPAMの欠損による疾患や異常症は、GPAMをコードする哺乳動物発現ベクターを作製して、これらのベクターを機械的手段によってGPAM欠損細胞に導入することによって治療する。in vivo 或いは ex vitro の細胞に用いる機械的な導入技術には、(i) 個々の細胞内へのDNAのマイクロインジェクション、(ii) 金粒子の打ち込み、(iii) リポソーム仲介性トランスフェクション、(iv) 受容体仲介性遺伝子導入、及び (v) DNAトランスポゾン (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. Recipon (1998) *Curr. Opin. Biotechnol.* 9:445-450) の使用が含まれる。

## 【0167】

GPAMの発現に影響を及ぼし得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ベクター (Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV (Stratagene, La Jolla CA)、PTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG (Clontech, Palo Alto CA) が含まれる。GPAMを発現させるために、(i) 恒常的に活性なプロモーター (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)、若しくは - アクチン遺伝子等)、(ii) 誘導性プロモーター (例えば、市販されているT-REXプラスミド (Invitrogen) に含まれている、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)、エクジソン誘導性プロモーター (市販されているプラスミドPVGRXR及びPINDに含まれている: Invitrogen)、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 (Rossi, F.M.V. and H.M. Blau, 前出)

、または ( i i i ) 正常な個体に由来する G P A M 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ることが可能である。

【 0 1 6 8 】

市販のリポソーム形質転換キット ( 例えば、Invitrogen が販売している P E R F E C T L I P I D 及び T R A N S F E C T I O N K I T ) 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である。別法では、リン酸カルシウム法 ( G r a h a m , F . L . a n d A . J . E b ( 1 9 7 3 ) V i r o l o g y 5 2 : 4 5 6 - 4 6 7 ) 若しくは電気穿孔法 ( N e u m a n n , B . 他 ( 1 9 8 2 ) E M B O J . 1 : 8 4 1 - 8 4 5 ) を用いて形質転換を行う。初代細胞に D N A を導入するためには、これらの標準的な哺乳動物トランスフェクションプロトコルを変更する必要がある。 10

【 0 1 6 9 】

本発明の別の実施例では、G P A M 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 ( L T R ) プロモーター若しくは独立したプロモーターのコントロール下で G P A M をコードするポリヌクレオチドと、( i i ) 好適な R N A パッケージングシグナルと、( i i i ) 追加のレトロウイルス・シス作用性 R N A 配列及び効率的なベクターの増殖に必要なコーディング配列を伴う R e v 応答性エレメント ( R R E ) 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、P F B 及び P F B N E O ) は S t r a t a g e n e 社から入手可能であり、公表データ ( R i v i e r e , I . 他 . ( 1 9 9 5 ) P r o c . N a t l . A c a d . S c i . U . S . A . 9 2 : 6 7 3 3 - 6 7 3 7 ) 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。このベクターは、V S V g ( A r m e n t a n o , D . 他 ( 1 9 8 7 ) J . V i r o l . 6 1 : 1 6 4 7 - 1 6 5 0 ; B e n d e r , M . A . 他 ( 1 9 8 7 ) J . V i r o l . 6 1 : 1 6 3 9 - 1 6 4 6 ; A d a m , M . A . a n d A . D . M i l l e r ( 1 9 8 8 ) J . V i r o l . 6 2 : 3 8 0 2 - 3 8 0 6 ; D u l l , T . 他 ( 1 9 9 8 ) J . V i r o l . 7 2 : 8 4 6 3 - 8 4 7 1 ; Z u f f e r e y , R . 他 ( 1 9 9 8 ) J . V i r o l . 7 2 : 9 8 7 3 - 9 8 8 0 ) 等の乱交雑エンベロープタンパク質若しく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を発現する好適なベクター産生細胞系 ( V P C L ) において増殖される。R I G G に付与された米国特許第 5 , 9 1 0 , 4 3 4 号 ( 「 M e t h o d f o r o b t a i n i n g r e t r o v i r u s p a c k a g i n g c e l l l i n e s p r o d u c i n g h i g h t r a n s d u c i n g e f f i c i e n c y r e t r o v i r a l s u p e r n a t a n t 」 ) 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、ある細胞集団 ( 例えば、C D 4 + T 細胞 ) の形質導入、並びに形質導入した細胞を患者に戻す方法は、遺伝子治療の分野では周知であり、多数の文献に記載されている ( R a n g a , U . 他 . ( 1 9 9 7 ) J . V i r o l . 7 1 : 7 0 2 0 - 7 0 2 9 ; B a u e r , G . 他 ( 1 9 9 7 ) B l o o d 8 9 : 2 2 5 9 - 2 2 6 7 ; B o n y h a d i , M . L . ( 1 9 9 7 ) J . V i r o l . 7 1 : 4 7 0 7 - 4 7 1 6 ; R a n g a , U . 他 ( 1 9 9 8 ) P r o c . N a t l . A c a d . S c i . U . S . A . 9 5 : 1 2 0 1 - 1 2 0 6 ; S u , L . ( 1 9 9 7 ) B l o o d 8 9 : 2 2 8 3 - 2 2 9 0 ) 。 20 30 40

【 0 1 7 0 】

別法では、アデノウイルス系遺伝子治療の送達系を用いて、G P A M の発現に関連する 1 或いは複数の遺伝子異常を有する細胞に G P A M をコードするポリヌクレオチドを送達する。アデノウイルス系ベクターの作製及びパッケージングは当分野では周知である。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の中に導入するために可変性であることが証明された ( C s e t e , M . E . 他 . ( 1 9 9 5 ) T r a n s p l a n t a t i o n 2 7 : 2 6 3 - 2 6 8 ) 。使用で 50

きる可能性のあるアデノウイルスベクターが、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また、Antinozzi, P. A. 他(1999) *Annu. Rev. Nutr.* 19:511-544; and Verma, I. M. and N. Somia (1997) *Nature* 18:389:239-242を参照し、引用することをもって本明細書の一部とする。

【0171】

別法では、ヘルペス系遺伝子治療の送達系を用いて、GPAMの発現に関連する1或いは複数の遺伝子異常を有する標的細胞にGPAMをコードするポリヌクレオチドを送達する。単純疱疹ウイルス(HSV)系のベクターは、HSV親和性の中枢神経細胞にGPAMを導入する際に特に重要である。ヘルペス系ベクターの作製及びパッケージングは当分野では周知である。複製適格性の単純疱疹ウイルス(HSV)I型系のベクターは、霊長類の眼にレポーター遺伝子を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は、DeLucaに付与された米国特許第5,804,413号(Herpes simplex virus swains for gene transfer)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。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コントロールの下で、細胞に導入される少なくとも1つの内在性遺伝子を含むゲノムからなる組換えHSV d92についての記載がある。また上記特許には、ICP4、ICP27及びICP22のために除去される組換えHSV株の作製及び使用方法が開示されている。HSVベクターについては、Goins, W. F. 他(1999) *J. Virol.* 73:519-532 and Xu, H. 他(1994) *Dev. Biol.* 163:152-161を参照し、引用することをもって本明細書の一部とする。クローニングされたヘルペスウイルス配列の操作や、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トランスフェクトした後の組換えウイルスの継代、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分野で周知の技術である。

【0172】

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GPAMをコードするポリヌクレオチドを標的細胞に送達する。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)の生物学的な研究が広範に行われ、遺伝子伝達ベクター(gene transfer vector)がSFVゲノムに基づいていることが分かった(Garoff, H. and K. - J. Li (1998) *Cun. Opin. Biotech.* 9:464-469)。ウイルスRNAの複製中に、通常はウイルスカプシ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。このサブゲノムRNAが完全長のゲノムRNAより高いレベルで複製されるため、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対してカプシドタンパク質が過剰に産生される。同様に、GPAMをコードする配列をウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数のGPAMをコードするRNAが産生され、高いレベルでGPAMが合成される。通常はウイルス感染は2~3日以内の細胞溶解に関係するが、シンドビスウイルス(SIN)の変異体を有するハムスターの正常な腎細胞(BHK-21)の持続的な感染を確立する能力は、ウイルスの溶解性の複製が遺伝子治療に適用できるように好適に変更することが可能であることを示唆している(Dryga, S. A. 他(1997) *Virology* 228:74-83)。様々な宿主にウイルスを導入できることから、様々なタイプの細胞にGPAMを導入することができる。ある集団における細胞のサブセットの特定の形質導入には、形質導入する前に細胞のソーティングを必要とする場合がある。ウイルスの感染性cDNAクローンの操作、ウイルスcDNA及びRNAのトランスフェクション、並びにウイルスの感染方法は当分野で周知である

。

## 【0173】

例えば開始部位から約 - 10 から約 + 10 までの転写開始部位に由来するオリゴヌクレオチドを用いて、遺伝子の発現を阻害することが可能である。同様に、三重らせん塩基対合法を用いて阻害することができる。三重らせん構造は、二重らせんがポリメラーゼ、転写因子、または調節分子の結合のために十分に広がるのを阻止するため有用である。三重鎖 DNA を用いる最近の治療の進歩は文献に記載されている（例えば、Gee, J. E. 等 (1994) *In: Huber, B. E. 及び B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY* を参照）。相補的な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 の翻訳を阻止するように設計できる。 10

## 【0174】

酵素性 RNA 分子であるリボザイムは、RNA の特異的切断を触媒するために用いることができる。リボザイム作用の機構には、相補的な標的 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が含まれ、ヌクレオチド鎖切断が続く。例えば、GPAM をコードする配列のヌクレオチド鎖切断を、特異的且つ効果的に触媒する組換え型のハンマーヘッド型リボザイム分子が含まれる。

## 【0175】

任意の潜在的 RNA 標的内の特異的なリボザイム切断部位が、後続の配列 GUA、GUU、及び GUC を含むリボザイム切断部位に対して、標的分子をスキャンニングすることによって初めに同定される。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 個から 20 個のリボヌクレオチドの短い RNA 配列を、オリゴヌクレオチドの機能を不全にする二次的な構造の特徴について評価することが可能である。候補標的の適合性も、リボヌクレアーゼ保護アッセイを用いて、相補的なオリゴヌクレオチドとのハイブリダイゼーションの容易性をテストすることによって評価することが可能である。 20

## 【0176】

本発明の相補的なリボ核酸分子及びリボザイムは、当分野で周知の方法を用いて、核酸分子の合成のために作製することができる。これらの方法には、固相ホスホラミダイト化合物などのオリゴヌクレオチドを化学的に合成する方法が含まれる。別法では、RNA 分子が *in vitro* 及び *in vivo* で GPAM をコードする DNA 配列の転写によって生成され得る、このような DNA 配列は T7 または SP6 等の好適な RNA ポリメラーゼプロモータを用いて、種々のベクターの中に組み入れることが可能である。別法では、相補的な RNA を構成的または誘導的に合成するこれらの cDNA 作製物は、細胞株、細胞、または組織の中に導入することができる。 30

## 【0177】

RNA 分子を修飾することによって、細胞内の安定性を高め、半減期を長くすることができる。可能な修飾には、分子の 5' 及び / または 3' 端部でのフランキング配列の追加、または分子のバックボーン内のホスホジエステル結合の代わりにホスホロチオネートまたは 2' Oメチルを用いる修飾が含まれるが、これらに限定されるものではない。PNA の生成に固有のこの概念は、内在性の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、及びウリジンのアセチル -、メチル -、チオ -、及び同様の修飾形態だけでなく、イノシン、キューエオシン (queosine)、及びワイプトシン (wybutosine) などの従来のものでない塩基を含めることによって、これらの分子の全体に拡大することができる。 40

## 【0178】

本発明の更なる実施例は、GPAM をコードするポリヌクレオチドの発現の変化に有効な化合物をスクリーニングする方法を含む。特定のポリヌクレオチドの発現の変化に有効な化合物には、限定するものではないが、特定のポリヌクレオチド配列と相互作用可能な非高分子化学物質、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 50

成オリゴヌクレオチド、転写因子やその他のポリペプチド転写調節因子が含まれる。有効な化合物は、ポリヌクレオチド発現のインヒビター或いはエンハンサーとして作用し、ポリヌクレオチドの発現を変化させ得る。従って、G P A M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、G P A M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、G P A M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、G P A M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。

#### 【0179】

特定のポリヌクレオチドの発現の変化の有効性を調べるために、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することができる。試験化合物は、有効な化合物の化学修飾を含む当分野で周知の任意の方法で得ることができる。このような方法は、ポリヌクレオチドの発現を変化させる場合、一般に市販されている或い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、標的ポリヌクレオチドの化学的及び/または構造的な特性に基づいて化合物を合理的にデザインする場合、更に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。G P A M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露して得る。サンプルには、例えば無傷細胞、透過処理した細胞、*in vitro*細胞遊離系または再構成生化学系が含まれ得る。G P A Mをコードするポリヌクレオチドの発現における変化は、当分野で周知の任意の方法でアッセイする。通常、G P A M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、特定の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の収量を定量し、その値が1或い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おける基準となり得る。試験化合物に曝露されるポリヌクレオチドの発現の変化が検出される場合は、ポリヌクレオチドの発現の変化に試験化合物が有効であることを示している。特定のポリヌクレオチドの発現の変化に有効な化合物を調べるために、例えば *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 細胞等のヒト細胞株 (Clarke, M.L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268:8-13) を用いてスクリーニングする。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を調べるための、各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、及び修飾オリゴヌクレオチド)の組み合わせライブラリのスクリーニングを含む (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)。

#### 【0180】

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でき、*in vivo*、*in vitro*、及び *ex vivo* での使用に等しく適している。*ex vivo* での治療の場合、患者から採取された肝細胞の中にベクターを導入して、自家移植で同じ患者に戻すためにクローニング増殖される。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運搬は、当分野で周知の方法を用いて実行することができる(例えば、Goldman, C.K. 他 (1997) *Nature Biotechnology* 15:462-66:を参照)。

#### 【0181】

上記したいかなる治療方法も、例えば、ヒト、イヌ、ネコ、ウシ、ウマ、ウサギ及びサルなどの哺乳動物を含む、治療が必要な全ての被験者に適用できる。

#### 【0182】

本発明の別の実施例は、上記した全ての治療効果のために、医学上認められる担体と共に医薬品或いは無菌組成物の投与に関連する。このような組成物は、G P A M、G P A Mの抗体、擬態、アゴニスト、アンタゴニスト、またはG P A Mのインヒビターなどからなる

。この組成物は、単体で、或いは安定剤などの1種類以上の別の薬剤と共に、無菌の生体適合性医薬品担体に投与することができる。このような医薬品担体には、生理食塩水、緩衝食塩水、ブドウ糖、及び水などが含まれるがこれらに限定されるものではない。この組成物は、単独或いは薬物またはホルモンなどの別の薬剤と共に投与することができる。

【0183】

本発明に用いられる組成物は、様々な経路を用いて投与するが可能である。この経路には、経口、静脈内、筋肉内、動脈内、骨髄内、クモ膜下、心室内、経皮、皮下、腹腔内、鼻腔内、腸内、局所、舌下、または直腸が含まれるがこれらに限定されるものではない。

【0184】

肺投与用の組成物は、液状または乾燥粉末状に調製することができる。このような組成物は通常、患者が吸入する直前にエアロゾル化する。小分子（例えば、従来の低分子量有機薬剤）の場合には、速効薬剤のエアロゾル輸送が当分野で周知である。高分子（例えばより大きなペプチドやタンパク質）の場合には、肺の肺胞領域を介する肺輸送の技術が近年向上したため、インスリン等の薬剤を実際に血中に輸送することが可能となった（Patton, J. S. 他, 米国特許第5,997,848号等を参照）。肺輸送は、針注射を用いなくて投与できるという点で優れており、潜在的に有毒な浸透エンハンサーが必要でなくなる。

10

【0185】

本発明に用いる好適な組成物には、目的を達成するため、効果的な量の活性処方成分を含む組成物が含まれる。当業者は、十分に自身の能力で効果的な服用量を決めることができる。

20

【0186】

G P A Mまたはその断片を含む高分子を直接細胞内に輸送するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内輸送を促進し得る。別法では、G P A Mまたはその断片をH I V T a t - 1タンパク質の陽イオンN末端部に結合することもできる。このようにして作製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されることが確認されている（Schwarze, S. R. 他（1999）Science 285:1569-1572）。

【0187】

どのような組成物であっても、治療に効果的な薬用量は、初めは、例えば腫瘍細胞の腫瘍細胞アッセイで、或いは動物モデルのどちらかで推定することができる。通常、動物モデルには、マウス、ウサギ、イヌ、サル、またはブタなどが用いられる。動物モデルはまた、好適な濃縮範囲及び投与の経路を決めるのに用いることができる。このような治療をもとに、ヒトへの有益な薬用量及び投与経路を決定することができる。

30

【0188】

医学的に効果的な薬用量は、症状や容態を回復させる、たとえばG P A Mまたはその断片、G P A Mの抗体、G P A Mのアゴニストまたはアンタゴニスト、インヒビターなどの活性処方成分の量に関連する。薬用有効度及び毒性は、たとえば、E D<sub>50</sub>（服用に対して集団の50%に医薬的效果がある用量）またはL D<sub>50</sub>（服用に対して集団の50%に致命的である用量）統計を計算するなど、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決定することができる。毒性効果と治療効果との薬用量比は治療指数であり、L D<sub>50</sub> / E D<sub>50</sub>と示すことができる。高い治療指数を示す組成物が望ましい。細胞培養アッセイ及び動物実験から得られたデータが、ヒトへの適用のために、薬用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薬用量は、毒性を殆ど或いは全く含まず、E D<sub>50</sub>を含む血中濃度の範囲であることが望ましい。薬用量は、用いられる投与形態及び患者の感受性、投与の経路によって、この範囲内で様々である。

40

【0189】

正確な薬用量は、治療が必要な患者に関する要素を考慮して、実務者によって決められるであろう。薬用量及び投与は、効果的なレベルの活性成分を与えるため或いは所望の効果

50

を維持するために調節される。薬用量の要素として考慮されるものには、疾患の重症度、患者の一般的な健康状態、年齢、体重、及び患者の性別、投与の時間及び頻度、併用する薬剤、反応感受性、及び治療に対する応答が含まれる。作用期間が長い組成物は、三日か四日に一度、一週間に一度、二週間に一度、特定の製剤の半減期及びクリアランス率によって左右され、投与され得る。

#### 【0190】

通常薬用量は投与の経路によって異なるが、約0.1~100,000 $\mu$ gまでの最大約1グラムまでである。特定の薬用量及び運搬の方法に関するガイダンスは文献に記載されており、一般に当分野の実務者は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とは異なったヌクレオチドの製剤を利用するであろう。同様に、ポリヌクレオチドまたはポリペプチドの運搬は、特定の細胞、状態、位置などに対して特異的であろう。

10

#### 【0191】

(診断)

別の実施例では、GPAMに特異的に結合する抗体が、GPAMの発現によって特徴付けられる疾患の診断、またはGPAMやGPAMのアゴニストまたはアンタゴニスト、インヒビター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用いられる。診断に有用な抗体は、治療のところで記載した方法と同じ方法で製剤される。GPAMの診断アッセイには、抗体及び標識を用いてヒトの体液或いは細胞や組織から採取されたものからGPAMを検出する方法が含まれる。これらの抗体は、修飾をして或いはしないで使用され、レポーター分子の共有結合性或いは非共有結合性の接着によって標識化され得る。当分野で周知の種々のレポーター分子が用いられるが、その内の幾つかは上記した。

20

#### 【0192】

GPAMを測定するためのELISA,RIA,及びFACSを含む種々のプロトコルは、当分野では周知であり、変わった或いは異常なレベルのGPAMの発現を診断する元となるものを提供する。正常或いは標準的なGPAMの発現の値は、複合体の形成に適した条件の下、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞とGPAMに対する抗体とを結合させることによって決定する。標準的な複合体形成の量は、測光法(photometric)などの種々の方法で定量され得る。被験者のGPAMの発現の量、制御及び疾患、生検組織からのサンプルが基準値と比較される。基準値と被験者との間の偏差が、診断の指標となる。

30

#### 【0193】

本発明の別の実施例によれば、GPAMをコードするポリヌクレオチドを診断のために用いることもできる。用いられるポリヌクレオチドには、オリゴヌクレオチド配列、相補的なRNA及びDNA分子、及びPNAが含まれる。このポリヌクレオチドを用いて、疾患と相関し得るGPAMを発現する生検組織における遺伝子の発現を検出し定量する。この診断アッセイを用いて、GPAMの存在の有無、更に過剰な発現を調べ、治療中のGPAM値の変化を監視する。

#### 【0194】

一実施形態では、GPAMまたは近縁の分子をコードする遺伝子配列を含むポリヌクレオチド配列を検出可能なPCRプローブを用いたハイブリダイゼーションによって、GPAMをコードする核酸配列を同定することが可能である。例えば5'調節領域である高度に特異的な領域か、例えば保存されたモチーフであるやや特異性の低い領域から作られているかのプローブの特異性と、ハイブリダイゼーション或いは増幅のストリンジェントは、プローブがGPAMをコードする自然界の配列のみを同定するかどうか、或いはアレルや関連配列コードする自然界の配列のみを同定するかどうかによって決まるであろう。

