

(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2005-508147

(P2005-508147A)

(43) 公表日 平成17年3月31日(2005.3.31)

|                           |                              |                   |           |                                                                              |  |
|---------------------------|------------------------------|-------------------|-----------|------------------------------------------------------------------------------|--|
| (51) Int.Cl. <sup>7</sup> |                              | F I               |           | テーマコード (参考)                                                                  |  |
| C 1 2 N 15/09             |                              | C 1 2 N 15/00     |           | Z N A A 2 G O 4 5                                                            |  |
| A 6 1 K 31/7088           |                              | A 6 1 K 31/7088   |           | 4 B O 2 4                                                                    |  |
| A 6 1 K 35/39             |                              | A 6 1 K 35/39     |           | 4 B O 6 3                                                                    |  |
| A 6 1 K 39/395            |                              | A 6 1 K 39/395    |           | D 4 B O 6 4                                                                  |  |
| A 6 1 K 48/00             |                              | A 6 1 K 39/395    |           | N 4 B O 6 5                                                                  |  |
|                           |                              | 審査請求 未請求 予備審査請求 有 |           | (全 310 頁) 最終頁に続く                                                             |  |
| (21) 出願番号                 | 特願2003-506477 (P2003-506477) | (71) 出願人          | 592221528 | バイオジェン・アイデック・エムエイ・インコーポレイテッド<br>アメリカ合衆国、マサチューセッツ 02142、ケンブリッジ、ケンブリッジ センター 14 |  |
| (86) (22) 出願日             | 平成14年6月24日 (2002.6.24)       | (74) 代理人          | 100078282 |                                                                              |  |
| (85) 翻訳文提出日               | 平成15年12月22日 (2003.12.22)     |                   | 弁理士 山本 秀策 |                                                                              |  |
| (86) 国際出願番号               | PCT/US2002/020275            | (74) 代理人          | 100062409 |                                                                              |  |
| (87) 国際公開番号               | W02003/000021                |                   | 弁理士 安村 高明 | (74) 代理人 100113413<br>弁理士 森下 夏樹                                              |  |
| (87) 国際公開日                | 平成15年1月3日 (2003.1.3)         |                   |           |                                                                              |  |
| (31) 優先権主張番号              | PCT/US01/19904               |                   |           | 最終頁に続く                                                                       |  |
| (32) 優先日                  | 平成13年6月22日 (2001.6.22)       |                   |           |                                                                              |  |
| (33) 優先権主張国               | 米国 (US)                      |                   |           |                                                                              |  |

(54) 【発明の名称】 P a n C A M核酸およびポリペプチド

## (57) 【要約】

P a n C A M ( G P 3 5 4 と名付けられる ) と名付けられた新規な免疫グロブリンスーパーファミリーのメンバーをコードする単離されたポリヌクレオチドが提供される。 P a n C A M は、予測された一回膜貫通ドメイン、および 5 つの免疫グロブリン ( I g ) ドメインを、タンパク質の細胞外部分に有する。 P a n C A M のタンパク質構造および組織分布は、 P a n C A M が、脾臓 (特に、β 島)、下垂体および中枢神経系 ( C N S ) における細胞-細胞認識、結合、シグナル伝達、および接着事象において役割を果たすことを示す。単離された P a n C A M に関連するポリヌクレオチドおよびポリペプチド、ベクターならびにこれらのいずれかを含む宿主細胞、 P a n C A M に対する抗体、このような抗体を産生する細胞、ならびに関連する診断方法および治療方法が、本発明によって提供される。

## 【特許請求の範囲】

## 【請求項 1】

配列番号 12 のアミノ酸残基 68～594を含む、単離されたポリペプチド。

## 【請求項 2】

配列番号 12 を含む、請求項 1 に記載のポリペプチド。

## 【請求項 3】

配列番号 26 を含む、請求項 1 に記載のポリペプチド。

## 【請求項 4】

配列番号 28 を含む、請求項 1 に記載のポリペプチド。

## 【請求項 5】

配列番号 8 を含む、請求項 1 に記載のポリペプチド。

## 【請求項 6】

さらに、異種配列を含む、請求項 1 に記載のポリペプチド。

## 【請求項 7】

配列番号 8 のアミノ酸残基 4～513を含む、単離されたポリペプチド。

## 【請求項 8】

配列番号 8 のアミノ酸残基 535～708を含む、単離されたポリペプチド。

## 【請求項 9】

配列番号 26 のアミノ酸残基 535～633を含む、単離されたポリペプチド。

## 【請求項 10】

請求項 1 に記載の単離されたポリペプチドをコードするヌクレオチド配列を含む、単離されたポリヌクレオチド。

## 【請求項 11】

配列番号 11 のヌクレオチド 202～1782を含む、請求項 10 に記載のポリヌクレオチド。

## 【請求項 12】

配列番号 11 を含む、請求項 10 に記載のポリヌクレオチド。

## 【請求項 13】

配列番号 25 を含む、請求項 10 に記載のポリヌクレオチド。

## 【請求項 14】

配列番号 27 を含む、請求項 10 に記載のポリヌクレオチド。

## 【請求項 15】

配列番号 7 のヌクレオチド 196～2319を含む、請求項 10 に記載のポリヌクレオチド。

## 【請求項 16】

配列番号 7 を含む、請求項 10 に記載のポリヌクレオチド。

## 【請求項 17】

請求項 10 に記載のポリヌクレオチドであって、該ヌクレオチド配列に作動可能に連結された転写調節配列をさらに含む、ポリヌクレオチド。

## 【請求項 18】

請求項 10 に記載のポリヌクレオチドを含む、ベクター。

## 【請求項 19】

前記ベクターが、プラスミドである、請求項 18 に記載のベクター。

## 【請求項 20】

前記ベクターが、ウイルスベクターである、請求項 18 に記載のベクター。

## 【請求項 21】

請求項 18 に記載のベクターであって、バキュロウイルス、アデノウイルス、パルボウイルス、ヘルペスウイルス、ポックスウイルス、アデノ随伴ウイルス、セムリキ森林ウイルス、ワクシニアウイルス、レンチウイルスおよびレトロウイルスからなる群より選択される、ベクター。

10

20

30

40

50

## 【請求項 22】

請求項 10 に記載のポリヌクレオチドを含む、宿主細胞。

## 【請求項 23】

前記宿主細胞が、細菌細胞、昆虫細胞、酵母細胞、植物細胞および哺乳動物細胞からなる群より選択される、請求項 22 に記載の宿主細胞。

## 【請求項 24】

前記宿主細胞がヒト細胞である、請求項 22 に記載の宿主細胞。

## 【請求項 25】

PanCAM ポリペプチドを産生する方法であって、該方法は、  
請求項 22 に記載の宿主細胞を、前記ヌクレオチドが発現される条件下で培地中で培養する工程、および  
該細胞からまたは該培養培地から該ポリペプチドを回収する工程、  
を包含する、方法。 10

## 【請求項 26】

請求項 1 または 7 に記載のポリペプチドに結合する、抗体。

## 【請求項 27】

前記抗体が、モノクローナル抗体である、請求項 26 に記載の抗体。

## 【請求項 28】

前記抗体が、ヒト化抗体または完全にヒト抗体である、請求項 26 に記載の抗体。

## 【請求項 29】

サンプル中の PanCAM コード配列の存在を決定する方法であって、該方法は、以下の工程：  
該サンプルと、高いストリンジェンシーのハイブリダイゼーション条件下で、配列番号 3、5、6、7、9、11、25 もしくは 27、または少なくとも 17 ヌクレオチドを有するそれらの一部分を含む核酸プローブとを接触させる工程、および  
該ポリヌクレオチドの、該サンプル中の核酸に対するハイブリダイゼーションを検出する工程であって、ここで、該ハイブリダイゼーションの存在が、該サンプル中の PanCAM コード配列の存在を示す、工程、  
を包含する、方法。 20

## 【請求項 30】

サンプル中の PanCAM タンパク質の存在を決定する方法であって、該方法は、以下の工程：  
該サンプルを、請求項 26 に記載の抗体と接触させる工程；および  
抗原に対する該抗体の特異的結合を検出する工程であって、ここで、該特異的結合の存在が、該サンプル中の PanCAM タンパク質の存在を示す、工程、  
を包含する、方法。 30

## 【請求項 31】

PanCAM タンパク質を結合する化合物を同定する方法であって、該方法は、以下の工程：  
PanCAM タンパク質を、試験化合物と接触させる工程；および  
該 PanCAM タンパク質と該試験化合物とによって形成される複合体を検出する工程であって、ここで、該複合体の存在が、該試験化合物が該 PanCAM タンパク質に結合することを示す、工程、  
を包含する、方法。 40

## 【請求項 32】

PanCAM タンパク質の活性を調節する化合物を同定する方法であって、該方法は、以下の工程：  
該 PanCAM タンパク質を、試験化合物と接触させる工程；および  
該 PanCAM タンパク質の活性に対する試験化合物の効果を決定する工程であって、ここで、該接触させる工程の後の該活性の変化が、該試験化合物が該 PanCAM タンパク 50

質の活性を調節することを示す、工程、  
を包含する、方法。

【請求項 33】

ヒト P a n C A M 遺伝子のホモログを同定する方法であって、該方法は、以下の工程：  
核酸ライブラリーを、中程度または高いストリンジェンシーのハイブリダイゼーション条件下で、配列番号 3、5、6、7、9、11、25 もしくは 27、または少なくとも 17  
ヌクレオチドからなるそれらの一部分を含む核酸プローブにハイブリダイズさせる工程、  
および  
該核酸プローブが該ライブラリー中の核酸配列にハイブリダイズするか否かを決定する工程、  
を包含し、ここで、このようにハイブリダイズした核酸配列が、ヒト P a n C A M 遺伝子のホモログである、方法。

10

【請求項 34】

被験体における疾患状態を診断する方法であって、該方法は、以下の工程：  
該被験体からの組織サンプル中の P a n C A M タンパク質の量または活性を、コントロールサンプル中の P a n C A M タンパク質の量または活性と比較する工程、  
を包含し、ここで、該コントロールサンプルと比較した該組織サンプル中の該 P a n C A M の量または活性の有意な相違が、該被験体が疾患状態を有することを示す、方法。

【請求項 35】

被験体における疾患状態を診断する方法であって、該方法は、以下の工程：  
該被験体からの組織サンプル中の P a n C A M m R N A の量を、コントロールサンプル中の P a n C A M m R N A の量と比較する工程、  
を包含し、ここで、該コントロールサンプルと比較した該組織サンプル中の該 P a n C A M m R N A の量の有意な相違が、該被験体が疾患状態を有することを示す、方法。

20

【請求項 36】

前記組織サンプルが、膵臓組織サンプルである、請求項 34 または 35 に記載の方法。

【請求項 37】

前記有意な相違が、有意な減少であり、そして前記疾患状態が、糖尿病に関連する、請求項 36 に記載の方法。

【請求項 38】

前記疾患状態が、中枢神経系に関連する、請求項 34 または 35 に記載の方法。

30

【請求項 39】

前記疾患状態が、下垂体に関連する、請求項 34 または 35 に記載の方法。

【請求項 40】

試験細胞中の遺伝的な病変または変異の存在または非存在を同定する方法であって、該遺伝的な病変または変異は、以下：(i) P a n C A M タンパク質をコードする遺伝子の異常な改変もしくは変異；(ii) P a n C A M タンパク質をコードする遺伝子の誤調節；および (iii) P a n C A M タンパク質の異常な翻訳後修飾、のうちの少なくとも 1 つによって特徴付けられ、該方法は、以下の工程：

該試験細胞からの核酸、および該遺伝的な病変または変異を欠く参照細胞からの核酸を、  
配列番号 3、5、6、7、9、11、25 もしくは 27、または少なくとも 17 ヌクレオチドを有するこれらの一部分を含む核酸プローブと、高いストリンジェンシーのハイブリダイゼーション条件下で別々にハイブリダイズさせる工程；および  
該核酸ハイブリッドを、高いストリンジェンシーの洗浄条件下で別々に洗浄し、該ハイブリッドを解離させる工程；および  
該核酸プローブが、該参照細胞の核酸と比較して、該試験細胞の核酸からより容易に解離されるか否かを決定する工程、  
を包含する、方法。

40

【請求項 41】

被験体における疾患状態を処置する方法であって、請求項 10 に記載のポリヌクレオチド

50

、または請求項 26 に記載の抗体を含む組成物を、該被験体に投与する工程を包含する、方法。

【請求項 42】

前記疾患状態が、膵臓に関連する、請求項 41 に記載の方法。

【請求項 43】

前記疾患状態が、膵臓β細胞に関連する、請求項 42 に記載の方法。

【請求項 44】

前記疾患状態が、糖尿病に関連する、請求項 43 に記載の方法。

【請求項 45】

前記疾患状態が、急性膵臓炎または慢性膵臓炎、膵臓の炎症、膵臓の壊死、外分泌機能不全、膵臓内分泌およびホルモン不均衡、膵臓腫瘍および関連の癌、ならびに膵臓に影響する自己免疫障害からなる群より選択される、請求項 41 に記載の方法。 10

【請求項 46】

前記疾患状態が、中枢神経系に関連する、請求項 41 に記載の方法。

【請求項 47】

前記状態が、アルツハイマー病、パーキンソン病、老年痴呆、偏頭痛、てんかん、神経炎、神経衰弱、ニューロパシー、神経変性および神経腫瘍からなる群より選択される、請求項 46 に記載の方法。

【請求項 48】

前記疾患状態が、下垂体ホルモン欠乏、成長ホルモン欠乏、および下垂体腫瘍からなる群より選択される、請求項 41 に記載の方法。 20

【請求項 49】

分散した膵臓細胞集団からβ細胞を単離する方法であって、該方法が、以下の工程：

PanCAMタンパク質に特異的な抗体を提供する工程、

該集団中のβ細胞を、該抗体に結合させ得る条件下で、該分散した膵臓細胞集団を、該抗体に接触させる工程、および

該抗体に結合しない細胞を、該抗体に結合したβ細胞から分離して、それによって、β細胞を単離する、工程、

を包含する、方法。

【請求項 50】

前記分散した膵臓細胞集団が、ウシ膵臓細胞集団またはブタ膵臓細胞集団である、請求項 49 に記載の方法。 30

【請求項 51】

糖尿病または糖尿病前状態の徴候または症状を示すヒトを処置する方法であって、該方法が、以下の工程：

請求項 49 に記載の方法によって単離されたβ細胞を提供する工程、および

該β細胞を該ヒトに導入する工程、

を包含する、方法。

【請求項 52】

前記β細胞が、皮下的に前記ヒトに導入される、請求項 51 に記載の方法。 40

【請求項 53】

前記β細胞が、生物学的に受容可能なキャリア中にカプセル化される、請求項 51 に記載の方法。

【請求項 54】

配列番号 31 を含む、単離されたポリペプチド。

【請求項 55】

配列番号 31 をコードするヌクレオチド配列を含む、単離されたポリヌクレオチド。

【請求項 56】

配列番号 32 を含む、請求項 55 に記載のポリヌクレオチド。

【発明の詳細な説明】

## 【技術分野】

## 【0001】

(関連出願)

本出願は、2001年6月22日に出願されたPCT/US01/19904から優先権を主張し、それは2000年6月22日に出願された米国仮特許出願番号第60/213,611号から優先権を主張する。PCT出願および仮特許出願の開示はともに、その全体が本明細書中で参考として援用される。

## 【0002】

(発明の分野)

本発明は、一般的には分子生物学の分野に関する。より詳細には、本発明は、免疫グロブリンのスーパーファミリーのメンバーに関する。 10

## 【背景技術】

## 【0003】

(発明の背景)

多数のタンパク質は、保存された構造のモチーフおよび生物学的機能に基づくスーパーファミリーに分類されてきた。スーパーファミリーは、特定の程度の配列相同性（通常、少なくとも15%）を共有するタンパク質のグループとして広く定義される。スーパーファミリーのメンバーによって共有される保存配列は、ドメインと称される緻密な三次構造の形成にしばしば寄与し、そしてしばしば特定のスーパーファミリーを特徴付けるドメインの全配列が単一のエクソンによってコードされる（例えば、Abbasら、CELLULAR AND MOLECULAR IMMUNOLOGY, W. B. Saunders Co., Philadelphia, PA. 1997を参照のこと）。スーパーファミリーのメンバーは、分岐進化によって共通の前駆遺伝子に由来するようであり、マルチドメインタンパク質は、1つのより多いスーパーファミリーに属し得る。タンパク質スーパーファミリーの例としては、以下が挙げられる：リガンドゲートイオンチャネルレセプタースーパーファミリー、電位依存性イオンチャネルレセプタースーパーファミリー、レセプターチロシンキナーゼスーパーファミリー、レセプタータンパク質チロシンホスファターゼスーパーファミリー、Gタンパク質共役型レセプタースーパーファミリー、および免疫グロブリン（Ig）スーパーファミリー。 20

## 【0004】

このIgスーパーファミリーは、部分的なアミノ酸配列の相同性ならびにIg重鎖およびIg軽鎖においてもともと同一であった三次構造的な特徴を共有するタンパク質を含む。Igスーパーファミリーの共通の構造モチーフは、いわゆる「Igドメイン」である。Igドメインは、約70～110個のアミノ酸残基および内部にCys-Cysジスルフィド結合を有する三次元的な球形構造である。これらのドメインは、2層のβプリーツシートを含み、各層は、5～10個のアミノ酸残基の3～5個の逆平行鎖から構成される。Igドメインは、IgVドメインまたはIgCドメインのいずれかに対する最も近隣な相同性に基づいてV様またはC様として分類される。一般的な見解として、例えば、Abbasら、前出を参照のこと。 30

## 【0005】

Igスーパーファミリーの最も同定されたメンバーは、細胞外部分および広く分岐する細胞質末端（cytoplasmic tail）においてIgドメインを有する内在性原形質膜タンパク質（通常、内因性酵素活性を持たない）である。Igスーパーファミリーメンバーの1つの反復特性は、（同じかまたは異なるアミノ酸配列の）異なるポリペプチド鎖におけるIgドメイン間の相互作用が、この分子の生物学的活性に必須であるということである。ヘテロフィリックな相互作用はまた、異なる細胞の表面上に発現される全体的に別個の分子におけるIgドメイン間で生じ得る。このような相互作用は、細胞-細胞結合を安定化する接着力を提供する。 40

## 【0006】

Igスーパーファミリーの多数のメンバーは、脊椎動物の免疫系における細胞認識、接着 50

および結合機能を媒介する細胞表面または可溶性分子である。I g スーパーファミリー分子を産生する2つの顕著な細胞型は、Bリンパ球およびTリンパ球である。免疫系において重要な例示的なI g スーパーファミリーメンバータンパク質としては、以下が挙げられる：抗体、T細胞レセプター、主要組織適合遺伝子複合体(MHC)クラスI分子および主要組織適合遺伝子複合体クラスII分子、CD2、CD3、CD4、CD5、CD8、CD28、CD20(B1)、CD32(FcγRII)、CD44、CD54(ICAM-1)、CD80(B7-1)、CD86(B7-2)、CD90(Thy-1)、CD102(ICAM-2)、CD106(VCAM-1)、CD121(IL-1R)、CD152(CTLA-4)、p-IgR、NCAM、およびCD140(PDGFR)(Abbasら、前出)。

10

#### 【0007】

いくつかのI g スーパーファミリーメンバーは、免疫系以外(例えば、神経系)で同定されてきた。これらの保存された構造モチーフおよび免疫系におけるこのようなモチーフの周知の機能に基づいて、これらのI g スーパーファミリーメンバーは、非免疫組織において同様に細胞の認識、結合および接着機能を実施するようである。特定の細胞型に局在する新規I g スーパーファミリーメンバーは、診断の目的のために有用な細胞マーカーおよび組織マーカーである。組織特異的I g スーパーファミリーメンバーはまた、特に組織発生の間または組織再生の間(例えば、組織損傷後または組織傷害後)の、組織における不適切な細胞-細胞の接着およびシグナル伝達に関連する異常な状態、障害および/または疾患を処置するための適切な治療用標的である。

20

#### 【発明の開示】

#### 【課題を解決するための手段】

#### 【0008】

#### (発明の要旨)

本発明は、少なくとも部分的には、これ以前に未知のI g スーパーファミリーメンバー(GP354と呼ばれる)をコードする遺伝子の発見に基づく。他に明示されない限り、下記の場合のgp354という名前は、本発明の新規の核酸のことを言い、それに対して、上記の場合のGP354という名前は、本発明の新規のポリペプチドのことを言う。

#### 【0009】

GP354は、主として膵臓において発現される内在性膜タンパク質である。このタンパク質はまた、PanCAMとも呼ばれ、膵臓細胞接着分子である。膵臓において、GP354(すなわちPanCAM)発現は、β細胞に局在される。GP354はまた、中枢神経系(CNS)(例えば、脳)において低いレベルで検出される。GP354は、予測される単一の膜スパンドメインおよび5つの細胞外I g ドメインを有する。ヒトGP354はヒトNEPH1の494個のアミノ酸フラグメントと約44%の同一性を共有し、そしてマウスNEPH1の368個のアミノ酸フラグメントと47%のアミノ酸同一性を共有する。2つのフラグメントはともにNEPH1の細胞外部分に存在する。GP354は、膵臓およびCNSでの細胞-細胞相互作用(例えば、接着)という役割を果たし、それらの器官の構造および機能に影響をおよぼすようである。

30

#### 【0010】

本発明は、脊椎動物GP354(例えば、哺乳動物(例えば、ヒトまたはマウス))をコードする単離されたポリヌクレオチドまたはその生物学的に活性な部分を提供する。本発明はまた、GP354をコードするポリヌクレオチドの検出について、プライマーまたはハイブリダイゼーションプローブとして使用するのに適したポリヌクレオチドフラグメントを提供する。この開示は、特定のGP354またはgp354配列を例示するが、これらの配列の全ての対立遺伝子変異体(例えば、ポリ多型)が本発明に含まれる。

40

#### 【0011】

幾つかの実施形態において、本発明は配列番号12のアミノ酸残基68~594を含む単離されたヒトGP354ポリペプチドを提供する。このような単離されたポリペプチドは、配列番号12、配列番号26、配列番号28または配列番号8を含み得る。このポリペ

50

プチドは、異種配列をさらに含み得る。本発明はまた、配列番号 31 を含む単離されたマウス PanCAM ポリペプチドを提供する。

#### 【0012】

本発明はまた、PanCAM タンパク質の細胞外ドメインを含む単離されたポリペプチドを提供し、単離されたポリペプチドは、例えば、配列番号 8 のアミノ酸残基 4～513；または少なくとも 10 個のアミノ酸（例えば、25 個または 50 個）残基を有するその部分を含む。このポリペプチドは、PanCAM 特異的抗体を誘導させるのに使用され得るかまたは PanCAM モジュレーターとして使用され得る。

#### 【0013】

本発明はまた、PanCAM タンパク質の細胞内ドメインを含む単離されたポリペプチドを提供し、単離されたポリペプチドは、例えば、配列番号 8 のアミノ酸残基 535～708 もしくは配列番号 26 のアミノ酸残基 535～633 または少なくとも 10 個のアミノ酸（例えば、25 個または 50 個）残基を有するその部分を含む。このポリペプチドは、ハイブリッド分子を形成するように異種の細胞外ドメインと結合し得る。このハイブリッド分子は、非 PanCAM 細胞外ドメインの刺激によって、PanCAM 様細胞内シグナルを伝達し得る。 10

#### 【0014】

本発明は、上記の単離されたポリペプチドをコードするヌクレオチド配列を含む単離されたポリヌクレオチドを提供する。例示的な単離されたポリヌクレオチドは、配列番号 11 の一部（例えば、ヌクレオチド 202～1782）または全部；配列番号 7 の一部（例えば、ヌクレオチド 196～2319）または全部；ならびに配列番号 1、配列番号 3、配列番号 5、配列番号 6、配列番号 9、配列番号 25、配列番号 27 および配列番号 32 の一部または全部を含むポリヌクレオチドである。幾つかの実施形態において、このポリヌクレオチドは、上記のヌクレオチド配列に作動可能に連結される転写調節配列をさらに含む。 20

#### 【0015】

また、これらの単離されたポリヌクレオチドを含むベクターが、本発明に含まれる。このベクターは、プラスミドベクター、または以下のウイルス由来のウイルス性のベクターであり得る；バキュロウイルス；アデノウイルス；パルボウイルス；ヘルペスウイルス；ポックスウイルス；アデノ関連ウイルス；セムリキ森林ウイルス；ワクシニアウイルス；レンチウイルスおよびレトロウイルス。 30

#### 【0016】

本発明はまた、本発明の単離されたポリヌクレオチドを宿す宿主細胞を含む。この宿主細胞は、例えば、細菌性細胞、昆虫細胞、酵母細胞、植物細胞または哺乳動物細胞（例えば、マウス細胞、またはヒト細胞）であり得る。幾つかの実施形態において、これらの宿主細胞は、PanCAM ポリペプチドを産生するために使用され得る。これを行うために、この宿主細胞は、PanCAM コード配列が発現される条件下で培養され、そして細胞または培養培地からポリペプチドを回収する。

#### 【0017】

本発明はまた、PanCAM タンパク質に特異的に結合する抗体を特徴とする。この抗体は、完全抗体またはそのフラグメント（例えば、Fab フラグメント、F(ab')<sub>2</sub> フラグメント）であり得る。それらはまた単一鎖抗体、キメラ抗体、ヒト化抗体および完全ヒト抗体であり得る。 40

#### 【0018】

本発明はまた、分散された膵臓細胞集団（例えば、ウシ膵臓細胞集団またはブタ膵臓細胞集団）からβ細胞を単離する方法を特徴とする。この方法は以下の工程を包含する；PanCAM タンパク質に特異的な抗体を提供する工程；集団中のβ細胞が抗体と結合することを可能にする条件下で、分散された膵臓細胞集団と抗体を接触させる工程；ならびに抗体に結合したβ細胞から抗体によって結合されていない細胞を分離、それによってβ細胞を単離する工程。 50



## 【0019】

このように単離されたβ細胞は、糖尿病または糖尿病前の状態の徴候または症状を示すヒトを処置するのに使用され得る。この処置方法は、単離されたβ細胞をヒトへ（例えば、皮下に）導入する工程を包含する。このβ細胞は、生物学的に受容可能なキャリア（例えば、アルギン酸塩球）中でカプセル化され得る（Ryuら、J. Clin. Invest. 2001、108（1）：31～3を参照のこと）。

## 【0020】

本発明はさらに、配列番号2（推定gp354 cDNAによってコードされる）；配列番号4（部分gp354膵臓cDNAによってコードされる）；配列番号8（誘導されたgp354 cDNAによってコードされる）；配列番号10（部分誘導されたgp354 cDNAによってコードされる）；または配列番号12（gp354膵臓cDNAによってコードされる、プラスミドFL86-3のインサート、これは、2002年6月11日にAmerican Type Culture Collection（ATCC）に寄託された）のアミノ酸配列；あるいは配列番号2、4、8、10、および12のいずれか1つの少なくとも1つのIgドメインに対して少なくとも80%（85%、95%、または98%）同一なアミノ酸配列を含むタンパク質をコードするヌクレオチド配列を含むポリヌクレオチドを特徴とする。

## 【0021】

いくつかの実施形態において、このポリヌクレオチドは、配列番号1の配列（gp354 cDNA）、または少なくとも17核酸単位（例えば、ヌクレオチド）を有するそのフラグメントを含む。このようなフラグメントの例は、配列番号3である。別の実施形態において、ポリヌクレオチドは、配列番号5の配列（gp354を含むゲノムDNA）、または少なくとも17核酸単位を有するそのフラグメントを含む。例示のフラグメントは、配列番号6（gp354上流ゲノムDNA）のフラグメントである。他の実施形態において、ポリヌクレオチドは、配列番号7の配列（誘導されたgp354 cDNA）、または少なくとも17核酸単位を有するそのフラグメントを含む。例示のフラグメントは、配列番号9（誘導されたgp354 cDNAのC末端フラグメント）である。他の実施形態において、ポリヌクレオチドは、配列番号11の配列（膵臓gp354 cDNA）、または少なくとも17核酸単位を有するそのフラグメントを含む。好ましいフラグメントは、GP354のすくなくとも1つの細胞外Igドメインおよび／または細胞内ドメインの一部または全てをコードする。

## 【0022】

本発明はまた、配列番号2のアミノ酸配列を含むポリペプチドの天然に存在する対立遺伝子改変体をコードするポリヌクレオチドを提供する。ここで、この核酸は、ストリンジェントな条件下で、配列番号1または配列番号11にハイブリダイズする。本発明はまた、配列番号4、8、10、または12のアミノ酸配列を含むポリペプチドの天然に存在する対立遺伝子改変体をコードするポリヌクレオチドを提供する。ここで、この核酸は、ストリンジェントな条件下で、配列番号1または配列番号11にハイブリダイズする。

## 【0023】

配列番号2、4、8、10、もしくは12のアミノ酸配列；またはそれらの配列のいずれか1つによりコードされるIgドメインと少なくとも80%（85%、95%、または98%）同一なアミノ酸を含む、単離されたGP354タンパク質もまた、本発明によって提供される。

## 【0024】

本発明はまた、配列番号1、3、5、7、9、もしくは11；または少なくとも1つのIgドメインをコードするそれらの配列のいずれか1つの一部と少なくとも約65%、好ましくは75%、85%、または95%同一な配列を含むポリヌクレオチドによってコードされる単離されたGP354タンパク質を提供する。配列番号1または11の配列を有する核酸に、ストリンジェントな条件下でハイブリダイズする配列を有するポリヌクレオチドによってコードされる単離されたGP354タンパク質もまた、提供される。