40

#### 【0195】

プローブはまた、関連する配列の検出に利用され、GPAMをコードする任意の配列と少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO: 3-4の配列、或いは

50

G P A M 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。

【0196】

G P A M をコードする D N A に対して特異的なハイブリダイゼーションプローブの作製方法には、G P A M 及び G P A M 誘導体をコードするポリヌクレオチド配列を m R N A プローブの作製のためのベクターにクローニングする方法がある。このようなベクターは市販されており、当業者には周知であり、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る。ハイブリダイゼーションプローブは、例えば<sup>32</sup>P 或いは<sup>35</sup>S などの放射性核種、或いはアビジン/ビオチン ( b i o t i n ) 結合系によってプローブに結合されたアルカリホスファターゼなどの酵素標識等の種々のレポーターの集団によって標識され得る。

【0197】

G P A M をコードするポリヌクレオチド配列を用いて、G P A M の発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患には細胞増殖異常及び自己免疫/炎症性の疾患が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、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、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、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、子宮の癌などが含まれ、自己免疫/炎症性の疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群 ( A I D S ) 及び副腎機能不全、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、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ ( 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的リンパ球減少症、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、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、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、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌の合併症、血液透析及び体外循環による炎症、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、蠕虫感染症、外傷による炎症が含まれる。G P A M をコードするポリヌクレオチド配列は、サザーン法やノーザン法、ドットプロット法、或いはその他の膜系の技術、P C R 法、ディップスティック ( d i p s t i c k )、ピン ( p i n )、E L I S A 式アッセイ、及び変異 G P A M の発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような質的或いは量的方法は、当分野では周知である。

【0198】

ある実施態様では、G P A M をコードするヌクレオチド配列は、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有用であろう。G P A M をコードするヌクレオチド配列は、標準的な方法で標識化され、ハイブリダイゼーション複合体の形成に好適な条件の下、患者から採取した体液或いは組織のサンプルに加えることができるであろう。好適な培養期間の後、サンプルを洗浄し、シグナルを定量して基準値と比較する。患者のサンプルのシグナルの量が、制御サンプルと比べて著しく変わっている場合は、サンプル内の G P A M をコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイを用いて、動物実験、臨床試験、或いは個人の患者の治療を監視における、特定の治療効果を推定することが可能である。

【0199】

G P A Mの発現に関連する疾患の診断の基準となるものを提供するために、発現の正常すなわち標準的なプロファイルが確立される。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の下、動物或いはヒトの何れかの正常な被験者から抽出された体液或いは細胞と、G P A Mをコードする配列或いはその断片とを結合させることにより達成され得る。標準的なハイブリダイゼーションは、正常な被験者から得た値と周知の量の実質的に精製されたポリヌクレオチドが用いられる実験からの値とを比較することによって定量可能である。正常なサンプルから得た標準的な値を、疾患の症状を示す被験者から得た値と比較可能である。基準値と被験者の値との偏差を用いて罹患しているかどうかを決定する。

#### 【0200】

疾患の存在が確定され、治療プロトコルが開始されると、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て、被験者における発現のレベルが正常な患者に示される値に近づき始めたかどうかを推定することが可能である。繰り返し行ったアッセイの結果を、数日から数ヶ月の期間の治療の効果を見るのに用いることができる。

#### 【0201】

癌では、個体からの生体組織における異常な量の転写物が、疾患の発生の素因を示し、また実際に臨床的症状が出る前に疾患を検出する方法を提供することが可能である。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或いは積極的な治療法を早くから利用して、癌の発生または進行を防ぐことが可能となる。

#### 【0202】

G P A Mをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、P C Rの利用が含まれ得る。このようなオリゴマーは、化学的な合成、酵素を用いた生成、或いは*in vitro*で生成され得る。オリゴマーは、好ましくはG P A Mをコードするポリヌクレオチドの断片、或いはG P A Mをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件の下、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェントな条件の下、近縁のD N A或いはR N A配列の検出及び/または定量のため用いることが可能である。

#### 【0203】

或る実施態様において、G P A M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、一塩基多型(S N P)を検出し得る。S N Pは、ヒトの先天性または後天性遺伝病の原因となる場合が多いヌクレオチドの置換、挿入及び欠失である。限定するものではないが、S N Pの検出方法には、一本鎖立体構造多型(S S C P)及び蛍光S S C P(f S S C P)法が含まれる。S S C Pでは、G P A M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たポリメラーゼ連鎖反応(P C R)でD N Aを増幅する。このD N Aは、例えば病変或いは正常な組織、生検サンプル、体液等に由来し得る。このD N A内のS N Pは、一本鎖形状のP C R産物の2次及び3次構造に差異を生じさせる。この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。f S C C Pでは、オリゴヌクレオチドプライマーを蛍光標識することによって、D N Aシーケンシング装置などのハイスループット機器でアンプリマー(amplimer)の検出をすることが可能になる。更に、インシリコS N P(in silico S N P: i s S N P)と呼ばれる配列データベース分析法は、共通のコンセンサス配列の構築に用いられる個々の重複するD N A断片の配列を比較することによって、多型を同定することができる。これらのコンピュータベースの方法は、D N A配列クロマトグラムの自動分析及び統計モデルを用いたシーケンシングエラーや研究室でのD N Aの調整に起因する配列のばらつきを排除する。別法では、例えばハイスループットのM A S S A R R A Yシステム(Sequenom, Inc., San Diego CA)を用いた質量分析によりS N Pを検出し、特徴付ける。

#### 【0204】

G P A Mの発現を定量するために用いられ得る方法には、ヌクレオチドの放射標識或いはビオチン標識、調節核酸の相互増幅(coamplification)、及び標準的な曲線に結果が加えられたものが含まれる(例えば、Melby, P. C.ら(1993

) J. Immunol. Methods, 159:235-44; Duplaa, C.ら(1993) Anal. Biochem. 229-236を参照)。多数のサンプルの定量速度は、ハイスループット型のアッセイを用いることで速くなるであろう。このアッセイでは、目的のオリゴマーやポリヌクレオチドが様々な希釈液中に含まれ、分光光度法或いは非色応答によって定量が迅速である。

#### 【0205】

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に由来する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標的として用いることができる。マイクロアレイを、上記したように多数の遺伝子の相対的な発現レベルを同時にモニタリングする転写イメージング技術に用いることができる。マイクロアレイはまた、遺伝子変異、突然変異及び多型の同定に用いることができる。この情報を用いて、遺伝子機能を決定し、疾患の遺伝的根拠を解明し、疾患を診断し、遺伝子発現に関連する疾病の進行/後退をモニタリングし、疾患の治療における治療薬の開発や活性のモニタリングを行うことができる。特に、患者にとって最適かつ有効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作成することができる。例えば、患者の薬理ゲノムプロファイルに基づいて、患者に対して極めて効果的でありながら副作用を殆ど示さない治療薬を選択することができる。

10

#### 【0206】

別の実施例では、GPAM、GPAMの断片、GPAMに特異的な抗体をマイクロアレイ上のエレメントとして用いることができる。マイクロアレイを用いて、上記のようにタンパク質間相互作用、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リング及び測定することが可能である。

20

#### 【0207】

特定の実施例は、或る組織または細胞型の転写イメージを生成する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型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型の転写物または逆転写物の全てに本発明のポリヌクレオチドまたはその相補配列をハイブリダイズすることにより、転写イメージが生成され得る。或る実施例では、本発明のポリヌクレオチドまたはその相補配列がマイクロアレイ上に複数のエレメントのサブセットを構成するハイスループット型でハイブリダイゼーションさせる。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となり得る。

30

#### 【0208】

転写イメージは、組織、細胞株、生検サンプル、またはその他の生体サンプルから単離した転写物を用いて生成し得る。従って、転写イメージは、組織または生検サンプルの場合には in vivo、または細胞株の場合には in vitro における遺伝子発現を反映する。

40

#### 【0209】

本発明のポリヌクレオチドの発現プロファイルを示す転写イメージはまた、合成化合物または天然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され得る。全ての化合物は、作用及び毒性の機構を示唆する、頻繁に分子フィンガープリント若しくは毒性シグネチャ(signature)と称されるような特徴的な遺伝子発現パターンを引き起こす(Nuwaysir, E. F. 他(1999) Mol. Carcinog. 24:153-159、Steiner, S. and N. L. Anderson(2000) Toxicol. Lett. 112-113:467-471、また言及することを以って本明細書の一部とする)。試験化合物が、毒性を有する既知の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

50

特性を共有している可能性が高い。フィンガープリンまたはシグネチャが、より多く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れば、より有用かつ正確になる。理想としては、発現のゲノム全域にわたって測定し、最高品質のシグネチャを提供することである。任意の試験化合物によっても発現が変化しない遺伝子も同様に重要である。それは、これらの遺伝子の発現レベルを用いて残りの発現データを標準化することができるためである。標準化処理は、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。毒性シグネチャのエLEMENTへの遺伝子機能を割り当てることは毒性機構の解明に役立つが、毒性の予測につながるシグネチャの統計的な一致には遺伝子機能の知識は必要ではない(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、毒性シグネチャを用いる毒性スクリーニングにおいて、全ての発現した遺伝子配列を含めることは重要でありまた望ましいことである。

#### 【0210】

一実施例では、試験化合物の毒性は、核酸を含有する生体サンプルをその試験化合物で処理して評価する。処理した生体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若しくは複数のプローブでハイブリダイズさせ、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することができる。処理した生体サンプル中の転写レベルを、非処理生体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が、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唆する。

#### 【0211】

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型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」という用語は、或る特定の組織または細胞型におけるタンパク質発現の包括的パターンを指す。プロテオームを構成する各タンパク質は、個々に更なる分析をすることができる。プロテオーム発現パターン即ちプロファイルは、所定の条件下で所定の時間に発現したタンパク質の数及びそれらの相対的な存在量を定量することにより分析する。従って、ある細胞のプロテオームのプロファイルは、特定の組織または細胞型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、このような分離は2次元ゲル電気泳動によって行う。この2次元ゲル電気泳動法では、まず、1次元の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、次に、2次元の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従って分離する(前出のSteiner and Anderson)。これらのタンパク質は、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染色剤を用いてゲルを染色して、分散した個別の位置にあるスポット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済みまたは未処理のいずれかの生体サンプルから得られる等位置にあ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調べる。スポット内のタンパク質は、例えば化学的または酵素的に切断した後、質量分析す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同一性は、好適には少なくとも5個の連続するアミノ酸残基であるその部分的な配列を、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。

#### 【0212】

プロテオームのプロファイルは、GPAMに特異的な抗体を用いてGPAM発現レベルを定量することによっても作成可能である。或る実施例では、マイクロアレイ上のELEMENTとして抗体を用い、マイクロアレイをサンプルに曝露して各アレイELEMENT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(Lueking, A. ら。(1999) Anal. Biochem. 270:103-111、Mendoze, L.G. ら。(1999) Biotechniques 27: 50

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、チオール反応性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、各アレイエレメントにおける蛍光結合の量を検出し得る。

【0213】

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して分析するべきである。或る組織における或るタンパク質では、転写物の存在量とタンパク質の存在量との相関性が低いことがあるため (Anderson, N. L. and J. Seilhamer (1997) *Electrophoresis* 18: 533 - 537)、プロテオーム毒性シグネチャは、転写イメージにはそれ程影響しないがプロテオームのプロファイルを変化させる化合物の分析において有用たり得る。更に、体液中での転写の分析は、mRNAが急速に分解するため困難である。しがたがって、このような場合にはプロテオームのプロファイル作成はより信頼でき、情報価値がある。

10

【0214】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理して評価する。処理された生体サンプル中で発現したタンパク質を分離して、各タンパク質の量が定量できるようにする。各タンパク質の量を、未処理生体サンプル中の対応するタンパク質の量と比較する。両サンプル中のタンパク質の量の差は、処理されたサンプル中の試験化合物に対する毒性反応を示唆する。個々のタンパク質は、それらのアミノ酸残基をシーケンシングし、これらの部分配列を本発明のポリペプチドと比較することで同定する。

20

【0215】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理することにより評価する。生体サンプルから得たタンパク質を、本発明のポリペプチドに特異的な抗体と共にインキュベートする。その抗体により認識されたタンパク質の量を定量する。処理された生体サンプル中のタンパク質の量を、未処理生体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が、処理サンプル中の試験化合物に対する毒性反応を示唆する。

【0216】

当分野で周知の方法でマイクロアレイを準備して使用し、分析する。(例えば、Brennan, T. M. 他 (1995) 米国特許第5,474,796号; Schena, M. 他 (1996) *Proc. Natl. Acad. Sci.* 93: 10614 - 10619; Baldeschweiler 他 (1995) PCT出願番号WO95/251116; Shalon, D. 他 (1995) PCT出願番号WO95/35505; Heller, R. A. 他 (1997) *Proc. Natl. Acad. Sci.* 94: 2150 - 2155; 及び Heller, M. J. 他 (1997) 米国特許第5,605,662号を参照)。様々なタイプのマイクロアレイが周知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, Londonに記載されている。また、この文献を引用することを以って本明細書の一部とする。

30

40

【0217】

本発明の別の実施例ではまた、GPAMをコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コーディング配列または非コーディング配列の何れかを用いることができるが、或る例では、コーディング配列より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間にコーディング配列が保存されていることにより、染色体マッピング時に望ましくない交差ハイブリダイゼーションが生じる可能性がある。この配列は、特定の染色体、染色体の特定領域または人工の染色体、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c

50

DNAライブラリに対してマッピングされる (Harrington, J. J. ら (1997) Nat Genet. 15:345-355、Price, C. M. (1993) Blood Rev. 7:127-134、Trask, B. J. (1991) Trends Genet. 7:149-154等を参照)。一度マッピングすると、本発明の核酸配列を用いて、例えば病状の遺伝と特定の染色体領域やまたは制限断片長多型 (RFLP) の遺伝とが関連するような遺伝子連鎖地図を作成可能である (Lander, E. S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

【0218】

in situ 蛍光ハイブリダイゼーション (FISH) は、他の物理的及び遺伝子地図データと関連し得る (例えば、Heinz-Ulrich, 他による (1995) in Meyers, 前出, pp. 965-968を参照)。遺伝子地図データの例は、種々の科学誌あるいは Online Mendelian Inheritance in Man (OMIM) のワールドワイドウェブのサイトで見つけることができる。物理的な染色体地図上の GPAM 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が、このような疾患と関連する DNA 領域の決定に役立つため、更なる位置を決定するクローニングが行われる。

【0219】

染色体標本の in situ ハイブリダイゼーション、及び確定した染色体マーカーを用いた結合分析などの物理的マッピング技術を用いて、遺伝子地図を拡張することもできる。マウスなどの別の哺乳動物の染色体上に遺伝子を配置させることにより、たとえ正確なヒト染色体の位置が分かっても、関連するマーカーが明らかになる場合が多い。この情報は、位置クローニング或いは別の遺伝子発見技術を用いて遺伝的疾患の研究をしている研究者にとって価値がある。疾患や症候群に關与する1つ或いは複数の遺伝子の位置が、例えば血管拡張性失調症の 11q22-23 などの特定の遺伝子領域に遺伝子結合によって大まかに決定されると、その領域に対するどの配列マッピングも、さらなる調査のための関連する遺伝子或いは調節遺伝子を表す (例えば、Gatti, R. A. 他による (1988) Nature 336:577-580を参照)。また、目的の本発明のヌクレオチド配列を用いて、正常者、保有者、即ち感染者の間の、転位置、反転などによる染色体位置の違いを検出することもある。

【0220】

本発明の別の実施例では、GPAM、その触媒作用断片或いは免疫原断片またはそのオリゴペプチドを、種々の任意の薬剤スクリーニング技術における化合物のライブラリのスクリーニングに用いることができる。このようなスクリーニングに用いる断片は、溶液に遊離、固体支持物に固定、細胞の表面上に保持、或いは細胞内に存在する。GPAM と検査する薬剤との結合による複合体の形成を測定してもよい。

【0221】

薬剤スクリーニングに用いる別の方法は、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のスクリーニング処理能力を高めるために用いられる (例えば、Geysen, 他による (1984) PCT 出願番号 WO 84/03564 を参照)。この方法では、相当な数の異なる小さな試験用化合物が、プラスチックピン或いは他の基板の上に合成される。試験用化合物は、GPAM、或いはその断片と反応してから洗浄される。次に、結合された GPAM が、当分野で周知の方法で検出される。精製された GPAM はまた、前記した薬剤をスクリーニングする技術に用いられるプレート上で直接被覆することもできる。別法では、非中和抗体を用いて、ペプチドを捕らえ、固体支持物に固定することもできる。

【0222】

別の実施例では、GPAM と結合可能な中和抗体が GPAM と結合するため試験用化合物と特に競合する、競合的薬剤スクリーニングアッセイを用いることができる。この方法では、抗体が、GPAM と1つ以上の抗原決定因子を共有するどのペプチドの存在も検出す

10

20

30

40

50

る。

【0223】

別の実施例では、発展途上の分子生物学技術にG P A Mをコードするヌクレオチド配列を用いて、限定はされないが、現在知られているトリプレット暗号及び特異的な塩基対相互作用などのヌクレオチド配列の特性に依存する新しい技術を提供することができる。

【0224】

当分野の技術者であれば、更なる説明がなくても前述の説明だけで最大限に本発明を利用できるであろう。したがって、以下に記載する実施例は、例示目的であって本発明を限定するものではない。

【0225】

前述した及び以下に記載する全ての特許出願、特許、刊行物、特に米国特許出願第60/091,631号に言及することをもって本明細書の一部とする。

【0226】

(実施例)

1 cDNAライブラリの作製

インサイトcDNAはL I F E S E Q G O L D データベース ( I n c y t e G e n o m i c s , P a l o A l t o C A ) に含まれているcDNAライブラリに由来し、表4の列5に示されている。組織の一部を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する一方、この組織の別の一部をホモジナイズしてフェノールに溶解するか、或いはT R I Z O L ( L i f e T e c h n o l o g i e s )、グアニジニウムイソチオシアネート及びフェノールの単相溶液などの好適な変性剤の混合液に溶解した。この溶解物を塩化セシウムにおいて遠心分離によって、或いはクロロホルムで抽出した。イソプロパノール或いは酢酸ナトリウムのどちらかとエタノール、或いは別の方法でこの溶解物からRNAを沈殿させた。

【0227】

RNAの純度を高めるために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、DNA分解酵素でRNAを処理する。殆どのライブラリでは、オリゴd(T)連結常磁性粒子(Promega)またはOLIGOTEXラテックス粒子(QIAGEN, Valencia CA)、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、POLY(A)PURE mRNA精製キット(Ambion, Austin TX)などの別のRNA単離キットを用いて組織溶解物から直接単離した。

【0228】

ある場合には、Stratagene社にRNAを提供し、Stratagene社が対応するcDNAライブラリを作製した。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)を用いて当分野で周知の推奨方法または類似の方法でcDNAを合成してcDNAライブラリを作製した。(例えば、Ausubel, 1997, 前出, ユニット5.1-6.6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鎖cDNAに結合させてから、好適な1つの制限酵素或いは複数の制限酵素でcDNAを消化した。殆どのライブラリでは、SEPHACRYL S 1000またはSEPHAROSE CL2B、SEPHAROSE CL4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、アガロースゲル電気泳動法によってcDNAの大きさ(300~1000bp)を選択した。PBLUESCRIPTプラスミド(Stratagene)または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)、pINCYプラスミド(Incyte Pharmaceuticals, Palo Alto CA)、またはそれらの誘導体などの好適なプラスミドのポリリンカーの適合性制限酵素部位にcDNAを結合させた

10

20

30

40

50



核生物のデータベース、およびBLOCKS、PRINTS、DOMO、PRODOM、およびPFAMなどの隠れマルコフモデル(HMM)を基にしたタンパク質ファミリーのデータベースから選択した配列に対して問合せた(HMMは、遺伝子ファミリーのコンセンサス主構造を分析する確率的手法である。例えば、Eddy, S. R. (1996) Curr. Opin. Struct. Biol. 6: 361-365を参照)。このような問合せは、BLAST、FASTA、BLIMPS、およびHMMRに基づいたプログラムを用いて行った。インサイトcDNA配列を組み立てて、完全長ポリヌクレオチド配列を作製した。或いは、GenBank cDNAs、GenBank EST、ステッチ配列(stitched sequence)、ストレッチ配列(stretched sequences)、またはGenscan-推定コード配列(実施例4および5を参照)を用いて、インサイトcDNA群を完全長の配列に伸長した。配列の組み立ては、Phred、Phrap、およびConsedに基づいたプログラムを用いて行い、GeneMark、BLAST、およびFASTAに基づいたプログラムを用いて、オープンリーディングフレームを決定するべくcDNA群をスクリーニングした。完全長のポリヌクレオチド配列を翻訳して対応する完全長ポリペプチド配列を得た。別法では、本発明のポリヌクレオチドは、完全長翻訳ポリペプチドの任意のメチオニン残基から始まり得る。次に、完全長ポリペプチド配列をGenBankタンパク質データベース(genepept)、SwissProt、BLOCKS、PRINTS、DOMO、PRODOM、Prosite、およびPFAMなどの隠れマルコフモデル(HMM)に基づいたタンパク質ファミリーデータベースに対して問合せて分析した。これらの完全長ポリヌクレオチド配列はまた、MACDNASIS PROソフトウェア(Hitachi Software Engineering, South San Francisco CA)およびLASERGENEソフトウェア(DNA STAR)を用いて分析した。ポリヌクレオチド配列およびポリペプチド配列のアラインメントを、アラインメントした配列間のパーセント同一性を計算するMEGALIGNマルチシークエンスアラインメントプログラム(DNA STAR)に組み込まれたCLUSTALアルゴリズムによって指定されたデフォルトパラメータを用いて作成した。

10

20

#### 【0233】

表7は、インサイトcDNAおよび完全長配列の組み立て、および組み立てた配列の分析に利用したツール、プログラム、およびアルゴリズム、並びにそれらの説明、引用文献、閾値パラメータを簡単に示す。表7の列1は用いたツール、プログラム、およびアルゴリズム、列2はそれらの簡単な説明、列3は引用することで本明細書の一部とした引用文献、列4の記載されている部分は2つの配列の一致の程度を評価するために用いたスコア、確率値、およびその他のパラメータを示す(スコアが高くなれば高くなるほど即ち確率値が低ければ低いほど、配列間の相同性が高くなる)。

30

#### 【0234】

完全長の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: 3-4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。

40

#### 【0235】

#### 4 ゲノムDNAからのコード配列の同定および編集

推定上のGタンパク質関連分子は、公共のゲノム配列データベース(例えば、gbpriやgbhtg)においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物に由来するゲノムDNA配列を分析するための汎用遺伝子同定プログラムである(Burge, C. および S. Karlin (1997) J. Mol. Biol. 268: 78-94、Burge, C. および S. Karlin (1998) Curr. Opin. Struct. Biol. 8: 346-354を参照)。このプログラムは推定エキソンを連結して、メチオニンから停止コドンまで伸長した組み立てcDNA配列を構築する。Genscanにより得られる

50

配列は、FASTAデータベースのポリヌクレオチド配列およびポリペプチド配列になる。Genscanによって一回で解析できる配列の最大長さは30kbに設定されている。これらのGenscan推定cDNA配列の内、どの配列がGタンパク質関連分子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PFAMモデルにおいてGタンパク質関連分子について問合せて分析した。潜在的なGタンパク質関連分子が、Gタンパク質関連分子としてアノテーションが付けられたインサイトcDNA配列に対する相同性を基に同定された。次に、これらの選択されたGenscan推定配列を、BLAST解析を用いてgeneptおよびgbpri公共データベースの配列と比較した。必要に応じて、Genscan推定cDNA配列を、geneptにおいてBLASTで最もヒットした配列と比較して、Genscan推定配列における余分なエキソンや省いてしまったエキソンなどのエラーを修正し、編集した。BLAST解析を用いてGenscan推定cDNA配列を含むインサイトcDNAまたは公共のcDNAを見つけ出すことにより、転写の証拠が得られる。インサイトcDNAがGenscan推定cDNA配列を含む場合、この情報を用いてGenscan推定配列を修正或いは確認できる。完全長ポリヌクレオチド配列は、実施例3に説明した組み立て方法でGenscan推定コード配列とインサイトcDNAおよび/または公共のcDNA配列を組み立てて作製した。別法では、完全長ポリヌクレオチド配列は、その全てが編集した或いは未編集のGenscan推定コード配列から作製した。

【0236】

#### 5 cDNA配列データを用いるゲノム配列データの組み立て

##### ステッチ配列 (Stitched Sequence)

部分的なcDNA配列を、実施例4に記載したGenscan遺伝子同定プログラムによって推定されたエキソンで伸長した。実施例3に記載されたように組み立てられた部分的なcDNAをゲノムDNAにマッピングし、関連するcDNAおよび1或いは複数のゲノム配列に由来する関連する推定Genscanエキソンを含む複数のクラスターに入れた。各クラスターを、グラフ理論および動的計画法に基づいたアルゴリズムを用いて、cDNAおよびゲノム情報を統合して分析し、後に確認される潜在的なスプライズバリエーションを生成し、編集或いは伸長して完全長の配列を作製した。或るクラスターの2つ以上の配列に或る区間の全長が存在する配列区間を同定し、推移(transitivity)により同定した区間を同等と考える。例えば、或る区間がcDNAおよび2つのゲノム配列のそれぞれに存在する場合、これら3つ全ての区間を同等と考える。この方法によって、関連しないが連続するゲノム配列をcDNA配列によって繋ぎ1つにする。このようにして同定された区間を、親配列(parent 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わせ、可能な最も長い配列および変異配列を作製する。或るタイプ(cDNAとcDNA、またはゲノム配列とゲノム配列)の親配列に沿って連結される区間と区間との繋ぎ合わせは、親配列のタイプが異なる(cDNAとゲノム配列)連結より好ましい。得られたステッチ配列を翻訳し、BLAST解析でgeneptおよびgbpri公共データベースにおける配列と比較した。Genscanによって推定された不適当なエキソンを、geneptにおいてBLASTで最もヒットした配列と比較して修正する。このような配列を更なるcDNA配列で伸長し、必要に応じてゲノムDNAで検査した。

【0237】

##### ストレッチ配列 (Stretched Sequence)

部分的なDNA配列をBLAST解析に基づいたアルゴリズムで完全長に伸長した。まず、実施例3に記載したように組み立てた部分的なcDNAを、BLASTプログラムを用いてGenBankの霊長類、げっ歯類、哺乳類、脊椎動物、および真核生物のデータベースなどの公共のデータベースに対して問い合わせた。次に、GenBankの相同性の最も高いタンパク質を、実施例4に記載したインサイトcDNA或いはGenscanエキソン推定配列の何れかと比較した。得られた複数の高スコアのセグメント対(HSP)を用いてキメラタンパク質を作製し、GenBankの相同タンパク質上に翻訳した配列

をマッピングした。元の GenBank の相同タンパク質に対して、キメラタンパク質に挿入や欠失が起こり得る。公共のヒトゲノムデータベースから相同ゲノム配列を探し出すために、GenBank の相同タンパク質およびキメラタンパク質の両方をプローブとして用いた。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。完全な遺伝子を含んでいるか得られたストレッチ配列を検査した。

【0238】

#### 6 GPAM 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: 3-4 を組み立て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、インサイト 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: 3-4 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。Stanford Human Genome Center (SHGC)、Whitehead Institute for Genome Research (WIGR) 及び Genethon などの公共の情報源から入手できる放射線ハイブリッド (radiation hybrid) 及び遺伝子マッピングのデータを用いて、クラスター化した配列がすでにマッピングされているかを調べる。クラスターにマッピングされた配列が含まれている場合は、そのクラスターの全ての配列 (特定の SEQ ID NO を含む) をそのマッピング位置に割り当てた。

10

【0239】

遺伝子地図の位置は、範囲、区間、またはヒト染色体によって表される。センチモルガンで示したマッピング位置の範囲は、染色体の短腕 (p) の末端から測定した (センチモルガン (cM) は、同一染色体上の遺伝子間の乗換え率に基づいた距離を表す単位である。平均すると、1 cM はヒトの染色体の 1 メガベースに概ね等しいが、組換え率の高い部分と低い部分があるため、大きく変化し得る)。距離 cM は、配列がそれぞれのクラスターに含まれている放射線ハイブリッドマーカーの境界を検出できる Genethon によってマッピングされた遺伝子マーカーに基づいている。NCBI 「GeneMap99」 (<http://www.ncbi.nlm.nih.gov/genemap>) などの公衆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。

20

【0240】

#### 7 ポリヌクレオチド発現の分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いられる実験用技術であり、特定の細胞種或いは組織からの RNA が結合されてい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を伴う (例えば、Sambrook, 前出, 7 章; 及び Ausubel, F.M., 他, 前出, 4 章及び 16 章を参照)。

30

【0241】

BLAST に用いる類似のコンピュータ技術を用いて、GenBank 或いは LIFESEQ (Incyte Pharmaceuticals) のような cDNA データベース内の同一或いは関連する分子を検索する。この分析は多くの膜系ハイブリダイゼーションより非常に速度が速い。さらにコンピュータ検索の感度を変更して、任意の特定の一致が、厳密な一致或いは相同的一致の何れかとして分類されるかを確定することができる。検索の基準は、

40

【0242】

【数 1】

(BLAST スコア × 配列一致率)

$5 \times (\text{長さ(配列1)}, \text{長さ(配列2)})$  の最小値

として定義される積スコアである。積スコアは、0 ~ 100 の標準化された値であり、以下のように求める。BLAST スコアにヌクレオチド配列の一致率を乗じ、その積を 2 つ

50

の配列の短い方の長さの5倍で除す。高スコアのセグメントの対(HSP)において一致する各塩基に+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(ギャップにより離隔される)。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、BLASTアラインメントの断片的重複と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%一致する場合にのみ得られる。積スコア70は、100%の同一性で重畳が70%であるか、或いは88%の同一性で重畳が100%であるかの何れかの場合である。積スコア50は、100%の同一性で重畳が50%であるか、或いは79%の同一性で重畳が100%であるかの何れかの場合である。

10

## 【0243】

或いは、GPAMをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば、ある完全長の配列は、少なくとも部分的にインサイトcDNA配列をオーバーラップさせて組み立てられる(実施例3を参照)。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、心血管系、結合組織、消化系、胚構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、肺、筋骨格系、神経系、脾臓、呼吸器系、感覚器官、皮膚、顎口腔系、分類不能/混合、または尿管などの1つの生物/組織のカテゴリーに分類される。各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。同様に、各ヒト組織は、癌、細胞系、発生、炎症、神経、外傷、心血管、プール(pool)などの1つの疾患/症状のカテゴリーに分類され、各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。得られるパーセンテージは、GPAMをコードするcDNAの疾患特異的な発現を反映する。cDNA配列およびcDNAライブラリ/組織の情報は、LIFESEQ GOLD データベース(Incyte Genomics, Palo Alto CA)から得ることができる。

20

## 【0244】

## 8 GPAM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は、完全長分子の好適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その完全長分子の好適な断片を伸長して作製した。一方のプライマーは既知の断片の5'の伸長を開始するために合成し、他方のプライマーは既知の断片の3'の伸長を開始するために合成した。開始プライマーは、OLIGO 4.06ソフトウェア(National Biosciences)或いは他の適切なプログラムを用いて、約22個から約30個のヌクレオチドの長さで約50%以上のGC含量を有し、かつ約68~72の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマー-プライマー二量体が生じないようにヌクレオチドを伸長した。

30

## 【0245】

選択されたヒトcDNAライブラリを用いてこの配列を伸長した。2段階以上の伸長が必要な場合、若しくは望ましい場合は、追加或いはネスト化プライマーの組を設計する。

## 【0246】

当分野で既知の方法を利用したPCR法で高い忠実度で増幅した。PCRはPTC-200 thermal cycler (MJ Research, Inc.)を用いて96ウェルブロックプレートで行った。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。

40

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

50

- ステップ4 68 で2分間  
 ステップ5 ステップ2、3、及び4を20回繰り返す  
 ステップ6 68 で5分間  
 ステップ7 4 で保管

別法では、プライマーの組、T7とSK+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 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、及び4を20回繰り返す  
 ステップ6 68 で5分間  
 ステップ7 4 で保管。

10

#### 【0247】

各ウェルのDNA濃度は、1X TE及び0.5µlの希釈していないPCR産物に溶解した100µlのPICO GREEN定量試薬(0.25% (v/v) PICO GREEN; 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。このプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンして、サンプルの蛍光を計測してDNAの濃度を定量化する。反応混合物の5~10µlのアリコットを1%のアガロースミニゲル上での電気泳動によって解析し、何れの反応物が配列を伸長することに成功したかを決定する。

20

#### 【0248】

伸長したヌクレオチドを脱塩及び濃縮してから384ウェルプレートに移し、CviJIコレラウイルスエンドヌクレアーゼ(Molecular Biology Research, Madison WI)で消化し、pUC18ベクター(Amersham Pharmacia Biotech)に再連結する前に音波処理またはせん断を行った。ショットガンシーケンシングのために、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に分離して断片を切断し、寒天をAgar ACE (Promega)で消化した。T4リガーゼ(New England Biolabs, Beverly MA)を用いて伸長したクローンをpUC18ベクター(Amersham Pharmacia Biotech)に再連結し、Pfu DNAポリメラーゼ(Stratagene)で制限部位の伸び出しを処理してコンピテント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。

30

#### 【0249】

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。

40

上記したようにPICO GREEN試薬(Molecular Probes)でDNAを定量化した。DNA回収率の悪いサンプルは、上記した条件で再び増幅した。サンプルを20%のジメチルサルホไซด์(dimethyl sulphoxide)(1:2, v/v)で希釈し、DYENAMIC DIRECTキット(Amersham Phar

50

macia Biotech)またはABI PRISM BIGDYE Terminator cycle sequencing ready reactionキット(Applied Biosystems)を用いてシーケンシングした。

#### 【0250】

同様に上述の手順で、完全長のポリヌクレオチド配列を検査したり、或いは完全長のポリヌクレオチド配列を利用して、この伸長のために設計したオリゴヌクレオチドと好適なゲノムライブラリを用いて5'調節配列を得た。

#### 【0251】

##### 9 個々のハイブリダイゼーションプローブの標識化及び使用法

SEQ ID NO: 3 - 4 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すが、より大きなcDNAフラグメントの場合でも基本的に同じ手順を用いる。オリゴヌクレオチドを、OLIGO 4.06ソフトウェア(National Bioscience)のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250  $\mu$ Ciの[<sup>32</sup>P]アデノシン三リン酸(Amersham, Chicago, IL)及びT4ポリヌクレオチドキナーゼ(DuPont NEN, Boston MA)とを組み合わせるにより標識する。標識されたオリゴヌクレオチドを、SEPHADEX G-25超精細排除デキストランビードカラム(Amersham Pharmacia Biotech)を用いて実質的に精製する。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、次のエンドヌクレアーゼ、Ase I、Bgl II、Eco RI、Pst I、Xba I或いはPvu II(DuPont NEN)の1つを用いて切断したヒトゲノム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用いる。

#### 【0252】

各切断物からのDNAを、0.7%アガロースゲル上で分画して、ナイロン製メンブラン(Nytran Plus, Schleicher & Schuell, Durham NH)に転写する。ハイブリダイゼーションは40°Cで16時間かけて行う。非特異的シグナルを取り除くため、例えば、最大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の条件の下、プロットを順次室温にて洗浄する。ハイブリダイゼーションパターンをオートラジオグラフィー或いは別のイメージ化手段で視覚化して比較する。

#### 【0253】

##### 10 マイクロアレイ

マイクロアレイ上のアレイエレメントの連結または合成は、フォトリソグラフィ、ピエゾプリント(インクジェットプリンター、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、均一な非多孔性の固体とするべきである(Schena(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。別法では、ドットプロット法またはスポットプロット法に類似のアレイを利用して、熱や紫外線、または化学的或いは機械的な結合手段で基板の表面にエレメントを配置して結合させることができる。通常のアレイは利用可能な方法や機械を用いて作製でき、任意の適正な数のエレメントを含めることができる(Schena, M. 他(1995) Science 270: 467-470、Shalon, D. 他(1996) Genome Res. 6: 639-645、Marshall, A. and J. Hodgson(1998) Nat. Biotechnol. 16: 27-31. を参照)。

#### 【0254】

完全長cDNA、発現遺伝子配列断片(EST)、或いはそれらの断片やオリゴマーが、マイクロアレイのエレメントとなり得る。ハイブリダイゼーションに好適な断片やオリゴマーを、LASERGENEソフトウェア(DNA STAR)などの当分野で周知のソフ

10

20

30

40

50

トウェアを用いて選択することが可能である。このアレイエレメントを、生体サンプル中のポリヌクレオチドとハイブリダイズさせる。生体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結合する。ハイブリダイゼーションの後、生体サンプルからハイブリダイズしなかった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けるハイブリダイゼーションを検出する。別法では、レーザー脱離及び質量スペクトロメトリー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程度及び相対的存在量は、算定することができる。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。

#### 【0255】

##### 組織または細胞サンプルの調製

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サンプルは、MMLV逆転写酵素、0.05 pg/μlのオリゴ(dT)プライマー(21mer)、1×第1鎖緩衝液、0.03単位/μlのRNアーゼインヒビター、500 μM dATP、500 μM dGTP、500 μM dTTP、40 μM dCTP、40 μM dCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。この逆転写反応は、GEMBRIGHTキット(Incycyte)を用いて、200 ngのポリ(A)<sup>+</sup>RNAを含む25 ml容量で行う。特異的なコントロールポリ(A)<sup>+</sup>RNAは、in vitro転写により非コーディング酵母ゲノムDNAから合成する。37°Cで2時間インキュベートした後、各反応サンプル(一方はCy3標識、他方は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を変性する。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピナラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。結合後、2つの反応サンプルを、1 mlのグリコーゲン(1 mg/ml)、60 mlの酢酸ナトリウム及び300 mlの100%エタノールを用いてエタノール沈殿させる。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。

#### 【0256】

##### 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挿入断片を含むベクターを含有する細菌性細胞から増幅する。PCR増幅は、cDNA挿入断片に隣接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)を用いて精製する。

#### 【0257】

精製したアレイエレメントを、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大量の蒸留水での洗浄と、0.1%のSDS及びアセトン中で超音波による洗浄を行う。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、蒸留水中で広範囲にわたって洗浄し、95%エタノール中の0.05%アミノプロピルシラン(Sigma)でコーティングする。コーティングしたスライドガラスは、110°Cの天火で硬化させる。

#### 【0258】

米国特許第5,807,522号に記載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。この特許に引用することを以って本明細書の一部

10

20

30

40

50

とする。平均濃度が  $100 \text{ ng} / \mu\text{l}$  のアレイエレメント DNA  $1 \mu\text{l}$  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 printing element) に充填する。次にこの装置が、スライド毎に約  $5 \text{ nl}$  のアレイエレメントサンプルを分注する。

#### 【0259】

マイクロアレイには、STRATALINKER UVクロスリンカー (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回洗浄し、蒸留水で3回洗浄する。非特異的な結合部位は、リン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) における  $0.2\%$  カゼイン中で60 で30分間マイクロアレイをインキュベートし、その後上述したように  $0.2\%$  SDS及び蒸留水で洗浄することによってブロックする。

10

#### 【0260】

##### ハイブリダイゼーション

ハイブリダイゼーション反応液は、 $5 \times \text{SSC}$ 、 $0.2\%$  SDSハイブリダイゼーション緩衝液にCy3及びCy5標識したcDNA合成産物を各  $0.2 \mu\text{g}$  含む  $9 \mu\text{l}$  のサンプル混合体を含めたものである。サンプル混合液を、65 で5分間加熱し、マイクロアレイ表面上に一定量分注してから  $1.8 \text{ cm}^2$  のカバーガラスで覆う。このアレイを、顕微鏡スライドより僅かに大きいキャピティを有する防水チェンバーに移す。チャンバーの角に  $140 \mu\text{l}$  の  $5 \times \text{SSC}$  を加えて、チャンバー内を湿度  $100\%$  に保持する。このアレイを含むチャンバーを、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 ( $1 \times \text{SSC}$ 、 $0.1\%$  SDS) において45 で10分間、第2洗浄緩衝液中 ( $0.1 \times \text{SSC}$ ) において45 で10分間それぞれ3回洗浄し、その後乾燥させる。

20

#### 【0261】

##### 検出

レポーター標識されたハイブリダイゼーション複合体は、Cy3を励起するための  $488 \text{ nm}$ 、及びCy3を励起するための  $632 \text{ nm}$  のスペクトル線を生成し得る Innova 70 混合ガス  $10 \text{ W}$  レーザー (Coherent, Inc., Santa Clara CA) を備えた顕微鏡で検出する。20倍の顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いて、アレイ上に励起レーザー光を集中させる。このアレイを含むスライドを顕微鏡のコンピュータ制御 X-Y ステージに置き、対物レンズを通してラスタスキャンする。本実施例で用いた  $1.8 \text{ cm} \times 1.8 \text{ cm}$  のアレイは、 $20 \mu\text{m}$  の解像度でスキャンする。

30

#### 【0262】

2つの異なるスキャンにおいて、混合ガスマルチラインレーザーは2つの蛍光体を連続的に励起する。放射された光は、波長に基づいて2つの蛍光体に対応する2つの光電子増倍管検出器 (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) に分割される。アレイと光電子増倍管との間に配設された好適なフィルタを用いて信号をフィルタリングする。用いる蛍光体の最大発光は、Cy3では  $565 \text{ nm}$ 、Cy5では  $650 \text{ nm}$  である。装置は両方の蛍光体からのスペクトルを同時に記録できるが、レーザー源に好適なフィルタを用いて、蛍光体1つにつき1回スキャンし、各アレイを通常2回スキャンする。