## 【0025】

本発明は、I g スーパーファミリーの他のメンバーをコードする核酸に対して、g p 3 5 4 核酸を特異的に検出する g p 3 5 4 ポリヌクレオチドを提供する。本発明はまた、本発明の g p 3 5 4 ポリヌクレオチドを含む核酸構築物（例えば、組換えベクター（例えば、クローニングベクター、標的化ベクター、または発現ベクター））を提供する。

## 【0026】

単離されたかまたは組換え G P 3 5 4 タンパク質およびポリペプチドが、本発明によって提供される。いくつかの実施形態において、G P 3 5 4 タンパク質およびポリペプチドは、以下の（重複する）生物学的活性の少なくとも1つを示す：（1）G P 3 5 4 タンパク質に天然で結合するリガンド（例えば、タンパク質レセプター、多糖など）との相互作用（例えば、結合）；（2）天然に存在するヒト G P 3 5 4 に対する自動抗体または、天然に存在するヒト G P 3 5 4 に対して惹起された抗体への結合；（3）膵臓機能（例えば、膵臓におけるシグナル伝達機能または膵臓の器官発達における段階）への関与；（4）神経機能（例えば、神経系におけるシグナル伝達機能または神経系の発達における段階）への関与；および（5）細胞間相互作用（例えば、認識、結合、および／または接着）の媒介。

10

## 【0027】

G P 3 5 4 タンパク質または生物学的に活性なその部分は、非 G P 3 5 4 ポリペプチド（例えば、異種アミノ酸配列（例えば、タンパク質の安定性、検出、精製、または標的細胞へのインビボ送達を容易にする配列））に作動可能に連結され、G P 3 5 4 融合タンパク質が形成され得る。

20

## 【0028】

本発明は、少なくとも1つの上記 g p 3 5 4 関連の単離されたポリヌクレオチド、G P 3 5 4 タンパク質もしくは生物学的に活性なその部分、抗体または融合タンパク質を含む薬学的組成物を提供する。この組成物は、薬学的に受容可能なキャリアを必要に応じて含む。このような組成物は、異常な G P 3 5 4 細胞局在、発現、および／または活性に患った被験体における状態を緩和する治療方法において有用である。

## 【0029】

従って、本発明はまた、本発明の g p 3 5 4 関連化合物または組成物を投与する工程を包含する、処置の方法を提供する。このような方法は、例えば、発達中または成体の膵臓および中枢神経系における細胞の認識、結合、シグナル伝達および接着の機能に相関する異常な状態、障害または疾患を処置するために有用である。

30

## 【0030】

膵臓で富化されるタンパク質として、G P 3 5 4 は、発達中および成体の膵臓における、特に、組織の発達の間ならびに（例えば、膵臓の損傷、外傷または変性状態の後の）組織の再生および／または治癒の間の、不適切な細胞-細胞結合、接着およびシグナル伝達に関連する異常な状態、障害および／または疾患を処置するための、適切な治療標的である。G P 3 5 4 は、免疫状態、自己免疫状態、および変性状態（I 型およびII 型の糖尿病が挙げられる）に関連する、膵臓の細胞死を阻害するための適切な治療標的であることもまた、企図される。G P 3 5 4 の神経形態は、同様に、中枢神経系における組織の異常を処置するため、組織の再生および修復のため、ならびに組織の変性および細胞死を阻害するための、適切な治療標的である。G P 3 5 4 関連組成物はまた、下垂体における疾患（例えば、下垂体ホルモン欠損、成長ホルモン欠損、および下垂体腫瘍）を処置するために使用され得る。

40

## 【0031】

本発明は、G P 3 5 4 活性を調節するための方法を提供する。この方法において、標的細胞が、G P 3 5 4 の活性または発現を調節（例えば、阻害または刺激）し、その結果、G P 3 5 4 の活性または発現が変化する薬剤と接触される。いくつかの実施形態において、この薬剤は、G P 3 5 4 に特異的に結合する抗体である。他の実施形態において、この薬剤は、g p 3 5 4 遺伝子の転写、g p 3 5 4 R N A のスプライシング、または g p 3 5

50

4 m R N A の翻訳を調節することによって、G P 3 5 4 の活性または発現を調節する。なお他の実施形態において、この薬剤は、g p 3 5 4 m R N A または g p 3 5 4 遺伝子のコード鎖に対してアンチセンスである配列を有する、核酸である。他の実施形態において、この薬剤は、G P 3 5 4 タンパク質、G P 3 5 4 タンパク質をコードする核酸、または G P 3 5 4 タンパク質のアンタゴニストもしくはアゴニスト（例えば、ペプチド、ペプチド模倣物、または他の低分子）であり得る。

【0032】

本発明はまた、G P 3 5 4 タンパク質に結合する化合物を同定するための方法を提供する。別の局面において、本発明は、G P 3 5 4 タンパク質の生物学的活性を調節する化合物を同定するための方法を提供し、この方法は、試験化合物の存在下および非存在下でこのタンパク質の生物学的活性または発現を測定する工程、およびこのタンパク質の活性を変化させる化合物を同定する工程を包含する。コンビネーションライブラリーを、これらの方法における候補化合物の供給源として使用し得る。 10

【0033】

本発明は、生物学的サンプル（例えば、患者由来の流体または組織サンプル）中の g p 3 5 4 ポリヌクレオチド、G P 3 5 4 タンパク質またはその活性の存在を検出するための方法を提供し、この方法は、このサンプルを、g p 3 5 4 ポリヌクレオチド配列、G P 3 5 4 タンパク質またはその活性の存在の指標を検出し得る薬剤と接触させることによる。

【0034】

g p 3 5 4 に関連する遺伝的病巣または変異の存在または非存在を同定するための診断アッセイが提供され、この病巣または変異は、以下の少なくとも1つによって特徴付けられる：(i) G P 3 5 4 タンパク質をコードする遺伝子の異常な改変または変異；(ii) G P 3 5 4 タンパク質をコードする遺伝子の誤調節（例えば、転写、スプライシングまたは翻訳）；および(iii) G P 3 5 4 タンパク質の異常な翻訳後修飾または局在；ここで、この遺伝子の野生型形態は、G P 3 5 4 の生物学的活性を有するタンパク質をコードする。 20

【0035】

本発明は、細胞の少なくともいくらかが本発明の単離されたポリヌクレオチドを含む、非ヒト動物（例えば、マウス、ラット、モルモット、ヒツジ、ヤギ、ウマまたはウシのような哺乳動物）を提供する。このような動物は、その体細胞および／または生殖細胞のいくらかのみが、このポリヌクレオチドを保有する場合に、キメラであり得る。あるいは、このような動物は、その体細胞および生殖細胞の全てがこのポリヌクレオチドを保有する場合に、トランスジェニックであり得る。 30

【0036】

本発明はまた、g p 3 5 4 遺伝子の内因性オルソログが遺伝子標的にによって破壊されている（すなわち、「ノックアウト」されている）非ヒト動物を提供する。g p 3 5 4 ポリヌクレオチドを含む細胞、組織および流体のような生物学的サンプル、ならびにこれらおよび上述の動物由来の G P 3 5 4 関連産物もまた、本発明の範囲内である。

【0037】

本発明は、本発明の核酸配列およびアミノ酸配列を格納する、コンピュータ読み取り可能手段を提供する。このコンピュータ読み取り可能手段の記録は、配列の読み取りおよび表示のため、ならびに本発明の配列の他の配列との比較、整列および並べのために、アクセスされ得る。 40

【0038】

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

【0039】

（発明の詳細な説明）

本発明は、これまで未知であったタンパク質、G P 3 5 4（すなわち、P a n C A M）をコードする新規ヒト遺伝子の発見に少なくとも一部基づく。この遺伝子（g p 3 5 4）を 50

、ヒトゲノムの（「探索用の（mining）」）公開されている核酸配列のコンピュータによる分析によって同定した。gp354遺伝子は、少なくとも14個のエキソンを含み、そして通常、ヒト第19染色体上に存在する。この遺伝子から転写されたmRNAは、1779個の塩基対（終止コドンを含む）のオープンリーディングフレームを有し、そして592個のアミノ酸残基であると予測されるタンパク質をコードする。この新規GP354タンパク質は、膵臓および脳において特異的に発現される。

#### 【0040】

（定義）

本明細書中で用いられる場合、「核酸」（「ポリヌクレオチド」ともいう）とは、DNA分子（例えば、cDNAまたはゲノムDNAまたは合成DNA）およびRNA分子（例えば、mRNA）を含む。この用語はまた、非天然ヌクレオチドアナログ、非天然ヌクレオシド間結合または両方を含む、DNAアナログまたはRNAのアナログを含むことが意図される。核酸は、任意のトポロジー構造であり得る。例えば、核酸は、一本鎖、二本鎖、三本鎖、四本鎖、部分的二本鎖、分枝、ヘアピン、環状、またはパドロック構造であり得る。例えば、Bannerら, Curr. Opin. Biotechnol. 12:11-15 (2001); Escudéら, Proc. Natl. Acad. Sci. USA 14;96(19):10603-7 (1999); Nilssonら, Science 265(5181):2085-8 (1994); Praseuthら, Biochim. Biophys. Acta. 1489(1):181-206 (1999); Fox, Curr. Med. Chem. 7(1):17-37 (2000); Kochetkovaら, Methods Mol. Biol. 130:189-201 (2000); Chanら, J. Mol. Med. 75(4):267-82 (1997)を参照のこと。

#### 【0041】

本明細書中で用いられる場合、「単離された核酸」（「単離されたポリヌクレオチド」ともいう）は、その核酸の天然の供給源に存在する他の核酸分子から分離されている核酸である。特に、単離された、非組換え天然染色体およびその500kbを超えるフラグメントが排除される。好ましくは、「単離された」核酸は、その核酸が由来する生物のゲノム中の核酸に天然で隣接する配列を実質的に含まない。例えば、好ましい単離されたgp354核酸は、単離された核酸が由来する細胞のゲノムDNA中の核酸に天然で隣接する、約10kb、5kb、4kb、3kb、2kb、1kb、0.5kbまたは0.1kb未満のヌクレオチド配列に隣接する。なおより好ましくは、この単離されたポリヌクレオチドは、5000塩基対以下、しばしば1000塩基対以下、500塩基対以下、100塩基対以下または50塩基対以下である。

#### 【0042】

しかし、「単離された」は、そのように記載される核酸またはポリヌクレオチド自体が天然の環境から物理的に取り出されることを必ずしも必要としない。例えば、生物のゲノム中の内因性核酸配列は、異種配列（すなわち、この内因性核酸配列に天然では隣接しない配列）が、その内因性核酸配列に隣接して配置され、この内因性核酸配列の発現が変化する場合に、本明細書中では「単離された」とみなされる。例として、非天然プロモーター配列が、（例えば、相同組換えによって）ヒトの細胞のゲノム中でgp354遺伝子の天然のプロモーターに代えて置換され得て、その結果、この遺伝子は、変化した発現パターンを有する。この遺伝子は、天然で隣接する配列の少なくともいくつかから分離されているので、今や「単離された」ものになる。

#### 【0043】

核酸はまた、対応する核酸がゲノム中に天然には生じない任意の改変を含む場合、「単離された」とみなされる。例えば、内因性のgp354コード配列は、例えば、ヒトの介入により、人為的に導入される挿入、欠損または点変異を含む場合、「単離された」とみなされる。「単離された核酸」はまた、宿主細胞の染色体中に非相同的部位で組みこまれた核酸、エピソームとして存在する核酸構築物、および宿主細胞染色体中に組みこまれた核

酸構築物を含む。さらに、「単離された核酸」は、組換え技術によって産生される場合に、他の細胞物質を実質的に含まないものであり得るか、もしくは培養培地を実質的に含まないものであり得るか、あるいは、化学的に合成される場合に、化学前駆物質もしくは他の化学物質を実質的に含まないものであり得る。

#### 【0044】

本発明のポリヌクレオチドは、全長G P 3 5 4タンパク質をコードし得る場合に、「全長」とみなされる。

#### 【0045】

本明細書中で用いられる場合に、句、参照核酸配列の「縮重改変体」とは、標準的な遺伝暗号に従って、翻訳されて、その参照核酸配列から翻訳された配列と同一のアミノ酸配列を提供し得る、核酸配列を含む。

10

#### 【0046】

本明細書中で用いられる場合に、用語「マイクロアレイ」（「核酸マイクロアレイ」ともいう）は、基材に結合した複数の核酸を示し、結合した核酸のそれぞれへのハイブリダイゼーションは、個別に検出可能である。その基材は、固体構成もしくは多孔性構成、平面構成もしくは非平面構成、単一構成もしくは分散構成、または他の任意の構成であり得る。

#### 【0047】

そう規定される場合に、用語「マイクロアレイ」は、S ch e n a（編）、DNA M i c r o a r r a y s : A P r a c t i c a l A p p r o a c h (P r a c t i c a l A p p r o a c h S e r i e s)、O x f o r d U n i v e r s i t y P r e s s (1999) (ISBN: 0199637768); Nature Genet. 21 (1) (補遺): 1-60 (1999); およびS ch e n a（編）、M i c r o a r r a y B i o c h i p : T o o l s a n d T e c h n o l o g y, E a t o n P u b l i s h i n g C o m p a n y / B i o T e c h n i q u e s B o o k s D i v i s i o n (2000) (ISBN: 1881299376); Brennerら、Pr oc. Natl. Acad. Sci. USA 97 (4): 1665-1670 (2000)において、このように呼ばれるかまたは同様に呼ばれる、全ての装置を含む。これら参考文献の全ての開示は、本明細書中において、その全体が参考として援用される。

20

#### 【0048】

核酸ハイブリダイゼーションに関して本明細書中で用いられる場合、用語「プローブ」（「核酸プローブ」または「ハイブリダイゼーションプローブ」ともいう）は、検出可能に標識されているか、または検出可能に標識されるように意図される、単離された既知の配列の核酸をいう。核酸マイクロアレイに関して本明細書中で用いられる場合、用語「プローブ」（等価には、「核酸プローブ」または「ハイブリダイゼーションプローブ」）は、基材に結合されているか、または基材に結合されるように意図されている、単離された核酸をいう。このような状況のいずれにおいても、用語「標的」は、配列相補性によって、プローブに結合することが意図される、核酸をいう。

30

#### 【0049】

他のように示されない限り、「配列番号Xを含む核酸」は、(i) 配列番号Xの配列または(ii) 配列番号Xに相補的な配列のいずれかを少なくとも一部が有する核酸をいう。この2つの間の選択は、状況によって、規定される。例えば、核酸がプローブとして用いられる場合、この2つの間の選択は、プローブが所望される標的に相補的であるという必要条件によって規定される。

40

#### 【0050】

本明細書中で目的のために、「高ストリンジェンシー条件」は、溶液相でのハイブリダイゼーションについて、65℃にて6×SSC（ここで、20×SSCは、3.0M NaClおよび0.3Mクエン酸ナトリウムを含む）、1% SDS中で、8-12時間水性（すなわち、ホルムアルデヒドを含まない）ハイブリダイゼーションを行い、その後65℃にて0.2×SSC、0.1% SDSで20分間、2回洗浄することとして規定さ

50

れる。65℃でのハイブリダイゼーションは、多数の要因（ハイブリダイズする配列の長さおよび同一性パーセントを含む）に依存して異なる速度で生じることが、当業者により理解される。

#### 【0051】

マイクロアレイに基づくハイブリダイゼーションについて、標準的な「高ストリンジェンシー条件」は、42℃の加湿オープン中において、50%ホルムアルデヒド、5×SSC、0.2μg/μlポリ(dA)、0.2μg/μlヒトcot1 DNAおよび0.5% SDS中での一晚のハイブリダイゼーション、その後55℃にて1×SSC、0.2% SDS中で5分間、次いで55℃にて0.1×SSC、0.2% SDS中で20分間マイクロアレイを連続洗浄することとして規定される。マイクロアレイに基づくハイブリダイゼーションについて、「中程度のストリンジェンシー条件」とは、構造的および機能的に関連したタンパク質をコードするmRNAへのクロスハイブリダイゼーションに適しており、高ストリンジェンシー条件についての条件と同じであるが、ハイブリダイゼーションの温度を下げ、室温（約25℃）で洗浄するように規定される。

10

#### 【0052】

本明細書中で用いられる場合、用語「タンパク質」、「ポリペプチド」および「ペプチド」は、交換可能に使用されて、天然に存在するかまたは合成されたアミノ酸のポリマーを長さに関わりなく示し、ここで、本明細書中のアミノ酸は、ペプチド結合に関与できる、天然に存在するアミノ酸、天然に存在するアミノ酸の構造改変体、および合成された天然に存在しないアナログを含む。用語「タンパク質」、「ポリペプチド」および「ペプチド」は、翻訳後および合成後の修飾（例えば、N末端アミノ酸またはC末端アミノ酸の切断反応およびグリコシル化）を、明示的に許容する。本明細書中において、用語「オリゴペプチド」は、25以下のアミノ酸残基を有する、タンパク質、ポリペプチドまたはペプチドを示す。

20

#### 【0053】

タンパク質、ポリペプチド、ペプチドまたはオリゴペプチドは、単離されたポリヌクレオチドによってコードされる場合；天然には見られない純度で存在する場合（ここで、純度は、他の細胞物質の存在に関して判断され得る）；および／あるいは天然には見出されないアミノ酸アナログもしくはアミノ酸誘導体、または標準的ペプチド結合以外の連結を含む場合に、「単離された」とみなされる。このように規定される場合「単離された」は、そのように記載されるタンパク質、ポリペプチド、ペプチドまたはオリゴペプチドが、天然の環境から物理的に取り出されたことを必ずしも必要としない。

30

#### 【0054】

タンパク質、ポリペプチド、ペプチドまたはオリゴペプチドは、少なくとも65%（例えば、少なくとも75%、85%または95%）の濃度（組成物中の総タンパク質に関する質量に基づいて測定される）で存在する場合に、本明細書中において「精製された」とみなされる。濃度が少なくとも85%である場合に、「実質的に精製された」とみなされる。

#### 【0055】

本明細書中において用いられる場合、用語「ホモログ」（「相同体」ともいう）は、「オルソログ」および「パラログ」を含む。「オルソログ」は、異なる生物種における、同じ遺伝子の別個の存在のことである。この別個の存在は、類似したアミノ酸配列または同一のアミノ酸配列を有し、ここで配列の類似性の程度は、その同じ遺伝子を有する共通の祖先からの、種の進化距離に部分的に依存する。「パラログ」は、1つの生物種における1つの遺伝子の別個の存在のことを示す。この別個の存在は、類似したアミノ酸配列または同一のアミノ酸配列を有し、ここで配列の類似性の程度は、この存在を生じる遺伝子複製事象からの、これらの別個の存在の進化距離に部分的に依存する。

40

#### 【0056】

「相同性」アミノ酸配列は、保存性アミノ酸置換を含むアミノ酸配列を含み、このポリペプチドは、実質的に同じ結合性および／または活性を有する。しかし、相同性アミノ酸配列は、他の既知のIgスーパーファミリーメンバーをコードするアミノ酸配列を含まない

50

。相同性（同一性%）は、例えば、デフォルト設定を用いるGAPプログラム（Wisconsin Sequence Analysis Package、Version 8 for Unix（登録商標）、Genetics Computer Group、University Research Park, Madison WI）によって、決定され得る（このGAPプログラムは、SmithおよびWaterman（Adv. Appl. Math.、2：482-489（1981）（これは本明細書中において、その全体が参考として援用される））のアルゴリズムを使用する）。

#### 【0057】

本明細書中で用いられる場合、用語「抗体」は、完全な抗体（2つの重鎖および2つの軽鎖からなる）またはそのフラグメントをいう。このようなフラグメントとしては、そのフラグメントが抗原に特異的に結合できる限り、以下が挙げられるが、これらに限定されない：種々のプロテアーゼを用いた消化によって生じるフラグメント、化学的切断および／または化学的解離によって生じるフラグメント、ならびに組換え産生されたフラグメント。これらのフラグメントには、Fabフラグメント、Fab'フラグメント、F(ab')<sub>2</sub>フラグメントおよび単鎖Fv（scFv）フラグメントがある。

10

#### 【0058】

配列が改変されているが、抗原に特異的に結合でき得る抗体もまた、用語「抗体」の範囲内にある。改変抗体の例は、種間キメラ抗体およびヒト化抗体；抗体融合物；ならびにヘテロマー抗体複合体（例えば、ディアボディ（diabody）（二重特異的抗体）、単鎖ディアボディおよびイントラボディ）である（例えば、Marasco（編）、Intracellular Antibodies：Research and Disease Applications、Springer-Verlag New York、Inc.（1998）（ISBN：3540641513）（この開示は、本明細書中においてその全体が参考として援用される）を参照のこと）。

20

#### 【0059】

「特異的結合」は、2つの分子が、環境中の他の分子への結合に優先して互いに結合する、能力をいう。典型的には、「特異的結合」は、反応中の偶然の結合より、少なくとも2倍、より典型的には少なくとも10倍、しばしば少なくとも100倍、識別する。典型的に、特異的結合反応のアフィニティーまたはアビディティは、少なくとも約 $10^{-7}$  M（例えば、少なくとも約 $10^{-8}$  Mまたは $10^{-9}$  M）である。

30

#### 【0060】

用語「領域」は、生体分子の1次構造の物理的に連続した位置を意味する。タンパク質の場合、領域は、そのタンパク質のアミノ酸配列の連続した位置によって規定される。

#### 【0061】

用語「ドメイン」は、生体分子の公知の機能または推測される機能に寄与する生体分子の構造をいう。ドメインは、それらの領域またはそれらの部分と同一の広がりであり得；ドメインはまた、生体分子の異なる非連続的な領域も含み得る。GP354タンパク質ドメインの例としては、細胞外Igドメイン（すなわち、N末端）、膜貫通型ドメイン、および細胞質ドメイン（すなわち、C末端）が挙げられるがこれらに限定されない。

#### 【0062】

本明細書中で使用される場合、用語「化合物」は、低分子、ペプチド、タンパク質、糖、ヌクレオチド、核酸、脂質などが挙げられるがこれらに限定されない任意の分子を意味し、このような化合物は、天然または合成であり得る。

40

#### 【0063】

他に規定がない限り、本明細書中で使用される全ての技術用語および科学用語は、本発明が属する分野の当業者によって一般に理解される意味と同じ意味を有する。例示的な方法および材料が、以下で記載されるが、本明細書中に記載される方法および材料と類似または等価の方法および材料もまた、本発明の実施において使用され得、そしてそれらは、当業者には明白である。本明細書中で言及される全ての刊行物および他の参考文献は、それら全体が参考として援用される。矛盾している場合、定義を含む本明細書が、支配する。

50

これらの材料、方法および実施例は、例示でのみあり、制限されることを意図しない。

【0064】

当業者に公知の組換えDNA技術の一般的な原理を示している標準的な参考としては、以下が挙げられる：Ausubelら、CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York (1998および2001に補遺)；Sambrookら、MOLECULAR CLONING: A LABORATORY MANUAL, 第2版, Cold Spring Harbor Laboratory Press, Plainview, New York (1989)；Kaufmanら、編, HANDBOOK OF MOLECULAR AND CELLULAR METHODS IN BIOLOGY AND MEDICINE, CRC Press, Boca Raton (1995)；McPherson, 編, DIRECTED MUTAGENESIS: A PRACTICAL APPROACH, IRL Press, Oxford (1991)。当業者に公知の免疫学の一般的な原理を示している標準的な参考資料としては、以下が挙げられる：HarlowおよびLane, ANTIBODIES: A LABORATORY MANUAL, 第2版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1999)；およびRoittら、IMMUNOLOGY, 第3版, Mosby-Year Book Europe Limited, London (1993)。当業者に公知の医学生理学および薬理学の一般的な原理を示している標準的な参考資料としては、以下が挙げられる：Harris 20  
on's PRINCIPLES OF INTERNAL MEDICINE, 第14版, (Anthony S. Fauciら編), McGraw-Hill Companies, Inc., 1998。

【0065】

(GP354 関連核酸)

gp354 遺伝子は、BACクローンのコンティグ38において、GenBank登録番号AC022315 (これは、2000年2月10日に登録された)で同定された。この寄託物は、gp354のヒトゲノム配列 (図4および配列番号5)を有し、これは、5' 30  
上流 (1~6278位) および3' 下流 (16490~20050) の非転写ゲノム配列を含む。

【0066】

本発明は、GP354タンパク質全体をコードする単離されたポリヌクレオチドを提供する。上記のように、本発明のこのような「全長」ポリヌクレオチドは、特に、全長GP354タンパク質を発現するために使用され得る。この全長ポリヌクレオチドはまた、核酸プローブとして (プローブとして) 使用され得、これらの実施形態の単離されたポリヌクレオチドは、gp354ポリヌクレオチドおよび関連ポリヌクレオチド配列にハイブリダイズする。

【0067】

好ましい実施形態において、本発明は、以下の一部または全てを含む単離されたポリヌクレオチドを提供する：(i) 配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号9、配列番号11、配列番号25、配列番号27または配列番号32のヌクレオチド配列、(ii) 配列番号1、配列番号3、配列番号5、配列番号6、配列番号7、配列番号9、配列番号11、配列番号25、配列番号27または配列番号32のヌクレオチド配列の縮重改変体、あるいは(iii) (i) または(ii) の相補体。 40

【0068】

配列番号1は、推定されるgp354 cDNA配列を示す。配列番号3は、RT-PCRから生じた例示的なgp354 cDNAフラグメントを示す。配列番号5は、5' 非転写領域および3' 非転写領域を含む、gp354コード配列を含むゲノムDNA配列を示す。配列番号6は、非コード領域の上流に含まれるゲノムDNA配列を示す。配列番号7は、gp354 cDNA配列の誘導された長型を示す。配列番号11は、別の例示的 50



な膵臓 g p 3 5 4 c D N A 配列を示す。配列番号 2 5 および配列番号 2 7 は、それぞれ、配列番号 7 の選択的スプライス変異体である P a n C A M - 1 および P a n C A M - 2 のコード配列を示す。

#### 【0069】

他の実施形態において、本発明は、以下を含む単離されたポリヌクレオチドを提供する：  
 (i) 配列番号 2、配列番号 4、配列番号 8、配列番号 10、配列番号 12、配列番号 26、配列番号 28 または配列番号 31 のアミノ酸配列を有するポリペプチドをコードするヌクレオチド配列、あるいは (i i) 配列番号 2、配列番号 4、配列番号 8、配列番号 10、配列番号 12、配列番号 26、配列番号 28 または配列番号 31 のアミノ酸配列を有するポリペプチドをコードするヌクレオチド配列の相補体。

10

#### 【0070】

配列番号 2、配列番号 4、配列番号 8、配列番号 10、配列番号 12、配列番号 26、配列番号 28 および配列番号 31 は、それぞれ、配列番号 1、配列番号 3、配列番号 7、配列番号 9、配列番号 11、配列番号 25、配列番号 27 および配列番号 32 によってコードされる G P 3 5 4 のアミノ酸配列を示す。例えば、図 1 ~ 10 を参照のこと。

#### 【0071】

他の実施形態において、本発明は、以下のヌクレオチド配列を有する単離されたポリヌクレオチドを提供する： (i) 配列番号 2、配列番号 8 または配列番号 12 の配列を有するポリペプチドをコードする、ヌクレオチド配列、 (i i) 保存的アミノ酸置換を有する配列番号 2、配列番号 8 または配列番号 12 の配列を有するポリペプチドをコードする、ヌクレオチド配列、あるいは (i i i) (i) または (i i) の相補体である、ヌクレオチド配列。ここで、配列番号 2 は、配列番号 1 の c D N A によってコードされる G P 3 5 4 のアミノ酸配列を示し、配列番号 8 は、配列番号 5 および配列番号 11 由来の配列によってコードされる G P 3 5 4 のアミノ酸配列を示し、そして配列番号 12 は、配列番号 11 の膵臓 c D N A によってコードされる G P 3 5 4 のアミノ酸配列を示す。

20

#### 【0072】

(G P 3 5 4 の一部をコードする核酸)

本発明はまた、G P 3 5 4 の選択部分をコードする単離されたポリヌクレオチドを提供する。本明細書中以下でさらに議論されるように、これらの「核酸分子」は、例えば、単独でかまたは融合タンパク質の要素として、G P 3 5 4 の特定の部分を発現するために使用され得る。核酸フラグメントはまた、領域特異的核酸プローブとしても使用され得る。

30

#### 【0073】

好ましい実施形態において、本発明は、以下を含む単離されたポリヌクレオチドを提供する： (i) 配列番号 3、配列番号 6 または配列番号 9 のヌクレオチド配列、 (i i) 配列番号 3、配列番号 6 または配列番号 9 のヌクレオチド配列の縮重改変体、あるいは (i i i) (i) または (i i) の相補体。配列番号 3 は、g p 3 5 4 膵臓 R N A 由来の 7 8 5 塩基対 R T - P C R フラグメントを示す。配列番号 6 は、g p 3 5 4 コード配列から上流のゲノム配列を示し、そして配列番号 9 は、g p 3 5 4 膵臓 R N A 由来の 1 7 8 2 塩基対 R T - P C R フラグメントを示す。