40

#### 【0263】

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コントロール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配列を含め、その位置におけるシグナルの強度がハイブリダイズする種の重量比  $1:100$ 、 $000$  に相関するようにする。異なる試料 (例えば検査細胞及びコントロール細胞を代表する) からの2つのサンプルを、各々異なる蛍光体で標識し、他と異なって発現する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させる場合には、較正は2つの蛍光体を有する較正するcDNAのサンプルを標識して、ハイブリダイゼーション混合液

50

に各々等量を加えて行う。

【0264】

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ビットRTI-835Hアナログ-デジタル(AID)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、リニア20色変換を用いてシグナル強度が青色(低シグナル)から赤色(高シグナル)までの擬似カラー範囲にマッピングされるイメージとして表示される。データはまた、定量的に分析される。2つの異なる蛍光体を同時に励起して測定する場合には、各蛍光体の発光スペクトルを用いて、先ずデータは蛍光体間の光学的漏話(重複発光スペクトルに起因する)に対して補正される。

10

【0265】

グリッドを蛍光シグナルイメージ上に重畳して、各スポットからのシグナルがグリッドの各エレメントに中央に位置するようにする。各エレメント内の蛍光シグナルを統合し、シグナルの平均強度に対応する数値を得る。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。

【0266】

1.1 相補的ポリヌクレオチド

GPAM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な配列は、天然のGPAMの発現を低下させるため即ち阻害するために用いられる。約15~約30個の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、より小さな或いはより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)及びGPAMのコーディング配列を用いて、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'配列から相補的な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、相補的なオリゴヌクレオチドを設計して、リボソームがGPAMをコードする転写物に結合するのを阻害する。

20

【0267】

1.2 GPAMの発現

GPAMの発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌でGPAMが発現するために、抗生物質耐性及び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、これらに限定されるものではない。組換えベクターを、BL21(DE3)など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)で誘発されるとGPAMを発現する。真核細胞でのGPAMの発現は、昆虫細胞株または哺乳動物細胞株に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)を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、GPAM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、強いポリヘドリン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。(例えば、Engelhard, E. K. 他(1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. 他(1996) Hum. Gene Ther. 7:1937-1945.を参照)。

30

40

【0268】

殆どの発現系では、GPAMが、例えばグルタチオンSトランスフェラーゼ(GST)、

50

またはFLAGや6-Hisなどのペプチドエピトープ標識で合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製が素早く1回で行うことができる。Schistosoma japonicumからの26キロダルトンの酵素GSTによって、タンパク質の活性及び抗原性を維持した状態で固定されたグルタチオンで融合タンパク質の精製が可能となる(Amersham Pharmacia Biotech)。精製の後、GST部分を特定の操作部位でGPAMからタンパク質的に切断できる。アミノ酸8個のペプチドであるFLAGで、市販のモノクローナル及びポリクローナル抗FLAG抗体(Eastman Kodak)を用いた免疫親和性の精製が可能となる。6個の連続するヒスチジン残基のストレッチである6-Hisによって、金属キレート樹脂(QIAGEN)で精製が可能となる。タンパク質の発現及び精製の方法は、Ausubel(1995, 前出, ch 10, 16)に記載されている。これらの方法で精製したGPAMを直接用いて以下の実施例16及び17のアッセイを行うことができる。

【0269】

### 1.3 機能のアッセイ

GPAMの機能は、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GPAMをコードする配列の発現によって評価する。cDNAを、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サブクローニングする。このようなベクターには、pCMV SPORT<sup>T</sup>M (Life Technologies.)及びpCR3.1 (Invitrogen, Carlsbad, CA)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含んでいる。5~10µgの組換えベクターを、例えば内皮由来か造血由来のヒト細胞株にリポソーム製剤或いは電気穿孔法によって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1~2µg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、形質移入された細胞と形質移入されていない細胞とを区別できる。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。このような標識タンパク質には、緑色蛍光タンパク質(GFP; Clontech)、及びCD64またはCD64-GFP融合タンパク質が含まれる。レーザー光学に基づいた技術を利用した自動流動細胞計測法(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。また、FCMで、先行した或いは同時の細胞死の現象を診断する蛍光分子の取り込みを検出して計量する。これらの現象には、プロピジウムヨウ化物でのDNAの染色によって計測される核DNA内容物の変化と、プロモデオキシウリジンの取り込み量の低下によって計測されるDNA合成の下方調節と、特異的な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のタンパク質の発現の変化と、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含まれる。流動細胞計測法は、Ormerod, M. G.による(1994) Flow Cytometry Oxford, New York, NY.に記載されている。

【0270】

遺伝子発現におけるGPAMの影響は、GPAMをコードする配列とCD64またはCD64-GFPのどちら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)の保存された領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgGがCD64に対する抗体のどちらかで被覆された磁気ビードを用いて分離することができる(DYNAL, Lake Success, NY)。mRNAは、当分野で周知の方法で細胞から精製することができる。GPAM及び目的の他の遺伝子をコードするmRNAの発現は、ノーザン分析やマイクロアレイ技術で分析することができる。

【0271】

### 1.4 GPAMに特異的な抗体の作製

ポリアクリルアミドゲル電気泳動法 (PAGE; 例えば、Harrington, M. G. (1990) *Methods Enzymol.* 1816-3088-495を参照) または他の精製技術で実質的に精製されたGPAMを用いて、標準的なプロトコルでウサギを免疫化して抗体を作り出す。

#### 【0272】

別法では、GPAMアミノ酸配列をLASERGENEソフトウェア(DNA STAR)を用いて解析して免疫原性の高い領域を決定し、対応するオリゴペプチドを合成してこれを用いて当業者に周知の方法で抗体を生産する。C末端付近の、或いは隣接する親水性領域内のエピトープなどの適切なエピトープの選択については、当分野で周知である(例えば、前出のAusubel, 1995, 11章を参照)。

10

#### 【0273】

通常、約15残基の長さのオリゴペプチドを、Applied BiosystemsのABI 431Aペプチドシンセサイザー(PE Biosystems)を用いてfmoc法のケミストリにより合成し、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)を用いた反応によりKLH(Sigma-Aldrich, St. Louis MO)に結合させて、免疫原性を高める(例えば、前出のAusubel, 1995を参照)。フロイントの完全アジュバ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GPAM活性を検査するには、ペプチドまたはGPAMを基板に結合し、1%BSAを用いてブロッキング処理し、ウサギ抗血清と反応させて洗浄し、さらに放射性ヨウ素標識されたヤギ抗ウサギIgGと反応させる。

20

#### 【0274】

##### 1.5 特異的抗体を用いる天然GPAMの精製

天然GPAM或いは組換えGPAMを、GPAMに特異的な抗体を用いるイムノアフィニティークロマトグラフィー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、CNBr-活性化SEPHAROSE(Amersham Pharmacia Biotech)のような活性化クロマトグラフィー用レジンと抗GPAM抗体とを共有結合させることにより形成する。結合の後、そのレジンを製造者の使用説明書に従ってブロッキング処理し洗浄する。

#### 【0275】

GPAMを含む培養液をイムノアフィニティークラムに通し、GPAMを優先的に吸着できる条件で(例えば、界面活性剤の存在下において高イオン強度のバッファーで)そのカラムを洗浄する。そのカラムを、抗体とGPAMとの結合を切るような条件で(例えば、pH2~3のバッファー、或いは高濃度の尿素またはチオシアン酸塩イオンのようなカオトロピックイオンで)溶出させ、GPAMを回収する。

30

#### 【0276】

##### 1.6 GPAMと相互作用する分子の同定

GPAMまたは生物学的に活性なその断片を、<sup>125</sup>I Bolton-Hunter 試薬(例えば、Bolton A.E. 及びW.M. Hunter (1973) *Biochem. J.* 133:529を参照)で標識する。マルチウェルプレートに予め配列しておいた候補の分子を、標識したGPAMと共にインキュベートし、洗浄して、標識したGPAM複合体を有する全てのウェルをアッセイする。様々なGPAM濃度で得られたデータを用いて、候補分子と結合したGPAMの数量及び親和性、会合についての値を計算する。

40

#### 【0277】

別法では、GPAMと相互作用する分子を、Fields, S. 及びO. Song (1989, *Nature* 340:245-246)に記載の酵母2-ハイブリッドシステム(yeast two-hybrid system)やMATCHMAKERシステム(Clontech)などの2-ハイブリッドシステムに基づいた市販のキットを用いて分析する。

50

## 【0278】

G P A Mはまた、ハイスルーブット型の酵母2ハイブリッドシステムを使用するP A T H C A L L I N Gプロセス(CuraGen Corp., New Haven CT)に用いて、遺伝子の2つの大きなライブラリによって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. 他(2000) 米国特許第6,057,101号)。

## 【0279】

## 17 G P A Mの活性の実証

G P A Mの活性のアッセイは、リガンド/受容体仲介性の細胞増殖の変化を調べるプロトタイプアッセイに基づいている。このアッセイでは、Swissマウス3T3細胞におけるDNAの合成の速度を測定する。G P A Mをコードするポリヌクレオチドを含むプラスミドを、当分野で周知のトランスフェクション法で静止状態の3T3培養細胞に加える。一過性にトランスフェクトされた細胞を、放射性DNA前駆体分子である[<sup>3</sup>H]チミジンの存在下でインキュベートする。次に、様々な量のG P A Mリガンドを培養細胞に加える。ラジオアイソトープカウンタを用いて、酸沈降性DNAに取り込まれた[<sup>3</sup>H]チミジンの量を適当な時間測定する。取り込まれた量が新規に合成されたDNAの量に正比例する。少なくとも100倍のG P A Mリガンドの濃度範囲に対して線形の線量効果曲線が、受容体活性を表す。1 ml当たりの活性1単位を、50%の反応レベルを引き起こすG P A Mの濃度と定義する。この場合、100%は、[<sup>3</sup>H]チミジンが酸沈降性DNAに最大に取り込まれたことを表す(McKay, I. および I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY, p. 73)。

10

20

## 【0280】

当業者であれば、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うことができるであろう。特定の好適な実施例に基づいて本発明を説明したが、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、分子生物学或いは関連する分野の専門家には明らかな、本明細書に記載の本発明の実施例の様々な改変は、特許請求の範囲に含まれる。

30

## 【0281】

(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に対する系統的な名称を示す。

## 【0282】

表2は、本発明のポリペプチドに最も近いGenBankの相同体のGenBankの識別番号およびアノテーションを示す。ポリペプチドとそのGenBankの相同体との間の一致を表す確率値スコアも示す。

## 【0283】

表3は、推定上のモチーフおよびドメインを含むポリペプチド配列の構造的な特徴、並びにポリペプチドの分析に用いた方法、アルゴリズム、および検索可能なデータベースを示す。

40

## 【0284】

表4は、ポリヌクレオチド配列の組み立てに用いたcDNA断片のリスト、並びにポリヌクレオチド配列の選択された断片のリストを示す。

## 【0285】

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。

## 【0286】

表6は、表5に示すcDNAライブラリの作製に用いた組織およびベクターを示す付録である。

## 【0287】

50

表 7 は、本発明のポリヌクレオチド及びポリペプチドの分析に用いたツール、プログラム、及びアルゴリズム、並びにその説明、引用文献、閾値パラメータを示す。

【表 1】

表 1

インサイトプロジェクト ID	ポリペプチド SEQ ID NO:	インサイトポリペプチド ID	ポリヌクレオチド SEQ ID NO:	インサイトポリヌクレオチド ID
4620239	1	4620239CD1	3	4620239CB1
5565933	2	5565933CD1	4	5565933CB1

10

20

30

40

【表 2】

表2

ポリペプチド SEQ ID NO:	インサイト ポリペプチドID	GenBank ID NO:	確率スコア	GenBank 相関係
1	4620239	g2933332	0.0	ect2 [ハツカネズミ] Miki, T. (1993) Nature 362:452-465
2	5565933	g2677836	1.50E-08	SEL-10 [線虫] Hubbard, E.J. (1997) Genes Dev. 22:3182-3193

10

20

30

40

【表 3】

表3-1

SEQ ID NO:	インサイトポリペプチド ID	アミノ酸残基数	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(signature)配列、ドメイン及びモチーフ	分析方法及びデータベース
1	4620239CD1	914	S406 T614 T645 T833 S834 T838 S15 S20 S40 S112 T222 S282 T287 S415 S504 S531 S588 T614 S622 S630 S680 T774 S868 S897 Y256 Y352 Y656 Y805 T11 T71 S158 S167 S211 S326 S376 T432 S439 S451 S630 S668 T713 S726 T846		RhoGEEドメイン: V456-E640 BRCA1 C末端 (BRCT)ドメイン: N178-Y354 Gds_Cdc24: L589-T614 シグナル切断: M1-A17 DM08580 P52735 1 do VAV; キナーゼ亜鉛 SH2; 491; S442-R707 PD115498:D641-I914 ECT2 タンパク質 オンコジーン グアニンスクレオチド放出因子 プロトオンコジーン PD009898:L195-L390 タンパク質 ECT2 オンコジーン グアニンスクレオチド放出因子 プロトオンコジーン T19E10 PD000777:V456-E640 タンパク質 因子 グアニンスクレオチド放出 スクレオチド グアニン交換 プロトオンコジーン 結合 SH3 PD124854:S391-Q455 ECT2 タンパク質 オンコジーン グアニンスクレオチド放出因子 プロトオンコジーン DM08397 A49307 26-564: T463-D641 BCR タンパク質	HMMER-PFAM MOTIFS SPSCAN BLAST_DOMO BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM

10

20

30

40

【表 4】

表3-2

SEQ ID NO:	インサイトポリペプチド ID	アミノ酸残基数	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(signature)配列、ドメイン及びモチーフ	分析方法及びデータベース
2	5565933CD1	459	T84 T122 T135 S345 S442 T23 S59 S98 T271 T290 T320 S339 S406	N389	WDドメイン, G-βリポド WD40:Y33-K432 G_βリポド I106-G120 L337-L351 シグナル切断: M1-L21 β-トランスフェュンファミリ- TRP-ASPリポド DM00005   P36037   208-251: H322-W349 P_値 0.0064 Go MS11; 膜; 反復性: YDR128W; DM00299   S61017   273-320: L309-R352 P_値 0.0064	HMMER_PFAM MOTIFS SPSCAN BLAST_DOMO BLAST_DOMO

10

20

30

40

表4

ポリスクレオチド SEQ ID NO:	インサイト ポリスクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
3	4620239CB1	4097	2591-3339, 1-25, 661-754, 312-403, 1248-1981	1272429F1 (TESTTUT02) SCJA00029V1 2347162H1 (TESTTUT02) 1317546T6 (BLADTUT02) 2347162F6 (TESTTUT02) 2951135F6 (KIDNFET01) 2708093H1 (PONSAT01) 2807188X304F1 (BLADTUT08) SCJA01465V1 SCJA00665V1 SCJA00864V1 1313418F1 (BLADTUT02) SCJA02003V1 2807188X308F1 (BLADTUT08) 2790863F6 (COLNTUT16) 4199368H1 (COLLITUT02) 5678392H1 (293TF2T01) 903697R6 (COLNNOT07) 1444644F1 (THYRNOT03) 3956567H1 (HEARFET02) 5502250H1 (BRABDIR01) 5532203H1 (HEARFET05) 3800458H1 (SPLNNT012) 4240334H1 (SYNWDIT01)	3040 2285 1 3075 6 3684 318 332 1093 1582 1181 3810 2483 675 1800 1046 535 804 1320 690 1 641 1232 300 261	3634 2887 232 3783 562 4094 621 809 1665 2139 1757 4097 3071 1154 2358 1322 804 1641 1307 278 908 1494 606 602
4	5565933CB1	1641	1-262			

10

20

30

40

表5

ポリヌクレオチド SEQ ID NO:	インサイト プロジェクトID	代表的ライブラリ
3	4620239CB1	BLADTUT02
4	5565933CB1	KIDNNOT05

10

20

30

【表 7】

表 8

ライブラリ	ベクター	ライブラリの説明
BLADTUT02	pINCY	このライブラリは、80歳白人女性の根治膀胱切除及びリンパ節切除の際に採取した膀胱腫瘍組織から単離したRNAを用いて作製した。病理学的には、グレード3の浸潤性移行上皮癌であった。家族歴には、腎不全、変形性関節症、及びアテローム性動脈硬化症が含まれていた。
KIDNIN05	PSPORT1	このライブラリは、脳無酸素症で死亡した生後2日のヒスバニック系女児の腎臓組織から単離したRNAを用いて作製した。家族歴には、先天性の心疾患があった。

【表 8】

10

20

30

表7-1

プログラム名	説明	引用文献	パラメーター閾値
ABI FACTURA	核酸配列においてベクター配列を除去して不特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及び注釈付け (annotation) に有用である。	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler BLAST	核酸配列を構築するプログラム。 Basic Local Alignment Search Tool は、アミノ酸及び核酸配列の配列類似性検索に有用であり、blastp及びblastn、blastx、tblastn、tblastxの5つのファンクションがある。	Applied Biosystems, Foster City, CA. Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10 以下
FASTA	Pearson及びLipmanアルゴリズムは、間合わせの配列と同種の配列群との類似性を検索する。FASTAは、fasta及びtfasta、fastx、tfastx、ssearchの少なくとも5つのファンクションを含む。	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E値=1.06E-6 構築された ESTs: fasta 同一性=95%以上、一致長さ=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	BLOCKS IMPROVED Searcher は、BLOCKS 及び PRINTS、DOMO、PRODOM、PFAM データベースにおける配列に対して、ある配列の一致性を調べ、遺伝子ファミリー及び配列相同性、構造的ファンクション領域を探索する。	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:5565-72, 1991. J.C. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; Artwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	確率値=1.0E-3 以下
HMMER	PFAM などのタンパク質ファミリーデータベース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間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; Sonnhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322. Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM ヒット: 確率値=1.0E-3 以下 シングルレベブプロトタイプスコア=0 以上

10

20

30

表7-2

プログラム名	説明	引用文献	パラメーター関連値
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における構造及び配列のモチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, 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≧特定の Prositeモチーフに対する GCG 指定 "HIGH" 値 一般に、スコア=1.4-2.1
Phred	高い感度及び精度で自動配列決定機のプロセスを調べる塩基誤出しアルゴリズムである。	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 プログラムであって、配列相同性の検索及び DNA 配列の構築に有用である。	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; Green, P., University of Washington, Seattle, WA.	スコア=120 以上 一致長さ=56 以上
Consed	Phrap で構築したものの表示及び編集をするためのグラフィックツールである。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスクランして分岐シグナルペプチドの存在を調べる重み付けマトリクス解析プログラムである。	Nielson, H. 他 (1997) Protein Engineering 10: 1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	スコア=3.5 以上
TMAP	重み付けマトリクスを用いたタンパク質配列上の隣置過セグメントの明確化および向きを決定するためのプログラム	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れマルコフモデル (HMM) を用いたタンパク質配列上の隣置過セグメントの明確化および向きを決定するためのプログラム。	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, pp. 175-182.	
Motifs	Prosite で定義された配列と一致したパターンについてアミノ酸配列を検索するプログラムである。	Bairoch 他、(1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

## 【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  
27 September 2001 (27.09.2001)

PCT

(10) International Publication Number  
WO 01/70807 A2

- (51) International Patent Classification: C07K 14/47, C12N 15/12, 15/66, 5/10, A01K 67/027, C07K 16/18, C12N 15/11, C12Q 1/68, A61K 38/17, A61P 37/00
- (21) International Application Number: PCT/US01/09220
- (22) International Filing Date: 22 March 2001 (22.03.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/191,631 23 March 2000 (23.03.2000) US
- (71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (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, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian 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, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

## Published:

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



- (72) Inventors; and  
Inventors/Applicants (for US only): LAL, Preeti [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US); YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US); LI, Dying, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (US).

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

WO 01/70807 A2

(54) Title: G-PROTEIN ASSOCIATED MOLECULES

(57) Abstract: The invention provides human G-protein associated molecules (GPAM) and polynucleotides which identify and encode GPAM. 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 GPAM.

WO 01/70807

PCT/US01/09220

**G-PROTEIN ASSOCIATED MOLECULES****TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of G-protein associated molecules and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative and autoimmune/inflammation disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of G-protein associated molecules.

**BACKGROUND OF THE INVENTION**

Signal transduction is the process of biochemical events by which cells respond to extracellular signals. Extracellular signals are transduced through a biochemical cascade that begins with the binding of a signal molecule such as a hormone, neurotransmitter, or growth factor, to a cell membrane receptor and ends with the activation of an intracellular target molecule. The process of signal transduction regulates a wide variety of cell functions including cell proliferation, differentiation, and gene transcription.

G-proteins bind guanine nucleotides, either directly or in association with other molecules. Guanine nucleotide binding proteins (GTP-binding proteins), are critical mediators of the signal transduction pathway. Extracellular ligands such as hormones, growth factors, neuromodulators, or other signaling molecules bind to transmembrane receptors, and the signal is propagated to effector molecules by intracellular signal transducing proteins. Many of these signal transduction proteins are GTP-binding proteins which regulate intracellular signaling pathways. GTP-binding proteins participate in a wide range of other regulatory functions including metabolism, growth, differentiation, cytoskeletal organization, and intracellular vesicle transport and secretion. Exchange of bound GDP for GTP followed by hydrolysis of GTP to GDP provides the energy that enables GTP-binding proteins to alter their conformation and interact with other cellular components. Two structurally distinct classes of GTP-binding proteins are recognized: heterotrimeric GTP-binding proteins, consisting of three different subunits, and monomeric, low molecular weight (LMW), GTP-binding proteins consisting of a single polypeptide chain.

Heterotrimeric GTP-binding proteins are composed of 3 subunits ( $\alpha$ ,  $\beta$  and  $\gamma$ ) which, in their inactive conformation, associate as a trimer at the inner face of the plasma membrane.  $G_\alpha$  binds GDP or GTP and contains the GTPase activity. The  $\beta\gamma$  complex enhances binding of  $G_\alpha$  to a receptor.  $G_\gamma$  is necessary for the folding and activity of  $G_\beta$ . (Neer, E.J. et al. (1994) Nature 371:297-300.) Multiple homologs of each subunit have been identified in mammalian tissues, and different

WO 01/70807

PCT/US01/09220

combinations of subunits have specific functions and tissue specificities. (Spiegel, A.M. (1997) *J. Inher. Metab. Dis.* 20:113-121.) G protein activity is triggered by seven-transmembrane cell surface receptors (G-protein coupled receptors) which respond to lipid analogs, amino acids and their derivatives, peptides, cytokines, and specialized stimuli such as light, taste, and odor. Activation of the receptor by its stimulus causes the replacement of the G protein-bound GDP with GTP.  $G_{\alpha}$ -GTP dissociates from the receptor and the  $\beta\gamma$  complex and each of these separated components can bind and regulate downstream effectors. The signaling stops when  $G_{\alpha}$  hydrolyzes its bound GTP to GDP and reassociates with the  $\beta\gamma$  complex (Neer, supra).

The  $\beta$  subunits of heterotrimeric G proteins, also known as G beta proteins or  $\beta$  transducins, contain seven tandem repeats of the WD-repeat sequence motif, a motif found in many proteins with regulatory functions. WD-repeat proteins contain from four to eight copies of a loosely conserved repeat of approximately 40 amino acids which participates in protein-protein interactions. Mutations and variant expression of  $\beta$  transducin proteins are linked with various disorders. Mutations in LIS1, a subunit of the human platelet activating factor acetylhydrolase, cause Miller-Dieker lissencephaly. RACK1 binds activated protein kinase C, and RbAp48 binds retinoblastoma protein. CstF is required for polyadenylation of mammalian pre-mRNA *in vitro* and associates with subunits of cleavage-stimulating factor. CD4, an integral membrane glycoprotein which functions as an HIV co-receptor for infection of human host cells is degraded by HIV-encoded Vpu in the endoplasmic reticulum. WD repeats of human beta TrCP molecule mediate the formation of the CD4-Vpu complex, inducing CD4 proteolysis (Neer, E.J. et al. (1994) *Nature* 371:297-300 and Margottin, F. et al. (1998) *Mol. Cell.* 1:565-574).

Irregularities in the GPCR signaling cascade may result in abnormal activation of leukocytes and lymphocytes, leading to the tissue damage and destruction seen in many inflammatory and autoimmune diseases such as rheumatoid arthritis, biliary cirrhosis, hemolytic anemia, lupus erythematosus, and thyroiditis. Abnormal cell proliferation, including cyclic AMP stimulation of brain, thyroid, adrenal, and gonadal tissue proliferation is regulated by G proteins. Mutations in  $G_{\alpha}$  subunits have been found in growth-hormone-secreting pituitary somatotroph tumors, hyperfunctioning thyroid adenomas, and ovarian and adrenal neoplasms (Meij, J.T.A. (1996) *Mol. Cell. Biochem.* 157:31-38; Aussen, C. et al. (1988) *J. Immunol.* 140:215-220).

LMW GTP-binding proteins regulate cell growth, cell cycle control, protein secretion, and intracellular vesicle interaction. They consist of single polypeptides which, like the alpha subunit of the heterotrimeric GTP-binding proteins, are able to bind to and hydrolyze GTP, thus cycling between an inactive and an active state. LMW GTP-binding proteins respond to extracellular signals from receptors and activating proteins by transducing mitogenic signals involved in various cell functions.

WO 01/70807

PCT/US01/09220

The binding and hydrolysis of GTP regulates the response of LMW GTP-binding proteins and acts as an energy source during this process (Bokoch, G. M. and Der, C. J. (1993) *FASEB J.* 7:750-759).

At least sixty members of the LMW GTP-binding protein superfamily have been identified and are currently grouped into the six subfamilies of ras, rho, arf, sar1, ran, and rab. Activated ras genes were initially found in human cancers and subsequent studies confirmed that ras function is critical in determining whether cells continue to grow or become differentiated. Other members of the LMW G-protein superfamily have roles in signal transduction that vary with the function of the activated genes and the locations of the GTP-binding proteins that initiate the activity. Rho GTP-binding proteins control signal transduction pathways that link growth factor receptors to actin polymerization, which is necessary for normal cellular growth and division. The rab, arf, and sar1 families of proteins control the translocation of vesicles to and from membranes for protein localization, protein processing, and secretion. Ran GTP-binding proteins are located in the nucleus of cells and have a key role in nuclear protein import, the control of DNA synthesis, and cell-cycle progression (Hall, A. (1990) *Science* 249:635-640; Barbacid, M. (1987) *Ann. Rev. Biochem.* 56:779-827; and Sasaki, T. and Takai, Y. (1998) *Biochem. Biophys. Res. Commun.* 245:641-645).

LMW GTP-binding proteins are GTPases which cycle between a GTP-bound active form and a GDP-bound inactive form. This cycle is regulated by proteins that affect GDP dissociation, GTP association, or the rate of GTP hydrolysis. Proteins affecting GDP association are represented by guanine nucleotide dissociation inhibitors and guanine nucleotide exchange factors (GEF). The best characterized is the mammalian homolog of the *Drosophila* Son-of-Sevenless protein. Proteins affecting GTP hydrolysis are exemplified by GTPase-activating proteins (GAP). Both GEF and GAP activity may be controlled in response to extracellular stimuli and modulated by accessory proteins such as RalBPI and POB1. The GDP-bound form is converted to the GTP-bound form through a GDP/GTP exchange reaction facilitated by guanine nucleotide-releasing factors. The GTP-bound form is converted to the GDP-bound form by intrinsic GTPase activity, and the conversion is accelerated by GAP (Ikeda, M. et al. (1998) *J. Biol. Chem.* 273:814-821; Quilliam, L. A. (1995) *Bioessays* 17:395-404.). Mutant Ras-family proteins, which bind but can not hydrolyze GTP, are permanently activated, and cause cell proliferation or cancer, as do GEFs that activate LMW GTP-binding proteins (Drivas, G. T. et al. (1990) *Mol. Cell. Biol.* 10:1793-1798; and Whitehead, I. P. et al. (1998) *Mol Cell Biol.* 18:4689-4697.)

The discovery of new G-protein associated molecules 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 and autoimmune/inflammation disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of G-

WO 01/70807

PCT/US01/09220

protein associated molecules.

#### SUMMARY OF THE INVENTION

5 The invention features purified polypeptides, G-protein associated molecules, referred to collectively as "GPAM" and individually as "GPAM-1," and "GPAM-2." 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:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence  
10 selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-2.

15 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:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the  
20 group consisting of SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-2. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:3-4.

25 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 of SEQ ID NO:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of  
30 SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. 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

WO 01/70807

PCT/US01/09220

consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. The method comprises a) 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.

10 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 selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2.

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:3-4, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 80% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, 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.

25 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 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 80% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, 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

WO 01/70807

PCT/US01/09220

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  
5 polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 80% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, 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  
10 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  
15 the group consisting of SEQ ID NO:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having  
20 an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional GPAM, 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  
25 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:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and  
30 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. 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

WO 01/70807

PCT/US01/09220

treating a disease or condition associated with decreased expression of functional GPAM, comprising administering to a patient in need of such treatment the composition.

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:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. 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 GPAM, 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:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. 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:1-2, b) a naturally occurring polypeptide comprising an amino acid sequence at least 80% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. 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

WO 01/70807

PCT/US01/09220

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 compound is indicative of a compound that modulates the activity of the polypeptide.

5 The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:3-4, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

10 The invention further provides a method for assessing toxicity of a test compound, said method comprising 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 selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, ii) a  
15 naturally occurring polynucleotide comprising a polynucleotide sequence at least 80% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, 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  
20 sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 80% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide  
25 of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment 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  
30 biological sample is indicative of toxicity of the test compound.

#### BRIEF DESCRIPTION OF THE TABLES

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

35 Table 2 shows the GenBank identification number and annotation of the nearest GenBank

WO 01/70807

PCT/US01/09220

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.

Table 4 lists the cDNA 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.

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

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

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.

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.

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 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

"GPAM" refers to the amino acid sequences of substantially purified GPAM obtained from any

WO 01/70807

PCT/US01/09220

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 GPAM. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other  
5 compound or composition which modulates the activity of GPAM either by directly interacting with GPAM or by acting on components of the biological pathway in which GPAM participates.

An "allelic variant" is an alternative form of the gene encoding GPAM. 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  
10 many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. 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 GPAM include those sequences with deletions,  
15 insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as GPAM or a polypeptide with at least one functional characteristic of GPAM. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding GPAM, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding GPAM.

The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of  
20 amino acid residues which produce a silent change and result in a functionally equivalent GPAM. 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 GPAM is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include  
25 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,  
30 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.

WO 01/70807

PCT/US01/09220

"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 GPAM. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of GPAM either by directly interacting with GPAM or by acting on components of the biological pathway in which GPAM 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 GPAM polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or 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 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 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 "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 sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having 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

WO 01/70807

PCT/US01/09220

functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic GPAM, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

5 "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. 10 Compositions comprising polynucleotide sequences encoding GPAM or fragments of GPAM 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; 15 SDS), and other components (e.g., Denhardt'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 20 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 25 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
30	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
35	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu

WO 01/70807

PCT/US01/09220

	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
5	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
10	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.

A "fragment" is a unique portion of GPAM or the polynucleotide encoding GPAM 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, 16, 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:3-4 comprises a region of unique polynucleotide sequence that

WO 01/70807

PCT/US01/09220

specifically identifies SEQ ID NO:3-4, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:3-4 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:3-4 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:3-4 and the region of SEQ ID NO:3-4 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:1-2 is encoded by a fragment of SEQ ID NO:3-4. A fragment of SEQ ID NO:1-2 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-2. For example, a fragment of SEQ ID NO:1-2 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-2. The precise length of a fragment of SEQ ID NO:1-2 and the region of SEQ ID NO:1-2 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 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

WO 01/70807

PCT/US01/09220

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/bl2.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*

15 *Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 11*

*Filter: on*

20 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, 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.

30 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,

WO 01/70807

PCT/US01/09220

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 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.

"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.

"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.

WO 01/70807

PCT/US01/09220

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 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, 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 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>3</sub>t or R<sub>3</sub>t analysis) or formed between one 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).

WO 01/70807

PCT/US01/09220

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.

"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 GPAM which is 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 GPAM 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, 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 GPAM. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of GPAM.

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.

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

"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.

"Post-translational modification" of an GPAM 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

WO 01/70807

PCT/US01/09220

cell type depending on the enzymatic milieu of GPAM.

"Probe" refers to nucleic acid sequences encoding GPAM, their complements, or fragments 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  
5 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 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).

10 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, 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  
15 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 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  
20 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 MA).

Oligonucleotides for use as primers are selected using software known in the art for such  
25 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  
30 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

WO 01/70807

PCT/US01/09220

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  
5 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  
10 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  
15 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.

20 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 (UTRs).  
25 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 radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and  
30 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.

WO 01/70807

PCT/US01/09220

The term "sample" is used in its broadest sense. A sample suspected of containing GPAM, nucleic acids encoding GPAM, 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.

5 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  
10 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  
15 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,  
20 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  
25 cell. Transformation may occur under natural or artificial conditions according to various methods 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"  
30 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

WO 01/70807

PCT/US01/09220

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 alternative 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 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

WO 01/70807

PCT/US01/09220

certain defined length of one of the polypeptides.

#### THE INVENTION

The invention is based on the discovery of new human G-protein associated molecules (GPAM), the polynucleotides encoding GPAM, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative and autoimmune/inflammation disorders.

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.

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.

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 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.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are G-protein associated molecules. For example, SEQ ID NO:1 is 93% identical, from residue M176 to residue I914, to mouse oncogene *ect2* (GenBank ID g293332) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The

WO 01/70807

PCT/US01/09220

BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a RhoGEF 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 MOTIFS and additional BLAST analyses provide further corroborative evidence that SEQ ID NO:1 is a guanine nucleotide exchange factor. SEQ ID NO:2 was analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-2 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:3-4 or that distinguish between SEQ ID NO:3-4 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 sequences in column 5 relative to their respective full length sequences.

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, 1272429F1 is the identification number of an Incyte cDNA sequence, and TESTTUT02 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., SCIA00029V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.) 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. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) 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

WO 01/70807

PCT/US01/09220

the relevant Incyte cDNA identification numbers are not shown.

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.

The invention also encompasses GPAM variants. A preferred GPAM 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 GPAM amino acid sequence, and which contains at least one functional or structural characteristic of GPAM.

The invention also encompasses polynucleotides which encode GPAM. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:3-4, which encodes GPAM. The polynucleotide sequences of SEQ ID NO:3-4, 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.

The invention also encompasses a variant of a polynucleotide sequence encoding GPAM. 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 GPAM. 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:3-4 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 ID NO:3-4. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of GPAM.

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 GPAM, 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 GPAM, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode GPAM and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring GPAM under appropriately selected

WO 01/70807

PCT/US01/09220

conditions of stringency, it may be advantageous to produce nucleotide sequences encoding GPAM 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 GPAM 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 GPAM and GPAM 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 GPAM 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:3-4 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 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 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.)

The nucleic acid sequences encoding GPAM may be extended utilizing a partial nucleotide

WO 01/70807

PCT/US01/09220

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.)

5 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 al.

10 (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

15 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 68°C to

20 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 into 5'

25 non-transcribed regulatory regions.

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

30 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.

WO 01/70807

PCT/US01/09220

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode GPAM may be cloned in recombinant DNA molecules that direct expression of GPAM, 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 GPAM.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter GPAM-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 Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of GPAM, 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 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 family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding GPAM 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, GPAM 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.

WO 01/70807

PCT/US01/09220

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 GPAM, 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  
5 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.)

10 In order to express a biologically active GPAM, the nucleotide sequences encoding GPAM 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  
15 encoding GPAM. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding GPAM. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding GPAM and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be  
20 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.*  
25 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding GPAM and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory  
30 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.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding GPAM. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with

WO 01/70807

PCT/US01/09220

yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); 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., T1 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 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* 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.)

15 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 GPAM. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding GPAM 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 GPAM 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 *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 GPAM are needed, e.g. for the production of antibodies, vectors which direct high level expression of GPAM may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

20

Yeast expression systems may be used for production of GPAM. A number of vectors 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*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

30

WO 01/70807

PCT/US01/09220

Plant systems may also be used for expression of GPAM. Transcription of sequences encoding GPAM 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. 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, 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 GPAM 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 infective virus which expresses GPAM 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 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 GPAM in cell lines is preferred. For example, sequences encoding GPAM 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 selective agent, and its presence allows growth and recovery of cells which successfully express the 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 *apr* 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

WO 01/70807

PCT/US01/09220

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 chlorosulfuron 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.)

5 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

10 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 GPAM is inserted within a marker gene sequence, transformed cells containing

15 sequences encoding GPAM can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding GPAM 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 GPAM and that express

20 GPAM 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 GPAM using either

25 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 GPAM is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, 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.)

30

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

WO 01/70807

PCT/US01/09220

or PCR probes for detecting sequences related to polynucleotides encoding GPAM include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding GPAM, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and  
5 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 ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as  
10 well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding GPAM 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 and/or the vector used. As will be understood by those of skill in the art, expression vectors containing  
15 polynucleotides which encode GPAM may be designed to contain signal sequences which direct secretion of GPAM 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 the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation,  
20 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 Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing  
25 of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding GPAM 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 GPAM protein containing a heterologous moiety that can be recognized by a commercially available antibody may  
30 facilitate the screening of peptide libraries for inhibitors of GPAM 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

WO 01/70807

PCT/US01/09220

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

5 located between the GPAM encoding sequence and the heterologous protein sequence, so that GPAM 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.

10 In a further embodiment of the invention, synthesis of radiolabeled GPAM 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.

GPAM of the present invention or fragments thereof may be used to screen for compounds  
15 that specifically bind to GPAM. At least one and up to a plurality of test compounds may be screened for specific binding to GPAM. 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 GPAM, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a  
20 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 GPAM 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 GPAM, either as a secreted  
25 protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing GPAM or cell membrane fractions which contain GPAM are then contacted with a test compound and binding, stimulation, or inhibition of activity of either GPAM or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is  
30 detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with GPAM, either in solution or affixed to a solid support, and detecting the binding of GPAM 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

WO 01/70807

PCT/US01/09220

libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

GPAM of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of GPAM. Such compounds may include agonists, antagonists, or partial  
5 or inverse agonists. In one embodiment, an assay is performed under conditions permissive for GPAM activity, wherein GPAM is combined with at least one test compound, and the activity of GPAM in the presence of a test compound is compared with the activity of GPAM in the absence of the test compound. A change in the activity of GPAM in the presence of the test compound is indicative of a compound that modulates the activity of GPAM. Alternatively, a test compound is combined with an  
10 in vitro or cell-free system comprising GPAM under conditions suitable for GPAM activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of GPAM 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 GPAM or their mammalian homologs may be  
15 "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 human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 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  
20 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 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).  
25 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 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 GPAM may also be manipulated in vitro in ES cells derived from  
30 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 into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding GPAM can also be used to create "knockin" humanized animals

WO 01/70807

PCT/US01/09220

(pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding GPAM is injected into animal ES cells, and the injected sequence 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 GPAM, e.g., by secreting GPAM in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

#### THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of GPAM and G-protein associated molecules. Therefore, GPAM appears to play a role in cell proliferative and autoimmune/inflammation disorders. In the treatment of disorders associated with increased GPAM expression or activity, it is desirable to decrease the expression or activity of GPAM. In the treatment of disorders associated with decreased GPAM expression or activity, it is desirable to increase the expression or activity of GPAM.

Therefore, in one embodiment, GPAM 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 GPAM. 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; and an autoimmune/inflammation disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, 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

WO 01/70807

PCT/US01/09220

sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, inflammation caused by hemodialysis and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and inflammation caused by trauma.

5 In another embodiment, a vector capable of expressing GPAM 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 GPAM including, but not limited to, those described above.

10 In a further embodiment, a composition comprising a substantially purified GPAM 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 GPAM including, but not limited to, those provided above.

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

15 In a further embodiment, an antagonist of GPAM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GPAM. Examples of such disorders include, but are not limited to, those cell proliferative and autoimmune/inflammation disorders described above. In one aspect, an antibody which specifically binds GPAM 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 GPAM.

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

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

30 An antagonist of GPAM may be produced using methods which are generally known in the art. In particular, purified GPAM may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind GPAM. Antibodies to GPAM 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

WO 01/70807

PCT/US01/09220

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 GPAM or with any fragment or oligopeptide thereof  
5 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.

10 It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to GPAM 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 GPAM amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric  
15 molecule may be produced.

Monoclonal antibodies to GPAM 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.  
20 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.  
25 Natl. 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 GPAM-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,  
30 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.)

WO 01/70807

PCT/US01/09220

Antibody fragments which contain specific binding sites for GPAM 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  
5 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  
10 immunoassays typically involve the measurement of complex formation between GPAM and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering GPAM 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  
15 may be used to assess the affinity of antibodies for GPAM. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of GPAM-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 GPAM epitopes, represents the average affinity, or avidity, of the antibodies for GPAM. The  $K_a$   
20 determined for a preparation of monoclonal antibodies, which are monospecific for a particular GPAM 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 GPAM-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  
25 ultimately require dissociation of GPAM, preferably in active form, from the 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 the quality and suitability of such preparations for certain downstream applications. For example, a  
30 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 GPAM-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 Coligan et al. *supra*.)