40

#### 【0074】

他の実施形態において、単離されたポリヌクレオチド、またはそれらの相補体は、G P 3 5 4 の N 末端細胞外部分に特徴的な I g ドメインを少なくとも 1 つ、好ましくは、2 つ、3 つ、4 つまたは 5 つ有するポリペプチドをコードする。詳細には、5 つの細胞外 I g ドメインは、それぞれ、配列番号 1 の g p 3 5 4 c D N A 配列のヌクレオチド 1 0 3 ~ 3 0 6、4 0 6 ~ 6 0 9、7 1 5 ~ 8 7 0、9 6 7 ~ 1 1 2 2 および 1 2 2 8 ~ 1 4 4 5 (図 1 参照)、ならびに、それぞれ、配列番号 8 の g p 3 5 4 c D N A 配列のヌクレオチド 3 0 7 ~ 5 1 0、6 1 0 ~ 8 1 3、9 1 9 ~ 1 0 7 4、1 1 7 1 ~ 1 3 2 6 および 1 4 3 2 ~ 1 6 5 9 (図 5 参照) によってコードされる。好ましい実施形態において、単離されたポリヌクレオチドは、少なくとも 1 コピーにおいて、少なくとも 2 つ、好ましくは 3

50

つ、より好ましくは4つ、そして最も好ましくは、5つ全てのドメインをコードする。

#### 【0075】

いくつかの使用について、このようなタンパク質産物、核酸フラグメント（またはそれらの相補体）は、膜を通るそのコードされたポリペプチドの輸送を媒介するシグナル分泌配列をコードする配列を含む。このようなシグナル分泌配列は、代表的には、膜を通る輸送が起こる場合、ポリペプチドから切断される。GP354シグナル分泌配列は、配列番号1のgp354 cDNA配列のヌクレオチド1～54（図1参照）、および配列番号8のgp354 cDNAのヌクレオチド1～57（図5参照）によってコードされる。より好ましくは、本発明の単離されたポリヌクレオチドのシグナル分泌配列は、gp354由来である。GP354のシグナル配列がまた、分泌の間に切断されると仮定した場合、成熟GP354ポリペプチド配列は、配列番号1のヌクレオチド55～57（図1参照）、および配列番号8のgp354 cDNAのヌクレオチド259～261（図5参照）によってコードされるN末端プロリン残基を有する。

10

#### 【0076】

本発明のポリヌクレオチドの他の好ましい実施形態は、GP354の膜貫通ドメインを有するポリペプチドをコードするポリヌクレオチド、またはその相補体である。上記の好ましい、単離されたポリヌクレオチドは、例えば、コードされたポリペプチドの膜への挿入が所望される場合に、必要に応じて、膜貫通ドメインをコードし得る。この膜貫通ドメインは、gp354配列によってコードされ得るか、または異種膜関連タンパク質の膜貫通ドメインをコードする異種遺伝子によってコードされ得る。gp354膜貫通ドメインは、配列番号1のgp354 cDNA配列のヌクレオチド1522～1590（図1参照）、および配列番号8のgp354 cDNAのヌクレオチド1726～1794（図5参照）によってコードされる。

20

#### 【0077】

所望の場合、本発明の単離されたポリヌクレオチドは、例えば、GP354結合相互作用に応答する特定のシグナル反応が望まれる場合、細胞内C末端ドメインをコードする配列（またはそれらの相補体）を含み得る。この細胞内ドメインは、gp354によってコードされ得るか（以下を参照）、または異種膜関連タンパク質の細胞内ドメインをコードする異種遺伝子によってコードされ得る。本発明の好ましいポリヌクレオチドは、GP354の（C末端）細胞内ドメインを有するポリペプチドをコードするポリヌクレオチド、またはそれらの相補体である。詳細には、GP354の1つの細胞内ドメインは、配列番号1のgp354 cDNA配列のヌクレオチド1591～1776（図1参照）によってコードされる。GP354の細胞内ドメインのより長い形態は、配列番号8のgp354 cDNA配列のヌクレオチド1795～2319（図5参照）によってコードされる。

30

#### 【0078】

本発明の1つの好ましい、単離されたポリヌクレオチドは、図3に示され（配列番号3を参照）、配列番号1のgp354 cDNA配列のヌクレオチド139～923（図1参照）を含む。このポリヌクレオチドは、プライマーGX1-218（配列番号16）およびGX1-219（配列番号17）を用いて、腭臓RNAから増幅されるRT-PCRフラグメントの配列を含む。実施例2を参照のこと。この好ましい、単離されたポリヌクレオチドは、配列番号2のアミノ酸47～307をコードし、すなわち、このポリヌクレオチドは、第1のN末端Igドメインのアミノ酸13～68（すなわち、このIgドメインの最初の12N末端アミノ酸を欠失している）をコードし、そしてGP354の第2のIgドメインおよび第3のIgドメインをコードする。

40

#### 【0079】

（クロスハイブリダイズする核酸）

核酸の実施形態の別のシリーズにおいて、本発明は、本発明の種々のgp354核酸にハイブリダイズする単離されたポリヌクレオチドを提供する。これらの「クロスハイブリダイズする核酸」は、特に、本発明のgp354に関連するタンパク質（さらに、アイソフォーム、ホモログ、パラログ、またはオルソログについても同様）に対するプローブとし

50

て、およびこれらのタンパク質の発現を駆動するために、使用され得る。

#### 【0080】

いくつかのこのような実施形態において、本発明は、高ストリンジェンシー条件下で、配列番号1、5、7、9または11；配列番号1、5、7、9または11の相補体；あるいは少なくとも17核酸ユニットを有するそれらのフラグメントを含むヌクレオチド配列を有するプローブにハイブリダイズする配列を含む、単離されたポリヌクレオチドを提供する。

#### 【0081】

(好ましい核酸)

脾臓組織または神経組織において発現される核酸（または、その相補体）は、上記の核酸の中で特に好ましい。上述のような、gp354の生物学的活性を有するポリペプチドをコードする核酸（または、その相補体）もまた、上記の核酸の中で特に好ましい。

#### 【0082】

(核酸フラグメント)

核酸の実施形態の別のシリーズにおいて、本発明は、特に、領域特異的核酸プローブとして、増幅プライマーとして、およびエピトープタンパク質フラグメントまたは免疫原性タンパク質フラグメントの発現または合成を指向するために有用性を示す、本発明の種々の単離されたポリヌクレオチドのフラグメントを提供する。

#### 【0083】

いくつかの実施形態において、本発明は、配列番号1、5、7、9または11から選択される連続する核酸配列の少なくとも17ヌクレオチド、少なくとも18ヌクレオチド、少なくとも20ヌクレオチド、少なくとも24ヌクレオチド、または少なくとも25ヌクレオチドを含む、単離されたポリヌクレオチドを提供する。

#### 【0084】

他の実施形態において、本発明は、(i) 配列番号2、4、8、10または12の少なくとも8個連続するアミノ酸の配列を有するポリペプチドをコードする、ヌクレオチド配列、(ii) 保存的アミノ酸置換を有する、配列番号2、4、8、10または12の少なくとも8個連続するアミノ酸の配列を有するポリペプチドをコードするヌクレオチド配列、あるいは(iii) (i) または(ii) の相補体である、ヌクレオチド配列、を含む単離された核酸を提供する。

#### 【0085】

(単一エキソンプローブ)

本発明はさらに、gp354遺伝子の1つのみのエキソンの一部を有する、ゲノム由来の単一エキソンプローブを提供する。このような単一エキソンプローブは、スプライス改変体の同定および特徴付けにおいて特定の有用性を有する。特に、このような単一エキソンプローブは、gp354の異なるアイソフォームの発現を同定および識別するために有用である。

#### 【0086】

いくつかの実施形態において、本発明は、配列番号1、5、7、9もしくは11、または配列番号1、5、7、9もしくは11の相補体の、以下のエキソン特異的部分の1つから選択されるヌクレオチド配列を含む、単離された核酸を提供する。ここで、この部分は、配列番号1、5、7、9もしくは11、またはそれらの相補体の部分のいずれか1つの少なくとも17個連続するヌクレオチド、少なくとも18個連続するヌクレオチド、少なくとも20個連続するヌクレオチド、少なくとも24個連続するヌクレオチド、少なくとも25個連続するヌクレオチド、または少なくとも50個連続するヌクレオチドを含む。

#### 【0087】

(表1：gp354 cDNA配列（配列番号1または2）およびゲノム配列（配列番号5）のエキソン座標)

#### 【0088】

【表1】

|         | <u>cDNA-1</u> | <u>cDNA-2</u> | <u>ゲノム</u>  |
|---------|---------------|---------------|-------------|
| エキソン 1  | 1-52          | 1-52          | 6483-6534   |
| エキソン 2  | 53-202        | 53-202        | 6699-6848   |
| エキソン 3  | 203-352       | 203-352       | 7762-7911   |
| エキソン 4  | 353-513       | 353-513       | 8058-8218   |
| エキソン 5  | 514-664       | 514-664       | 8835-8985   |
| エキソン 6  | 665-770       | 665-770       | 9651-9756   |
| エキソン 7  | 771-919       | 771-919       | 9873-10021  |
| エキソン 8  | 920-1047      | 920-1041      | 10263-10390 |
| エキソン 9  | 1048-1180     | 1042-1180     | 10476-10608 |
| エキソン 10 | 1181-1281     | 1181-1281     | 10895-10995 |
| エキソン 11 | 1282-1501     | 1282-1501     | 11159-11378 |
| エキソン 12 | 1502-1606     | 1502-1606     | 11847-11951 |
| エキソン 13 | 1607-1710     | 1607-1716     | 12287-12390 |
| エキソン 14 | 1711-1779     | 1717-1782     | 14002-14067 |

10

20

(表 2 : g p 3 5 4 c D N A - 4 配列 (配列番号 1 1) およびゲノム配列 (配列番号 5) のエキソン座標)

【 0 0 8 9 】

【表 2】

|         | <u>cDNA</u> | <u>ゲノム</u>  |
|---------|-------------|-------------|
| エキソン 1  | 1-256       | 6278-6534   |
| エキソン 2  | 257-406     | 6699-6848   |
| エキソン 3  | 407-556     | 7762-7911   |
| エキソン 4  | 557-717     | 8058-8218   |
| エキソン 5  | 718-868     | 8835-8985   |
| エキソン 6  | 869-974     | 9651-9756   |
| エキソン 7  | 975-1123    | 9873-10021  |
| エキソン 8  | 1124-1245   | 10263-10390 |
| エキソン 9  | 1246-1384   | 10476-10608 |
| エキソン 10 | 1385-1485   | 10895-10995 |
| エキソン 11 | 1486-1705   | 11159-11378 |
| エキソン 12 | 1706-1810   | 11847-11951 |
| エキソン 13 | 1811-1920   | 12281-12390 |
| エキソン 14 | 1921-1986   | 14002-14067 |
| エキソン 15 | 1987-2959   | 15511-16483 |

30

40

50

## (転写制御核酸)

別の局面において、本発明は、g p 3 5 4 遺伝子の転写を制御する核酸配列エレメントを含む、ゲノム由来の単離されたポリヌクレオチドを提供する。これらの核酸は、特に、組み換え構築物中の異種コード領域の発現を駆動するために使用され得、従って、このような異種コード領域に、ネイティブ g p 3 5 4 遺伝子の発現パターンを与える。これらの核酸はまた、逆に、g p 3 5 4 のゲノム遺伝子座に異種転写制御エレメントを標的化するために使用され得、g p 3 5 4 遺伝子自体の発現パターンを変化させる。

## 【0090】

このような実施形態の第1のシリーズにおいて、本発明は、配列番号5のヌクレオチド1～6 4 8 3；配列番号5のヌクレオチド1 4 8 3～6 4 8 2；配列番号5のヌクレオチド2 4 8 3～6 4 8 2；配列番号5のヌクレオチド3 4 8 3～6 4 8 2；配列番号5のヌクレオチド4 4 8 3～6 4 8 2；配列番号5のヌクレオチド5 4 8 3～6 4 8 2；もしくは配列番号5のヌクレオチド5 9 8 3～6 4 8 2；またはこのような配列の相補体を含む、単離されたポリヌクレオチドを提供する。 10

## 【0091】

他の実施形態において、本発明は、配列番号5のヌクレオチド1～6 4 8 3；配列番号5のヌクレオチド1 4 8 3～6 4 8 2；配列番号5のヌクレオチド2 4 8 3～6 4 8 2；配列番号5のヌクレオチド3 4 8 3～6 4 8 2；配列番号5のヌクレオチド4 4 8 3～6 4 8 2；配列番号5のヌクレオチド5 4 8 3～6 4 8 2；もしくは配列番号5のヌクレオチド5 9 8 3～6 4 8 2；またはこのような配列の相補体の、少なくとも17、18、20、24、または25ヌクレオチドを含む、単離されたポリヌクレオチドを提供する。 20

## 【0092】

配列番号5のヌクレオチド1～6 4 8 3；配列番号5のヌクレオチド1 4 8 3～6 4 8 2；配列番号5のヌクレオチド2 4 8 3～6 4 8 2；配列番号5のヌクレオチド3 4 8 3～6 4 8 2；配列番号5のヌクレオチド4 4 8 3～6 4 8 2；配列番号5のヌクレオチド5 4 8 3～6 4 8 2；もしくは配列番号5のヌクレオチド5 9 8 3～6 4 8 2；またはこのような配列の相補体を含む単離されたポリヌクレオチドの各々は、g p 3 5 4 遺伝子の発達特異的および組織特異的な発現および調節を媒介する転写制御配列を有する。このような転写制御配列は、それらに作動可能に連結された異種核酸配列に、このような発達特異的および組織特異的な発現パターンを与えるために有用である。 30

## 【0093】

## (g p 3 5 4 核酸分子の他の規定された特徴)

本明細書中に詳細に示される全ての核酸配列は、デオキシリボヌクレオチドの配列として示される。しかし、示される配列は、ポリヌクレオチド組成に適するように解釈されることが意図される：例えば、単離された核酸がRNAで構成される場合、示される配列は、リボヌクレオチドを意図し、チミジンがウリジンで置換される。

## 【0094】

多型（例えば、単一ヌクレオチド多型（SNP））は、真核生物ゲノムにおいて頻繁に発生する。1, 400, 000より多いSNPが、ヒトゲノムにおいてすでに同定されており（International Human Genome Sequencing Consortium, Nature 409: 860-921 (2001)）、そして 40  
1つの種の1つの個体から決定された配列は、その集団に存在する他の対立遺伝子形態と異なり得る。あるいは、単一ヌクレオチド多型ではない小さな欠失および挿入は、全体的な集団においてまれではなく、しばしばそのタンパク質の機能を変更しない。

## 【0095】

従って、本発明は、本明細書中に詳細に記載されるポリヌクレオチド（例えば、配列番号1、3、5、6、7、9、および11）と配列において同一な単離されたポリヌクレオチドを提供するだけでなく、それらの詳細に記載される核酸配列の対立遺伝子改変体である単離されたポリヌクレオチドもまた提供することが特に強調される。さらに、本発明は、配列番号1、3、5、6、7、9、および11または少なくとも1つのIgドメインをコ 50

ードするこれらの配列のいずれか1つの部分と配列において少なくとも約65%同一であり、代表的には、本明細書中に詳細に記載されるポリヌクレオチドと配列において少なくとも約70%、75%、80%、85%、または90%同一であり、有用には、本明細書中に詳細に記載されるポリヌクレオチドと配列において少なくとも約91%、92%、93%、94%、または95%同一であり、より有用には本明細書中に詳細に記載されるポリヌクレオチドと配列において少なくとも約96%、97%、98%、または99%同一であり、そして最も保存的には、本明細書中に詳細に記載されるポリヌクレオチドと配列において少なくとも約99.5%、99.6%、99.7%、99.8%、または99.9%同一である、gp354のホモログ（例えば、パラログおよびオルソログ）を提供する。これらの配列改変体は天然に存在し得るか、または無作為変異誘発または特異的変異誘発によるようなヒトの介入により生じ得る。 10

#### 【0096】

核酸配列改変体は、例えば、配列番号7に存在する配列の252位、703位、770位、1249位、および1811～1816位に生じることが見出された。

#### 【0097】

本明細書における目的のために、2つの核酸配列のパーセント同一性は、Tatianaら、「Blast 2 sequences—a new tool for comparing protein and nucleotide sequences」、FEMS Microbiol Lett. 174:247–250 (1999)の手順を用いて決定される。この手順は、National Center for Biotechnology Informationにてオンラインで利用可能なコンピュータプログラムBLAST2 SEQUENCESにより達成される。 20

#### 【0098】

核酸配列のパーセント同一性を評価するために、BLAST2 SEQUENCESのBLASTNモジュールが、以下のデフォルト値で使用され、そして両方の配列は、それらの全体において入力される：(i)リワード フォー ア マッチ (reward for a match):1；(ii)ペナルティー フォー ア ミスマッチ (penalty for a mismatch):−2；(iii)オープンギャップ (open gap) 5ペナルティーおよびエクステンションギャップ (extension gap) 2ペナルティー；(iv)ギャップX\_\_ドロップオフ50エクスペクト (gap X\_\_drop off 50 expect) 10ワードサイズ11フィルター。 30

#### 【0099】

本発明の単離されたポリヌクレオチドは、GP354タンパク質およびタンパク質フラグメントの発現に有用であり、従って、本発明は、GP354タンパク質およびその一部をコードする単離されたポリヌクレオチドを提供する（本明細書中に詳細に記載されるポリヌクレオチドと配列において同一なポリヌクレオチドだけでなく、それらの縮重改変体も同様に提供する）。周知のように、遺伝暗号は縮重し、そして最適な発現のためのコドンの選択は種間で変化する。

#### 【0100】

また、周知のように、アミノ酸置換は、天然の対立遺伝子改変体の間で頻繁に生じ、しばしば保存的置換は、タンパク質機能において最小の (de minimis) 変化しか生じない。 40

#### 【0101】

従って、本発明は、本明細書中に詳細に記載されるポリヌクレオチドと配列において同一のポリヌクレオチドだけでなく、保存的アミノ酸置換または中程度の保存的アミノ酸置換を有する、GP354およびその一部分をコードするポリヌクレオチドもまた提供する。

#### 【0102】

保存的アミノ酸置換と呼ぶために種々の基準が存在するが（主に、進化上の関連タンパク質間で観察される変化か、または予想される化学的類似性かのいずれかに基づく）、本明細書における目的のためには、保存的置換は、本明細書の以下に複写したPAM250対 50

数尤度行列 (log-likelihood matrix) において正の値を有する任意の変化である (Gonnet ら、Science 256 (5062) : 1443-5 (1992) を参照のこと)。

【0103】

【化1】

|   | A  | R  | N  | D  | C  | Q  | E  | G  | H  | I  | L  | K  | M  | F  | P  | S  | T  | W  | Y  | V  |
|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| A | 2  | -1 | 0  | 0  | 0  | 0  | 0  | 0  | -1 | -1 | -1 | 0  | -1 | -2 | 0  | 1  | 1  | -4 | -2 | 0  |
| R | -1 | 5  | 0  | 0  | -2 | 2  | 0  | -1 | 1  | -2 | -2 | 3  | -2 | -3 | -1 | 0  | 0  | -2 | -2 | -2 |
| N | 0  | 0  | 4  | 2  | -2 | 1  | 1  | 0  | 1  | -3 | -3 | 1  | -2 | -3 | -1 | 1  | 0  | -4 | -1 | -2 |
| D | 0  | 0  | 2  | 5  | -3 | 1  | 3  | 0  | 0  | -4 | -4 | 0  | -3 | -4 | -1 | 0  | 0  | -5 | -3 | -3 |
| C | 0  | -2 | -2 | -3 | 12 | -2 | -3 | -2 | -1 | -1 | -2 | -3 | -1 | -1 | -3 | 0  | 0  | -1 | 0  | 0  |
| Q | 0  | 2  | 1  | 1  | -2 | 3  | 2  | -1 | 1  | -2 | -2 | 2  | -1 | -3 | 0  | 0  | 0  | -3 | -2 | -2 |
| E | 0  | 0  | 1  | 3  | -3 | 2  | 4  | -1 | 0  | -3 | -3 | 1  | -2 | -4 | 0  | 0  | 0  | -4 | -3 | -2 |
| G | 0  | -1 | 0  | 0  | -2 | -1 | -1 | 7  | -1 | -4 | -4 | -1 | -4 | -5 | -2 | 0  | -1 | -4 | -4 | -3 |
| H | -1 | 1  | 1  | 0  | -1 | 1  | 0  | -1 | 6  | -2 | -2 | 1  | -1 | 0  | -1 | 0  | 0  | -1 | 2  | -2 |
| I | -1 | -2 | -3 | -4 | -1 | -2 | -3 | -4 | -2 | 4  | 3  | -2 | 2  | 1  | -3 | -2 | -1 | -2 | -1 | 3  |
| L | -1 | -2 | -3 | -4 | -2 | -2 | -3 | -4 | -2 | 3  | 4  | -2 | 3  | 2  | -2 | -2 | -1 | -1 | 0  | 2  |
| K | 0  | 3  | 1  | 0  | -3 | 2  | 1  | -1 | 1  | -2 | -2 | 3  | -1 | -3 | -1 | 0  | 0  | -4 | -2 | -2 |
| M | -1 | -2 | -2 | -3 | -1 | -1 | -2 | -4 | -1 | 2  | 3  | -1 | 4  | 2  | -2 | -1 | -1 | -1 | 0  | 2  |
| F | -2 | -3 | -3 | -4 | -1 | -3 | -4 | -5 | 0  | 1  | 2  | -3 | 2  | 7  | -4 | -3 | -2 | 4  | 5  | 0  |
| P | 0  | -1 | -1 | -1 | -3 | 0  | 0  | -2 | -1 | -3 | -2 | -1 | -2 | -4 | 8  | 0  | 0  | -5 | -3 | -2 |
| S | 1  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | -2 | -2 | 0  | -1 | -3 | 0  | 2  | 2  | -3 | -2 | -1 |
| T | 1  | 0  | 0  | 0  | 0  | 0  | 0  | -1 | 0  | -1 | -1 | 0  | -1 | -2 | 0  | 2  | 2  | -4 | -2 | 0  |
| W | -4 | -2 | -4 | -5 | -1 | -3 | -4 | -4 | -1 | -2 | -1 | -4 | -1 | 4  | -5 | -3 | -4 | 14 | 4  | -3 |
| Y | -2 | -2 | -1 | -3 | 0  | -2 | -3 | -4 | 2  | -1 | 0  | -2 | 0  | 5  | -3 | -2 | -2 | 4  | 8  | -1 |
| V | 0  | -2 | -2 | -3 | 0  | -2 | -2 | -3 | -2 | 3  | 2  | -2 | 2  | 0  | -2 | -1 | 0  | -3 | -1 | 3  |

10

20

30

本明細書における目的のために、「中程度の保存的」置換は、本明細書の上に複写した PAM 250 対数尤度行列においてマイナスではない値を有する任意の変化である。

【0104】

生物学的活性の重度の減少または除去を回避するために、種々の種の GP 354 タンパク質間または Ig ファミリーメンバー間で保存されるアミノ酸残基は、遺伝子操作時に変更されない (保存的置換を除く)。例えば、GP 354 の Ig ドメインを維持するために、システイン残基は保存されるべきである。

【0105】

ポリヌクレオチドの関連性はまた、機能試験、2つのポリヌクレオチドが規定のハイブリダイゼーションのストリンジェンシーで互いに対合する能力を使用して特徴付けられ得る。従って本発明は、特に本明細書中で記載されるポリヌクレオチドに対して配列が同一である単離ポリヌクレオチドを提供するだけでなく、また、高ストリンジェンシー条件 (本明細書中で規定される通り) の下で本発明の単離 gp 354 ポリヌクレオチド (「参照核酸」) の全てまたは種々の部分にハイブリダイズする単離ポリヌクレオチド (「交差ハイブリダイズ性の核酸」) も提供する。

40

【0106】

このような交差ハイブリダイズ性の核酸は有用であり、とりわけ、選択的スプライシングの改変体およびホモログ (例えば、オルソログおよびパラログ) のような、本発明のタンパク質に関係するタンパク質に対するプローブとして有用であり、そしてこのタンパク質の発現を駆動するために有用である。特に有用なオルソログは、他の霊長類種由来 (例え

50

ば、チンパンジー、カニクイザル、ヒヒ、オラウータン、およびゴリラ) ; げっ歯類由来 (例えば、ラット、マウス、モルモット) ; ウサギ目 (例えば、ウサギ) および家畜由来 (例えば、ウシ、ブタ、ヒツジ、ウマ、ヤギ) のオルソログである。

#### 【0107】

参照核酸のハイブリダイズ部分は、代表的に、少なくとも15ヌクレオチド長であり、そしてしばしば、少なくとも17ヌクレオチド長、20ヌクレオチド長、25ヌクレオチド長、30ヌクレオチド長、35ヌクレオチド長、40ヌクレオチド長、または50ヌクレオチド (nt) 長である。参照核酸のより大きな部分、例えば、少なくとも50ntまで、100ntまで、150ntまで、200ntまで、250ntまで、300ntまで、350ntまで、400ntまで、450ntまで、500nt以上までの部分、および参照核酸の完全な長さを含む部分に対してハイブリダイズする、交差ハイブリダイズ核酸もまた、有用である。 10

#### 【0108】

交差ハイブリダイズ核酸のハイブリダイズ部分は、参照核酸の少なくとも一部に対して、配列で少なくとも75%同一である。代表的に、交差ハイブリダイズ核酸のハイブリダイズ部分は、参照核酸の少なくとも一部の配列に対して、少なくとも80%、しばしば少なくとも85%、86%、87%、88%、89%、またさらに少なくとも90%の配列で同一である。しばしば、交差ハイブリダイズ核酸のハイブリダイズ部分は、参照核酸配列の少なくとも一部の配列に対して、少なくとも、91%、92%、93%、94%、95%、96%、97%、98%、または99%の配列で同一である。時々、交差ハイブリダイズ核酸のハイブリダイズ部分は、参照核酸の少なくとも一部に対して、少なくとも99.5%の配列で同一である。 20

#### 【0109】

本発明はまた、本発明の単離されたポリヌクレオチドまたは核酸の種々のフラグメントを提供する。参照核酸の「フラグメント」によって、本明細書中で、単離されたポリヌクレオチドまたは核酸が意図され、どのように獲得されるにせよ、これは参照核酸配列の一部に同一な核酸配列を有し、ここで、この参照核酸の部分は少なくとも17ヌクレオチドであり、そして参照核酸の完全長よりも少ない。

#### 【0110】

理論上、17ヌクレオチドのオリゴヌクレオチドは、30億塩基のヒトゲノム中で1回よりも少ない頻度でランダムに生じるのに十分長さであり、従って、哺乳動物のゲノムの複雑なものの核酸混合物中の参照配列を一意的に同定し得る核酸プローブを提供するのに十分な長さである。さらなる特異性は、ゲノムより低度の (subgenomic) 複雑なものの核酸サンプルを探ることによって、そして/または例えばポリメラーゼ連鎖反応 (PCR) によるような、収集して核酸の増幅を開始する (prime) ために、17ヌクレオチド長ほどの複数のフラグメントを使用することによって、獲得され得る。 30

#### 【0111】

本発明の核酸プローブを使用して、ホモログまたは同一タンパク質をコードするRNA転写物またはゲノム配列を検出し得る。このプローブは、それに結合された標識基 (例えば、放射性同位体、蛍光化合物、酵素または酵素補因子) を含み得る。このようなプローブは、(i) GP354タンパク質を誤って発現する (例えば、異所性のスプライシング、異常なmRNAレベル) 細胞または組織または(ii) gp354遺伝子中に変異 (例えば、欠失、挿入、もしくは点変異) を有する細胞または組織を同定するための診断キットの一部として使用され得る。このような診断キットは、好ましくは、標識された試薬および使用のための指示書の組み込みを含む。 40

#### 【0112】

本発明の単離ポリヌクレオチドはまた、PCR、プライマー伸長などにおけるプライマーとして使用され得る。プライマーとして有用であるために、このポリヌクレオチドは、例えば、少なくとも6ヌクレオチド長 (例えば、少なくとも7ヌクレオチド長、8ヌクレオチド長、9ヌクレオチド長または10ヌクレオチド長) であり得る。これらのプライマー 50



は、例えば、g p 3 5 4 mRNAまたはcDNAの増幅のために、g p 3 5 4 遺伝子のエキソン配列にハイブリダイズし得る。あるいは、これらのプライマーは、転写されない、例えば、g p 3 5 4 遺伝子のゲノム構造の調節部分を利用するために、g p 3 5 4 遺伝子のイントロン配列、または上流もしくは下流の調節配列にハイブリダイズし得る。

#### 【0113】

本発明の核酸プライマーをまた使用して、例えば、SNP検出のための1塩基伸長（SBE）を開始（prime）し得る（例えば、米国特許第6,004,744号（この開示は本明細書中で参考としてその全体が援用される）を参照のこと）。等温増幅アプローチ（例えば、環状増幅を周回（roll）させること）もまたここで良く記載される。例えば、Schweitzerら、Curr. Opin. Biotechnol. 12（1） 10  
：21-7（2001）；米国特許第5,854,033号および同5,714,320号、ならびにWO 97/19193およびWO 00/15779（これらの開示は本明細書中で参考としてその全体が援用される）を参照のこと。環状増幅を周回させることは、SNP検出を容易にするために他の技術と組み合わせられ得る。例えば、Lizardiら、Nature Genet. 19（3）225-32（1998）を参照のこと。

#### 【0114】

以下に記載されるように、少なくとも6連続アミノ酸をコードする核酸フラグメント（すなわち、18ヌクレオチド以上のフラグメント）が、参照核酸によってコードされるタンパク質のエピトープをマッピングする際に有用性を有する、ペプチドの発現または合成を指向するのに有用である。例えば、Geysenら、Proc. Natl. Acad. Sci. USA 81：3998-4002（1984）ならびに米国特許第4,708,871号および同5,595,915号を参照のこと。 20