WO 01/70807

PCT/US01/09220

In another embodiment of the invention, the polynucleotides encoding GPAM, 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 encoding GPAM.

5 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 GPAM. (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 sequences into appropriate target cells can be used. Antisense sequences can be delivered  
10 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 vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood*  
15 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.* 25(14):2730-2736.)

20 In another embodiment of the invention, polynucleotides encoding GPAM 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 immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency  
25 (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), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii)  
30 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), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides

WO 01/70807

PCT/US01/09220

*brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in GPAM expression or regulation causes disease, the expression of GPAM from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

5 In a further embodiment of the invention, diseases or disorders caused by deficiencies in GPAM are treated by constructing mammalian expression vectors encoding GPAM and introducing these vectors by mechanical means into GPAM-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vivo* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (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 GPAM include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), 15 PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). GPAM 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. USA* 20 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 25 gene encoding GPAM from a normal individual.

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 30 (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 GPAM expression are treated by constructing a retrovirus vector consisting of (i) the

WO 01/70807

PCT/US01/09220

polynucleotide encoding GPAM 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 PFBNEO) are

5 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

10 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 Number 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

15 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

20 polynucleotides encoding GPAM to cells which have one or more genetic abnormalities with respect to the expression of GPAM. 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

25 described in U.S. Patent Number 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 389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver

30 polynucleotides encoding GPAM to target cells which have one or more genetic abnormalities with respect to the expression of GPAM. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing GPAM 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 ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has

WO 01/70807

PCT/US01/09220

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 Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV  
5 d92 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 herpesvirus  
10 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  
15 deliver polynucleotides encoding GPAM 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, resulting  
20 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 GPAM into the alphavirus genome in place of the capsid-coding region results in the production of a large number of GPAM-coding RNAs and the synthesis of high levels of GPAM in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent  
25 infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) 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 GPAM 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 manipulating  
30 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 be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes

WO 01/70807

PCT/US01/09220

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.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A  
5 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,  
10 engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding GPAM.

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,  
15 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 by  
20 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 GPAM. 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 constructs that  
25 synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

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  
30 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

WO 01/70807

PCT/US01/09220

compound which is effective in altering expression of a polynucleotide encoding GPAM. 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 GPAM expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding GPAM may be therapeutically useful, and in the treatment of disorders associated with decreased GPAM expression or activity, a compound which specifically promotes expression of the polynucleotide encoding GPAM 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 GPAM 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 GPAM 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 GPAM. 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 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.

WO 01/70807

PCT/US01/09220

Patent No. 6,022,691).

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.

5 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.)

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  
10 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  
15 Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of GPAM, antibodies to GPAM, and mimetics, agonists, antagonists, or inhibitors of GPAM.

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,  
20 sublingual, or rectal means.

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),  
25 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 needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active  
30 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.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising GPAM or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the

WO 01/70807

PCT/US01/09220

macromolecule. Alternatively, GPAM 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 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).

5 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 route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

10 A therapeutically effective dose refers to that amount of active ingredient, for example GPAM or fragments thereof, antibodies of GPAM, and agonists, antagonists or inhibitors of GPAM, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose  
15 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  $LD_{50}/ED_{50}$  ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are 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  $ED_{50}$  with little or no toxicity.  
20 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 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  
25 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 biweekly depending on the half-life and clearance rate of the particular formulation.

30 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.

WO 01/70807

PCT/US01/09220

**DIAGNOSTICS**

In another embodiment, antibodies which specifically bind GPAM may be used for the diagnosis of disorders characterized by expression of GPAM, or in assays to monitor patients being treated with GPAM or agonists, antagonists, or inhibitors of GPAM. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for GPAM include methods which utilize the antibody and a label to detect GPAM 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 GPAM, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of GPAM expression. Normal or standard values for GPAM expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to GPAM under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of GPAM 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 GPAM 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 GPAM may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of GPAM, and to monitor regulation of GPAM levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding GPAM or closely related molecules may be used to identify nucleic acid sequences which encode GPAM. 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 probe identifies only naturally occurring sequences encoding GPAM, 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 GPAM 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:3-4 or from genomic sequences including promoters, enhancers, and introns of the GPAM gene.

WO 01/70807

PCT/US01/09220

Means for producing specific hybridization probes for DNAs encoding GPAM include the cloning of polynucleotide sequences encoding GPAM or GPAM 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

5 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 GPAM may be used for the diagnosis of disorders associated with expression of GPAM. Examples of such disorders include, but are not limited to, a cell

10 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,

15 ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an autoimmune/inflammation disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune

20 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

25 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, inflammation caused by hemodialysis and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. The

30 polynucleotide sequences encoding GPAM 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 GPAM expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding GPAM may be useful in assays that

WO 01/70807

PCT/US01/09220

detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding GPAM 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 GPAM 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 GPAM, 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 GPAM, 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 days to months.

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 GPAM 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 GPAM, or a fragment of a polynucleotide complementary to the polynucleotide encoding GPAM, and will be employed under optimized conditions for identification of a specific gene or condition.

WO 01/70807

PCT/US01/09220

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 GPAM may be used to detect single nucleotide polymorphisms (SNPs). SNPs are  
5 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 GPAM are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal  
10 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 amplicons in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP  
15 (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  
20 throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of GPAM 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. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be  
25 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.

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  
30 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 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

WO 01/70807

PCT/US01/09220

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 display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

5 In another embodiment, GPAM, fragments of GPAM, or antibodies specific for GPAM 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.

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  
10 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 Seifhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 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  
15 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 resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies,  
20 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 may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental  
25 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) *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.  
30 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 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

WO 01/70807

PCT/US01/09220

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 example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released 5 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 containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated 10 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 indicative of a toxic response caused by the test compound in the treated sample.

15 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 analyzed by quantifying the number of expressed proteins and their relative abundance under given 20 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 sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are 25 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 30 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.

WO 01/70807

PCT/US01/09220

A proteomic profile may also be generated using antibodies specific for GPAM to quantify the levels of GPAM 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 (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; 5 Mendoze, 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 10 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.

15 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 20 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 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 25 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.* 30 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.

WO 01/70807

PCT/US01/09220

In another embodiment of the invention, nucleic acid sequences encoding GPAM 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 GPAM 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, 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.

In another embodiment of the invention, GPAM, 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

WO 01/70807

PCT/US01/09220

solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between GPAM and the agent being tested may be measured.

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 GPAM, or fragments thereof, and washed. Bound GPAM is then detected by methods well known in the art. Purified GPAM can also be coated directly onto plates for use in the aforementioned drug screening techniques. 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 GPAM specifically compete with a test compound for binding GPAM. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with GPAM.

In additional embodiments, the nucleotide sequences which encode GPAM 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.

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, in particular U.S. Ser. No. 60/091,631, are expressly incorporated by reference herein.

#### EXAMPLES

##### 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.

WO 01/70807

PCT/US01/09220

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  
5 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 SUPERSOFT plasmid system (Life Technologies), using the  
10 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  
15 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., PBLUESOFT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells  
20 including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

## II. Isolation of cDNA Clones

Plasmids obtained as described in Example 1 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  
25 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, 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  
30 high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal 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

WO 01/70807

PCT/US01/09220

(Labsystems Oy, Helsinki, Finland).

**III. Sequencing and Analysis**

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  
5 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  
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).  
10 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  
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,  
15 *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 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  
20 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 databases such as PFAM.  
(HMM is a probabilistic approach which analyzes consensus primary structures of gene families.  
See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) The queries were  
25 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 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 Consed, and cDNA assemblages  
30 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 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,

WO 01/70807

PCT/US01/09220

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, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are  
5 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 Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold  
10 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 match between two sequences (the higher the score or the lower the probability value, the greater the identity  
15 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:3-4. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

#### 20 IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative G-protein associated molecules were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and S. Karlin (1998) *Curr. Opin. Struct. Biol.* 8:346-354). The program concatenates predicted exons to form an  
25 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 associated molecules, the encoded polypeptides were analyzed by querying  
30 against PFAM models for G-protein associated molecules. Potential G-protein associated molecules were also identified by homology to Incyte cDNA sequences that had been annotated as G-protein associated molecules. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri 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

WO 01/70807

PCT/US01/09220

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

##### 10 "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 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.

##### 30 "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

WO 01/70807

PCT/US01/09220

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.

#### VI. Chromosomal Mapping of GPAM Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:3-4 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:3-4 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, or 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 (<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

WO 01/70807

PCT/US01/09220

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:

5

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum}(\text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}))}$$

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 GPAM 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, 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 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 GPAM. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ

WO 01/70807

PCT/US01/09220

GOLD database (Incyte Genomics, Palo Alto CA).

#### VIII. Extension of GPAM 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  
5 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  
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  
10 structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one 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  
15 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 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  
20 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 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  
25 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  
(Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the  
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.

30 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

WO 01/70807

PCT/US01/09220

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:3-4 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 µCi 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 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 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

WO 01/70807

PCT/US01/09220

compared.

#### X. Microarrays

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 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) *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 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 the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

#### 25 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 reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM 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

WO 01/70807

PCT/US01/09220

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  
5 then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 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  
10 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 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope  
15 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 US  
20 Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, 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).  
25 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 µl of sample mixture consisting of 0.2 µg each of Cy3 and  
30 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  
35 140 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for

WO 01/70807

PCT/US01/09220

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  
5 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  
10 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  
15 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  
20 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  
25 fluorophores and 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  
30 signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and  
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  
35 to obtain a numerical value corresponding to the average intensity of the signal. The software used

WO 01/70807

PCT/US01/09220

for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

#### XI. Complementary Polynucleotides

Sequences complementary to the GPAM-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring GPAM. 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 GPAM. 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 GPAM-encoding transcript.

#### XII. Expression of GPAM

Expression and purification of GPAM is achieved using bacterial or virus-based expression systems. For expression of GPAM 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 (tac)* 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 GPAM upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of GPAM 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 GPAM 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, GPAM 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, 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 GPAM at

WO 01/70807

PCT/US01/09220

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 Ausubel (1995, *supra*, ch. 10 and 16). Purified GPAM obtained by these methods can be used directly in the assays shown in Examples XVI and XVII where applicable.

### XIII. Functional Assays

GPAM function is assessed by expressing the sequences encoding GPAM 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 GPAM on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding GPAM and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions 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

WO 01/70807

PCT/US01/09220

of mRNA encoding GPAM and other genes of interest can be analyzed by northern analysis or microarray techniques.

#### XIV. Production of GPAM Specific Antibodies

GPAM 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 GPAM 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 anti-peptide and anti-GPAM activity by, for example, binding the peptide or GPAM 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 GPAM Using Specific Antibodies

Naturally occurring or recombinant GPAM is substantially purified by immunoaffinity chromatography using antibodies specific for GPAM. An immunoaffinity column is constructed by covalently coupling anti-GPAM 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 GPAM are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of GPAM (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/GPAM 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 GPAM is collected.

#### XVI. Identification of Molecules Which Interact with GPAM

GPAM, or biologically active fragments thereof, are 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.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled GPAM, washed, and any wells with labeled GPAM complex are assayed. Data obtained using different concentrations of

WO 01/70807

PCT/US01/09220

GPAM are used to calculate values for the number, affinity, and association of GPAM with the candidate molecules.

Alternatively, molecules interacting with GPAM 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).

GPAM 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).

10 **XVII. Demonstration of GPAM Activity**

An assay for GPAM 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 GPAM 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 GPAM 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 GPAM ligand concentration range is indicative of receptor activity. One unit of activity per milliliter is defined as the concentration of GPAM producing a 50% response level, where 100% represents maximal incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY, p. 73.)

25 Various modifications and variations of the described methods and systems of the invention 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 following claims.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
4520319	1	4520319D1	3	4520319C1
565933	2	565933D1	4	565933C1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
1	4620239	g293332	0.0	ect2 [Mus musculus] Miki, T. et al. (1993) Nature 362:452-465
2	5565933	g2677836	1.50E-08	SEU-10 [Caenorhabditis elegans] Hubbard, E.J. (1997) Genes Dev. 22:3182-3193

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1.	462033CD1	914	S406 T614 T645 T633 S834 T638 S15 S20 S40 S112 T222 S282 T287 S415 S504 S531 S588 T614 S622 S630 S680 T774 S868 S897 Y256 Y352 Y656 Y605 T11 T71 S158 S167 S211 S326 S376 T432 S439 S451 S630 S668 T713 S726 T846		RhoGEF domain: Y456-E640 BRCA1 C terminus (BRCT) domain: N178-Y354 G4s C6c24: L589-T614 Signal cleavage: M1-A17 DM08580 P52735 1 do VAV; KINASE; ZINC; SH2; 491: S442-R707 P1115498-D641-I914 ECT2 PROTEIN ONCOGENE GUANINENUCLEOTIDE RELEASING FACTOR PROTOONCOGENE P1009898-L195-L390 PROTEIN ECT2 ONCOGENE GUANINENUCLEOTIDE RELEASING FACTOR PROTOONCOGENE T19E1D P1000777-Y456-E640 PROTEIN FACTOR GUANINENUCLEOTIDE RELEASING NUCLEOTIDE GUANINE EXCHANGE PROTOONCOGENE BINDING SH3 P031485(L-S191-L455) ECT2 PROTEIN ONCOGENE GUANINENUCLEOTIDE RELEASING FACTOR PROTOONCOGENE P089337 P49307 26-564: P463-D641 PCR PROTEIN	HMMER-PPAM MOTIFS SPSCAN BLAST_DOMO BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM

WO 01/70807

PCT/US01/09220

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
2	5565933CD1	459	T84 T122 T135 S345 S442 T23 S59 S98 T271 T290 T320 S339 S406	N399	WD domain, G-beta repeat WD40-Y33-K432 G Beta Repeats I106-G120 L337-L351 Sigma1_cleaveage:MI-E21 SEPA-TRANSUCIN FAMILY TRP-ASP REPEATS DM00005 P36037 268-251:H322-W549 E2 value 0.0064 G-MS1A: MEMBRANE, REPETITIVE, YDR128W; DM00299 S61017 273-320:L309-R353 E2 value 0.0064	EMBR, PFAM MOTIFS SPSCAN ELAST_DOMO ELAST_DOMO

Table 4

Polynucleotide Seq ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
3	462023CB1	4097	2561-3339, 1-23, 661-754, 312-403, 1248-1981	127123DH1 (TESPTT02) SC7A0203V1 2347165H1 (TESPTT02) 1317346T6 (BLADTT02) 2347162F6 (TESPTT02) 6 2951135F6 (KJDNFT01) 2708093H1 (PONSAT01) 318 2807188X304F1 (BLADTT08) 332 SC7A01465V1 SC7A00655V1 SC7A00864V1 1313418F1 (BLADTT02) 3810 SC7A0203V1 2807188X308F1 (BLADTT08) 675 2790863F6 (COLMPT16) 1800 4199369H1 (COLMPT16) 1046 5678392H1 (293TF201) 535 203687F6 (COLMPT01) 1861 1865 3954555H1 (BLADTT02) 570 5502250H1 (BLADTT01) 641 5522203H1 (BLADTT05) 1232 3800458H1 (SPLNNT012) 300 4240334H1 (SYANDTT01) 261	3040 288 289 3075 6 3684 621 1093 1665 1582 1181 4097 2483 1154 2358 1322 1861 1861 570 308 1494 606 602	1654 287 232 3783 562 4094 621 809 1665 2139 1757 4097 3071 1154 2358 1322 1861 1861 570 308 1494 606 602
4	556593CB1	1641	1-262	4199369H1 (COLMPT16) 1800 5678392H1 (293TF201) 535 203687F6 (COLMPT01) 1861 1865 3954555H1 (BLADTT02) 570 5502250H1 (BLADTT01) 641 5522203H1 (BLADTT05) 1232 3800458H1 (SPLNNT012) 300 4240334H1 (SYANDTT01) 261	1800 1046 535 1861 1861 570 308 1494 606 602	1861 1861 570 308 1494 606 602

WO 01/70807

PCT/US01/09220

Table 5

Polymer SEQ. ID NO:	Incyte Project ID	Representative Library
3	462039CB1	BLAD/VV102
4	556593CB1	KIDN/OT05

WO 01/70807

PCT/US01/09220

Table 6

Library	Vector	Library Description
BLADVT02	pLNCY	Library was constructed using RNA isolated from bladder tumor tissue removed from an 80-year-old Caucasian female during a radical cystectomy and lymph node excision. Pathology indicated grade 3 invasive transitional cell carcinoma. Family history included acute renal failure, osteoarthritis, and atherosclerosis.
KIDNMT03	PSFORT1	Library was constructed using RNA isolated from the kidney tissue of a 2-day-old Hispanic male, who died from cerebral anoxia. Family history included congenital heart disease.

WO 01/70807

PCT/US01/09220

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.	
ABIPARACEL.FDF	A Fast Data Finder useful in comparing and associating 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) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 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, tfastn, fastx, tfastx, and search.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci. USA</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183:63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 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 scores=100 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) <i>Nucleic Acids Res.</i> 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-105; and Atwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 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) <i>J. Mol. Biol.</i> 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) <i>Nucleic Acids Res.</i> 26:320-322; Durbin, R. et al. (1998) <i>Our World View, in a Nutshell</i> , Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Scores= 0 or greater

WO 01/70807

PCT/US01/09220

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; Baroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores:CCG-specified "HIGH" 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.	Scores=120 or greater; Match length=56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Godton, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nicolson, H. et al. (1997) Protein Engineering 10:1-6; Clavette, J.M. and S. Audic (1997) CABIOS 12:431-439.	Scores=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.	Baroch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 01/70807

PCT/US01/09220

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
  - a) a polypeptide comprising an amino acid sequence selected from the group consisting of  
5 SEQ ID NO:1-2,
  - b) a naturally occurring polypeptide comprising an amino acid sequence at least 80%  
identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2,
  - c) a biologically active fragment of a polypeptide having an amino acid sequence selected  
from the group consisting of SEQ ID NO:1-2, and
  - 10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from  
the group consisting of SEQ ID NO:1-2.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-2.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID  
20 NO:3-4.
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.
9. A method for producing a polypeptide of claim 1, the method comprising:
  - 30 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 01/70807

PCT/US01/09220

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide selected from the group consisting of:
- 5 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4,  
b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 80% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4,  
c) a polynucleotide complementary to the polynucleotide of a),  
d) a polynucleotide complementary to the polynucleotide of b), and  
10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
- 15 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, 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  
20 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.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
- 25 15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and  
30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
- 35

WO 01/70807

PCT/US01/09220

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-2.
18. A method for treating a disease or condition associated with decreased expression of functional GPAM, comprising administering to a patient in need of such treatment the composition of claim 16.
19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:
- 10 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and  
b) detecting agonist activity in the sample.
20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.
- 15 21. A method for treating a disease or condition associated with decreased expression of functional GPAM, comprising administering to a patient in need of such treatment a composition of claim 20.
- 20 22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and  
b) detecting antagonist activity in the sample.
- 25 23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.
24. A method for treating a disease or condition associated with overexpression of functional GPAM, comprising administering to a patient in need of such treatment a composition of claim 23.
- 30 25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:
- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and  
35 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a

WO 01/70807

PCT/US01/09220

compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- 5 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  
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  
10 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.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method

- 15 comprising:  
a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,  
b) detecting altered expression of the target polynucleotide, and  
c) comparing the expression of the target polynucleotide in the presence of varying amounts of  
20 the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

- 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  
25 least 20 contiguous nucleotides of a polynucleotide of claim 11 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 11 or fragment thereof;  
c) quantifying the amount of hybridization complex; and  
30 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.

WO 01/70807

PCT/US01/09220

29. A diagnostic test for a condition or disease associated with the expression of GPAM in a biological sample comprising the steps of:

- a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and  
5 b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

- a) a chimeric antibody,  
10 b) a single chain antibody,  
c) a Fab fragment,  
d) a F(ab')<sub>2</sub> fragment, or  
e) a humanized antibody.

15 31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of GPAM in a subject, comprising administering to said subject an effective amount of the composition of claim  
20 31.

33. A composition of claim 31, wherein the antibody is labeled.

34. A method of diagnosing a condition or disease associated with the expression of GPAM in a subject, comprising administering to said subject an effective amount of the composition of claim  
25 33.

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from  
30 the group consisting of SEQ ID NO:1-2, or an immunogenic fragment thereof, under conditions to elicit an antibody response;  
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

WO 01/70807

PCT/US01/09220

group consisting of SEQ ID NO:1-2.

36. An antibody produced by a method of claim 35.

5 37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

- 10 a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
- b) isolating antibody producing cells from the animal;
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;
- 15 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 ID NO:1-2.

39. A monoclonal antibody produced by a method of claim 38.

20

40. A composition comprising the antibody of claim 39 and a suitable carrier.

41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.

25

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2 in a sample, comprising the steps of:

- 30 a) incubating the antibody of claim 10 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:1-2 in

WO 01/70807

PCT/US01/09220

the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2 from a sample, the method comprising:
- 5 a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
- 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:1-2.
- 10 45. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.
46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.
47. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:3.
- 15 48. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:4.

WO 01/70807

PCT/US01/09220

<110> INCYTE GENOMICS, INC.  
 LAL, Preeti  
 YUE, Henry  
 LU, Dyoung Aina M.

<120> G-PROTEIN ASSOCIATED MOLECULES

<130> PF-0762 PCT

<140> To Be Assigned  
 <141> Herewith

<150> 60/091.631  
 <151> 2000-03-23

<160> 4  
 <170> PERL Program

<210> 1  
 <211> 914  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4620239CD1

<400> 1  
 Met Ala Glu Asn Ser Val Leu Thr Ser Thr Thr Gly Arg Thr Ser  
 1 5 10 15  
 Leu Ala Asp Ser Ser Ile Phe Asp Ser Lys Val Thr Glu Ile Ser  
 20 25 30  
 Lys Glu Asn Leu Leu Ile Gly Ser Thr Ser Tyr Val Glu Glu Glu  
 35 40 45  
 Met Pro Gln Ile Glu Thr Arg Val Ile Leu Val Gln Glu Ala Gly  
 50 55 60  
 Lys Gln Glu Glu Leu Ile Lys Ala Leu Lys Thr Ile Lys Ile Met  
 65 70 75  
 Glu Val Pro Val Ile Lys Ile Lys Glu Ser Cys Pro Gly Lys Ser  
 80 85 90  
 Asp Glu Lys Leu Ile Lys Ser Val Ile Asn Met Asp Ile Lys Val  
 95 100 105  
 Gly Phe Val Lys Met Glu Ser Val Glu Glu Phe Glu Gly Leu Asp  
 110 115 120  
 Ser Pro Glu Phe Glu Asn Val Phe Val Val Thr Asp Phe Gln Asp  
 125 130 135  
 Ser Val Phe Asn Asp Leu Tyr Lys Ala Asp Cys Arg Val Ile Gly  
 140 145 150  
 Pro Pro Val Val Leu Asn Cys Ser Gln Lys Gly Glu Pro Leu Pro  
 155 160 165  
 Phe Ser Cys Arg Pro Leu Tyr Arg Thr Ser Met Met Asn Leu Val  
 170 175 180  
 Leu Cys Phe Thr Gly Phe Arg Lys Lys Glu Glu Leu Val Arg Leu  
 185 190 195  
 Val Thr Leu Val His His Met Gly Gly Val Ile Arg Lys Asp Phe  
 200 205 210  
 Ser Ser Lys Val Thr His Leu Val Ala Asn Cys Thr Gln Gly Glu  
 215 220 225  
 Lys Phe Arg Val Ala Val Ser Leu Gly Thr Pro Ile Met Lys Pro  
 230 235 240  
 Glu Trp Ile Tyr Lys Ala Trp Glu Arg Arg Asn Glu Gln Asp Phe  
 245 250 255  
 Tyr Ala Ala Val Asp Asp Phe Arg Asn Glu Phe Lys Val Pro Pro  
 260 265 270  
 Phe Gln Asp Cys Ile Leu Ser Phe Leu Gly Phe Ser Asp Glu Glu  
 275 280 285  
 Lys Thr Asn Met Glu Glu Met Thr Glu Met Gln Gly Gly Lys Tyr

WO 01/70807

PCT/US01/09220

290	295	300
Leu Pro Leu Gly Asp	Glu Arg Cys Thr His	Leu Val Val Glu
305	310	315
Asn Ile Val Lys Asp	Leu Pro Phe Glu Pro	Ser Lys Lys Leu
320	325	330
Val Val Lys Gln Glu	Trp Phe Trp Gly Ser	Ile Gln Met Asp
335	340	345
Arg Ala Gly Glu Thr	Met Tyr Leu Tyr Glu	Lys Ala Asn Thr
350	355	360
Glu Leu Lys Lys Ser	Val Ser Met Leu Ser	Leu Asn Thr Pro
365	370	375
Ser Asn Arg Lys Arg	Arg Arg Leu Lys Glu	Thr Leu Ala Gln
380	385	390
Ser Arg Glu Thr Asp	Val Ser Pro Phe Pro	Pro Arg Lys Arg
395	400	405
Ser Ala Glu His Ser	Leu Ser Ile Gly Ser	Leu Leu Asp Ile
410	415	420
Asn Thr Pro Glu Ser	Ser Ile Asn Tyr Gly	Asp Thr Pro Lys
425	430	435
Cys Thr Lys Ser Ser	Lys Ser Ser Thr Pro	Val Pro Ser Lys
440	445	450
Ser Ala Arg Trp Gln	Val Ala Lys Glu Leu	Tyr Gln Thr Glu
455	460	465
Asn Tyr Val Asn Ile	Leu Ala Thr Ile Ile	Gln Leu Phe Gln
470	475	480
Pro Leu Glu Glu Glu	Gly Gln Arg Gly Gly	Pro Ile Leu Ala
485	490	495
Glu Glu Ile Lys Thr	Ile Phe Gly Ser Ile	Pro Asp Ile Phe
500	505	510
Val His Thr Lys Ile	Lys Asp Asp Leu Glu	Asp Leu Ile Val
515	520	525
Trp Asp Glu Ser Lys	Ser Ile Gly Asp Ile	Phe Leu Lys Tyr
530	535	540
Lys Asp Leu Val Lys	Thr Tyr Pro Pro Phe	Val Asn Phe Phe
545	550	555
Met Ser Lys Glu Thr	Ile Ile Lys Cys Glu	Lys Gln Lys Pro
560	565	570
Phe His Ala Phe Leu	Lys Ile Asn Gln Ala	Lys Pro Glu Cys
575	580	585
Arg Gln Ser Leu Val	Glu Leu Leu Ile Arg	Pro Val Gln Arg
590	595	600
Pro Ser Val Ala Leu	Leu Leu Asn Asp Leu	Lys Lys His Thr
605	610	615
Asp Glu Asn Pro Asp	Lys Ser Thr Leu Glu	Lys Ala Ile Gly
620	625	630
Leu Lys Glu Val Met	Thr His Ile Asn Glu	Asp Lys Arg Lys
635	640	645
Glu Ala Gln Lys Gln	Ile Phe Asp Val Val	Tyr Glu Val Asp
650	655	660
Cys Pro Ala Asn Leu	Leu Ser Ser His Arg	Ser Leu Val Gln
665	670	675
Val Glu Thr Ile Ser	Leu Gly Glu His Pro	Cys Asp Arg Gly
680	685	690
Gln Val Thr Leu Phe	Leu Phe Asn Asp Cys	Leu Glu Ile Ala
695	700	705
Lys Arg His Lys Val	Ile Gly Thr Phe Arg	Ser Pro His Gly
710	715	720
Thr Arg Pro Pro Ala	Ser Leu Lys His Ile	His Leu Met Pro
725	730	735
Ser Gln Ile Lys Lys	Val Leu Asp Ile Arg	Glu Thr Glu Asp
740	745	750
His Asn Ala Phe Ala	Leu Leu Val Arg Pro	Pro Thr Glu Gln
755	760	765
Asn Val Leu Leu Ser	Phe Gln Met Thr Ser	Asp Glu Leu Pro
770	775	780
Glu Asn Trp Leu Lys	Met Leu Cys Arg His	Val Ala Asn Thr
785	790	795

WO 01/70807

PCT/US01/09220

Cys Lys Ala Asp Ala Glu Asn Leu Ile Tyr Thr Ala Asp Pro Glu  
 800 805 810  
 Ser Phe Glu Val Asn Thr Lys Asp Met Asp Ser Thr Leu Ser Arg  
 815 820 825  
 Ala Ser Arg Ala Ile Lys Lys Thr Ser Lys Lys Val Thr Arg Ala  
 830 835 840  
 Phe Ser Phe Ser Lys Thr Pro Lys Arg Ala Leu Arg Arg Ala Leu  
 845 850 855  
 Met Thr Ser His Gly Ser Val Glu Gly Arg Ser Pro Ser Ser Asn  
 860 865 870  
 Asp Lys His Val Met Ser Arg Leu Ser Ser Thr Ser Ser Leu Ala  
 875 880 885  
 Gly Ile Pro Ser Pro Ser Leu Val Ser Leu Pro Ser Phe Phe Glu  
 890 895 900  
 Arg Arg Ser His Thr Leu Ser Arg Ser Thr Thr His Leu Ile  
 905 910

&lt;210&gt; 2

&lt;211&gt; 459

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5565933CD1

&lt;400&gt; 2

Met Leu Arg Trp Leu Ile Gly Gly Gly Arg Glu Pro Gln Gly Leu  
 1 10 15  
 Ala Glu Lys Ser Pro Leu Gln Thr Ile Gly Glu Glu Gln Thr Gln  
 20 25 30  
 Asn Pro Tyr Thr Glu Leu Leu Val Leu Lys Ala His His Asp Ile  
 35 40 45  
 Val Arg Phe Leu Val Gln Leu Asp Asp Tyr Arg Phe Ala Ser Ala  
 50 55 60  
 Gly Asp Asp Gly Ile Val Val Val Trp Asn Ala Gln Thr Gly Glu  
 65 70 75  
 Lys Leu Leu Glu Leu Asn Gly His Thr Gln Lys Ile Thr Ala Ile  
 80 85 90  
 Ile Thr Phe Pro Ser Leu Glu Ser Cys Glu Glu Lys Asn Gln Leu  
 95 100 105  
 Ile Leu Thr Ala Ser Ala Asp Arg Thr Val Ile Val Trp Asp Gly  
 110 115 120  
 Asp Thr Thr Arg Gln Val Gln Arg Ile Ser Cys Phe Gln Ser Thr  
 125 130 135  
 Val Lys Cys Leu Thr Val Leu Gln Arg Leu Asp Val Trp Leu Ser  
 140 145 150  
 Gly Gly Asn Asp Leu Cys Val Trp Asn Arg Lys Leu Asp Leu Leu  
 155 160 165  
 Cys Lys Thr Ser His Leu Ser Asp Thr Gly Ile Ser Ala Leu Val  
 170 175 180  
 Glu Ile Pro Lys Asn Cys Val Val Ala Ala Val Gly Lys Glu Leu  
 185 190 195  
 Ile Ile Phe Arg Leu Val Ala Pro Thr Glu Gly Ser Leu Glu Trp  
 200 205 210  
 Asp Ile Leu Glu Val Lys Arg Leu Leu Asp His Gln Asp Asn Ile  
 215 220 225  
 Leu Ser Leu Ile Asn Val Asn Asp Leu Ser Phe Val Thr Gly Ser  
 230 235 240  
 His Val Gly Glu Leu Ile Ile Trp Asp Ala Leu Asp Trp Thr Met  
 245 250 255  
 Gln Ala Tyr Glu Arg Asn Phe Trp Asp Pro Ser Pro Gln Leu Asp  
 260 265 270  
 Thr Gln Gln Glu Ile Lys Leu Cys Gln Lys Ser Asn Asp Ile Ser  
 275 280 285  
 Ile His His Phe Thr Cys Asp Glu Glu Asn Val Phe Ala Ala Val  
 290 295 300

WO 01/70807

PCT/US01/09220

Gly Arg Gly Leu Tyr Val Tyr Ser Leu Gln Met Lys Arg Val Ile  
 305 310 315  
 Ala Cys Gln Lys Thr Ala His Asp Ser Asn Val Leu His Ile Ala  
 320 325 330  
 Arg Leu Pro Asn Arg Gln Leu Ile Ser Cys Ser Glu Asp Gly Ser  
 335 340 345  
 Val Arg Ile Trp Glu Leu Arg Glu Lys Gln Gln Leu Ala Ala Glu  
 350 355 360  
 Pro Val Pro Thr Gly Phe Phe Asn Met Trp Gly Phe Gly Arg Val  
 365 370 375  
 Ser Lys Gln Ala Ser Gln Pro Val Lys Lys Gln Gln Glu Asn Ala  
 380 385 390  
 Thr Ser Cys Ser Leu Glu Leu Ile Gly Asp Leu Ile Gly His Ser  
 395 400 405  
 Ser Ser Val Glu Met Phe Leu Tyr Phe Glu Asp His Gly Leu Val  
 410 415 420  
 Thr Cys Ser Ala Asp His Leu Ile Ile Leu Trp Lys Asn Gly Glu  
 425 430 435  
 Arg Glu Ser Gly Leu Arg Ser Leu Arg Leu Phe Gln Lys Leu Glu  
 440 445 450  
 Glu Asn Gly Asp Leu Tyr Leu Ala Val 455

<210> 3  
 <211> 4097  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4620239CB1

<400> 3  
 gccgcgggga ggaatggggg tatttggtag aggagtcggc gtttgaagag gtggaactcc 60  
 tagggctttt ttgagagtgc tgatttagaa gaatacaaat catggctgaa aatagtgat 120  
 taacttcaac taatgggagg actagctctc cagactcttc cattttgat tctaaagtta 180  
 ctgagatttc caaggaaac ttacttattg gatctacttc atatgtagaa gaagagatgc 240  
 ctcaagttga aacaagagtg atattggttc aagaagctgg aaaacaagaa gaactataa 300  
 aagccttaaa gactattaaa ataatggaag tccctgttat aaagataaaa gaaagtgtc 360  
 ctggaaaatc ggatgaaaaa ttaataaaaa gtgttattaa tatggacatt aaagtgggct 420  
 ttgtaaatg gtagtcggtg gaagaatttg aaggtttgga ttctccggaa ttgaaaaatg 480  
 tatttttagt cacggacttc caggattcag tctttaaata cctctacaag gctgatgta 540  
 gagttattgg acacagagtt gatttaaat gtctacaaaa aggaagcct ttggcatttt 600  
 catgtctgcc gttgtatagt acaagatgga tgaatctagt actatgcttt acggattta 660  
 ggaaaaaaga agaactagtc aggttgggtga cattggtcca tcacatgggt ggagttattc 720  
 gaaaagactt tagtcaaaa gttacacatt tgggtggcaa ttgtacacaa ggagaaaaat 780  
 tcagggttgc tgtgagtcta ggtactccaa ttatgaagcc agaatggatt tataaagctt 840  
 gggaaagcgc gaatgaacag gatttctatg cagcagttga tgactttaga aatgaattta 900  
 aagttctctc atttcaagat tgtattttaa gtttctctgg atttccagat gaagagaaaa 960  
 ccaatattga agaatgact gaantgcaag gaggtaaata ttaccgctt ggagatgaaa 1020  
 gatgcactca cctgtagtt gaagagaata tagtaaaaga tcttcccttt gaacctcaa 1080  
 agaaacttta tggttcaag caagagtggt tctggggaag cattcaaatg gatcccgag 1140  
 ctggagaaac tatgtattta tatgaaagc caaatactcc tgagctcaag aaatcagttg 1200  
 caatgctttc tctaaatacc cctaacagca atcgcaaac agctcgttta aaagaaacac 1260  
 ttgctcagct ttcaagagag acagacgtgt caccatttcc accccgtaag gcacctcag 1320  
 ctgagcattc cctttccata ggytcactcc tagatatctc caacacacca gagtctagca 1380  
 ttaactatgg agacacccca aagtcttgta ctaagtcttc taaaagotcc actccagttc 1440  
 ctccaagca gtcagcaagg tggcaagttg caaaagagct ttatcaaat gaaagtaatt 1500  
 atgttaatat attggaaca atattcagtt tatttcaagt accattggaa gaggaaagac 1560  
 aacgtggtgg acctatcctt gcaccagagg agattaagac tatttttggt agcattccag 1620  
 atatctttga tgtacacact aagataaagg atgatcttga agacctata gttaatggc 1680  
 atgagagcaa aagcattggt gacatttttc tgaataatcc aaaaatttg gtaaaaacct 1740  
 accctccctt tgaactttc ttgaaatga gcaagyaaac aattattaaa tgtgaaaaac 1800  
 agaaaccaag atttcatgct tttctcaaga taaaccaaac aaaaaccgaa ttgggacggc 1860  
 agagccttgt tgaacttctt atccgaccag tacagaggtt acccaggtt gcattacttt 1920  
 taaatgatct taagaagcat acagctgatg aaaaatccaga caaaagcact ttgaaaaaag 1980  
 ctattggatc actgaaggaa gtaatgacgc atattaatga ggataagaga aaaacagaag 2040

WO 01/70807

PCT/US01/09220

```

ctcaaaagca aatTTTTgat gttgtttatg aagtagatgg atgcccagct aatcttttat 2100
ctcttcaccg aagcttagta cagcgggttg aaacaatttc tctaggtgag caccctctgt 2160
acagaggaga acaagtaact ctcttccctc tcaatgattg cctagagata gcaagaaaac 2220
ggcacagggt tattggcact tctaggagtc ctcatggcca aaccgcagcc ccagctctcc 2280
tcaagataat tcaactaatg cctctttctc agatbaagaa ggtattggag ataaagagaa 2340
cagaagattg coataatgct tttgccttgc tttgtagggc accaacagag cagcgaatg 2400
tgctactcag tttccagatg acatcagatg aecttccaaa agaaaactgg ctaaagatgc 2460
tgtgtgcaca tgtagctaac accatttcta aagcagatgc tgagaatctt atttatactg 2520
ctgatccaga atcctttgaa gtaaatcaaa aagatattga cagtagatct agtagagcat 2580
caagagcaat aaaaaagact tcaaaaaaag ttacaagagc attctcttcc tccaaaactc 2640
caaaaagagc tttcgaaggg gctcttatga catcccaggg ctcatgtggg ggaagaaagtc 2700
ctctccagcaa tgataagcat gtaatgagtc gtcttctctg cacatcaca ttagcaggtta 2760
tcccactctc ctcccttctc agccttctct cctctcttgg aagggagaat ctaacgttaa 2820
tgatgctctc aactcatttg atatgaagcg ttacaaaact cttaaalatt agaaatata 2880
agacccctca tactcaataa agaaaactgac ttaaatggta ctgttaatta goacttgggtg 2940
aaagctggaa ggaagataaa taacactaaa ctatgctatt tgattttctc tcttgaaga 3000
gtaaggttta cctgtttacat tttcaagtta attcatgtaa aaaaatgata tgattttgat 3060
gtaatttacc tcttgtttga atctgtcatt caaaggccaa taatttaagt tgcctacgc 3120
tgatattagt agctttgcaa ccttgataga gtaaaataat tttatgggtg ggtgccaat 3180
actgctgtga atctatttgt atagtatcca tgaaatgaa tttggaaata gatatctgtg 3240
cagctcaatt tatgcagaga ttaaatgaca tcaataact ggtgaaaac ttgcataagc 3300
ttctgattaa atagtgggtc tgtttccact gtcagtttg aagtatttaa ataaccctc 3360
ctttcacagt ttttttctt ctaacggctt tcaagatctc agcatgtgga ttttaaaaga 3420
tttgcctca ttaacaagaa taacatttaa aggagattgt tcaaaaat ttttgaact 3480
tgagataagg acagaagaat tgagaacaat tgtatatttt gcaaaaacaa gatgtttgta 3540
gctgtttcag agagagtacg gtatatttat ggttaattta tccactagca aatcttgatt 3600
tagtttgata gtgtgtggaa ttttattttg aaggataaga ccatgggaaa atgttggtaa 3660
agactgtttg taaccttcat gaaataactc tggagtggc atcagttta cttaactct 3720
tgaaatgca tagatatgct catgttcaac ttttbatgt ggtcttataa ttaaatgtaa 3780
aatgaaaat tcaatttctg tttcaagatg tgatattctt cacaatagcc tttttatagt 3840
cagtaattca gaataatcaa gttcattgg ataaatgcat ttttattcc tatttctta 3900
gggagtgcta caaatgtttg tcacttaaat tcaagtttc tgttttaata gtttaactgac 3960
tatagattgt tttctatgcc atgtatgtgc cactctogag agtagtaaat gactctttgc 4020
tacattttaa aagcaattgt attagtaaga actttgtaaa taaataccta aaaccnagt 4080
gtaaaaaaa aaaaaaa 4097

```

```

<210> 4
<211> 1641
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5565933CB1

```

```

<400> 4
gccccgccg tggccgccg caccccaagc gactgcccac actaagcctc cgtggctggg 60
tacgggagcg ctttggggac aaaaattctc cctcaactgt ggtctgcatt cctctggccc 120
gtgggtgtag ctggggcggg aagtattagc gtctcaagttg cctctgagcc ggggaggaag 180
gagggagccg agcctggggc ggagtttggg ctgactgggg ctggaccggg caagacggcg 240
ccgtctgccg gatgtggcga tggctgatcg gggagggcgg agaaaccgag gaaactctct 300
agaaatctcc tttacagaca ataggtggaag aacaaaacca gaatccctac actgactgac 360
tagactgaa ggctcactat gatattgtac gatttctggt acagttagat gactacagat 420
tgcatctgc tgggatgat ggaatgttag ttgtgtggaa tgcccagaca ggggaaaaac 480
ttttagaact gaatggacac actcaaaaga taacagctat tattacattc cctctcttgc 540
aatctgtgta agagaaaaat caactcactc tgaacagcctc tgctgataga acagttattg 600
tgtgggatgg tgatactacc agacaagttc agagaatato atgcttccag tctactgtaa 660
agttcttaac tttctctcag agactagatg tttggcttcc tggtaggaat gacctgtgtg 720
tgtggaaccg aaaaatagat ctctgtgtga agactagcca ccttctgcat acaggtatta 780
gtgctttgtg tgaatacctc aagaactgtg tigtggcagc agttggcaaa gaactgata 840
ttttcaggtt ggtagcacc cagaagatg cactagaatg ggtatctct gaagttaagc 900
gectccttga tcaaccagat aatattctct catgtatgaa tgtcaatgat ttgagttttg 960
tcaccggctc ccacgtcgga gagctgatca tctgggatgc cctggactgg acaatgagc 1020
cctatgaacg caactcttgg gaccatctcc cacaaactgga caccacaaca gaaataaac 1080
tctgtcaaaa atcaaatgac atttctatc atcatttcc acgtgatgaa gagaatgata 1140
tctgtcagtt tggaaagggg ttatacgtgt atagccttca aatgaaagctg gtagttgctc 1200
gccagaaaac tgcacatgac tccaatgttc tgcacattgc cagacttcca aacaggcagt 1260

```

WO 01/70807

PCT/US01/09220

```
taatctcatg otcagaagat ggcagtgac gcatttggga gttaagagaa aaacagcagc 1320
ttgcagctga gcoctgtacca acagggtttt ttaacatgtg gggatttggg agagtcagca 1380
aacaagccag ccaacctggtt aaaagcagc agaaaaatgc tacttcatgt tcactggagc 1440
ttattggaga ttgatggga cactcatcat ctgtggagat gttttatatac ttggaagtc 1500
atggactagt gacgtgtctc gctgacatc tcattatctt gtggaaaat ggagagcag 1560
aatctggatt ggcaggttta agattatttc aaaattaga ggagaatggt gactatacc 1620
ttgctgtcta gtttaaggaa t 1641
```

## 【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  
27 September 2001 (27.09.2001)

PCT

(10) International Publication Number  
WO 01/70807 A3

- (51) International Patent Classification: C07K 14/47, C12N 15/12, 15/66, 5/10, A01K 67/027, C07K 16/18, C12N 15/11, C12Q 1/68, A61K 38/17, A61P 37/00
- (21) International Application Number: PCT/US01/09220
- (22) International Filing Date: 22 March 2001 (22.03.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/191,631 23 March 2000 (23.03.2000) 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): LAL, Preeti [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LU, Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (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, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published: with international search report
- (88) Date of publication of the international search report: 21 February 2002
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/70807 A3

(54) Title: G-PROTEIN ASSOCIATED MOLECULES

(57) Abstract: The invention provides human G-protein associated molecules (GPAM) and polynucleotides which identify and encode GPAM. 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 GPAM.