#### 【0115】

そして、以下に記載されるように、少なくとも8連続アミノ酸をコードする核酸（すなわち、24ヌクレオチド以上のフラグメント）が、免疫原として有用性を有するペプチドの発現または合成を指向する際に有用である。例えば、Lener、「Tapping the immunological repertoire to produce antibodies of predetermined specificity」Nature 299：592-596（1982）；Shinnickら、Annu. Rev. Microbiol. 37：425-46（1983）；Sutcliffeら、Science 219：660-6（1983）を参照のこと。 30

#### 【0116】

従って、本発明の核酸フラグメントは、少なくとも17ヌクレオチド長であり、代表的には18ヌクレオチド長であり、そしてしばしば少なくとも24ヌクレオチド長、25ヌクレオチド長、30ヌクレオチド長、35ヌクレオチド長、40ヌクレオチド長または45ヌクレオチド（nt）長である。もちろん、当業者によって認識されるように、少なくとも50nt、100nt、150nt、200nt、250nt、300nt、350nt、400nt、450nt、500nt以上を有するより大きなフラグメントもまた有用であり、そして時には好ましい。

#### 【0117】

本発明が、発現されたメッセージの調査に基づくのではなく、むしろゲノム配列の模倣物に基づいているので、さらに、g p 3 5 4 遺伝子の一部を含む、単離されたゲノム由来のポリヌクレオチドまたは核酸を提供する。本発明は特に、ゲノム由来の単一エキソプローブを提供し、これらは少なくともエキソン（「参照エキソン」）の一部を含み、そして高ストリンジェンシー条件下で、参照エキソンを含む転写物由来の核酸に、検出可能にハイブリダイズし得る。しかし、この単一エキソプローブは、参照エキソンを欠くが、ゲノム中で参照エキソンに隣接して見出される1つ以上のエキソンを含む核酸には、高ストリンジェンシー条件下で検出可能にハイブリダイズしない。 40

#### 【0118】

本発明はまた、単離されたゲノム由来のポリヌクレオチドまたは核酸を提供し、これらは 50

g p 3 5 4 遺伝子の転写を制御する核酸配列エレメントを含む。転写制御配列は、例えば、プロモーター、エンハンサー、オペレーター、ターミネーター、サイレンサーなどを含む。

#### 【0119】

転写もしくは翻訳のアンチセンス阻害における使用について、またはリボザイムのような酵素性核酸分子のアンチセンス媒介性標的化について所望される場合、本発明の単離されたポリヌクレオチドまたは核酸は、1つ以上の修飾塩基（以下を参照のこと）および／または1つ以上の修飾されたかもしくは変更された内部ヌクレオシドの結合を有用に含み得、これらはしばしば、ヌクレアーゼ耐性を提供する。Hartmannら（編）、Manual of Antisense Methodology (Perspective in Antisense Science)、Kluwer Law International (1999) (ISBN: 079238539X); Steinら（編）、Applied Antisense Oligonucleotide Technology、Wiley-Liss（表紙（1998）(ISBN: 0471172790); Chadwickら（編）、Oligonucleotides as Therapeutic Agents—Symposium 第209、John Wiley & Sons Ltd (1997) (ISBN: 0471972797)を参照のこと。このような変更された塩基および内部ヌクレオシドの結合はしばしば、Gamberら、Nucleic Acids Res. 28 (21): 4332-9 (2000)（この開示は本明細書中に参考としてその全体が援用される）に記載されるように、本発明の単離された核酸が標的化遺伝子の修正のために使用されることである場合にまた、所望される。

#### 【0120】

本発明のアンチセンス核酸分子（およびアンチセンスによって標的化される酵素的核酸）は、治療設定において使用され得る。これらの分子は、作動可能に連結された転写調節配列を含む発現ベクターから発現され得、その活性は、ベクターが導入された細胞型によって決定され得る。アンチセンス遺伝子を使用する遺伝子発現の調節の考察については、Weintraubら、Antisense RNA as a molecular tool for genetic analysis、Trends. Genet. 1 (1) (1986)を参照のこと。

#### 【0121】

本発明のアンチセンス核酸は、リボザイムであり得る。リボザイムは、リボヌクレアーゼ活性を有する触媒的RNA分子であり、これは単鎖核酸（例えば、mRNA）を切断し得、この単鎖核酸に対してリボザイムは相補的な領域を有する。従って、リボザイムを使用して、g p 3 5 4 mRNA転写産物を触媒的に切断し得、これらによってg p 3 5 4 mRNAの翻訳を阻害し得る。g p 3 5 4をコードする核酸に対して特異性を有するリボザイムは、本明細書中に開示されるg p 3 5 4ポリヌクレオチドのヌクレオチド配列（配列番号1～3）に基づき設計され得る。

#### 【0122】

g p 3 5 4のオリゴヌクレオチド模倣物（例えば、ペプチド核酸（PNA））は、治療適用および診断適用において使用され得る。例えば、Hyrupら、Bioorg. Med. Chem. Lett. 4: 5-23 (1996)を参照のこと。PNA化合物において、核酸のホスホジエステル骨格は、特に、アミド結合で連結される反復性のN-（2-アミノエチル）グリシン単位によってアミド含有骨格に置き換えられる。例えば、PNAは、例として転写停止もしくは翻訳停止、または複製阻害を誘導することによって、遺伝子発現の配列特異的な調節のための、アンチセンス薬剤またはアンチジーン（antigene）薬剤として使用され得る。g p 3 5 4のPNAはまた、例えば、PNA指向性のPCRクランピングによって、例えば遺伝子中の1塩基対変異の分析において；他の酵素（例えば、S1ヌクレアーゼ）と組み合わせて使用される場合に人工制限酵素として；またはDNA配列およびハイブリダイゼーションに対するプローブもしくはプライマーとして、使用され得る（Hyrupら、前出；およびPerry-O'Keefe、前出）。g

p 3 5 4 の P N A を改変して、例えば、脂質もしくは他の P N A に対する補助基 (h e l p e r g r o u p) を結合することによってか、P N A - D N A キメラの形成によってか、または当該分野で公知の薬物送達のリポソームもしくは他の技術 (以下) の使用によって、それらの安定性または細胞取り込みを増強し得る。

#### 【0123】

本発明のオリゴヌクレオチドは、(例えば、インビボで宿主細胞レセプターを標的化するために) ペプチドまたは細胞膜もしくは血液脳関門を横切る輸送を促進する薬剤のような、他の付加基を含み得る。さらに、オリゴヌクレオチドは、ハイブリダイゼーション誘発性の切断薬剤または挿入薬剤を用いて改変され得る。この目的に向けて、オリゴヌクレオチドは別の分子 (例えば、ペプチド、ハイブリダイゼーション誘発性の架橋薬剤、輸送薬剤、ハイブリダイゼーション誘発性の切断薬剤など (以下を参照のこと)) と結合体化され得る。

10

#### 【0124】

天然に見出される核酸構成物との差異 (例えば、非天然塩基、内部ヌクレオシド連結の変更、合成後修飾) は、g p 3 5 4 ポリヌクレオチドの長さ全域に存在し得るか、またはその分散した部分に有用に配置され得る。後者の例として、分散した D N A ドメインおよび R N A ドメインを有するキメラ核酸が合成され得、そして米国特許第 5, 7 6 0, 0 1 2 号および同 5, 7 3 1, 1 8 1 号 (これらの開示は、本明細書中で参考としてその全体が援用される) にさらに記載されるように、標的化遺伝子修復のための利用を実証した。D N A および P N A の両方を含むキメラ核酸は、改変 P C R 反応において有用性を有することが実証されている。M i s r a ら、B i o c h e m. 3 7 : 1 9 1 7 - 1 9 2 5 (1 9 9 8) を参照のこと; また、F i n n ら、N u c l. A c i d s R e s. 2 4 : 3 3 5 7 - 3 3 6 3 (1 9 9 6) を参照のこと (これらは、本明細書中で参考として援用される)。

20

#### 【0125】

本発明のポリヌクレオチドおよび核酸はまた、基材に有用に結合され得る。この基材は、多孔性または固体、平板もしくは非平板、統一型もしくは分散型であり得; この結合は共有性であるかまたは非共有性であり得る。本発明の核酸は、基質に結合されると、非標識状態でプローブとして使用され得る。例えば、本発明の核酸は、多孔性基材 (通常は膜であり、代表的にはニトロセルロース、ナイロンまたは正に荷電した誘導体化ナイロン) に有用に結合され得; そのように結合されると、本発明の核酸を使用して、例えば、リバースドットプロットによって、標識された核酸サンプル (ゲノム核酸のサンプルまたは転写物由来の核酸のサンプルのいずれか) 中に存在する g p 3 5 4 核酸を検出し得る。

30

#### 【0126】

本発明の核酸はまた、固体支持体 (例えば、ガラス) に有用に結合され得るが、他の固体材料 (例えば、アモルファスシリコン、結晶シリコンまたはプラスチック) もまた使用され得る。本発明の核酸は、固体支持体表面に共有結合で結合され得るか、または推定の非共有結合の相互作用またはその何らかの組み合わせによって、変性および支持を容易にするカオトロピック剤において、誘導化された表面に適用され得る。

#### 【0127】

本発明の核酸は、複数の他の核酸と一緒に結合される基材に結合し得る (結合した複数の核酸の各々へのハイブリダイゼーションが、別々に検出可能である)。例えば低密度では、多孔性膜上のこれらの基材結合型コレクションは、代表的にマクロアレイと称され、高密度では、代表的に固体支持体 (例えば、ガラス) 上でこれらの基材に結合した複数の核酸のコレクションは、口語的にマイクロアレイと称される。本明細書中で使用される場合、用語マイクロアレイは全ての密度のアレイを含む。従って、本発明は本発明の核酸を含むマイクロアレイを提供する。

40

#### 【0128】

本発明の単離された核酸をハイブリダイゼーションのプローブとして使用して、ゲノムサンプルおよび転写物由来の核酸サンプルの両方において、g p 3 5 4 核酸を検出、特徴付

50

け、および定量し得、そしてそれらから g p 3 5 4 核酸を単離し得る。このようなプローブは、溶液中で遊離する場合、代表的に（しかし不可変ではなく）、検出可能な標識である；マイクロアレイとして基材に結合される場合、このようなプローブは代表的に（しかし不可変ではなく）、未標識である。

#### 【0129】

例えば、本発明の単離された核酸をプローブとして使用して、クロモソームの広がりに対する蛍光インサイチュハイブリダイゼーション（FISH）を介して、g p 3 5 4 ゲノム遺伝子座中の全体的な変化（例えば、g p 3 5 4 ゲノム遺伝子座の欠失、挿入、転座、および重複）を検出および特徴付け得る。例えば、Andreeffら（編）、Introduction to Fluorescence In Situ Hybridization: Principles and Clinical Applications、John Wiley & Sons（1999）（ISBN：0471013455（この開示は本明細書中でその全体が参考として援用される）を参照のこと。本発明の単離された核酸をプローブとして使用して、例えば制限フラグメント長多型のサザンブロット検出を使用して、より小さいゲノム変化を評価し得る。本発明の単離された核酸をプローブとして使用して、本発明の核酸を含むゲノムクローンを単離し得る。これらはその後、配列レベルでの欠失、挿入、転座および置換（1ヌクレオチド多型（SNP））を同定するために、制限マップを作成され得、そして配列決定され得る。

#### 【0130】

本発明の単離された核酸をまたプローブとして使用して、転写物由来の核酸サンプル中の g p 3 5 4 核酸の検出、特徴付け、および定量し得、そしてこれらから g p 3 5 4 核酸を単離し得る。例えば、本発明の単離された核酸をハイブリダイゼーションプローブとして使用して、総RNAサンプルまたはポリA<sup>+</sup>選択されたRNAサンプルのノーザンブロットによって、g p 3 5 4 mRNAを検出、長さによる特徴付け、および定量し得る。本発明の単離された核酸をまたハイブリダイゼーションプローブとして使用して、組織切片へのインサイチュハイブリダイゼーションによって、g p 3 5 4 のメッセージを検出、局在による特徴付け、および定量し得る（例えば、Schwarzhacherら、In Situ Hybridization、Springer-Verlag New York（2000）（ISBN：0387915966）（この開示は参考としてその全体が援用される）を参照のこと）。

#### 【0131】

さらに、本発明の核酸をハイブリダイゼーションプローブとして使用して、cDNAライブラリー中の g p 3 5 4 クローンの発現量を測定し得る。例えば、本発明の単離された核酸をハイブリダイゼーションプローブとして使用して、cDNAライブラリーから g p 3 5 4 核酸を単離し得、g p 3 5 4 RNAメッセージの配列レベルの特性（欠失、挿入、削減（選択的スプライシングされた形態でのエキソンの欠失、挿入および削減を含む）ならびに1ヌクレオチド多型の同定を含む）を認める。

#### 【0132】

本明細書中の以下の実施例に記載されるように、本発明の単離された核酸をまた使用して、転写物由来のサンプル中の g p 3 5 4 核酸を検出および定量し、g p 3 5 4 遺伝子の発現量を測定し得る。g p 3 5 4 発現の測定は、本明細書中の以下の実施例でさらに詳細に記載されるように、g p 3 5 4 が通常発現される膵臓組織および神経組織において、そして通常発現される様式で、ならびにこれが誤発現され得る組織においてのいずれかにおいて、異常な g p 3 5 4 発現と関連した状態、障害、および疾患のための診断アッセイにおいて、特に有用性を有する。

#### 【0133】

従って、当業者に容易に明らかとなるように、各 g p 3 5 4 核酸プローブ（標識されるか、基材に結合されるか、またはその両方）は、通常、膵臓組織および神経組織（発現が既に確定されている）における g p 3 5 4 発現のレベルを測定するためのツールとしての使用のために、利用可能である。

10

20

30

40

50

## 【0134】

g p 3 5 4を発現することが未だ実証されていない組織に関して、本発明のg p 3 5 4核酸は通常、g p 3 5 4核酸の存在を検出するため、例えば、臍臓もしくは神経系における異常なg p 3 5 4の細胞内発現、または他の組織における異常な組織分布と関係する状態、障害または疾患を示す患者の組織におけるg p 3 5 4 R N A発現を検出するために、そのような組織を調査するためツールとして、利用可能である。

## 【0135】

上記のように、本発明の核酸プローブはマイクロアレイを構築する際に有用であり；次いで、このマイクロアレイは、例えば、薬物発見および標的バリデーションプログラムにおいて遺伝子発現を測定および調査するのに有用である、工業製品である。各々のg p 3 5 4核酸プローブは、マイクロアレイ上に含まれる場合、プローブ中に含まれるg p 3 5 4遺伝子のその一部を検出するのに特異的に有用なマイクロアレイを構成し、このようにして、このマイクロアレイデバイス上で、シグナルを検出する能力を付与する（このようなプローブの非存在ではシグナルがないことが報告されている）。 10

## 【0136】

g p 3 5 4の発現レベルの変化は、有用性を有する発現量を測定のために観察される必要はない。遺伝子発現分析を使用して細胞に対する化学薬剤の毒性を評価する場合、例えば、この薬剤が遺伝子の発現レベルを変化し得ないことは、この薬物が、その遺伝子の発現タンパク質が一部である経路に、影響しそうではないことの証拠である。同様に、遺伝子発現分析を使用して薬理学的薬剤の副作用を評価する場合（リード化合物の発見であれ、 20 それに続くリード化合物誘導体のスクリーニングであれ）、その薬剤が、遺伝子発現レベルを変化し得ないことは、その薬物がこの遺伝子発現タンパク質が一部である経路に影響しない証拠である。W O 9 9 / 5 8 7 2 0（本明細書中でその全体が参考として援用される）は、遺伝子の個性または機能（この発現が計算において使用される）に関係なく、第1遺伝子発現プロファイルおよび第2遺伝子発現プロファイルの関連性を定量するための方法、ならびに複数の遺伝子発現プロファイルの関連性を整理するための方法を提供する。

## 【0137】

本発明の、ゲノム由来の単一のエキソンプローブおよびゲノム由来の単一のエキソンプローブマイクロアレイは、本発明の核酸のスプライシング改変体のハイスループット検出を 30 可能とする、さらなる有用性を有する。

## 【0138】

プロモーターを伴うポリヌクレオチド挿入物に隣接するベクターのような核酸構築物中に挿入された本発明のポリヌクレオチドは、本発明の核酸のいずれかの鎖に相補的なR N Aのインビトロ発現を駆動するために用いられ得る。このR N Aは、c D N A－m R N Aサブトラクションにおいてか、またはインビトロ翻訳のために一本鎖プローブとして用いられ得る。G P 3 5 4タンパク質またはその一部をコードするようなこれらのポリヌクレオチドは、さらに、単独であるか、または融合タンパク質の一部分のいずれかとして、G P 3 5 4タンパク質またはタンパク質フラグメントを発現するために用いられ得る。発現は、ゲノムから、または本発明の転写物由来のポリヌクレオチドからであり得る。 40

## 【0139】

タンパク質発現がゲノムD N Aから行われる場合、代表的には、発現は、最初のR N A転写物からイントロンをスプライシングし得る、真核生物細胞、代表的には哺乳動物細胞で行われる。発現は、エピソーム性ベクターから駆動され得るか、または宿主細胞染色体中に組み込まれたゲノムD N Aから駆動され得る。以下に記載されるように、発現が、本発明の転写物由来（またはそうでなければイントロンのない）ポリヌクレオチドからである場合、発現は、広範な種類の原核生物細胞または真核生物細胞中で行われ得る。

## 【0140】

インビトロで発現されて、タンパク質、タンパク質フラグメント、またはタンパク質融合体は、その後単離されて、本発明のタンパク質またはタンパク質イソ型に特異的なイムノ 50

アッセイにおける標準として用いられ得；治療薬剤として用いられ得、例えば、本発明のタンパク質が欠損している個体における受動置換療法として投与され得；ワクチンとして投与され得；特異的抗体のインビトロ産生のために用いられ得、その後、この抗体は、例えば、本発明のタンパク質の検出および定量的ための分析薬剤として、または免疫療法薬剤として用いられ得る。

#### 【0141】

本発明の単離されたポリヌクレオチドおよび核酸はまた、本発明のタンパク質のインビボ発現を駆動するために用いられ得る。インビボ発現は、遺伝子療法を目的にするベクター—代表的にはウイルスベクター、しばしば、複製不能レンチウイルス、レトロウイルス、アデノウイルス、またはアデノ随伴ウイルス（AAV）を基礎にしたベクター—から駆動され得る。インビボ発現は、この核酸に対して内因性または（例えばベクターからの）外因性の発現制御シグナルから駆動され得る。本発明の他のウイルスベクターとしては、例えば、バキュロウイルス、アデノウイルス、パルボウイルス、ヘルペスウイルス、ポックスウイルス、アデノ随伴ウイルス、セムリキ森林ウイルス、ワクシニアウイルス、およびレトロウイルス由来のベクターが挙げられる。

10

#### 【0142】

本発明の単離されたgp354ポリヌクレオチドの種々の形態（例えば、ゲノムまたはcDNA）を、雄もしくは雌の前核中にマイクロインジェクションし得るか、または胚性幹（ES）細胞中に組み込んで、本発明のタンパク質を産生し得るトランスジェニック非ヒト動物を作製し得る。

20

#### 【0143】

本発明のゲノム核酸はまた、被験体中のgp354遺伝子座への相同的組み換えを標的化するために用いられ得る。例えば、米国特許第6,187,305号；同第6,204,061号；同第5,631,153号；同第5,627,059号；同第5,487,992号；同第5,464,764号；同第5,614,396号；同第5,527,695号および同第6,063,630号；ならびにKmiecら（編）、Gene Targeting Protocols, 133巻、Humana Press（2000）（ISBN:0896033600）；Joyner（編）、Gene Targeting: A Practical Approach, Oxford University Press, Inc.（2000）（ISBN:0199637938）；Sedivyら、Gene Targeting, Oxford University Press（1998）（ISBN:071677013X）；Tymmsら（編）、Gene Knockout Protocols, Humana Press（2000）（ISBN:0896035727）；Makら（編）、The Gene Knockout Facts Book, 第2巻、Academic Press, Inc.（1998）（ISBN:0124660444）；Torresら、Laboratory Protocols for Conditional Gene Targeting, Oxford University Press（1977）（ISBN:019963677X）；Vega（編）、Gene Targeting, CRC Press, LLC（1994）（ISBN:084938950X）を参照のこと。これらの開示はその全体が参考として本明細書に援用される。

30

40

#### 【0144】

ゲノム領域が、転写調節エレメントを含む場合、相同組換えを用いて、ヒト細胞からのGP354タンパク質のインビトロ産生の目的、および遺伝子療法の目的のための両方に、GP354の発現を改変し得る。例えば、それらの開示の全体が参考として本明細書に援用される、米国特許第5,981,214号、同第6,048,524号；同第5,272,071号を参照のこと。相同組換えに代表的に用いられるようなものより小さな本発明のポリヌクレオチドのフラグメントもまた、おそらく、相同組換えの間に関与する機構とは異なる細胞の機構によって、標的化された遺伝子矯正または改変のために用いられ得る。例えば、米国特許第5,945,339号、同第5,888,983号、同第5,8

50

71, 984号、同第5, 795, 972号、同第5, 780, 296号、同第5, 760, 012号、同第5, 756, 325号、同第5, 731, 181号；およびCulverら、「Correction of chromosomal point mutations in human cells with bifunctional oligonucleotides」Nature Biotechnol. 17(10):989-93(1999)；Gamberら、Nucl. Acids Res. 28(21):4332-9(2000)を参照のこと。これらの開示は、本明細書に参考として援用される。

#### 【0145】

本発明のポリヌクレオチドは、標準的な技法により、ゲノムライブラリー、cDNAライブラリー、およびmRNAサンプルのような核酸サンプルをプローブするために、本発明の標識プローブを用いることによって得られ得る。本発明のポリヌクレオチドはまた、本明細書中以下の実施例1にさらに示されるように、本発明の核酸プライマーを用いる増幅によっても得られ得る。本発明のポリヌクレオチドはまた、特に、約100ヌクレオチドより少ない場合、代表的には、市販の自動化合成機を用いる固相合成により、化学的に合成され得る。

#### 【0146】

(ベクターおよび宿主細胞)

(A. 核酸構築物)

本発明は、1つ以上の単離された本発明のポリヌクレオチドを含む、ベクターのような核酸構築物、およびこのようなベクターが導入された宿主細胞を提供する。

#### 【0147】

このようなベクターは、宿主細胞において本発明のポリヌクレオチドを増やすため（クローニングベクター）、異なる生物由来の宿主細胞間で本発明のポリヌクレオチドをシャトルするため（シャトルベクター）、宿主細胞染色体中に本発明のポリヌクレオチドを挿入するため（挿入ベクター）、インビトロまたは宿主細胞内で本発明のポリヌクレオチドのセンスRNA転写物またはアンチセンスRNA転写物を発現するため、および単独または異種ポリペプチドへの融合として、本発明のポリヌクレオチドによりコードされたポリヌクレオチドを発現するため（発現ベクター）に用いられ得る。本発明のベクターは、しばしば、このような使用のいくつかに好適である。

#### 【0148】

ベクター類は、今や、当該分野で周知であり、そして、特に、Jonesら（編）、Vectors: Cloning Applications: Essential Techniques (Essential Techniques Series)、John Wiley & Son Ltd 1998 (ISBN: 047196266X)；Jonesら（編）、Vectors: Expression Systems: Essential Techniques (Essential Techniques Series)、John Wiley & Son Ltd 1998 (ISBN: 0471962678)；Gacesaら、Vectors: Essential Data、John Wiley & Sons, 1995 (ISBN: 0471948411)；Cid-Arregui（編）、Viral Vectors: Basic Science and Gene Therapy、Eaton Publishing Co.、2000 (ISBN: 188129935X)；Sambrookら、Molecular Cloning: A Laboratory Manual (第3版)、Cold Spring Harbor Laboratory Press、2001 (ISBN: 0879695773)；Ausubelら（編）、Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology (第4版)、John Wiley & Sons、1999 (ISBN: 047132938X)に記載されている。これらの開示はそれらの全体が本明細書中に参考

10

20

30

40

50

として援用される。広範な種類のベクターが市販されている。現存するベクターの使用および改変は当該分野の技術内に十分ある。

#### 【0149】

代表的には、ベクターは、ウイルス、プラスミド、原核生物または真核生物の染色体エレメント、またはそれらのいくつかの組み合わせ由来であり、そして、いくつかの組み込みベクターは、染色体が改変される宿主中で機能的である起点を欠き、そしていくつかのベクターは、選択マーカを欠いているが、少なくとも1つの複製起点、少なくとも1つの異種核酸の挿入のための部位（代表的には、複数の緊密にクラスターとなった単一の切断制限部位を備えたポリリンカーの形態である）、および少なくとも1つの選択マーカを含む。本発明のベクターは、さらに、少なくとも1つの位置でベクター中に挿入された少なくとも1つの単離された本発明のポリヌクレオチド核酸を含む。存在する場合、複製起点および選択マーカは、所望の宿主細胞を基にして選択され；次に、宿主細胞は、所望の適用を基にして選択される。

10

#### 【0150】

例えば、原核生物細胞、代表的には *E. coli* は、代表的にはクローニング、すなわち、宿主細胞中のポリヌクレオチド配列の増幅のために選択される。このような場合、ベクター複製は、（ $\lambda$ ファージ、M13、T7、T3およびP1のような）大腸菌に感染するファージの複製戦略、または自律複製エピソーム、著名には、pBR322プラスミドおよびpUCシリーズプラスミドを含む、ColE1プラスミドおよびその後の誘導体の複製起点を基にすることを意味する。*E. coli* が宿主として用いられる場合、選択マーカは、同様に、グラム陰性細菌中の選択性について選択される：例えば、代表的なマーカは、アンピシリン、テトラサイクリン、クロラムフェニコール、カナマイシン、ストレプトマイシン、ゼオシンのような抗生物質に対する耐性を付与する；栄養要求性マーカもまた用いられ得る。

20

#### 【0151】

別の例として、代表的には *S. cerevisiae* である酵母細胞が、真核生物遺伝子研究のため、（例えば、ツーハイブリッドシステムの使用を介する）相互作用タンパク質成分の同定のため、およびタンパク質発現のために特に選択される。酵母における使用のための本発明のベクターは、代表的には、いつもではないが、酵母における使用のために適切な複製起点、および酵母において機能的である選択マーカを含む。

30

#### 【0152】

適切な酵母ベクターの例は、組み込みYIpベクター、セントロメア配列CENを含む複製エピソームのYE pベクター、および自律複製配列ARSを含む。YACは、インビボでテロメア（TEL）として機能する同種または異種のDNA配列を含み、かつ酵母ARS（複製起点）セグメントおよびCEN（セントロメア）セグメントを含む、YLpと称される酵母線状プラスミドに基づく。

#### 【0153】

酵母ベクターにおける選択マーカとしては、特異的な自己栄養性変異（例えば、*ura3-52*、*his3-D1*、*leu2-D1*、*trp1-D1*および*lys2-201*）を補完する種々の自己栄養性マーカ（これらの（*Saccharomyces cerevisiae*における）最も一般的なものは、URA3、HIS3、LEU2、TRP1およびLYS2である）が挙げられる。このURA3酵母遺伝子およびLYS2酵母遺伝子は、原栄養体系統の増殖を阻害するが、それぞれ、*ura3*変異体および*lys2*変異体の増殖を許容する特異的なインヒビターである、5-フルオロオロチン酸（FOA）および $\alpha$ -アミノアジピン酸（ $\alpha$ AA）に、それぞれ基づくネガティブ選択をさらに可能にする。他の選択マーカは、例えば、ゼオシン（*zeocin*）に対する耐性を与える。

40

#### 【0154】

昆虫細胞は、高い効率のタンパク質発現について、しばしば選択される。宿主細胞が、*Spodoptera frugiperda*（例えば、Sf9細胞株およびSf21細胞

50



株、ならびに  $\text{expressSF}^{\text{TM}}$  細胞 (Protein Sciences Corp., Meriden, CT, USA) に由来する場合、ベクター複製戦略は、代表的には、バキュロウイルス生活環に基づく。代表的には、バキュロウイルス移入ベクターを使用して、野生型  $\text{AcMNPV}$  ポリヘドリン遺伝子を目的の異種遺伝子で置き換える。野生型ゲノムにおいてこのポリヘドリン遺伝子に隣接する配列は、この移入ベクター上の発現カセットの 5' および 3' に位置される。 $\text{AcMNPV}$  DNA を用いる同時トランスフェクション後、相同組換え現象がこれらの配列の間で起こり、この目的の遺伝子、およびこのポリヘドリンプロモーターまたは  $\text{p10}$  プロモーターを保有する組換えウイルスを生じる。選択は、 $\text{lacZ}$  融合活性についてのビジュアルスクリーニングに基づき得る。

10

#### 【0155】

哺乳動物細胞は、薬学的因子として意図されるタンパク質の発現についてしばしば選択され、また、タンパク質もしくは生理学的経路の潜在的なアゴニストおよびアンタゴニストのスクリーニングのための宿主細胞としても選択される。哺乳動物細胞における自律的染色体外複製について意図されるベクターとしては、代表的には、ウイルス起源（例えば、（ラージ T 抗原を発現する細胞株（例えば、COS1 細胞および COS7 細胞）における複製のための）SV40 起源、パピローマウイルス起源、または長期間のエピソーム複製のための EBV 起源（例えば、EBV EBNA-1 遺伝子産物およびアデノウイルス E1A を構成的に発現する 293-EBNA 細胞における使用のための））が挙げられる。哺乳動物染色体の一部として組み込み、従って複製を意図されるベクターは、哺乳動物細胞内で機能的な（例えば、SV40 起源の）複製起点を含み得るが、必ずしもこれらを含むに及ばない。ウイルス（例えば、レンチウイルス、アデノウイルス、アデノ随伴ウイルス、ワクシニアウイルス、および様々な哺乳動物レトロウイルス）に基づくベクターは、代表的には、ウイルス複製戦略に従って複製する。

20

#### 【0156】

哺乳動物細胞での用途のための選択マーカーは、ネオマイシン (G418)、プラスチジン、ハイグロマイシンおよびゼオシンに対する耐性、ならびに HAT 培地を使用するプリンサルベージ経路に基づく選択を含む。

#### 【0157】

植物細胞はまた、代表的には、植物ウイルス（例えば、カリフラワーモザイクウイルス (CaMV)；タバコモザイクウイルス (TMV)）由来のベクターレプリコン、および植物における適切性について選択される選択マーカーを用いる、発現のために使用され得る。

30

#### 【0158】

プラスミドまたはウイルス由来のベクターにおいて容易に適応され得るポリヌクレオチドより大きな、本発明のポリヌクレオチドの増殖のために、本発明は、 $\text{gp354}$  核酸、しばしばゲノム核酸を含む、人工染色体 (BAC、YAC、および HAC) をさらに提供する。

#### 【0159】

プラスミドまたはウイルス由来のベクターにおいて容易に適応され得るポリヌクレオチドより大きな、本発明のポリヌクレオチドの増殖のために、本発明は、 $\text{gp354}$  核酸（しばしば、ゲノム核酸）を含む、人工染色体 (BAC、YAC、および HAC) をさらに提供する。例えば、Shizuya ら, Keio J. Med. 50 (1): 26-30 (2001)；Shizuya ら, Proc. Natl. Acad. Sci. USA 89 (18): 8794-7 (1992)；Kuroiwa ら, Nature Biotechnol. 18 (10): 1086-90 (2000)；Hennings ら, Proc. Natl. Acad. Sci. USA 96 (2): 592-7 (1999)；Harrington ら, Nature Genet. 15 (4): 345-55 (1997)（これらの開示は、本明細書中で参考として援用される）を参照のこと。

40

#### 【0160】

50

本発明のベクターはまた、しばしば、挿入された異種核酸からのRNAのインビトロ転写を可能にするエレメントを含む。このようなベクターは、代表的に、ファージプロモーター（例えば、核酸挿入物に隣接するT7、T3またはSP6由来のもの）を含む。しばしば、このようなプロモーターの異なる2つは、挿入された核酸に隣接し、センス鎖およびアンチセンス鎖の両方の別個のインビトロ産生を分離し得る。

#### 【0161】

挿入された異種核酸からポリペプチドの発現を駆動する本発明の発現ベクターは、しばしば、タンパク質コード異種核酸挿入物に作動可能に連結される他の種々の遺伝的エレメント（代表的には、転写を駆動し、そして調節する遺伝的エレメント（例えば、プロモーターエレメントおよびエンハンサーエレメント）、RNAプロセッシングを容易にする遺伝的エレメント（例えば、転写終止、スプライシングシグナルおよび／またはポリアデニル化シグナル）、ならびに翻訳を容易にする遺伝的エレメント（例えば、リボソームコンセンサス配列）を含む。他の転写制御配列としては、例えば、オペレーター、サイレンサーなどが挙げられる。このような発現制御エレメント（誘導性発現、および発生発現または組織調節発現を与える発現制御エレメントを含む）の使用は、当該分野で周知である。

10

#### 【0162】

脾臓、神経系または乳腺においてGP354を発現することが可能な組織特異的調節エレメントは、特に有用であり得、そして当該分野で公知である（例えば、神経特異的神経フィラメントプロモーター（ByrneおよびRuddle（1989）Proc. Natl. Acad. Sci. USA 86:5473-5477）、脾臓特異的プロモーター（Edlundら（1985）Science 230:912-916）、および乳腺特異的プロモーター（例えば、乳清プロモーター；米国特許第4,873,316号および欧州特許出願公開番号264,166））。発生的に調節されるプロモーターもまた選択され得、これらとしては、以下が挙げられるが、これらに限定されない：マウスhoxプロモーター（KesselおよびGruss（1990）Science 249:374-379）およびα-フェトプロテインプロモーター（CampesおよびTilghman（1989）Genes Dev. 3:537-546）。膨大な種々の誘導性プロモーターが、公知であり、そして具体的な適用に基づいて選択され得る。

20

#### 【0163】

発現ベクターは、精製および／または視覚化を容易にする小タンパク質タグに、発現されたポリペプチドを融合するように設計され得る。多くのこのようなタグが公知であり、そして利用可能である。発現ベクターはまた、精製タグおよび／または同定タグより大きなポリペプチドに、異種核酸挿入物によってコードされるタンパク質を融合するように設計され得る。有用なタンパク質融合物としては、コードされるタンパク質の提示をファージまたは細胞の表面上に可能にするタンパク質融合物、内因性蛍光タンパク質（例えば、ルシフェラーゼまたは緑色蛍光タンパク質（GFP）様発色団を有するタンパク質）への融合物、IgG Fc領域または他の免疫グロブリン型の定常ドメインへの融合物、およびツウハイブリット選択系における使用のための融合物が挙げられる。

30

#### 【0164】

発現されるタンパク質の分泌のために、広範な種々のベクターが利用可能であり、これらは、リーダーペプチドのような分泌シグナルをコードする適切な配列を含む。ファージディスプレイ、酵母ディスプレイ、および哺乳動物ディスプレイのために設計されるベクターは、例えば、N末端細胞表面標的化シグナルおよびC末端膜貫通アンカー化ドメインを使用して、組換えタンパク質を標的化する。

40

#### 【0165】

異種核酸にコードされるタンパク質を、基質非依存性の、元来蛍光発光性のAequorea victoria由来の緑蛍光タンパク質（「GFP」）ならびにその多くのカラーシフト改変体および／または安定化改変体の発色団に融合する、広範な種々のベクターが現在存在する。

#### 【0166】

50

FcRnレセプター（これはまた、FcRプレセプターおよびBrambellレセプター、FcRbと呼称される）との相互作用を通してタンパク質薬学的製品の血清半減期を増大するために、IgG Fc領域に対する異種配列の融合を可能にするベクターはまた、広範に利用可能である。

#### 【0167】

本発明のタンパク質、タンパク質融合物およびタンパク質フラグメントの長期間の高収量組換え産生のために、安定発現が好ましい。安定発現は、（好ましくは、選択可能マーカーを有する）ベクターの、宿主細胞ゲノムへの組み込み、その後の組み込み体の選択によって、容易に達成される。

#### 【0168】

##### （B. 宿主細胞）

本発明は、細胞内でエピソーム的存在するか、または宿主細胞染色体に全体的もしくは部分的に取り込まれるかのいずれかにある、本発明のベクターのような核酸構築物を含む、宿主細胞（原核生物（細菌）細胞または真核生物細胞（例えば、酵母細胞、昆虫細胞、植物細胞および動物細胞）のいずれか）をさらに含む。

#### 【0169】

他の考慮（その考慮のいくつかは、上記である）の中で、宿主細胞株は、発現タンパク質を所望の様式でプロセスする能力について選択され得る。ポリペプチドのこのような翻訳後修飾としては、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、およびアセチル化が挙げられるが、これらに限定されず、そしてこれは、このような翻訳後修飾を有するGP354タンパク質を提供するための本発明の局面である。

#### 【0170】

適切な宿主細胞の代表的な、非限定的な例としては、細菌細胞（例えば、*E. coli*、*Caulobacter crescentus*、*Streptomyces species*、および*Salmonella typhimurium*）；酵母細胞（例えば、*Saccharomyces cerevisiae*、*Schizosaccharomyces pombe*、*Pichia pastoris*、*Pichia methanolic*）；昆虫細胞株（例えば、*Spodoptera frugiperda*由来の細胞（例えば、Sf9細胞株およびSf21細胞株、ならびにexpressSF<sup>TM</sup>細胞（Protein Sciences Corp.、Meriden、CT、USA））、*Drosophila* S2細胞、および*Trichoplusia ni* High Five（登録商標）細胞（Invitrogen、Carlsbad、CA、USA））；ならびに哺乳動物細胞が挙げられる。代表的な哺乳動物細胞としては、COS1細胞およびCOS7細胞、チャイニーズハムスター卵巣（CHO）細胞、NIH 3T3細胞、293細胞、HEPG2細胞、HeLa細胞、L細胞、HeLa、MDCK、HEK293、WI38、マウスES細胞株（例えば、129/SV株、C57/BL6株、DBA-1株、129/SVJ株由来）、K562、Jurkat細胞、ならびにBW5147が挙げられる。他の有用な哺乳動物細胞株が、周知であり、そしてAmerican Type Culture Collection (ATCC) (Manassas, VA, USA) およびNational Institute of General medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Cell Repositories (Camden, NJ, USA) から容易に入手可能である。

#### 【0171】

本発明のベクターおよび核酸を、宿主細胞に導入する方法は、当該分野で周知であり；技術の選択は、導入される特異的なベクターおよび選択される宿主細胞に主に依存する。

#### 【0172】

##### （GP354タンパク質、GP354ポリペプチドおよびGP354フラグメント）

本発明は、抗原としての使用（例えば、エピトープマッピング）、免疫原としての使用（例えば、抗体を惹起するため、またはワクチンとして）、および治療組成物における使用

10

20

30

40

50

のために適切な G P 3 5 4 タンパク質およびこれらの種々のフラグメントを提供する。異種ポリペプチドに対する G P 3 5 4 ポリペプチドおよびフラグメントの融合物、ならびに本発明のこのタンパク質、フラグメントおよび融合物の他の部分に対する（例えば、キャリアタンパク質に対する、発色団に対する）結合体がまた、提供される。

#### 【0173】

いくつかの実施形態において、本発明は、全長 g p 3 5 4 c D N A（配列番号 1、配列番号 7 または配列番号 11）、または変性改変体にコードされるアミノ酸配列を含む、単離された G P 3 5 4 ポリペプチドを提供する。本発明はまた、必要に応じて 1 つ以上の保存的アミノ酸置換を有する、全長 g p 3 5 4 c D N A（配列番号 1、配列番号 7 または配列番号 11）にコードされるアミノ酸を有する、単離された G P 3 5 4 ポリペプチドを提供する。

10

#### 【0174】

本発明はまた、g p 3 5 4 c D N A（配列番号 1、配列番号 7 または配列番号 11）のヌクレオチド配列の一部分または全部を有するプローブに高度にストリンジェントな条件下でハイブリダイズするポリヌクレオチド配列によってコードされるアミノ酸配列を含む、単離された G P 3 5 4 ポリペプチドを提供する。好ましくは、ストリンジェントにまたは中程度にストリンジェントにクロスハイブリダイズする本発明のポリヌクレオチドによってコードされる単離された G P 3 5 4 ポリペプチドは、少なくとも 1 つの G P 3 5 4 の生物学的活性を有する。

#### 【0175】

別の一連の実施形態において、本発明は、必要に応じて 1 つ以上の保存的アミノ酸置換を有する、配列番号 2、配列番号 8 または配列番号 12 の G P 3 5 4 アミノ酸配列を含む、単離された G P 3 5 4 ポリペプチドを提供する。1 つ以上の保存的アミノ酸置換を必要に応じて有する、配列番号 2、配列番号 8 または配列番号 12 の G P 3 5 4 ポリペプチド配列によってコードされるアミノ酸配列を有する、単離された G P 3 5 4 ポリペプチドがまた、提供される。本発明は、上記の単離されたポリペプチドの各々のフラグメントをさらに提供し、特に、配列番号 2、配列番号 8 または配列番号 12 に与えられる配列の、少なくとも 6 アミノ酸、8 アミノ酸、15 アミノ酸から配列全体までを有するフラグメントを提供する。

20

#### 【0176】

上記の単離されたポリペプチドの各々は、N末端の 18 アミノ酸または 21 アミノ酸のシグナル配列を含み、このシグナル配列は、代表的には、膜を通過するタンパク質の挿入の際に除去される。従って、本発明は、N末端のシグナル配列が取り除かれた上記の単離された G P 3 5 4 ポリペプチドを提供する。切断は、配列番号 2 の 18～19 位または配列番号 8 の 21～22 位に、G 残基と P 残基との間に生じると推定される。

30

#### 【0177】

従って、本発明は、G P 3 5 4 の推定成熟 N 末端細胞外ドメインの全てまたは一部分を含む単離された G P 3 5 4 ポリペプチドを提供する（G P 3 5 4 ドメインおよび G P 3 5 4 配列について、図 1 および図 5 ならびに配列番号 2 および配列番号 8 を参照のこと）。G P 3 5 4 の推定成熟 N 末端細胞外ドメイン（すなわち、分泌シグナル配列を欠いている）は、配列番号 2 のアミノ酸 19-507、または配列番号 8 のアミノ酸 22～510 からなる。特定された配列の、少なくとも、6 アミノ酸、8 アミノ酸、15 アミノ酸から配列全体までを有する上記の配列のフラグメントもまた、含まれる。

40

#### 【0178】

本発明はまた、G P 3 5 4 の N 末端の細胞外ドメインの全てまたは一部分を含むかまたは有する単離された G P 3 5 4 ポリペプチドを提供する（G P 3 5 4 ドメインおよび G P 3 5 4 配列について、図 1 および図 5 ならびに配列番号 2 および配列番号 8 を参照のこと）。G P 3 5 4 の N 末端の細胞外ドメインは、配列番号 2 のアミノ酸 1～507、または配列番号 8 のアミノ酸 1～510 からなる。特定された配列の、少なくとも、6 アミノ酸、8 アミノ酸、15 アミノ酸から配列全体までを有する上記の配列のフラグメントもまた、

50

含まれる。

【0179】

好ましい実施形態において、単離されたG P 3 5 4ポリペプチドは、G P 3 5 4の細胞外ドメインの全体を有するか、またはそれを含み、そして機能的G P 3 5 4膜貫通ドメインを欠く。この膜貫通ドメインが、除外されるか、欠失されるか、または変異されて、非機能的となり得る。G P 3 5 4の膜貫通ドメインは、配列番号2のアミノ酸508～アミノ酸530、または配列番号8のアミノ酸511～533からなる。

【0180】

他の好ましい実施形態において、単離されたG P 3 5 4ポリペプチドは、異種タンパク質ドメインに融合されるG P 3 5 4 N末端細胞外ドメインの一部または全体からなる。好ましくは、単離されたG P 3 5 4ポリペプチドは、少なくとも1つの細胞外I gドメイン、より好ましくは、2つのG P 3 5 4細胞外I gドメインを含み、そして最も好ましくは、3つ、4つ、または5つのG P 3 5 4細胞外I gドメインを含む。

10

【0181】

G P 3 5 4の膜貫通ドメインおよびG P 3 5 4のC末端細胞質領域からなる群より選択されるG P 3 5 4フラグメントを含む、単離されたG P 3 5 4ポリペプチドもまた好ましい。他の好ましい実施形態において、単離されたG P 3 5 4ポリペプチドは、異種タンパク質ドメインに融合される、G P 3 5 4細胞質ドメインまたは膜貫通ドメインの一部分または全体からなる。

【0182】

本発明のG P 3 5 4フラグメントは、ネイティブG P 3 5 4タンパク質の連続した部分であり得る。しかし、本明細書中で提供されるようなG P 3 5 4遺伝子およびタンパク質配列の知見が、ネイティブG P 3 5 4タンパク質において連続的でない種々のドメインの再結合を可能にすることが理解される。

20

【0183】

本発明はまた、G P 3 5 4および関連するタンパク質の選択部分を含むポリペプチドを提供する。本明細書中の以下でさらに議論されるように、特に異種タンパク質フラグメントに結合される場合、これらのタンパク質フラグメントは、例えば、タンパク質-タンパク質相互作用を介して特定の細胞型に因子を標的化するため；I gドメイン含有タンパク質間のタンパク質-タンパク質相互作用を阻害するため；競合的結合アッセイのため；およびフラグメント特異的G P 3 5 4抗体を惹起するために使用され得る。

30

【0184】

このような実施形態の第1連において、タンパク質フラグメントは、少なくとも1つのコピーにおいて、1つ、2つ、3つ、4つ、または5つのG P 3 5 4のN末端細胞外部分に特徴的なI gドメインを含む。具体的には、5つの細胞外I gドメインが、それぞれ、配列番号2のG P 3 5 4アミノ酸配列のアミノ酸35～102、136～203、239～290、323～374、および410～485（図1参照のこと）にコードされ、そして、それぞれ、配列番号8のG P 3 5 4アミノ酸配列のアミノ酸38～109、139～206、242～293、326～377および413～488（図5参照のこと）にコードされる。好ましい実施形態において、このタンパク質フラグメントは、少なくとも1つのコピーの中で、少なくとも2つ、好ましくは3つ、より好ましくは4つ、そして最も好ましくは5つ全てのドメインをコードする。

40

【0185】

好ましくは、このタンパク質フラグメントは、膜を介するこのポリペプチド輸送を媒介するN末端シグナル分泌配列を含む。このG P 3 5 4シグナル分泌配列は、配列番号2のG P 3 5 4アミノ酸配列のアミノ酸1～18（図1を参照のこと）および配列番号8のG P 3 5 4アミノ酸配列のアミノ酸1～21（図5を参照のこと）によってコードされる。より好ましくは、このタンパク質フラグメントのシグナル分泌配列は、G P 3 5 4由来である。

【0186】

50

上記の好ましいタンパク質フラグメントは、このポリペプチドの膜への挿入が非常に所望される場合、必要に応じて膜貫通ドメインを含み得る。この膜貫通ドメインは、G P 3 5 4 ドメイン（以下を参照のこと）であり得るか、または異種膜結合タンパク質の膜貫通ドメインをコードする異種遺伝子によってコードされ得る。

#### 【0187】

そのように所望される場合、上記の好ましいタンパク質フラグメントは、G P 3 5 4 結合相互作用に対する応答において特定のシグナル伝達反応が、所望される場合に、細胞内C末端ドメインをさらに含み得る。この細胞内ドメインは、G P 3 5 4 に由来し得る（以下を参照のこと）か、または、異種膜結合タンパク質の細胞内ドメインをコードする異種遺伝子にコードされ得る。

10

#### 【0188】

本発明のタンパク質フラグメントの他の好ましい実施形態は、G P 3 5 4 の膜貫通ドメインを含むフラグメントである。具体的には、このG P 3 5 4 膜貫通ドメインは、配列番号2のG P 3 5 4 アミノ酸配列のアミノ酸508～530（図1を参照のこと）にコードされる。

#### 【0189】

上記のタンパク質フラグメントのさらに他の好ましい実施形態は、G P 3 5 4 のC末端細胞内ドメインを有する。具体的には、G P 3 5 4 の1つの細胞内ドメインは、配列番号2のG P 3 5 4 アミノ酸配列のアミノ酸531～592（図1参照のこと）にコードされる。G P 3 5 4 の細胞内ドメインの別の形態は、配列番号8のG P 3 5 4 アミノ酸配列のアミノ酸534～708（図5を参照のこと）によってコードされる。これらの異なった細胞内ドメイン形態は、選択的スプライシングによって生成され得ると考えられている。

20

#### 【0190】

本発明の好ましいタンパク質フラグメントは、配列番号1のg p 3 5 4 c D N A 配列のヌクレオチド139～923にコードされる（図1を参照のこと）。これは、プライマー、G X 1-218（配列番号16）およびG X 1-219（配列番号17；実施例2を参照のこと）を使用して腭臓R N A から増幅したR T-PCRフラグメントによってコードされ、そして、配列番号2のアミノ酸47～307からなり、すなわち、これは、G P 3 5 4 の第1のN末端I g ドメイン（68アミノ酸のうち最初の12アミノ酸を欠く）ならびに第2および第3のI g ドメインのほとんどをコードしている。

30

#### 【0191】

上述のように、本発明は、好ましいG P 3 5 4 活性を実質的に保持する配列を有する、挿入もしくは欠失の様式だけ、保存的置換または中程度に保存的置換の様式だけ、ハイブリダイゼーション関連タンパク質として、またはクロスハイブリダイズするタンパク質として、特に、上で参照される配列識別子（配列番号）（S E Q I D N O s）における特徴を有する記載されたタンパク質と、配列において異なるタンパク質をさらに提供する。上で議論されるように、本発明は、異種ポリペプチドに対する、本明細書中に記載されるポリペプチド、タンパク質およびタンパク質フラグメントの融合物をさらに提供する。

#### 【0192】

免疫原として使用される場合、本発明の様々なタンパク質の実施形態は、特に、G P 3 5 4 タンパク質の種々のエピトープに結合する抗体を惹起するために使用され得る。

40

#### 【0193】

（G P 3 5 4 タンパク質の特徴の他の定義付け）

図1は、g p 3 5 4 c D N A クローン（配列番号1）によってコードされる推定アミノ酸配列（配列番号2）を示す。同様に、配列番号4、配列番号8、配列番号10、および配列番号12において提供されるアミノ酸配列は、それぞれ、配列番号3、配列番号7、配列番号9、および配列番号11で提供されるヌクレオチド配列から推定される。別段指摘されない限り、本発明のタンパク質のアミノ酸配列は、核酸配列からの推定翻訳として決定された。従って、本明細書中で示される任意のアミノ酸配列は、上記で詳細に記載したようなこの核酸配列における誤差に起因する誤差を含み得る。さらに、一塩基多型（S

50

NP)は、真核生物ゲノムにおいて頻繁に生じ(ヒトゲノムにおいて140万を超えるSNPが既に同定されている(International Human Genome Sequencing Consortium, Nature 409:860-921(2001))、そして、1つの種の1個体から決定された配列は、集団の中に存在する、他の対立遺伝子形態とは異なり得る。タンパク質の機能を変更しない少量の欠失および挿入は、頻繁に見出され得る。

#### 【0194】

従って、本発明は、本明細書中の特徴を有する記載されたポリペプチドと配列において同一なGP354ポリペプチドのみでなく、本明細書中の特徴を有する記載されたポリペプチドと配列において少なくとも約80%同一な単離されたタンパク質、本明細書中の特徴を有する記載されたポリペプチドと配列において代表的には少なくとも約85%、90%、91%、92%、93%、94%または95%同一な単離されたタンパク質、本明細書中の特徴を有する記載されたポリペプチドに配列において通常は少なくとも約96%、97%、98%、または99%同一な単離されたタンパク質、および最も保存的には、本明細書中の特徴を有する記載されたポリペプチドに配列において少なくとも約99.5%、99.6%、99.7%、99.8%および99.9%同一な単離されたタンパク質が提供される。これらの配列改変体は、天然に存在し得るか、または無作為突然変異誘発または指向された突然変異誘発による人の介入により生じ得る。

#### 【0195】

本明細書中の目的のために、2つのアミノ酸配列のパーセント同一性は、Tatianaらの「Blast 2 sequences—a new tool for comparing protein and nucleotide sequence」, FEMS Microbiol Lett. 174:247-250(1999)の手順を使用して決定される。この手順は、コンピュータプログラムBlast 2 SEQUENCESによって達成され、National Center for Biotechnology Informationからオンラインで利用可能である。

#### 【0196】

アミノ酸配列のパーセント同一性を評価するために、Blast 2 SEQUENCESのBlast Pモジュールが、以下のデフォルト値を用いて使用される：(i) BLOSUM62マトリクス(Henikoffら, Proc. Natl. Acad. Sci USA 89(22):10915-9(1992))；(ii) オープンギャップ(open gap)11ペナルティおよびエクステンション1ギャップペナルティ；ならびに(iii) gap x\_dropoff50 期待値(expect)10ワードサイズ3フィルター、そして両方の配列は全体が入力された。

#### 【0197】

周知のように、アミノ酸置換は、天然の対立遺伝子改変体において頻繁に生じ、保存的置換がしばしば、タンパク質機能における最小の変化のみを引き起こす。従って、本発明は、本明細書中の特徴を持つ記載されるタンパク質と配列において同一なタンパク質のみではなく、保存的アミノ酸置換を有する、GP354タンパク質の配列またはそれらの一部分を有する単離されたタンパク質、を提供する。中程度の保存的アミノ酸置換を有する、GP354タンパク質の配列およびその一部分を有する単離されたタンパク質もまた、提供される。これらの保存的な置換された改変体または中程度に保存的に置換された改変体は、天然に存在し得るか、または人の介入によって生じ得る。

#### 【0198】

対立遺伝子改変は、例えば、195位、196位、539位、および540位での、配列番号2と配列番号8との間のアミノ酸配列における差異の原因であり得る。スプライス改変体(例えば、差次的な5'スプライス部位選択、または3'スプライス部位選択)はまた、配列番号2のC末端アミノ酸配列と配列番号8のC末端アミノ酸配列との間の差異の原因であり得る。

#### 【0199】

10

20

30

40

50

当該分野で周知のように、タンパク質の関連性はまた、機能試験、所定のハイブリダイゼーションストリンジェンシーで互いに塩基対を形成するコード核酸の能力を使用して、特徴付けられる。従って、本明細書中の特徴を有する記載されたタンパク質と配列において同一である単離されたタンパク質を提供するのみではなく、様々な本発明の単離されたポリヌクレオチド（「参照核酸」）の全てまたは一部分に（本明細書中の上記で定義されるような）高度なストリンジェンシー条件下でハイブリダイズする核酸によりコードされる単離されたタンパク質（「ハイブリダイゼーション関連タンパク質」）を提供することも、本発明の別の局面である。

#### 【0200】

ハイブリダイゼーション関連タンパク質は、本発明のGP354タンパク質の代替的なアイソフォーム、ホモログ、パラログ、およびオルソログであり得る。特に有用なオルソログは、チンパンジー、アカゲマカクザル、ヒヒ、オランウータン、およびゴリラのような他の霊長類由来；ラット、マウス、モルモットのようなげっ歯類由来；ウサギ（rabbit）のようなウサギ目の動物（lagomorph）由来；およびウシ、ブタ、ヒツジ、ウマ、ヤギのような家畜由来のオルソログである。

#### 【0201】

タンパク質の関連性はまた、第2の機能的試験（第2のタンパク質が抗体に結合するのを競合的に阻害する第1のタンパク質の能力）を使用して特徴付けられ得る。従って、本発明の別の局面は、本明細書中に詳細に記載されるタンパク質と配列が同一な単離されたタンパク質を提供するだけでなく、本発明の種々の単離されたGP354タンパク質（「参照タンパク質」）の全てまたは一部分への抗体の結合を競合的に阻害する単離されたタンパク質（「交差反応性タンパク質」）もまた提供することである。このような競合的阻害は、当業者に周知の免疫アッセイを使用して容易に決定され得る。

#### 【0202】

10%までの非同一性を生じる欠失および挿入を有するタンパク質、保存的または中程度に保存的な置換を有するタンパク質、ハイブリダイゼーション関連タンパク質、および交差反応性タンパク質を含む、本明細書中に詳細に記載されるタンパク質とアミノ酸配列において異なる本発明のタンパク質の中で、1つ以上のGP354活性を実質的に保持するタンパク質が好ましい（前出を参照のこと）。

#### 【0203】

変化に寛容であるが機能を保持する残基は、アラニン走査変異誘発、Cunninghamら、Science 244 (4908) : 1081-5 (1989) ; トランスポゾンリンカー走査変異誘発、Chenら、Gene 263 (1-2) : 39-48 (2001) ; ホモログ走査変異誘発およびアラニン走査変異誘発の組み合わせ、Jinら、J. Mol. Biol. 226 (3) : 851-65 (1992) ; 機能的アッセイが後に続くコンビナトリアルアラニン走査、Weissら、Proc. Natl. Acad. Sci. USA 97 (16) : 8950-4 (2000) のような当該分野で公知の方法を使用して、既知の残基においてタンパク質を変更することによって同定され得る。トランスポゾンリンカー走査キットは、市販されている（New England Biolabs, Beverly, MA, USA, カタログ番号E7-102S ; EZ::TN<sup>TM</sup> In-Frame Linker Insertion Kit, カタログ番号EZI04KN, Epicentre Technologies Corporation, Madison, WI, USA）。

#### 【0204】

以下にさらに記載されるように、本発明の単離されたタンパク質は、GP354タンパク質、これらのアイソフォーム、ホモログ、パラログ、および／またはオルソログを特異的に認識する抗体を惹起する特異的な免疫原として容易に使用され得る。これらの抗体は、次いで、とりわけ、本発明のGP354タンパク質についてアッセイ（例えば、タンパク質流体サンプル（例えば、血清）の検出のためにELISAによるか、組織サンプル中のタンパク質の検出のために免疫組織化学またはレーザー走査サイトメトリーによるか、ま



たは細胞懸濁液中の細胞内タンパク質の検出のためのフローサイトメトリーによる) するために、G P 3 5 4 タンパク質の特異的抗体媒介性の単離および／または精製(例えば、免疫沈降による) のために、およびG P 3 5 4 作用の特異的なアンタゴニストまたはアゴニストとしての使用のために特異的に使用され得る。

#### 【0205】

本発明の単離されたタンパク質はまた、本発明のG P 3 5 4 タンパク質の濃度および／または量の特異的に決定するために使用されるアッセイにおける特異的な標準としての使用のために直ちに利用可能である。周知であるように、タンパク質分析物の検出および定量のためのE L I S Aキットは、代表的に、測定標準としての使用のための既知の濃度の単離かつ精製されたタンパク質を含む(例えば、ヒトインターフェロン $\gamma$  O p t E I Aキット、カタログ番号555142、Pharmingen, San Diego, CA USAは、バキュロウイルス生成のヒト組換え $\gamma$ インターフェロンを含む)。