## 【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

		International Application No. PCT/US 01/09220	
A. CLASSIFICATION OF SUBJECT MATTER			
IPC 7	C07K14/47 C07K16/18	C12N15/12 C12N15/11	C12N15/66 C12Q1/68
		C12N5/10 A61K38/17	A01K67/027 A61P37/00
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 A61K			
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, PAJ, BIOSIS, EMBL			
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] EMBL:A1916675, 30 July 1999 (1999-07-30) Database accession no. A1916675 XP002178391 tu89e03.x1 NCI CGAP Gas4 Homo sapiens cDNA clone IMAGE:2258236~3' similar to TR:Q07139 Q07139 ECT2 ONCOGENE ; , mRNA sequence. EST. "Cancer Genome Anatomy Project (CGAP)" abstract  ---  -/--		1-15, 45-48
<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 claim or other special reason (as specified) "O" document relating 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. "Z" document member of the same patent family	
Date of the actual completion of the international search		Date of mailing of the international search report	
27 September 2001		08 10 01	
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL -2280 HV Rijswijk Tel: (+31-70) 340-2040, Tx. 31 651 epo rx. Fax: (+31-70) 340-3016		Authorized officer  Mateo Rosell, A.M.	

Form PCT/ISA210 (second sheet) (July 1992)

## INTERNATIONAL SEARCH REPORT

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		International Application No. PCT/US 01/09220
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL [Online] EMBL:HS1212277, 13 May 1997 (1997-05-13) Database accession no. AA416834 XP002178392 zu08f03.r1 Soares testis NHT Homo sapiens cDNA clone IMAGE:731261 5', mRNA sequence. EST. Homo sapiens abstract</p> <p>---</p>	1-15, 45-48
X	<p>DATABASE EMBL [Online] EMBL:AK001323, 22 February 2000 (2000-02-22) Database accession no. AK001323 XP002178393 Homo sapiens cDNA FLJ10461 fis, clone NT2RP1001482, Isogai et al., "WEDO human cDNA sequencing project" abstract</p> <p>---</p>	1-15, 45-48
X	<p>DATABASE EMBL [Online] EMBL:AK001766, 22 February 2000 (2000-02-22) Database accession no. AK001766 XP002178394 Homo sapiens cDNA FLJ10904 fis, clone OVARC1000013, weakly similar to APOPTOTIC PROTEASE ACTIVATING FACTOR 1. Homo sapiens. abstract</p> <p>---</p>	1-15, 45-48
X	<p>DATABASE EMBL [Online] SWALL:ECT2 MOUSE, 15 July 1999 (1999-07-15) Database accession no. Q07139 XP002178395 cited in the application ECT2 PROTEIN (ECT2 ONCOGENE). Mus musculus. abstract</p> <p>---</p>	1-15, 45-48
A	<p>WO 98 57990 A (ONYX PHARMA INC) 23 December 1998 (1998-12-23) page 3, line 15 -page 4, line 2 page 6, line 20 -page 28, line 27</p> <p>---</p>	1-48
A	<p>WO 98 37196 A (LUDWIG INST CANCER RES) 27 August 1998 (1998-08-27) page 1, line 1 -page 2, line 11 page 21, line 24 -page 33, line 30</p> <p>---</p> <p style="text-align: center;">-/--</p>	1-48

6

Form PCT/ISA210 (continuation of second sheet) (July 1992)

## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US 01/09220

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SASAKI TAKUYA ET AL: "The rho small G protein family-Rho GDI system as a temporal and spatial determinant for cytoskeletal control." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 245, no. 3, 28 April 1998 (1998-04-28), pages 641-645, XP002178390 ISSN: 0006-291X cited in the application the whole document	1
A	OVERBECK A F ET AL: "GUANINE NUCLEOTIDE EXCHANGE FACTORS: ACTIVATORS OF RAS SUPERFAMILY PROTEINS" MOLECULAR REPRODUCTION AND DEVELOPMENT, NEW YORK, NY, US, vol. 42, 1995, pages 468-476, XP000910820 the whole document	1
P,X	DATABASE EMBL [Online] SWALL:Q9H8V3, 1 March 2001 (2001-03-01) Database accession no. Q9H8V3 XP002178396 CDNA FLJ13205 FIS, CLONE NT2RP3004534, HIGHLY SIMILAR TO MOUSE ONCOGENE (ECT2) MRNA (FRAGMENT), Homo sapiens. Isogai et al., "NEDO human cDNA sequencing project" abstract	1-15, 45-48
P,X	DATABASE EMBL [Online] SWALL:Q9NV62, 1 October 2000 (2000-10-01) Database accession no. Q9NV62 XP002178399 HYPOTHETICAL 51.7 KDA PROTEIN. Homo sapiens. Isogai et al., "NEDO human cDNA sequencing project." abstract	1-15, 45-48
P,X	EP 1 074 617 A (HELIX RES INST) 7 February 2001 (2001-02-07) abstract page 3 (SEQ.ID.N. 871, 3842, 5425, 11405, 11406, 11810, 12323, 12324, 12398, 12399, 13647, 13648, 15238, 15239, 15954, 15955). & DATABASE EMBL [Online] 26 June 2001 (2001-06-26) Database accession nos: AAH04036;AH07007;AAH08590;AAH14169; AAB92847;AAH14386;AAH14655;AAB93279; AAH14693;AAB93316;AAH15428;AAB93830; AAH16330;AAB94519;AAH16753;AAB94812. ---	1-15, 45-48
	-/--	

6

Form PCT/ISA210 (continuation of second sheet) (July 1992)

## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US 01/09220

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>DATABASE EMBL [Online]            SWALL:Q9NSV8, 1 October 2000 (2000-10-01)            Database accession no. Q9NSV8            XP002178398            HYPOTHETICAL 33.4 KDA PROTEIN (FRAGMENT).            DKFZP434C0523. Homo sapiens            abstract</p> <p style="text-align: center;">---</p>	1-15, 45-48
T	<p>DATABASE EMBL [Online]            SWALL:BAB55317, 14 June 2001 (2001-06-14)            Database accession no. BAB55317            XP002178400            CDNA FLJ14807 FIS, CLONE NT2RP4001760,            WEAKLY SIMILAR TO PUTATIVE            DE RHO/RAC GUANINE NUCLEOTIDE EXCHANGE            FACTOR.            Homo sapiens. Isogai et al., "NEDO human            cDNA sequencing project."            abstract</p> <p style="text-align: center;">-----</p>	1-15, 45-48

6

Form PCT/ISA210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 01/09220
<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:		
1.	<input checked="" type="checkbox"/>	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Although claims 18, 21, 24, 32 and 34 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.
2.	<input checked="" type="checkbox"/>	Claims Nos.: 20, 23 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210
3.	<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).
<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:		
1.	<input type="checkbox"/>	As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
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.
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.:
4.	<input 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.:
Remark on Protest	<input type="checkbox"/>	The additional search fees were accompanied by the applicant's protest.
	<input type="checkbox"/>	No protest accompanied the payment of additional search fees.

International Application No. PCT/US 01/09220

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20, 23

Claims 20 and 23 refer to an antagonist and agonist of the polypeptides without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results 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 SEARCH REPORT				International Application No. PCT/US 01/09220	
Information on patent family members					
Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9857990	A	23-12-1998	AU 7966498	A	04-01-1999
			CN 1268954	T	04-10-2000
			EP 0996638	A2	03-05-2000
			WO 9857990	A2	23-12-1998
WO 9837196	A	27-08-1998	AU 6661298	A	09-09-1998
			EP 0981613	A1	01-03-2000
			US 6083721	A	04-07-2000
			WO 9837196	A1	27-08-1998
EP 1074617	A	07-02-2001	AU 6180800	A	19-02-2001
			AU 6180900	A	19-02-2001
			AU 6181000	A	19-02-2001
			AU 6181100	A	19-02-2001
			AU 6181200	A	19-02-2001
			AU 6181300	A	19-02-2001
			AU 6181400	A	19-02-2001
			AU 6181500	A	19-02-2001
			AU 6181600	A	19-02-2001
			AU 6315800	A	19-02-2001
			EP 1074617	A2	07-02-2001
			WO 0109315	A1	08-02-2001
			WO 0109345	A1	08-02-2001
			WO 0109316	A1	08-02-2001
			WO 0109349	A1	08-02-2001
			WO 0109317	A1	08-02-2001
			WO 0109318	A1	08-02-2001
			WO 0109319	A1	08-02-2001
			WO 0109346	A1	08-02-2001
			WO 0109320	A1	08-02-2001
WO 0109321	A1	08-02-2001			
WO 0109322	A1	08-02-2001			
WO 0109323	A1	08-02-2001			

## 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 1/16	A 6 1 P 1/18	4 C 0 8 4
A 6 1 P 1/18	A 6 1 P 3/10	4 C 0 8 5
A 6 1 P 3/10	A 6 1 P 5/14	4 H 0 4 5
A 6 1 P 5/14	A 6 1 P 5/38	
A 6 1 P 5/38	A 6 1 P 7/00	
A 6 1 P 7/00	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 9/10	
A 6 1 P 9/10	A 6 1 P 11/00	
A 6 1 P 11/00	A 6 1 P 13/12	
A 6 1 P 13/12	A 6 1 P 17/00	
A 6 1 P 17/00	A 6 1 P 17/06	
A 6 1 P 17/06	A 6 1 P 17/16	
A 6 1 P 17/16	A 6 1 P 19/00	
A 6 1 P 19/00	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 19/06	
A 6 1 P 19/06	A 6 1 P 19/10	
A 6 1 P 19/10	A 6 1 P 21/04	
A 6 1 P 21/04	A 6 1 P 25/28	
A 6 1 P 25/28	A 6 1 P 29/00	
A 6 1 P 29/00	A 6 1 P 29/00	1 0 1
A 6 1 P 31/04	A 6 1 P 31/04	
A 6 1 P 31/10	A 6 1 P 31/10	
A 6 1 P 31/12	A 6 1 P 31/12	
A 6 1 P 31/18	A 6 1 P 31/18	
A 6 1 P 33/00	A 6 1 P 33/00	
A 6 1 P 33/02	A 6 1 P 33/02	
A 6 1 P 33/10	A 6 1 P 33/10	
A 6 1 P 35/00	A 6 1 P 35/00	
A 6 1 P 35/02	A 6 1 P 35/02	
A 6 1 P 37/00	A 6 1 P 37/00	
A 6 1 P 37/04	A 6 1 P 37/04	
A 6 1 P 37/08	A 6 1 P 37/08	
A 6 1 P 43/00	A 6 1 P 43/00	1 0 5
C 0 7 K 14/705	A 6 1 P 43/00	1 1 1
C 0 7 K 16/28	C 0 7 K 14/705	
C 1 2 N 1/15	C 0 7 K 16/28	
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 1/68	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 33/53

M

G 0 1 N 33/566

C 1 2 N 5/00

A

A 6 1 K 49/02

C

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,S,G,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

(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 9 4 0 8 7 ・ サニーベイル・ルイスアベニュー 8 2 6

(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 9 5 1 2 3 ・ サンノゼ・コイドライブ 2 3 3

F ターム(参考) 2G045 AA25 AA40 BA13 BB03 BB20 CB01 CB21 DA12 DA13 DA14  
 DA36 DA77 FB02 FB03 FB04 FB07  
 4B024 AA01 AA11 BA63 CA04 DA02 DA06 EA02 HA11 HA17  
 4B063 QA13 QA18 QA19 QQ08 QQ43 QQ79 QR32 QR48 QR55 QR62  
 QR77 QS25 QS34 QX07  
 4B064 AG20 AG27 CA19 CA20 CC24 DA01 DA13  
 4B065 AA90X AA93Y AB01 CA24 CA44 CA46  
 4C084 AA17 NA14 ZA021 ZA151 ZA361 ZA451 ZA511 ZA551 ZA591 ZA661  
 ZA681 ZA751 ZA811 ZA891 ZA941 ZA961 ZA971 ZB021 ZB111 ZB131  
 ZB151 ZB211 ZB261 ZB271 ZB321 ZB331 ZB351 ZB371 ZB381 ZC061  
 ZC081 ZC311 ZC351 ZC412 ZC551  
 4C085 HH17 JJ01 KA03 KA29 KB52 KB82 LL15  
 4H045 AA10 AA11 AA20 BA10 CA40 DA50 DA75 EA20 EA50 FA74

专利名称(译)	G蛋白相关分子		
公开(公告)号	<a href="#">JP2004501610A</a>	公开(公告)日	2004-01-22
申请号	JP2001569007	申请日	2001-03-22
[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	ラルプリーティ ユエヘンリー リュデュングアイナエム		
发明人	ラル、プリーティ ユエ、ヘンリー リュ、デュング・アイナ・エム		
IPC分类号	G01N33/50 A61K45/00 A61K51/00 A61P1/04 A61P1/06 A61P1/16 A61P1/18 A61P3/10 A61P5/14 A61P5/38 A61P7/00 A61P7/06 A61P9/00 A61P9/10 A61P11/00 A61P13/12 A61P17/00 A61P17/06 A61P17/16 A61P19/00 A61P19/02 A61P19/04 A61P19/06 A61P19/10 A61P21/04 A61P25/28 A61P29 /00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/00 A61P37/04 A61P37/08 A61P43/00 C07K14/47 C07K14/705 C07K16/28 C12N1 /15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/68 G01N33 /15 G01N33/53 G01N33/566		
CPC分类号	A01K2217/05 A61P1/04 A61P1/06 A61P1/16 A61P1/18 A61P3/10 A61P5/14 A61P5/38 A61P7/00 A61P7/06 A61P9/00 A61P9/10 A61P11/00 A61P13/12 A61P17/00 A61P17/06 A61P17/16 A61P19/00 A61P19/02 A61P19/04 A61P19/06 A61P19/10 A61P21/04 A61P25/28 A61P29/00 A61P31/04 A61P31 /10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/00 A61P37/04 A61P37/08 A61P43/00 C07K14/47		
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P1/04 A61P1/06 A61P1/16 A61P1/18 A61P3/10 A61P5/14 A61P5 /38 A61P7/00 A61P7/06 A61P9/00 A61P9/10 A61P11/00 A61P13/12 A61P17/00 A61P17/06 A61P17 /16 A61P19/00 A61P19/02 A61P19/04 A61P19/06 A61P19/10 A61P21/04 A61P25/28 A61P29/00 A61P29/00.101 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/00 A61P37/04 A61P37/08 A61P43/00.105 A61P43/00.111 C07K14 /705 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/68.A C12Q1/68.Z G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K49/02.C		
F-TERM分类号	2G045/AA25 2G045/AA40 2G045/BA13 2G045/BB03 2G045/BB20 2G045/CB01 2G045/CB21 2G045 /DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/DA77 2G045/FB02 2G045/FB03 2G045/FB04 2G045/FB07 4B024/AA01 4B024/AA11 4B024/BA63 4B024/CA04 4B024/DA02 4B024/DA06 4B024 /EA02 4B024/HA11 4B024/HA17 4B063/QA13 4B063/QA18 4B063/QA19 4B063/QQ08 4B063/QQ43 4B063/QQ79 4B063/QR32 4B063/QR48 4B063/QR55 4B063/QR62 4B063/QR77 4B063/QS25 4B063 /QS34 4B063/QX07 4B064/AG20 4B064/AG27 4B064/CA19 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/CA24 4B065/CA44 4B065/CA46 4C084 /AA17 4C084/NA14 4C084/ZA021 4C084/ZA151 4C084/ZA361 4C084/ZA451 4C084/ZA511 4C084 /ZA551 4C084/ZA591 4C084/ZA661 4C084/ZA681 4C084/ZA751 4C084/ZA811 4C084/ZA891 4C084 /ZA941 4C084/ZA961 4C084/ZA971 4C084/ZB021 4C084/ZB111 4C084/ZB131 4C084/ZB151 4C084 /ZB211 4C084/ZB261 4C084/ZB271 4C084/ZB321 4C084/ZB331 4C084/ZB351 4C084/ZB371 4C084 /ZB381 4C084/ZC061 4C084/ZC081 4C084/ZC311 4C084/ZC351 4C084/ZC412 4C084/ZC551 4C085 /HH17 4C085/JJ01 4C085/KA03 4C085/KA29 4C085/KB52 4C085/KB82 4C085/LL15 4H045/AA10 4H045/AA11 4H045/AA20 4H045/BA10 4H045/CA40 4H045/DA50 4H045/DA75 4H045/EA20 4H045 /EA50 4H045/FA74		

摘要(译)

本发明提供了鉴定和编码GPAM的人G蛋白相关分子 ( GPAM ) 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GPAM异常表达有关的疾病的方法。

特表2004-5016

(P2004-501610)

(43) 公表日 平成16年1月22日(2004.1.22)

(51) Int. Cl. 7		F I	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	45/00	A 6 1 K 45/00		4 B 0 2 4
A 6 1 K	51/00	A 6 1 P 1/04		4 B 0 6 3
A 6 1 P	1/04	A 6 1 P 1/06		4 B 0 6 4
A 6 1 P	1/06	A 6 1 P 1/16		4 B 0 6 5
		審査請求 未請求 予備審査請求 有	(全 178 頁)	最終頁に続く
(21) 出願番号	特願2001-569007 (P2001-569007)	(71) 出願人	301005050	
(86) (22) 出願日	平成13年3月22日 (2001. 3. 22)		インサイト・ゲノミックス・インコーポ イテッド	
(85) 翻訳文提出日	平成14年9月24日 (2002. 9. 24)		アメリカ合衆国カリフォルニア州9431 4・パロアルト・ボータードライブ 3 60	
(86) 国際出願番号	PCT/US2001/009220	(74) 代理人	100089266	
(87) 国際公開番号	W02001/070807		弁理士 大島 陽一	
(87) 国際公開日	平成13年9月27日 (2001. 9. 27)	(72) 発明者	ラル、ブリーティ	
(31) 優先権主張番号	60/191, 631		アメリカ合衆国カリフォルニア州950 6・サンタクララ・ピーオーボックス 142	
(32) 優先日	平成12年3月23日 (2000. 3. 23)			
(33) 優先権主張国	米国 (US)			

最終頁に続く