#### 【0206】

本発明の単離されたタンパク質はまた、タンパク質-タンパク質相互作用の表面強化レーザー脱離イオン化(S E L D I) 検出のための特異的な生体分子捕捉プローブとしての使用のために直ちに利用可能である(W O 9 8 / 5 9 3 6 2 ; W O 9 8 / 5 9 3 6 0 ; W O 9 8 / 5 9 3 6 1 ; およびMerchantら、Electrophoresis 21 (6) : 1164-77 (2000) (この開示は、その全体が参考として本明細書中に援用される))。同様に、本発明の単離されたタンパク質はまた、B I A C O R E表面プラズモン共鳴プローブ上の特異的な生体分子捕捉プローブとしての使用のために直ちに利用可能である。Weinbergerら、Pharmacogenomics 1 (4) : 395-416 (2000) ; Malmqvist, Biochem. Soc. Trans. 27 (2) 335-40 (1999) を参照のこと。

#### 【0207】

本発明の単離されたタンパク質はまた、G P 3 5 4 の産生またはG P 3 5 4 活性において特定の欠損があると診断された患者における、治療補充剤として有用である。

#### 【0208】

本発明はまた、本発明の種々のタンパク質のフラグメントを提供する。このタンパク質フラグメントは、G P 3 5 4 の抗原性フラグメントおよび免疫原性フラグメントとして有用である。タンパク質の「フラグメント」とは、本明細書中では、どのように入手されようが、参照アミノ酸配列の一部と同一であるアミノ酸配列を有する(この一部分は、少なくとも6アミノ酸であり、かつ参照核酸の全体よりも少ない) 単離されたタンパク質(同様に、ポリペプチド、ペプチド、オリゴペプチド) が意図される。このように定義されるように、「フラグメント」は、参照タンパク質の物理的フラグメント化によって得られる必要はない。しかし、それによりこのような起源は除外されない。

#### 【0209】

少なくとも6連続するアミノ酸のフラグメントは、参照タンパク質のB細胞エピトープおよびT細胞エピトープをマッピングする際に有用である。例えば、Geysenら、「Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid」、Proc. Natl. Acad. Sci. USA 81 : 3998-4002 (1984) ならびに米国特許第4, 708, 871号および同第5, 595, 915号(これらの開示は、その全体が参考として本明細書中に援用される) を参照のこと。このフラグメントは、このようなエピトープマッピングにおいて有用であるために、それ自体が免疫優性エピトープの免疫原性的一部分である必要はなく、またネイティブな抗体によって認識すらされる必要はないので、本発明のタンパク質の少なくとも6アミノ酸の全てのフラグメントは、このような研究において有用性を有する。

#### 【0210】

少なくとも8連続するアミノ酸のフラグメント(しばしば、少なくとも15連続するアミ

ノ酸)は、本発明のタンパク質を認識する抗体を惹起するための免疫原としての有用性を有する。例えば、Lerner「Tapping the immunological repertoire to produce antibodies of predetermined specificity」、Nature 299:592-596 (1982); Shinnickら、「Synthetic peptide immunogens as vaccines」、Annu. Rev. Microbiol. 37:425-46 (1983); Sutcliffeら、「Antibodies that react with predetermined sites on proteins」、Science 219:660-6 (1983) (これらの開示は、その全体が参考として本明細書中に援用される)を参照のこと。上記の参考文献にさらに記載されるように、キャリア(例えば、タンパク質)に結合体化された実質上全ての8マーが、免疫原性であると判明する(すなわち、結合体化ペプチドについての抗体を惹起し得ると判明する);従って、本発明のタンパク質の少なくとも8アミノ酸の全てのフラグメントは、免疫原性としての有用性を有する。 10

#### 【0211】

少なくとも8、9、10、または12連続するアミノ酸のフラグメントはまた、タンパク質全体もしくはその一部分の、抗体(エピトープマッピングにおけるような)、および天然の結合パートナー(例えば、多量体の複合体におけるサブユニット、または係るタンパク質のレセプターまたはリガンド)への結合の競合的インヒビターとして有用である;この競合的な阻害は、目的のタンパク質に特異的に結合する分子の同定および分離を可能にする(米国特許第5,539,084号および同第5,783,674号(これらは、その全体が参考として本明細書中に参考として援用される))。 20

#### 【0212】

従って、本発明のタンパク質またはタンパク質フラグメントは、少なくとも6アミノ酸長、代表的には少なくとも8、9、10、または12アミノ酸長、そしてしばしば少なくとも15アミノ酸長である。しばしば、本発明のタンパク質またはそのフラグメントは、少なくとも20、25、30、35、または50アミノ酸長以上である。少なくとも75、100、150以上のアミノ酸を有するより大きいフラグメントもまた有用であり、時々好まれる。

#### 【0213】

本発明はさらに、本発明のGP354タンパク質およびタンパク質フラグメントの各々の、異種ポリペプチドへの融合物を提供する。融合とは、本発明のタンパク質またはタンパク質フラグメントが、アミノ酸のペプチド結合ポリマーまたはアミノ酸アナログのペプチド結合ポリマーにおいて、異種ポリペプチドに直鎖状に連続していることが、本明細書中で意図される;「異種ポリペプチド」とは、本発明のタンパク質またはタンパク質フラグメントと天然では連続して存在しないポリペプチドが、本明細書中で意図される。そのように定義される場合、その融合物は、変更された配置にて、GP354タンパク質の複数のフラグメントから全体的になり得る;このような場合において、GP354フラグメントのいずれかは、融合タンパク質において、他のGP354フラグメントと異種であると考えられ得る。しかし、より代表的には、異種ポリペプチドは、GP354タンパク質自体から引き出されない。 30 40

#### 【0214】

本発明の融合タンパク質は、本発明のタンパク質の少なくとも1つのフラグメントを含み、このフラグメントは、少なくとも6、代表的に少なくとも8、しばしば少なくとも15、および有用には少なくとも16、17、18、19、または20アミノ酸長である。融合物に含まれる本発明のタンパク質のフラグメントは、有用には、少なくとも25、50、75、100、または150アミノ酸長であり得る。本発明のGP354タンパク質全体、または機能的ドメイン(例えば、N末端GP354 IgドメインおよびC末端細胞内ドメイン)を含む融合物は、特定の有用性を有する。GP354 Igドメインを含む融合物は、他のIgドメイン含有分子を認識する融合タンパク質および表面にこれらを提 50

示する細胞を操作する際に有用である。これは、順に、G P 3 5 4 または結合パートナーを発現した脾臓細胞または C N S 由来細胞；または G P 3 5 4 または結合パートナーを発現する脾臓細胞または C N S 由来細胞に由来する細胞表面分子の全てまたは一部分に、異種配列（例えば、毒素または治療薬）を標的化するために有用であり得る。

#### 【0215】

本発明の融合タンパク質内に含まれる異種ポリペプチドは、少なくとも6アミノ酸長、しばしば少なくとも8アミノ酸長、そして好ましくは少なくとも15、20、25アミノ酸長である。より大きいポリペプチド（例えば、I g G F c 領域）、およびタンパク質全体（例えば、ルシフェラーゼまたは G F P 発色団含有タンパク質）さえも含む融合物は、特定の有用性を有する。

10

#### 【0216】

本発明のベクターおよび発現ベクターの記載において上述したように、本発明の融合タンパク質に含まれる異種ポリペプチドは、有用には、組換え的に発現されるタンパク質の精製および／または可視化を容易にするように設計されたポリペプチドを含む。精製タグはまた、化学的に合成される融合物に組み込まれ得るが、化学合成は、代表的に、H P L C によるさらなる精製で十分である十分な純度を提供する；しかし、上記のような可視化タグは、タンパク質が化学合成によって生成される場合でさえそれらの有用性を保持し、そしてそのように含まれる場合、本発明の融合タンパク質は、G P 3 5 4 の存在の直接的な検出可能なマーカーとして有用になる。

#### 【0217】

また上に議論されるように、本発明の融合タンパク質に含まれる異種ポリペプチドは、有用には、分泌シグナルおよび／またはリーダー配列の組み込みを介して、原核生物宿主については細胞膜周辺腔または細胞外環境、真核生物については培養培地中に、組換え的に発現されるタンパク質の分泌を容易にするポリペプチドを含み得る。

20

#### 【0218】

本発明の他の有用なタンパク質融合物は、酵母ツーハイブリッドシステム中の餌として本発明のタンパク質の使用を可能にする融合物を含む。Bartelら（編）、The Yeast Two-Hybrid System, Oxford University Press (1997) (ISBN: 0195109384)；Zhuら、Yeast Hybrid Technologies, Eaton Publishing, (2000) (ISBN 1-881299-15-5)；Fieldsら、Trends Genet. 10 (8)：286-92 (1994)；Mendelsohnら、Curr. Opin. Biotechnol. 5 (5)：482-6 (1994)；Lubanら、Curr. Opin. Biotechnol. 6 (1)：59-64 (1995)；Allenら、Trends Biochem. Sci. 20 (12)：511-6 (1995)；Dress, Curr. Opin. Chem. Biol. 3 (1)：64-70 (1999)；Topcuら、Pharm. Res. 17 (9)：1049-55 (2000)；Fashenaら、Gene 250 (1-2)：1-14 (2000)（これらの開示は、その全体が参考として本明細書中に援用される）を参照のこと。代表的に、このような融合物は、E. coli LexAまたは酵母GAL4 DNA結合ドメインのいずれかに対する。核局在シグナルに融合された餌を発現する関連した餌プラスミドは、利用可能である。

30

40

#### 【0219】

他の有用なタンパク質融合物としては、上記のような、ファージまたは細胞の表面上にコードされたタンパク質の提示を可能にする融合物、内因的に検出可能なタンパク質（例えば、蛍光タンパク質または光放射タンパク質）への融合物、および安定なタンパク質ドメイン（例えば、I g G F c 領域のような免疫グロブリン重鎖ドメイン）への融合物が挙げられる。

#### 【0220】

本発明のタンパク質およびタンパク質フラグメントはまた、本発明のタンパク質に結合す

50

るかまたはこれを組み込む細胞の特異的切除を行なうために、タンパク質毒素（例えば、Pseudomonas 外毒素 A、ジフテリア毒素、Shiga 毒素 A、炭疽菌毒素致死因子、リシン、または他の生物学的に有害な部分）に有用に融合され得る。

#### 【0221】

本発明の単離されたタンパク質、タンパク質フラグメント、およびタンパク質融合物は、ネイティブなペプチド結合によって結合される天然のアミノ酸から構成され得るか、またはタンパク質の長さ全体にわたってかもしくはその1つ以上の部分に局在するかのいずれかで、非天然のアミノ酸アナログ、非ネイティブな結合、および合成後（翻訳後）改変のいずれかもしくは全てを含み得る。

#### 【0222】

当該分野で周知のように、単離されたタンパク質が使用される場合（例えば、エピトープマッピングについて）、このような非天然のアナログ、非ネイティブな残基間結合、または合成後改変の範囲は、ペプチドの抗体への結合を可能にする範囲に限定される。非ヒト宿主（例えば、マウス）における抗体の調製のための免疫原として使用される場合、このような非天然のアナログ、非ネイティブな残基間結合、または合成後改変の範囲は、タンパク質の免疫原性に干渉しない範囲に限定される。単離されたタンパク質が治療剤（例えば、ワクチン）としてかまたは補充療法のために使用される場合、このような変化の範囲は、単離されたタンパク質に毒性を与えない範囲に限定される。

#### 【0223】

固相化学合成の間にかまたは組換え方法によって非天然のアミノ酸を組み込むための技術は、当該分野において十分に確立されている。手順は、特に、Chanら（編）、Fmoc Solid Phase Peptide Synthesis: A Practical Approach (Practical Approach Series), Oxford Univ. Press (2000年3月) (ISBN: 0199637245); Jones, Amino Acid and Peptide Synthesis (Oxford Chemistry Primers, No7), Oxford Univ. Press (1992年8月) (ISBN: 0198556683); および Bodanszky, Principles of Peptide Synthesis (Springer Laboratory), Springer Verlag (1993年12月) (ISBN: 0387564314)（これらの開示は、その全体が参考として本明細書中に援用されている）において記載されている。

#### 【0224】

天然のアミノ酸のD-エナンチオマーは、化学的ペプチド合成の間に容易に組み込まれ得る；D-アミノ酸から構築されるペプチドは、タンパク質分解攻撃に対してより耐性である；D-エナンチオマーの組み込みはまた、ペプチドに対して特定の三次元コンホメーションを与えるために使用され得る。化学合成の間に一般に付加される他のアミノ酸アナログとしては、オルニチン、ノルロイシン、リン酸化アミノ酸（代表的に、ホスホセリン、ホスホトレオニン、ホスホチロシン）、L-マロニルチロシン（ホスホチロシンのハイブリダイズ不可能なアナログ）（Koleら、Biochem. Biophys. Res. Com. 209: 817-821 (1995)）および種々のハロゲン化フェニルアラニン誘導体が挙げられる。

#### 【0225】

検出可能な標識を有するアミノ酸アナログはまた、標識されたポリペプチドを提供するために、合成の間に有用に組み込まれ得る。例えば、ビオチンは、ビオチノイル-（9-フルオレニルメトキシカルボニル）-L-リジン（Fmocビオサイチン）（Molecular Probes, Eugene, OR, USA）を使用して付加され得る。（ビオチンはまた、E. coli BirA基質ペプチドの融合タンパク質中への組み込みによって酵素的に付加され得る。）ダブシル（dabcyl）-L-リジンのFmocおよびtBOC誘導体（Molecular Probes, Inc., Eugene, OR, USA）を使用して、合成の間にペプチド配列中の選択された部位にダブシル発色団を

10

20

30

40

50

組み込み得る。アミノナフタレン誘導体 EDANS（蛍光共鳴エネルギー移動（FRET）システムにおけるダブシルクエンチャーと対形成するための最も一般的な発色団）は、EDANS- $\gamma$ -FMOC-L-グルタミン酸または対応する tBOC 誘導体（共に Molecular Probes, Inc. Eugene, OR, USA）を使用することによって、自動化ペプチド合成の間に導入され得る。テトラメチルローダミン発色団は、（FMOC）- $\gamma$ -TMR-L-リジン（Molecular Probes, Inc. Eugene, OR, USA）を使用するペプチドの自動化 FMOC 合成の間に組み込まれ得る。

#### 【0226】

化学合成の間に組み込まれ得る他の有用なアミノ酸アナログとしては、アスパラギン酸、グルタミン酸、リジン、およびアリル側鎖保護を有するチロシンアナログ（Applied Biosystems, Inc., Foster City, CA, USA）が挙げられる；アリル側鎖は、環状ペプチド、分枝鎖ペプチド、スルホン化ペプチド、グリコシル化ペプチド、およびリン酸化ペプチドの合成を可能にする。化学合成の間に組み込み可能である、大多数の他の FMOC 保護された非天然アミノ酸アナログは、例えば、The Peptide Laboratory（Richmond, CA, USA）から市販されている。

#### 【0227】

非天然のアミノ酸残基はまた、所望の非天然のアミノ酸との化学的アミノアシル化によって、サプレッサー tRNA（代表的には、UAG 停止コドンを認識する tRNA）を操作することにより生合成的に付加され得る。従来の部位特異的変異誘発は、タンパク質遺伝子中の目的の部位に選択された停止コドン UAG を導入するために使用される。アシル化サプレッサー tRNA および変異遺伝子が、インビトロ転写／翻訳系において組み合わせられる場合、非天然のアミノ酸は、特定化された位置にそのアミノ酸を含むタンパク質を与えるように、UAG コドンに応答して組み込まれる。Liu ら、Proc. Natl Acad. Sci. USA 96（9）：4780-5（1999）；Wang ら、Science 292（5516）：498-500（2001）。

#### 【0228】

本発明の単離された GP354 タンパク質、タンパク質フラグメントおよび融合タンパク質はまた、環状形態および分枝形態を生じる結合を含む、非ネイティブの残基間結合を含み得る。本発明の単離された GP354 タンパク質およびタンパク質フラグメントはまた、タンパク質の長さ全体にわたってかまたは 1 つ以上のその部分に局在してかのいずれかで、翻訳後改変および合成後改変を含み得る。

#### 【0229】

例えば、真核生物細胞における組換え発現によって生成される場合、本発明の単離されたタンパク質、フラグメントおよび融合タンパク質は、代表的に、N 連結および／または O 連結グリコシル化を含み、これらのパターンは、タンパク質配列におけるグリコシル化部位の利用可能性および宿主細胞の同一性の両方を反映する。グリコシル化パターンのさらなる改変は、酵素的に行なわれ得る。別の例として、本発明の組換えポリペプチドはまた、宿主媒介性プロセスから生じるいくつかの場合、改変された開始メチオニン残基を含み得る。

#### 【0230】

本発明のタンパク質、タンパク質フラグメント、およびタンパク質融合物が、化学合成によって生成される場合、合成後改変が、樹脂からの脱保護および切断の前か、または脱保護および切断の後に、行なわれ得る。合成タンパク質の脱保護および切断の前の改変は、しばしば、例えば、樹脂結合合成ペプチドの N 末端への改変部分の標的化を可能にすることによって、より大きい制御を可能にする。有用な合成後（および翻訳後）改変は、発光団のような検出可能な標識への結合体化を含む。一方で、リジン残基の N 末端アミノ基および  $\epsilon$  アミノ基と、そして他方で、システイン残基の遊離チオール基と、非変性条件下で反応する、広範な種々のアミン反応性およびチオール反応性の発色団誘導体が、合成され

ている。

#### 【0231】

種々のアミン反応性またはチオール反応性の発色団へのタンパク質の結合体化を可能にするキットが市販されている；Molecular Probes, Inc. (Eugene, OR, USA) は、例えば、Alexa Fluor 350、Alexa Fluor 430、Fluorescein-EX、Alexa Fluor 488、Oregon Green 488、Alexa Fluor 532、Alexa Fluor 546、Alexa Fluor 546、Alexa Fluor 568、Alexa Fluor 594、およびTexas Red-Xへのタンパク質の結合体化のためのキットを提供する。Alexa Fluor (登録商標) 350、Alexa Fluor (登録商標) 488、Alexa Fluor (登録商標) 532、Alexa Fluor (登録商標) 546、Alexa Fluor (登録商標) 568、Alexa Fluor (登録商標) 594、Alexa Fluor (登録商標) 647 (モノクローナル抗体標識キット)、BODIPY色素、Cascade Blue、Cascade Yellow、Dansyl、リサミンローダミンB、Marina Blue、Oregon Green 488、Oregon Green 514、Pacific Blue、ローダミン6G、ローダミングリーン、ローダミンレッド、テトラメチルローダミン、Texas Redを含む、広範な種々の他のアミン反応性およびチオール反応性の発色団が、市販されている (Molecular Probes, Inc., Eugene, OR, USA)。

#### 【0232】

本発明のポリペプチドはまた、二官能性連結試薬を使用して、発色団、他のタンパク質、および他の高分子に結合体化され得る。一般的なホモ二官能性試薬としては、例えば、APG、AEDP、BASED、BMB、BMDB、BMH、BMOE、BM[PEO] 3、BM[PEO] 4、BS3、BSOCOES、DFDNB、DMA、DMP、DMS、DPDPB、DSG、DSP (Lomant 試薬)、DSS、DST、DTBP、DTME、DTSSP、EGS、HBVS、Sulfo-BSOCOES、Sulfo-DST、Sulfo-EGS (全てPierce, Rockford, IL, USAから入手可能) が挙げられる；一般的なヘテロ二官能性架橋剤としては、ABH、AMAS、ANB-NOS、APDP、ASBA、BMPA、BMPH、BMPS、EDC、EMCA、EMCH、EMCS、KMUA、KMUH、GMBS、LC-SMCC、LC-SPDP、MBS、M2C2H、MPBH、MSA、NHS-ASA、PDPH、PMPH、SADP、SAED、SAND、SANPAH、SASD、SATP、SBAP、SFAD、SIA、SIAB、SMCC、SMPB、SMPH、SMPT、SPDP、Sulfo-EMCS、Sulfo-GMBS、Sulfo-HSAB、Sulfo-KMUS、Sulfo-LC-SPDP、Sulfo-MBS、Sulfo-NHS-LC-ASA、Sulfo-SADP、Sulfo-SANPAH、Sulfo-SIAB、Sulfo-SMCC、Sulfo-SMPB、Sulfo-LC-SMPT、SVSB、TFCS (全てPierce, Rockford, IL, USAから入手可能) が挙げられる。

#### 【0233】

本発明のタンパク質、タンパク質フラグメント、およびタンパク質融合物は、このような架橋剤を使用して、アミン反応性またはチオール反応性ではない発色団に結合体化され得る。本発明のタンパク質、タンパク質フラグメント、および融合タンパク質に有用に結合体化され得る他の標識としては、放射性標識、内視鏡的超音波検査造影剤、およびMRI造影剤が挙げられる。本発明のタンパク質、タンパク質フラグメント、およびタンパク質融合物はまた、抗GP354抗体を惹起するための免疫原性を増加させるために、キャリアタンパク質 (例えば、KLH、ウシサイログロブリン、およびウシ血清アルブミン (BSA)) に対してさえも架橋剤を使用して有用に結合体化され得る。

#### 【0234】

本発明のGP354タンパク質、タンパク質フラグメントおよびタンパク質融合物はまた

、通常、ポリエチレングリコール（PEG）に結合され得る；PEG化は、代償療法のために静脈内に投与されるタンパク質の血清半減期を増大する。Delgadoら, Crit. Rev. Ther. Drug Carrier Syst. 9 (3-4) : 249-304 (1992) ; Scottら, Curr. Pharm. Des. 4 (6) : 423-38 (1998) ; DeSantisら, Curr. Opin. Biotechnol. 10 (4) : 324-30 (1999) (これらの全体が、参考として本明細書中に援用される)。PEGモノマーは、穏やかな条件下で直接的な結合を可能するトレシルクロリド (tresyl chloride) (2, 2, 2-トリフルオロエタンスルホニルクロリド) で活性化されたPEGモノマーを使用するPEG化で、直接的に、またはリンカーを介してタンパク質に結合され得る。

10

#### 【0235】

本発明の単離されたGP354タンパク質（その融合物を含む）は、典型的には、上記されるような本発明の発現ベクターを使用して組換え発現によって、または特に約100アミノ酸よりも少ない場合、必要に応じて化学的合成によって（典型的には、固相合成）、および時には、インビトロ翻訳によって産生され得る。

#### 【0236】

本発明の単離されたタンパク質の産生の後には、必要に応じて精製を行い得る。組換え発現されたタンパク質の精製は、今や当該分野の技術範囲内である。例えば、Thorne (編), Applications of Chimeric Genes and Hybrid Proteins, Part A: Gene Expression and Protein Purification (Methods in Enzymology, 第326巻), Academic Press (2000), (ISBN: 0121822273) ; Harbin (編), Cloning, Gene Expression and Protein Purification: Experimental Procedures and Process Rationale, Oxford Univ. Press (2001) (ISBN: 0195132947) ; Marshakら, Strategies for Protein Purification and Characterization: A Laboratory Course Manual, Cold Spring Harbor Laboratory Press (1996) (ISBN: 0-87969-385-1) ; および Roe (編), Protein Purification Applications, Oxford University Press (2001) (これらの開示は、その全体が参考として本明細書中に援用される) を参照のこと。従って、ここで詳述される必要はない。

20

30

#### 【0237】

しかし、簡潔には、精製タグが、このようなタグを付加した発現ベクターの使用を介して融合される場合、精製は、タグに適切である手段（例えば、ポリヒスチジンタグのための固定化金属アフィニティークロマトグラフィーの使用）によって、少なくとも部分的にもたらされ得る。当該分野において通常の他の技術としては、硫酸アンモニウム分画、免疫沈降、タンパク質高速液体クロマトグラフィー (FPLC)、高速液体クロマトグラフィー (HPLC) および分取ゲル電気泳動が挙げられる。化学的に合成されたペプチドの精製は、例えば、HPLCによって容易にもたらされ得る。

40

#### 【0238】

従って、純粋または実質的に純粋な形態において本発明の単離されたGP354タンパク質を提供することが、本発明の局面である。本発明の精製されたタンパク質は、上記のように単離されたタンパク質であり、これは、組成物中の総タンパク質に関して質量基準 (w/w) において測定される場合、少なくとも95%の濃度で存在する。このような精製物は、しばしば、例えば、HPLCによるようなさらなる精製を伴わずに化学合成の間に得られ得る。本発明の精製されたタンパク質は、96%、97%、98%および99%の濃度でさえ存在し得る（組成物中の総タンパク質に関して質量基準において測定される）

50

。本発明のタンパク質はなお、H P L Cによるような精製後に、99.5%、99.6% および99.7%、99.8%または99.9%のレベルでさえ存在し得る。

#### 【0239】

本発明の単離されたタンパク質は、治療薬として（例えば、ワクチンとしてまたは代償療法のために）使用される場合、高レベルの純度が、好ましいが、本発明の単離されたタンパク質はまた、低純度でも有用である。例えば、本発明の部分的に精製されたタンパク質を、免疫原として使用して、実験動物において抗体を惹起し得る。

#### 【0240】

従って、本発明は、実質的に精製された形態において本発明の単離されたタンパク質を提供する。本発明の「実質的に精製されたタンパク質」は、組成物中の総タンパク質に関して質量基準において測定される場合、少なくとも70%の濃度で存在する、上記のような単離されたタンパク質である。通常、実質的に精製されたタンパク質は、組成物中の総タンパク質に関して質量基準において測定される場合、少なくとも75%、80%、または少なくとも85%、90%、91%、92%、93%、94%、94.5%でさえまたは少なくとも94.9%の濃度でさえ存在する。

#### 【0241】

好ましい実施形態において、本発明の精製されたタンパク質および実質的に精製されたタンパク質は、検出可能な両性電解質、アクリルアミドモノマー、ビスアクリルアミドモノマーおよびポリアクリルアミドを欠いた組成物中にある。

#### 【0242】

本発明のG P 3 5 4タンパク質、フラグメントおよび融合物は、通常、基質に結合され得る。基質は、多孔性、実質的に非多孔性（例えば、プラスチック）または中実；平面または非平面であり；結合は、共有結合または非共有結合であり得る。多孔性基質（通常は、膜）は、典型的に、ニトロセルロース、ポリビニリデンフルオライド（P V D F）または陽イオン誘導体化された親水性P V D Fを含み；このように結合されると本発明のタンパク質、フラグメントおよび融合物を使用して、本発明の固定化タンパク質に特異的に結合する、例えば、血清中の抗体を検出および定量し得る。本発明のタンパク質、フラグメントおよび融合物は、プラスチックのような、実質的に非多孔性基質に結合する場合、本発明の固定化タンパク質に特異的に結合する、例えば、血清中の抗体を検出および定量するために使用され得る。

#### 【0243】

本発明のタンパク質、フラグメントおよび融合物はまた、表面増強レーザー脱着イオン化供給源（surface enhanced laser desorption ionization source）としての使用に適切である基質に結合され得る；このように結合されると本発明のタンパク質、フラグメントまたは融合物は、第2のタンパク質を結合し、次いで検出するのに有用であり、このタンパク質は、表面結合タンパク質にそれらの間の生物学的な相互作用を示す十分な親和性またはアビディティで結合する。

#### 【0244】

本発明のタンパク質、フラグメントおよび融合物はまた、表面プラズモン共鳴検出における使用に適切である基質に結合され得る。このように結合されると本発明のタンパク質、フラグメントまたは融合物は、第2のタンパク質を結合し、次いで検出するのに有用であり、このタンパク質は、表面結合タンパク質に2つの間の有意な生物学的な相互作用を示すのに十分な親和性またはアビディティで結合する。

#### 【0245】

（抗体および抗体産生細胞）

本発明は、本発明のG P 3 5 4タンパク質およびタンパク質フラグメントに特異的に結合するか、または本発明の単離されたG P 3 5 4核酸によってコードされる1つ以上のタンパク質およびタンパク質フラグメントに結合する抗体（そのフラグメントおよび誘導体を含む）を提供する。抗体は、ネイティブコンホメーションのタンパク質に存在するか、またはいくつかの場合、例えば、S D S中での可溶化によって変性されるタンパク質に存在

10

20

30

40

50



するか、いずれかのようなタンパク質またはタンパク質フラグメントの直線のエピトープ、接触しないエピトープまたはコンホメーションエピトープに特異的であり得る。

【0246】

本発明はまた、抗体の結合が、本発明の1つ以上のGP354タンパク質およびタンパク質フラグメントによって、または本発明の単離されたgp354ポリヌクレオチドによってコードされる1つ以上のタンパク質およびタンパク質フラグメントによって競合的に阻害され得る抗体（そのフラグメントおよび誘導体を含む）を提供する。

【0247】

第1の一連の抗体の実施形態において、本発明は、配列番号2、4、8、10または12における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、抗体（ポリクローナルおよびモノクローナルの両方）ならびにそのフラグメントおよび誘導体を提供する。 10

【0248】

このような抗体は、種々のインビトロ免疫アッセイ（例えば、ウエスタンブロッティングおよびELISA）において有用である。このような抗体はまた、免疫沈降、免疫アフィニティークロマトグラフィーまたは磁気ビーズ媒介精製（magnetic bead-mediated purification）によってGP354タンパク質（関連する交差反応性タンパク質を含む）を単離および精製するのに有用である。このような方法は、当該分野で周知である。

【0249】

第2の一連の抗体の実施形態において、本発明は、抗体の特異的な結合が、本発明の単離されたタンパク質およびポリペプチドによって競合的に阻害され得る抗体（ポリクローナルおよびモノクローナルの両方）ならびにそのフラグメントおよび誘導体を提供する。 20

【0250】

他の実施形態において、本発明はさらに、検出可能に標識された上記の抗体を提供し、そしてなお他の実施形態において、基質に結合される上記の抗体を提供する。

【0251】

本発明の抗PanCAM抗体（実施例11および19、下記）を使用して、膵臓β細胞（例えば、大きな動物（例えば、ウシ、ブタ、およびウマ）由来；または霊長類（例えば、サルおよびチンパンジー）由来）を単離し得る。そのように得られたβ細胞は、糖尿病患者（例えば、ヒト患者）中の制限された生存期間のカプセル化（従って、免疫保護化）異種移植片を作製するのに有用である。 30

【0252】

いくつかの実施形態において、本発明の抗体を使用して、β島に対する治療的薬剤を標的化する。例えば、I型糖尿病は、β島へのT細胞侵入によって標識される。従って、これらのT細胞（例えば、抗-CD3抗体、抗-CD4抗体または抗-CD8抗体）に特異的である細胞傷害性薬剤（すなわち、細胞の生存度および／または機能を損なう薬剤）は、抗PanCAM抗体に結合体化し得、それによって疾患部位に指向され得る。別の例において、アンチセンス療法（例えば、MHC II型分子に対するアンチセンスオリゴ、これは、β細胞を攻撃するTリンパ球に高く発現される；Chaturvediら、J Immunol 164（12）：6610-20（2000））は、抗PanCAM抗体 40に結合し得、そしてβ島に指向され得る。

【0253】

本発明の抗体はまた、複数の病巣である傾向があるので、処置することが難しいインスリノーマの部位に細胞毒を指向するために、細胞傷害性薬剤（例えば、放射性薬剤のような非特異的薬剤；リシンおよびコレラ毒素のような標的可能な毒素；およびドキソルビシンのような化学療法剤）に結合体化され得る。抗体をまた使用して、腫瘍の挙動を阻害するかまたは他に改変するような遺伝子治療（例えば、応答可能なオンコジーンまたはそのインシュリン遺伝子のいずれか、またはインシュリン放出を調節するシグナル伝達機構における遺伝子にアンチセンスオリゴヌクレオチドまたはDNA構築物を送達するため）としてインスリノーマにDNAを送達し得る。

## 【0254】

本明細書中において使用される場合、用語「抗体」は、第1の分子種に特異的に結合し得る、少なくとも一部が、少なくとも1つの免疫グロブリン遺伝子によってコードされるポリペプチドおよびこのような特異的な結合が可能なままのそのフラグメントまたは誘導体をいう。

## 【0255】

「特異的に結合する」および「特異的結合」とは、本明細書中で、抗体が、抗体および第1の分子種が混合される他の分子種に結合するよりむしろ第1の分子種に結合する能力を意図する。抗体が、第1の分子種に特異的に結合し得る場合、この抗体は特に、この第1の分子種を「認識する」と言われる。

10

## 【0256】

当該分野で周知であるように、抗体が、混合物中の分子種間を識別し得る程度は、混合物中の種のコンホメーションの関連性に部分的に依存する；典型的に、本発明の抗体は、少なくとも2倍、さらに典型的には少なくとも5倍、典型的には10倍、25倍、50倍、75倍より多くまでおよびしばしば100倍より多く、ならびに時として500倍または1000倍より多く非GP354タンパク質への偶発的な結合を越えて識別する。本発明のタンパク質またはタンパク質フラグメントを検出するために使用される場合、ヒト臍臓および神経組織由来のサンプルにおける本発明のタンパク質の存在を決定するために使用され得る際、本発明の抗体は、十分に特異的である。

## 【0257】

典型的には、本発明のGP354タンパク質またはタンパク質フラグメントに対する本発明の抗体（または、IgMペンタマーの場合、抗体マルチマー）の親和性あるいはアビディティは、少なくとも約 $1 \times 10^{-6}$ モル濃度（M）、典型的には少なくとも約 $5 \times 10^{-7}$  M、通常少なくとも約 $1 \times 10^{-7}$  Mであり、少なくとも $1 \times 10^{-8}$  M、 $5 \times 10^{-9}$  Mおよび $1 \times 10^{-10}$  Mの親和性およびアビディティが、特に有用であることがわっている。

20

## 【0258】

本発明の抗体は、任意の哺乳動物種由来の天然に存在する形態（例えば、IgG、IgM、IgD、IgEおよびIgA）であり得る。

## 【0259】

ヒト抗体は、まれではあるが、ヒトドナーまたはヒト細胞から直接的に引き出され得る。このような場合、本発明のタンパク質に対する抗体は、典型的に、本発明のタンパク質またはタンパク質フラグメントでの、偶発的な免疫化（例えば、自己免疫化）から生じる。このような抗体は、典型的に、不変ではないが、ポリクローナルである。

30

## 【0260】

ヒト抗体は、ヒト免疫グロブリン遺伝子を発現するトランスジェニック動物を使用してさらに頻繁に得られる。このトランスジェニック動物は、本発明のGP354タンパク質免疫原で肯定的に免疫され得る。ヒト抗体を産生し得るヒトIgトランスジェニックマウスおよび特異的な免疫化によってトランスジェニックマウス由来のヒト抗体を産生する方法は、当該分野で周知である。例えば、米国特許第6,162,963号；同第6,150,584号；同第6,114,598号；同第6,075,181号；同第5,939,598号；同第5,877,397号；同第5,874,299号；同第5,814,318号；同第5,789,650号；同第5,770,429号；同第5,661,016号；同第5,633,425号；同第5,625,126号；同第5,569,825号；同第5,545,807号；同第5,545,806号および同第5,591,669号（これらの開示は、それらの全体が、参考として本明細書中に開示される）を参照のこと。このような抗体は、典型的には、モノクローナルであり、そして典型的には、マウス抗体の産生に関して開発された技術を使用して産生される。

40

## 【0261】

本発明の抗体が、インビボ診断薬または治療薬としてヒトに投与されるべきである場合、

50

投与された抗体へのレシピエント免疫応答が、しばしば、マウスのような別の種に由来する抗体の投与によって引き起こされるレシピエント免疫応答より実質的に低いので、ヒト抗体は、特に有用であり、そしてしばしば、好ましい。

#### 【0262】

本発明の Ig G、Ig M、Ig D、Ig E および Ig A 抗体はまた、有用なことに、げっ歯類（典型的には、マウス、だけでなくラット、モルモットおよびハムスター）、ウサギ目の動物（典型的には、ウサギ）およびより大きな哺乳動物（例えば、ヒツジ、ヤギ、ウシおよびウマ）を含む他の哺乳動物種から得られる。このような場合、ヒト抗体を産生するトランスジェニック非ヒト哺乳動物を用いる場合と同様に、偶発的な免疫化は、必要とされず、そして非ヒト哺乳動物は、典型的に、本発明のタンパク質またはタンパク質フラグメントを用いて、標準的な免疫プロトコルに従って、肯定的に免疫される。 10

#### 【0263】

上記されるように、他のところで上記される（これらの考察は、本明細書中に参考として援用される）ような二機能性のリンカーを都合よく使用して、キャリア、典型的には、ウシサイログロブリン、キーホールカサガイヘモシアニンまたはウシ血清アルブミンのようなタンパク質に結合体化される場合、本発明のタンパク質の8個以上連続するアミノ酸の実質的に全てのフラグメントは、免疫原として効果的に使用され得る。

#### 【0264】

免疫原性はまた、他の部分と本発明のタンパク質およびタンパク質フラグメントの融合によって与えられ得る。本発明のペプチドは、例えば、分枝したポリリジンコアマトリクスにおける固相合成によって産生され得る；これらの多重抗原ペプチド（MAP）は、高い純度、増大したアビディティ、正確な化学的定義およびワクチン開発において改善した安全性を提供する。Tamら, Proc. Natl. Acad. Sci. USA 85: 5409-5413 (1988)；Posnettら, J. Biol. Chem. 263, 1719-1725 (1988)。 20

#### 【0265】

非ヒト哺乳動物を免疫するためのプロトコルは、当該分野において非常に確立されている（Harlowら（編）, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory (1998) (ISBN: 0879693142)；Coliganら（編）, Current Protocols in Immunology, John Wiley & Sons, Inc. (2001) (ISBN: 0-471-52276-7)；Zola, Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Background to Bench), Springer Verlag (2000) (ISBN: 0387915907)（これらの開示は、本明細書中に参考として援用される））。 30

#### 【0266】

非ヒト哺乳動物由来の抗体は、ポリクローナルまたはモノクローナルであり得、ここでポリクローナル抗体は、本発明のタンパク質の免疫組織化学的検出において特定の利点を有し、そしてモノクローナル抗体は、本発明のタンパク質の特定のエピトープを識別および区別する際に利点を有する。 40

#### 【0267】

免疫の後に、本発明の抗体は、任意の当該分野で受け入れられた技術を使用して産生され得る。このような技術は、当該分野で周知である（Coliganら（編）, Current Protocols in Immunology, John Wiley & Sons, Inc. (2001) (ISBN: 0-471-52276-7)；Zola, Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Backgr 50

ound to Bench), Springer Verlag (2000) (ISBN: 0387915907); Howardら(編), Basic Methods in Antibody Production and Characterization, CRC Press (2000) (ISBN: 0849394457); Harlowら(編), Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory (1998) (ISBN: 0879693142); Davis(編), Monoclonal Antibody Protocols, 第45巻, Humana Press (1995) (ISBN: 0896033082); Delves(編), Antibody Production: Essential Techniques, John Wiley & Son Ltd (1997) (ISBN: 0471970107); Kenney, Antibody Solution: An Antibody Methods Manual, Chapman & Hall (1997) (ISBN: 0412141914) (これらの全体が、本明細書中に参考として援用される)。従って、ここで詳述する必要はない。

#### 【0268】

本発明の抗体のフラグメントまたは誘導体が、所望される場合、宿主細胞における組換え発現は、特に有用である。組換え抗体(抗体全体、抗体フラグメントまたは抗体誘導体のいずれか)産生のための宿主細胞は、原核生物または真核生物であり得る。

#### 【0269】

原核生物宿主は、本発明のファージディスプレイされた抗体を産生するのに特に有用である。ファージディスプレイされた抗体の技術(ここで、抗体可変領域フラグメントは、例えば、線維状のファージ(例えば、M13)の表面上でのディスプレイに関して遺伝子I I Iタンパク質(p I I I)または遺伝子V I I Iタンパク質(p V I I I)に融合される)は、現在までに十分に確立され(Sidhu, Curr. Opin. Biotechnol. 11(6): 610-6(2000); Griffithsら, Curr. Opin. Biotechnol. 9(1): 102-8(1998); Hoogenboomら, Immunotechnology, 4(1): 1-20(1998); Raderら, Curr. Opin. Biotechnol. 8: 503-508(1997); Aujaumeら, Hum. Antibodies 8: 155-168(1997); Hoogenboom, Trends in Biotechnol. 15: 62-70(1997); de Kruijffら, 17: 453-455(1996); Barbasら, Trends Biotechnol. 14: 230-234(1996); Winterら, Annu. Rev. Immunol. 433-455(1994))、そしてこのようなライブラリー由来の抗体フラグメントを産生、成長、スクリーニング(パニング(pan))および使用するために必要とされる技術およびプロトコルは、最近以下に編集されている: Barbasら, Phage Display: A Laboratory Manual, Cold Spring Harbor Laboratory Press (2001) (ISBN 0-87969-546-3); Kayら(編), Phage Display of Peptides and Proteins: A Laboratory Manual, Academic Press, Inc. (1996); Abelsonら(編), Combinatorial Chemistry, Methods in Enzymology vol. 267, Academic Press (1996年5月) (これらの開示は、その全体が、参考として本明細書中に援用される)。典型的には、ファージディスプレイされた抗体フラグメントは、sc FvフラグメントまたはFabフラグメントである; 所望される場合、全長抗体は、ディスプレイするファージから完全な抗体に可変領域をクローニングし、そしてさらなる原核宿主細胞または真核宿主細胞において全長抗体を発現することによって産生され得る。

#### 【0270】

真核細胞はまた、本発明の抗体、抗体フラグメントおよび抗体誘導体の発現に有用である。例えば、本発明の抗体フラグメントは、Pichia pastoris, Takah

ashira, Biosci. Biotechnol. Biochem. 64 (10) : 2138-44 (2000); Freyreら, J. Biotechnol. 76 (2-3) : 157-63 (2000); Fischerら, Biotechnol. Appl. Biochem. 30 (Pt 2) : 117-20 (1999); Pennellら, Res. Immunol. 149 (6) : 599-603 (1998); Eldinら, J. Immunol. Methods. 201 (1) : 67-75 (1997); および Saccharomyces cerevisiae, Frenkenら, Res. Immunol. 149 (6) : 589-99 (1998); Shustaraら, Nature Biotechnol. 16 (8) : 773-7 (1998) (これらの開示は、その全体が、本明細書中に援用される) において産生され得る。

10

**【0271】**

本発明の抗体 (抗体フラグメントおよび誘導体を含む) はまた、昆虫細胞において産生され得る。Liら, Protein Expr. Purif. 21 (1) : 121-8 (2001); Ailorら, Biotechnol. Bioeng. 58 (2-3) : 196-203 (1998); Hsuら, Biotechnol. Prog. 13 (1) : 96-104 (1997); Edelmanら, Immunology 91 (1) : 13-9 (1997); および Nesbittら, J. Immunol. Methods. 151 (1-2) : 201-8 (1992) (これらの開示は、その全体が、本明細書中に援用される)。

20

**【0272】**

本発明の抗体ならびにそのフラグメントおよび誘導体はまた、植物細胞において産生され得る。Giddingsら, Nature Biotechnol. 18 (11) : 1151-5 (2000); Gavilondora, Biotechniques 29 (1) : 128-38 (2000); Fischerら, J. Biol. Regul. Homeost. Agents 14 (2) : 83-92 (2000); Fischerら, Biotechnol. Appl. Biochem. 30 (Pt 2) : 113-6 (1999); Fischerら, Biol. Chem. 380 (7-8) : 825-39 (1999); Russell, Curr. Top. Microbiol. Immunol. 240 : 119-38 (1999); および Maら, Plant Physiol. 109 (2) : 341-6 (1995) (これらの開示は、その全体が、本明細書中に援用される)。

30

**【0273】**

本発明の抗体、抗体フラグメントおよび抗体誘導体の組換え発現に有用な哺乳動物細胞としては、CHO細胞、COS細胞、293細胞および骨髓腫細胞が挙げられる。Vermaraら, J. Immunol. Methods 216 (1-2) : 165-81 (1998) は、抗体の発現について細菌、酵母、昆虫および哺乳動物発現系を概説および比較する。

**【0274】**

本発明の抗体はまた、Merkら, J. Biochem. (Tokyo). 125 (2) : 328-33 (1999) および Ryabovara, Nature Biotechnol. 15 (1) : 79-84 (1997) にさらに記載されるように無細胞翻訳によって調製され得、そして Pollackら, J. Immunol. Methods 231 (1-2) : 147-57 (1999) (これらの開示は、その全体が、本明細書中に参考として援用される) にさらに記載されるようにトランスジェニック動物の乳汁に調製され得る。

40

**【0275】**

本発明は、本発明の1つ以上のGP354タンパク質およびタンパク質フラグメント、本発明の単離されたgp354ポリヌクレオチドによってコードされる1つ以上のタンパク質およびタンパク質フラグメントに特異的に結合する抗体フラグメント、あるいは本発明の1つ以上のタンパク質およびタンパク質フラグメント、または本発明の単離されたポリ

50

ヌクレオチドによってコードされる1つ以上のタンパク質およびタンパク質フラグメントによって結合が競合的に阻害され得る抗体フラグメントをさらに提供する。

【0276】

このような抗体フラグメントのうちでも有用なフラグメントは、Fab、Fab'、Fv、F(ab')<sub>2</sub>、および単鎖Fv(scFv)フラグメントである。他の有用なフラグメントは、Hudson, Curr. Opin. Biotechnol. 9(4):395-402(1998)に記載される。従って、本発明は、本発明の1つ以上のGP354タンパク質およびタンパク質フラグメント、本発明の単離された核酸によってコードされる1つ以上のタンパク質およびタンパク質フラグメントに特異的に結合する抗体誘導体、あるいは本発明の1つ以上のタンパク質およびタンパク質フラグメント、または本発明の単離されたポリヌクレオチドによってコードされる1つ以上のタンパク質およびタンパク質フラグメントによって結合が、競合的に阻害され得る抗体誘導体を提供する。

10

【0277】

このような誘導体のうちでも有用な誘導体は、キメラ抗体、霊長類化抗体およびヒト化抗体である；このような誘導体は、従って、非ヒト哺乳動物種由来の改変されていない抗体よりヒトにおいて、免疫原性が低く、インビボ投与にさらに適切である。

【0278】

キメラ抗体は、典型的に別の種（典型的にはヒト）の定常領域に融合した、1種（典型的にはマウス）の免疫グロブリンの重鎖および／または軽鎖可変領域（CDRとフレームワーク残基の両方を含む）を含む。例えば、米国特許第5,807,715号；Morrissonら, Proc. Natl. Acad. Sci. USA. 81(21):6851-5(1984)；Sharonら, Nature 309(5966):364-7(1984)；Takedaら, Nature 314(6010):452-4(1985)（これらの開示は、その全体が、参考として本明細書中に援用される）を参照のこと。

20

【0279】

霊長類化抗体およびヒト化抗体は、典型的には、非ヒト霊長類に移植されたマウス抗体由来の重鎖および／または軽鎖CDRまたはヒト抗体V領域フレームワークを含む（通常、ヒト定常領域をさらに含む）（Riechmannら, Nature 332(6162):323-7(1988)；Coら, Nature 351(6326):501-2(1991)；米国特許第6,054,297号；同第5,821,337号；同第5,770,196号；同第5,766,886号；同第5,821,123号；同第5,869,619号；同第6,180,377号；同第6,013,256号；同第5,693,761号；および同第6,180,370号（これらの開示は、その全体が、参考として本明細書中に援用される））。

30

【0280】

本発明の他の有用な抗体誘導体は、ヘテロマー抗体複合体および抗体融合体（例えば、二抗体（二重特異性抗体）、単鎖二抗体および内抗体）を含む。

【0281】

本発明の抗体（そのフラグメントおよび誘導体を含む）は、有用なためには、標識され得る。従って、本発明の別の局面は、本発明の1以上のタンパク質およびタンパク質フラグメント、本発明の単離されたポリヌクレオチドによってコードされる1以上のGP354タンパク質およびタンパク質フラグメントに特異的に結合する標識された抗体、または本発明の1以上のタンパク質およびタンパク質フラグメント、または本発明の単離されたポリヌクレオチドによってコードされる1以上のタンパク質およびタンパク質フラグメントによって結合が競合的に阻害され得る標識された抗体を提供することである。

40

【0282】

標識の選択は、一部、所望される用途に依存する。本発明の抗体が、組織サンプルの免疫組織化学的染色のために使用される場合、標識は、有用には、検出可能な生成物の生成および局所的堆積を触媒する酵素であり得る。代表的に抗体に結合体化されて、その免疫組織化学的可視化を可能にする酵素は、周知であり、そして、これらとしては、アルカリホ

50

スファターゼ、 $\beta$ -ガラクトシダーゼ、グルコースオキシダーゼ、西洋ワサビペルオキシダーゼ (HRP) およびウレアーゼが挙げられる。本発明の抗体はまた、コロイド状金を使用して標識され得る。

#### 【0283】

可視的に検出可能な生成物、発光標識および蛍光標識の生成ならびに堆積のための多数の代表的な物質もまた、周知であり、そしてさらに記載される必要はない。例えば、Thorpeら、Methods Enzymol. 133:331-53 (1986); Krickaら、J. Immunoassay 17(1):67-83 (1996); およびLundqvistら、J. Biolumin. Chemilumin. 10(6)353-9 (1995) (これらの開示は、その全体が本明細書中で参考として援用される) を参照のこと。増強された化学発光検出 (ECL) のためのキットは、市販されている。

#### 【0284】

本発明の抗体が、例えば、フローサイトメトリー検出、走査型レーザーサイトメトリー検出、または蛍光イムノアッセイのために使用される場合、これらは、発蛍光団で有用に標識され得る。本発明の抗体に有用に結合し得る広範な種々の発蛍光団標識が存在する。多くは、例えば、Molecular Probes, Inc., Eugene, OR, USAから入手可能である。

#### 【0285】

標識されたアビジン、ストレプトアビジン、カプタビジン (cap tavidin) またはニュートラビジン (neutravidin) を使用する二次検出のために、本発明の抗体は、ビオチンで有用に標識され得る。

#### 【0286】

本発明の抗体が、たとえば、ウェスタンブロット適用のために使用される場合、これらの抗体は、有用には、放射性同位体 (例えば、 $^{33}\text{P}$ 、 $^{32}\text{P}$ 、 $^{35}\text{S}$ 、 $^3\text{H}$  および  $^{125}\text{I}$ ) で標識され得る。別の例としては、本発明の抗体が放射免疫治療のために使用される場合、この標識は、有用なためには、 $^{228}\text{Th}$ 、 $^{227}\text{Ac}$ 、 $^{225}\text{Ac}$ 、 $^{223}\text{Ra}$ 、 $^{213}\text{Bi}$ 、 $^{212}\text{Pb}$ 、 $^{212}\text{Bi}$ 、 $^{211}\text{At}$ 、 $^{203}\text{Pb}$ 、 $^{194}\text{Os}$ 、 $^{188}\text{Re}$ 、 $^{186}\text{Re}$ 、 $^{153}\text{Sm}$ 、 $^{149}\text{Tb}$ 、 $^{131}\text{I}$ 、 $^{125}\text{I}$ 、 $^{111}\text{In}$ 、 $^{105}\text{Rh}$ 、 $^{99\text{m}}\text{Tc}$ 、 $^{97}\text{Ru}$ 、 $^{90}\text{Y}$ 、 $^{90}\text{Sr}$ 、 $^{88}\text{Y}$ 、 $^{72}\text{Se}$ 、 $^{67}\text{Cu}$  または  $^{47}\text{Sc}$  であり得る。別の例として、本発明の抗体が、インビボの診断用途のために使用される場合、これらの抗体は、MRI造影剤 (例えば、ガドリニウムジエチレントリアミン五酢酸 (DTPA)、Laufferら、Radiology 207(2):529-38 (1998)) への結合体化によって、または放射性同位体標識によって検出可能にされ得る。当業者に理解されるように、上記の任意の標識の使用は、言及された適用に限定されない。

#### 【0287】

本発明の抗体 (そのフラグメントおよび誘導体を含む) はまた、本発明のタンパク質を、提示および/または発現する細胞に対して、毒素の除去 (ablative) 作用を標的化するために、生物学的に有害な部分 (例えば、毒素) に結合体化され得る。一般に、このような免疫毒素中の抗体は、Pseudomonas 外毒素 A、ジフテリア毒素、志賀毒素 A、炭疽毒素致死因子、またはリシンに結合体化される。Hall (編)、Immunotoxin Methods and Protocols (Methods in Molecular Biology、166巻)、Humana Press (2000) (ISBN: 0896037754); およびFrankelら (編)、Clinical Applications of Immunotoxins、Springer-Verlag New York, Incorporated (1998) (ISBN: 3540640975) を参照のこと。これらの開示は、総説に関して、その全体が本明細書中で参考として援用される。

#### 【0288】

本発明の抗体は、有用には、基板に結合され得る。従って、本発明は、本発明の1以上のGP354タンパク質およびタンパク質フラグメント、本発明の単離されたポリヌクレオチドによってコードされる1以上のタンパク質およびタンパク質フラグメントに対して特異的に結合する抗体を提供する。すなわち、その結合は、基材に結合した本発明の1以上のタンパク質およびタンパク質フラグメント、または本発明の単離されたポリヌクレオチドによってコードされる1以上のタンパク質およびタンパク質フラグメントによって競合的に阻害され得る。基材は多孔性または非多孔性、平面または非平面であり得る。

#### 【0289】

例えば、本発明の抗体は、有用には、免疫親和性クロマトグラフィーの目的のために、濾過媒体（例えば、NH<sub>2</sub>S活性化SepharoseまたはCNBr-活性化Sepharose）に結合体化され得る。 10

#### 【0290】

本発明の抗体はまた、有用なためには、代表的にはビオチン-スト렙トアビジン相互作用によって、常磁性ミクロスフェアに結合体化され得、次いで、このミクロスフェアは、本発明のタンパク質を発現または提示する細胞の単離のために使用され得る。別の例として、本発明の抗体は、有用なためには、ELISAのためのマイクロタイタープレートの表面に結合され得る。

#### 【0291】

上記のように、本発明の抗体は、原核生物細胞および真核生物細胞において産生され得る。従って、本発明はまた、本発明の抗体を発現する細胞（本発明の抗体を発現するように組換え的に改変されたハイブリドーマ細胞、B細胞、形質細胞、および宿主細胞を含む）を提供する。 20

#### 【0292】

本発明はまた、本発明の1以上のGP354タンパク質およびタンパク質フラグメント、本発明の単離されたポリヌクレオチドによってコードされる1以上のタンパク質およびタンパク質フラグメントに対して特異的に結合するように進化したアプタマーを提供する。すなわち、その結合は、本発明の1以上のタンパク質およびタンパク質フラグメント、本発明の単離されたポリヌクレオチドによってコードされる1以上のタンパク質およびタンパク質フラグメントによって競合的に阻害され得る。

#### 【0293】

（薬学的組成物および治療方法）

GP354は、主に脾臓で、そして神経組織（例えば、CNS）でより少ない量で発現される、免疫グロブリン（Ig）スーパーファミリーの新規のメンバーである。GP354、および内在性（integral）細胞表面膜タンパク質は、細胞-細胞認識および接着反応を媒介する他のファミリーのメンバーにおいて既知の、その細胞外部分において、5つのシグネチャー（signature）Igドメインを有する。Igスーパーファミリーのメンバーとして、GP354は、脾臓、神経およびGP354が発現する潜在的な他の組織における細胞-細胞認識機能、結合機能および接着機能を媒介するために重要であるようである。 30

#### 【0294】

脾臓に豊富なタンパク質として、GP354は、脾臓において、特に組織発達の間、ならびに組織再生および／または組織治癒の間（例えば、脾臓の損傷、外傷もしくは変性状態の後）、不適切な細胞-細胞結合、接着およびシグナル伝達に関連する異常な状態、障害および／または疾患を処置するための適切な治療標的である。GP354は、免疫、自己免疫、および変性状態に関連する脾臓の細胞死を阻害するために有用であることもまた想定される。GP354の神経の形態が、CNS組織における、組織の再生および修復についての、ならびに変性および細胞死を抑制するための同様に適切な治療標的であることが想定される。 40

#### 【0295】

従って、本発明は、本発明の核酸、タンパク質および抗体、ならびにGP354活性の模 50



倣物、アゴニスト、アンタゴニストまたは調節剤を含む薬学的組成物を提供し、これは、G P 3 5 4 の誤った発現 (mis-expression) に関連する障害、状態または疾患の処置 (すなわち、寛解) のための薬学的薬剤として、または G P 3 5 4 関連分子および細胞認識経路に関与する他の成分の異常な発現もしくは活性を克服するために、投与され得る。G P 3 5 4 発現は、膵臓において比較的集中しているので、G P 3 5 4 の誤った発現が、膵臓障害もしくは膵臓疾患、および／または機能の膵臓発達における先天的な欠損に関連し得ることが予測される。

#### 【0296】

本発明の組成物の投与が有効であり得る膵臓の障害および疾患としては、急性膵炎 (しばしば、しかしいつもではなく、異常な膵臓の外分泌の機能 (例えば、血清、腹水および／もしくは胸膜液アミラーゼレベルの増大、または異常なリパーゼまたはトリプシノーゲンレベル) を示す) が挙げられる。膵臓炎症および壊死はまた、急性膵炎および慢性膵炎、ならびに外分泌の機能不全に関連する。種々の膵臓内分泌腫瘍が特徴付けられ、そして膵臓に影響する自己免疫障害もまた記載されている。膵臓の障害および疾患の診断および処置のより詳細な説明については、Harrison's PRINCIPLES OF INTERNAL MEDICINE, 第14版 (Anthony S. Fauciら、編)、McGraw-Hill Companies, Inc.、1998、Part Eleven, 第3節 (この開示は、その全体において参考として援用される) を参照のこと。

10

#### 【0297】

G P 3 5 4 発現はまた、膵臓において検出されるよりも低いレベルではあるが、神経 C N S 組織においても検出される。従って、G P 3 5 4 の誤った発現は、神経の機能不全、障害もしくは疾患、または C N S の異常な発達に関連し得る。本発明の組成物を用いて処置することによって回復され得る神経障害の例としては、限定されないが、アルツハイマー病、パーキンソン病、老年性痴呆、片頭痛、てんかん、神経炎、神経衰弱、ニューロパシー、および G P 3 5 4 媒介神経移動、神経変性 (例えば、G P 3 5 4 媒介自己免疫疾患 (例えば、多発性硬化症の特定の形態)) に関連する任意の他の疾患、ならびに神経腫瘍 (例えば、神経膠腫、神経膠星状芽細胞腫および神経膠星状細胞腫) が挙げられる。

20

#### 【0298】

本発明の組成物が有用であり得るいくつかの他の障害としては、内分泌およびホルモン問題 (例えば、糖尿病)、膵臓の疾患、癌 (特に、膵臓癌) などが挙げられる。このような疾患を有するか、またはこのような疾患を発症する危険性を有する個体を処置するための、G P 3 5 4 調節因子 (G P 3 5 4 アンチセンス試薬、G P 3 5 4 リガンドおよび抗 G P 3 5 4 抗体を含む) の使用は、本発明の局面である。

30

#### 【0299】

本発明の組成物は、薬学的に受容可能なキャリア中に、代表的に、約 0.1 ~ 90 重量% (例えば、1 ~ 20 % または 1 ~ 10 %) の本発明の治療剤を含む。経口投与のための組成物の固体処方物は、適切なキャリアまたは賦形剤 (例えば、コーンスターチ、ゼラチン、ラクトース、アカシア、スクロース、微結晶性セルロース、カオリン、マンニトール、リン酸二カルシウム (dicarcium phosphate)、炭酸カルシウム、塩化ナトリウムまたはアルギン酸) を含み得る。使用され得る崩壊剤としては、限定ではなく、微結晶性セルロース、コーンスターチ、グリコール酸ナトリウムデンプン、およびアルギン酸が挙げられる。使用され得る錠剤結合剤としては、アカシア、メチルセルロース、カルボキシメチルセルロースナトリウム、ポリビニルピロリドン (Povidone<sup>T M</sup>)、ヒドロキシプロピルメチルセルロース、スクロース、デンプンおよびエチルセルロースが挙げられる。使用され得る潤沢剤としては、ステアリン酸マグネシウム、ステアリン酸、シリコーン流体、タルク、蠟、油およびコロイド状シリカが挙げられる。

40

#### 【0300】

水または他の水性ビヒクル中に調製される経口投与のための組成物の液体処方物は、種々の懸濁剤 (例えば、メチルセルロース、アルギン酸塩、トラガカント、ペクチン、ケルギ

50

ン (k e l g i n)、カラゲナン、アカシア、ポリビニルピロリドンおよびポリビニルアルコール)を含み得る。液体処方物はまた、活性化合物、湿潤剤、甘味剤ならびに着色剤および矯味・矯臭剤と一緒に含む、溶液、乳濁液、シロップ、およびエリキシルを含み得る。種々の液体処方物および粉末処方物は、処置されるべき哺乳動物の肺への吸入のために、従来の方法によって調製され得る。

#### 【0301】

組成物の注射可能な処方物は、種々のキャリア(例えば、植物油、ジメチルアセトアミド、ジメチルホルムアミド、乳酸エチル、炭酸エチル、ミリスチン酸イソプロピル、エタノール、ポリオール(グリセロール、プロピレングリコール、液体ポリエチレングリコールなど)を含み得る。静脈内注射のために、この化合物の水溶性型が点滴法によって投与され得、それによって抗真菌剤および生理学的に受容可能な賦形剤を含む薬学的処方物が注入される。生理学的に受容可能な賦形剤としては、例えば、5%のデキストロース、0.9%の生理食塩水、リンゲル溶液または他の適切な賦形剤が挙げられ得る。筋内調製物(例えば、化合物の適切な可溶性塩形態の滅菌処方物)は、薬学的賦形剤(例えば、注射用蒸留水、0.9%の生理食塩水、または5%のグルコース溶液)中に溶解され得、そして投与され得る。化合物の適切な不溶性形態は、水性の基剤または薬学的に受容可能な油基剤(例えば、長鎖脂肪酸のエステル(例えば、オレイン酸エチル))中の懸濁物として、調製および投与され得る。

10

#### 【0302】

局所的な半固体軟膏処方物は、代表的に、薬学的なクリーム状基剤のようなキャリア中に、約1~20%、例えば、5~10%の濃度の活性成分を含む。局所的使用のための種々の処方物としては、活性成分ならびに種々の支持体およびビヒクルを含む、ドロップ、チンキ剤、ローション、クリーム、溶液および軟膏が挙げられる。各薬学的処方物中の治療剤の最適な割合は、処方自体ならびに特定の病理および関連の治療レジメンにおいて所望される治療的効果に従って変化する。

20

#### 【0303】

吸入および経皮的処方物もまた、容易に調製され得る。

#### 【0304】

薬学的処方法は、十分確立された分野であり、そして以下においてさらに記載される: G e n n a r o (編)、R e m i n g t o n : T h e S c i e n c e a n d P r a c t i c e o f P h a r m a c y、第20版、L i p p i n c o t t、W i l l i a m s & W i l k i n s (2000) (ISBN: 0683306472); A n s e lら、P h a r m a c e u t i c a l D o s a g e F o r m s a n d D r u g D e l i v e r y S y s t e m s、第7版、L i p p i n c o t t、W i l l i a m s & W i l k i n s P u b l i s h e r s (1999) (ISBN: 0683305727); および K i b b e (編)、H a n d b o o k o f P h a r m a c e u t i c a l E x c i p i e n t s、A m e r i c a n P h a r m a c e u t i c a l A s s o c i a t i o n、第3版、(2000) (ISBN: 091733096X) (これらの開示は、その全体が本明細書中で参考として援用される)。医薬分野の当業者に公知の従来の方法は、患者に薬学的処方物を投与するために使用され得る。

40

#### 【0305】

代表的には、この薬学的処方物は、この薬学的処方物を含む経皮的パッチを患者の皮膚に適用し、そしてこのパッチを患者の皮膚に接触させたままにする(一般に、パッチ当たり、1~5時間)ことによって、患者に投与される。他の経皮的経路の投与(例えば、局所的に適用されるクリーム、軟膏などの使用による)は、従来の技術を適用することによって使用され得る。これらの薬学的処方物はまた、標準的な技術を使用することによって、他の従来の経路(例えば、経腸的な、皮下の、肺内の、経粘膜的、腹腔内の、子宮内の、舌下の、鞘内の、または筋内の経路)を介して投与され得る。さらに、この薬学的処方物は、例えば、1ヶ月、3ヶ月または6ヶ月の蓄積投与可能な材料および方法、または生分解性材料および方法を使用することによって、注射可能な蓄積経路の投与を介して患者に

50

投与され得る。

【0306】

投与経路にかかわらず、治療用タンパク質または抗体薬剤は、代表的に、患者の体重1 kg当たり、0.01 mg～30 mgの毎日の投薬量（例えば、1 mg/kg～5 mg/kg）で投与される。この薬学的処方物は、所望の場合、所望される毎日の総用量を達成するために、1日当たり複数の用量で投与され得る。処置方法の有効性は、障害の既知の徴候または症状について患者をモニタリングすることによって評価され得る。

【0307】

本発明の薬学的組成物は、容器、パッケージまたはディスペンサー中に、単独でかまたは投与のための標識および手引書を有するキットの一部として、含まれ得る。

10

【0308】

（トランスジェニック動物および細胞）

別の局面において、本発明は、gp354アイソフォーム核酸を含むトランスジェニック細胞およびトランスジェニック非ヒト生物、ならびにヒトgp354遺伝子の内因性オルソログの破壊を標的化されたトランスジェニック細胞およびトランスジェニック非ヒト生物を提供する。これらの細胞は、胚性幹細胞または体細胞であり得る。これらのトランスジェニック非ヒト生物は、キメラ、非キメラヘテロ接合体、および非キメラホモ接合体であり得る。

【0309】

本発明の宿主細胞は、非ヒトトランスジェニック動物を作製するために使用され得る。例えば、いくつかの実施形態において、本発明の宿主細胞は、gp354ヌクレオチド配列が導入されている受精卵または胚性幹細胞である。このような宿主細胞は、そのゲノムに外因性gp354配列が導入されている非ヒトトランスジェニック動物を作製するために使用され得るか、または動物中の関連の内因性gp354配列を変更もしくは置換するために使用され得る。

20

【0310】

本明細書中で使用される場合、「トランスジェニック動物」は、1以上の動物細胞が導入遺伝子を含んでいる、非ヒト動物、好ましくは哺乳動物、より好ましくはウシ、ヤギ、ヒツジまたはげっ歯類（例えば、ラット、マウス）である。トランスジェニック動物の他の例としては、非ヒト霊長類、イヌ、ニワトリ、両生類などが挙げられる。

30

【0311】

本明細書中で使用される場合、「導入遺伝子」は、トランスジェニック動物が発生する細胞のゲノム中に組み込まれ、そして成熟動物のゲノム中に保持される外因性DNAであり、それによって、トランスジェニック動物の1以上の細胞型または組織においてコードされた遺伝子産物の発現を指示する。

【0312】

本明細書中で使用される場合、「相同組換え動物」は、内因性gp354遺伝子が、内因性遺伝子と、動物の発生前に動物の細胞（例えば、動物の胚性幹細胞）中に導入された外因性DNA分子との間の相同組換えによって変更されている、非ヒト動物、好ましくは哺乳動物、より好ましくはマウスである。

40

【0313】

本発明の非ヒトトランスジェニック動物は、gp354の機能および／または活性を研究するため、ならびにgp354活性の調節因子を同定および／または評価するために有用である。これらはまた、GP354タンパク質またはポリペプチドフラグメントを生成するための方法（すなわち、ここで、このタンパク質は、非ヒト哺乳動物の乳腺において生成される）において有用である。

【0314】

本発明のトランスジェニック動物は、例えば、微量注入、レトロウイルス感染によって、受精卵母細胞の雄性前核中にgp354コード核酸を導入し、そして偽妊娠雌性親動物中で卵母細胞を発生させることによって作製され得る。配列番号1、3、5、6、7、9、

50

または11のヒトgp354DNA配列を含むか有するポリヌクレオチドは、導入遺伝子として、非ヒト動物のゲノム中に導入され得る。あるいは、本発明の単離されたポリヌクレオチドに対するハイブリダイゼーションによって単離されるヒトgp354遺伝子の非ヒトホモログ（例えば、マウスgp354遺伝子）は、導入遺伝子として使用され得る。異種転写制御配列の配列、イントロン配列、ポリアデニル化シグナルなどもまた、導入遺伝子に作動可能に連結されて、レシピエント宿主動物において、導入遺伝子の発現の効率性を増大させ得るか、またはさもなければ導入遺伝子の発現を調節し得る（例えば、発生または組織特異的な様式で）。

#### 【0315】

胚操作および微量注入によって、トランスジェニック動物、特にマウスのような動物を作製するための方法は、当該分野で慣習的になっており、そして例えば、以下において記載される：米国特許第4,736,866号；同第4,870,009号；および同第4,873,191号；ならびにHogan 1986、MANIPULATING THE MOUSE EMBRYO、Cold Spring Harbor Laboratory Press、Cold Spring Harbor、N. Y.。類似の方法は、他のトランスジェニック動物の作製のために使用される。トランスジェニック創始動物は、この動物の組織または細胞においてゲノム中のgp354導入遺伝子の存在および／またはgp354 mRNAの発現に基づいて、同定され得る。次いで、トランスジェニック創始動物は、導入遺伝子を保有するさらなる動物を繁殖させるために使用され得る。さらに、gp354をコードする導入遺伝子を保有するトランスジェニック動物は、他の導入遺伝子を保有する他のトランスジェニック動物とさらに交配され得る。

#### 【0316】

相同組換え動物を作製するために、gp354遺伝子の少なくとも一部を含むベクターが調製され、このgp354遺伝子の少なくとも一部分には、欠失、付加または置換が導入されており、それによってgp354遺伝子を変更する（例えば、機能的に破壊する）。gp354遺伝子は、ヒト遺伝子（例えば、配列番号1、5、9、または11）であり得るが、より好ましくはヒトgp354遺伝子の非ヒトホモログである。例えば、配列番号1、5、9、または11のヒトgp354遺伝子のマウスホモログを使用して、マウスゲノム中で内因性gp354遺伝子を変更するために適切な相同組換えベクターを構築し得る。

#### 【0317】

いくつかの実施形態において、このベクターは、相同組換えの際に、内因性gp354遺伝子が機能的に破壊される（すなわち、もはや機能的なタンパク質をコードしない；「ノックアウト」ベクターとも呼ばれる）ように設計される。あるいは、このベクターは、相同組換えの際に、内因性gp354遺伝子に変異またはそうでなければ変更されているが、なお機能的タンパク質をコードする（例えば、上流の調節領域が変更されて、これによって内因性GP354タンパク質の発現を変更し得る）ように設計され得る。相同組換えベクターにおいて、gp354遺伝子の変更された部分は、ベクターによって保有される外因性gp354遺伝子と胚性幹細胞中の内因性gp354遺伝子との間で相同組換えが生じることを可能にするために、gp354遺伝子のさらなる核酸によってその5'末端および3'末端が隣接される。さらなる隣接gp354核酸は、内因性遺伝子との首尾よい相同組換えのために十分な長さである。代表的には、数千ベースの隣接DNA（5'末端および3'末端の両方で）が、ベクターに含まれる。例えば、相同組換えベクターの例示的な記載については、ThomasらCell 51:503（1987）を参照のこと。

#### 【0318】

このベクターは、（例えば、エレクトロポレーションによって）胚性幹細胞株に導入され、そして導入されたgp354遺伝子が内因性gp354遺伝子で相同的に組換えられた細胞が選択される（例えば、Liら、Cell 69:915（1992）を参照のこと）。次いで、選択された細胞は動物（例えば、マウス）の胚盤胞に注入されて、集合キメ

ラを形成する。例えば、Bradley 1987、TERATO CARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH、Robertson編、IRL、Oxford、113-152頁を参照のこと。次いで、キメラ胚を、適切な偽妊娠雌性親動物中に移植し、胚を出産間近にする。生殖細胞に相同的に組換えられたDNAを有する子孫を使用して、動物の全細胞が導入遺伝子の生殖系列伝達によって相同的に組換えられたDNAを含む動物を繁殖させ得る。

#### 【0319】

相同組換えベクターおよび相同組換え動物を構築するための方法は、以下にさらに記載される：Bradley Curr. Opin. Biotechnol. 2: 823-829 (1991)；PCT国際公開番号WO90/11354；WO91/01140；WO92/0968；およびWO93/04169。 10

#### 【0320】

本明細書中に記載される非ヒトトランスジェニック動物のクローンもまた、WilmutらNature 385: 810-813 (1997)に記載の方法に従って作製し得る。簡潔には、トランスジェニック動物由来の細胞（例えば、体細胞）を、単離し得、そして増殖周期を抜け出しG0期に入るように誘導し得る。次いで、この静止期の細胞を、例えば、電気パルスの使用によって、静止期の細胞が単離されたのと同種の動物由来の除核した卵母細胞に融合させ得る。次いで、再構築された卵細胞を、桑実胚または未分化胚芽細胞に発生するように培養し、次いで、偽妊娠雌性親動物に移す。この雌性親動物の生まれた子孫は、細胞（例えば、体細胞）が単離された動物のクローンである。 20

#### 【0321】

インビボでの導入遺伝子の制御された発現は、制御可能な組換え系（例えば、cre/loxPレコンビナーゼ系（例えば、Laksora、Proc. Natl. Acad. Sci. USA 89: 6232-6236 (1992)を参照のこと）およびFLP組換え系（O'Gormanら、Science 251: 1351-1355 (1991)））を使用して達成され得る。cre/loxPレコンビナーゼ系が、導入遺伝子の発現を制御するために使用される場合、Creレコンビナーゼおよび選択されたタンパク質の両方をコードする導入遺伝子を含む動物が必要とされる。系の両方のエレメントを含むトランスジェニック動物は、例えば、2匹のトランスジェニック動物を交配させることによって得られ、各々は、選択されたタンパク質をコードする導入遺伝子またはレコンビナーゼをコードする導入遺伝子のいずれかを含む。 30

#### 【0322】

（アンチセンス試薬および方法）

（A. アンチセンス）

本発明の単離されたポリヌクレオチドの多くは、gp354ポリヌクレオチドを認識し、そのgp354ポリヌクレオチドにハイブリダイズするアンチセンスポリヌクレオチドである。全長およびフラグメントのアンチセンスポリヌクレオチドが提供される。本発明のフラグメントアンチセンス分子は、(i) gp354 RNAを特異的に認識し、そしてそのgp354 RNAに特異的にハイブリダイズする分子（GP354をコードするDNAと、他の公知の分子をコードするDNAとの配列比較によって決定される）を含む。 40  
GP354をコードするポリヌクレオチドに独特な配列の同定は、任意の公に利用可能な配列データベースの使用によって、そして／または市販の配列比較プログラムの使用によって推定され得る。所望の配列の同定の後に、制限消化による単離または当該分野において周知の種々のポリメラーゼ連鎖反応技術のいずれかを使用する増幅が実施され得る。アンチセンスポリヌクレオチドは、gp354 mRNAを発現する細胞によってGP354の発現を調節することに特に関する。

#### 【0323】

GP354をコードするヌクレオチド配列由来の配列番号1、3、5、6、7、9または11に記載されるヌクレオチド配列のアンチセンスオリゴヌクレオチド、またはフラグメント、あるいは、それらに相補的であるかまたは相同な配列は、種々の組織における遺伝 50

子発現を探索するための診断ツールとして有用である。例えば、組織は、この酵素のネイティブな発現またはそれに関連する病理学的な状態を調査するために、従来のオートラジオグラフィ技術によって検出可能な基を有するオリゴヌクレオチドプローブを用いてインサイチュで探索され得る。特定の局面において、少なくとも約10ヌクレオチド、約25ヌクレオチド、約50ヌクレオチド、約100ヌクレオチド、約250ヌクレオチド、もしくは約500ヌクレオチド、またはg p 3 5 4コード鎖全体あるいはそれらの一部のみに相補的な配列を含むアンチセンス核酸分子が提供される。配列番号2、4、8、10、または12のG P 3 5 4タンパク質のフラグメント、ホモログ、誘導体およびアナログをコードする核酸分子、配列番号1、3、5、6、7、9または11のG P 3 5 4核酸配列に相補的なアンチセンス核酸が、さらに提供される。

10

#### 【0324】

本発明のアンチセンス核酸分子は、G P 3 5 4をコードするヌクレオチド配列のコード鎖の「コード領域」または非コード領域に対するアンチセンスであり得る。用語「コード領域」とは、アミノ酸残基に翻訳されるコドンを含むヌクレオチド配列の領域をいう（例えば、ヒトG P 3 5 4のタンパク質コード領域は、配列番号1、7または11に示されるコード領域に対応する）。

#### 【0325】

アンチセンスオリゴヌクレオチドは、好ましくは、配列番号1、7または11のヌクレオチド配列の調節領域、あるいはそれらに対応するmRNA（開始コドン、T A T Aボックス、エンハンサー配列などを含むが、これらに限定されない）に関する。アンチセンス核酸分子は、g p 3 5 4のコード領域全体または非コード領域全体に相補的であり得るが、より好ましくは、g p 3 5 4 mRNAのコード領域または非コード領域の一部のみに対するアンチセンスであるオリゴヌクレオチドである。例えば、アンチセンスオリゴヌクレオチドは、g p 3 5 4 mRNAの翻訳開始部位の周りの領域に相補的であり得る。アンチセンスオリゴヌクレオチドは、例えば、約5ヌクレオチド長、約10ヌクレオチド長、約15ヌクレオチド長、約20ヌクレオチド長、約25ヌクレオチド長、約30ヌクレオチド長、約35ヌクレオチド長、約40ヌクレオチド長、約45ヌクレオチド長または約50ヌクレオチド長であり得る。

20

#### 【0326】

本発明のアンチセンス核酸は、当該分野で公知の手順を使用して、化学合成または酵素連結反応を使用して構築され得る。例えば、アンチセンス核酸（例えば、アンチセンスオリゴヌクレオチド）は、天然に存在するヌクレオチドまたは種々の改変されたヌクレオチド（分子の生物学的安定性を増加させるため、またはアンチセンス核酸とセンス核酸との間に形成される二重鎖の物理的安定性を増加させるために設計される（例えば、ホスホロチオエート誘導体およびアクリジン置換ヌクレオチドが使用され得る））を使用して化学的に合成され得る。

30

#### 【0327】

あるいは、アンチセンス核酸は、核酸がアンチセンス方向にサブクローニングされた発現ベクターを使用して生物学的に産生され得る（すなわち、挿入された核酸から転写されるRNAは、以下のサブセクションにおいてさらに記載されるように、目的の標的核酸に対してアンチセンス方向である）。

40

#### 【0328】

本発明のアンチセンス核酸分子（好ましくは、10～20ヌクレオチドの長さのオリゴヌクレオチド）は、代表的に、G P 3 5 4タンパク質をコードする細胞mRNAおよび／またはゲノムDNAにハイブリダイズするかまたは結合して、それによって、例えば、転写および／または翻訳を阻害することによって、タンパク質の発現を阻害するように、被験体に投与されるかまたはインサイチュで作製される。転写レベルまたは翻訳レベルのいずれかでのg p 3 5 4発現の抑制は、異常なg p 3 5 4発現によって特徴付けられる疾患／状態に関する細胞モデルまたは動物モデルを作製するために有用である。

#### 【0329】

50

ハイブリダイゼーションは、安定な二重鎖を形成するための従来のヌクレオチド相補性によって、または例えば、DNA二重鎖に結合するアンチセンス核酸分子の場合、二重鎖らせんの主溝における特異的な相互作用によってであり得る。ホスホロチオエートおよびメチルホスホネートアンチセンスオリゴヌクレオチドは、本発明による治療的使用に特に意図される。アンチセンスオリゴヌクレオチドは、これらの5'末端に、ポリーLーリジン部分、トランスフェリン部分、ポリリジン部分、またはコレステロール部分を付加することによってさらに改変され得る。

#### 【0330】

本発明のアンチセンス核酸分子の投与の経路の例としては、組織部位での直接注入が挙げられる。あるいは、アンチセンス核酸分子は、選択された細胞を標的化するために改変され得、次いで、全身的に投与され得る。例えば、全身投与のために、アンチセンス分子は、例えば、アンチセンス核酸分子を、細胞表面レセプターまたは抗原に結合するペプチドまたは抗体に連結することによって、選択された細胞表面に発現されるレセプターまたは抗原に特異的に結合するように改変され得る。アンチセンス核酸分子はまた、本明細書中に記載されるベクターを使用して細胞に送達され得る。アンチセンス分子の十分な細胞内濃度を達成するために、アンチセンス核酸分子が強力なpol IIプロモーターまたはpol IIIプロモーターの制御下に配置されるベクター構築物が好ましい。

#### 【0331】

なお他の実施形態において、本発明のアンチセンス核酸分子は、aーアノマー核酸分子である。aーアノマー核酸分子は、通常のbーユニットとは異なり、鎖が互いに平行に並ぶ(run parallel)相補的なRNAを有する特定の二本鎖ハイブリッドを形成する(Gaultierら、Nuc l. Acids Res 15:6625-6641 (1987))。アンチセンス核酸分子はまた、2'ーOーメチルリボヌクレオチド(Inoueら、Nuc l. Acids Res 15:6131-6148 (1987))またはキメラRNA-DNAアナログ(Inoueら、FEBS Lett 215:327-330 (1987))を含み得る。

#### 【0332】

(B. リボザイムおよび触媒核酸)

なお別の一連の実施形態において、本発明のアンチセンス核酸は、gp354特異的リボザイムの一部(または、改変として、「ヌクレオザイム」)である。リボザイムは、相補的な領域を有する、一本鎖核酸(例えば、mRNA)を切断し得るリボヌクレアーゼ活性を有する触媒性RNA分子である。従って、リボザイム(例えば、ハンマーヘッド、ヘアピン、グループIイントロンリボザイムなど)は、gp354 mRNA転写物を触媒的に切断するために使用されて、それによってgp354 mRNAの翻訳を阻害し得る。gp354コード核酸に対する特異性を有するリボザイムは、本明細書中に開示されるgp354ポリヌクレオチドのヌクレオチド配列(配列番号1、3、5、6、7、9または11)に基づいて設計され得る。例えば、米国特許第5,116,742号;同第5,334,711号;同第5,652,094号;および同第6,204,027号(それらの全体が本明細書中において参考として援用される)を参照のこと。

#### 【0333】

例えば、活性部位のヌクレオチド配列がGP354コードmRNAにおいて切断されるヌクレオチド配列に相補的である、テトラヒメナL-19 IVS RNAの誘導体が、構築され得る。例えば、Cechら、米国特許第4,987,071号;およびCechら、米国特許第5,116,742号を参照のこと。あるいは、gp354 mRNAは、RNA分子のプールから特異的なリボヌクレアーゼ活性を有する触媒性RNAを選択するために使用され得る。例えば、Bartelら、Science 261:1411-1418 (1993)を参照のこと。

#### 【0334】

gp354遺伝子の発現は、gp354の調節領域(例えば、gp354プロモーターおよび/またはエンハンサー)に相補的なヌクレオチド配列を標的化して、標的細胞におけ

10

20

30

40

50

る g p 3 5 4 遺伝子の転写を妨げる三重らせん構造を形成することによって阻害され得る。一般的に、Helene、Anticancer Drug Des. 6: 569-84 (1991); Heleneら、Ann. N. Y. Acad. Sci. 660: 27-36 (1992); および Maher、Bioassays 14: 807-15 (1992) を参照のこと。

#### 【0335】

(C. ペプチド核酸 (PNA))

他の好ましいオリゴヌクレオチド模倣物 (特に、インビボ投与に有用である) において、糖およびヌクレオシド間連結の両方は、新規な基で置換される (例えば、ペプチド核酸 (PNA))。例えば、Hyrupら、Bioorg. Med. Chem. Lett. 4: 5-23 (1996) を参照のこと。PNA化合物において、核酸のホスホジエステル骨格は、特に、アミド結合によって連結されるN-(2-アミノエチル) グリシン単位を繰り返すことによって、アミド含有骨格で置換される。核酸塩基は、代表的には、メチレンカルボニル連結によって、骨格のアミド部分のアザ窒素原子に直接的または間接的に結合される。PNAオリゴマーの合成は、Hyrupら、上記; およびPerry-O'Keefeら、Proc. Natl. Acad. Sci. USA 93: 14670-675 (1996) に記載されるように、標準的な固相ペプチド合成プロトコルを使用して実施され得る。

10

#### 【0336】

g p 3 5 4 の PNA は、治療的適用および診断的適用において使用され得る。例えば、PNA は、例えば、転写停止または翻訳停止を誘導することによってかあるいは複製を阻害することによって、遺伝子発現の配列特異的調節のためのアンチセンスまたは抗遺伝子剤として使用され得る。g p 3 5 4 の PNA はまた、例えば、PNA 指向 PCR クランピングによって、例えば遺伝子における単一塩基対変異の分析において; 他の酵素 (例えば、S1 ヌクレアーゼ) との組み合わせにおいて使用される場合に、人工制限酵素として; または DNA 配列およびハイブリダイゼーションに対するプローブもしくはプライマーとして使用され得る (Hyrupら、上記; および Perry-O'Keefe、上記)。

20

#### 【0337】

他の実施形態において、g p 3 5 4 の PNA は、例えば、それらの安定性または細胞取り込みを増強するために、親油性または他のヘルパー基を PNA に取り付けることによって、PNA-DNA キメラの形成によって、あるいはリポソームまたは当該分野で公知の薬物送達の他の技術の使用によって、改変され得る。例えば、g p 3 5 4 の PAN-DNA キメラが作製され得、これは、PNA および DNA の有利な性質を組み合わせ得る。このようなキメラによって、DNA 認識酵素 (例えば、RNase H および DNA ポリメラーゼ) が、DNA 部分と相互作用し得、一方、PNA 部分は、高い結合親和性および特異性を提供する。PNA-DNA キメラは、塩基スタッキング (stacking)、核酸塩基間の結合の数、および方向 (Hyrup、上記) の点で、選択された適切な長さのリンカーを使用して、連結され得る。PNA-DNA キメラの合成は、Hyrup、上記および Finnら、Nucleic Acids Res. 24: 3357-63 (1996) に記載されるように実施され得る。

30

40

#### 【0338】

例えば、DNA 鎖は、標準的なホスホルアミダイト結合化学を使用して固体支持体上で合成され得、そして改変ヌクレオシドアナログ (例えば、5'-(4-メトキシトリチル) アミノ-5'-デオキシチミジンホスホルアミダイト) は、PNA と DNA の 5' 末端との間に使用され得る (Magら、Nucleic Acids Res. 17: 5973-88 (1989))。次いで、PNA モノマーを、逐次様式で結合して、5' PNA セグメントおよび 3' DNA セグメントを有するキメラ分子を作製する (Finnら、上記)。あるいは、キメラ分子は、5' DNA セグメントおよび 3' PNA セグメントを用いて合成され得る。Petersenら、Bioorg. Med. Chem. Lett. 5: 1119-11124 (1975) を参照のこと。

50



## 【0339】

他の実施形態において、オリゴヌクレオチドは、ペプチドのような他の付属基（例えば、インビボで宿主細胞レセプターを標的化するため）、あるいは細胞膜（例えば、L e t s i n g e r ら、P r o c . N a t l . A c a d . S c i . U S A 86:6553-6556 (1989) ; L e m a i t r e ら、P r o c . N a t l . A c a d . S c i . U S A 84:648-652 (1987) ; P C T 公開番号 W O 88 / 09810 を参照のこと）または血液-脳関門（例えば、P C T 公開番号 W O 89 / 10134 を参照のこと）を横切る輸送を容易にする薬剤を含み得る。さらに、オリゴヌクレオチドは、ハイブリダイゼーション誘引 ( t r i g g e r e d ) 切断剤（例えば、K r o l ら、B i o T e c h n i q u e s 6:958-976 (1988) を参照のこと）、またはインターカレート剤（例えば、Z o n、P h a r m . R e s . 5:539-549 (1988) を参照のこと）を用いて改変され得る。この目的のために、オリゴヌクレオチドは、別の分子（例えば、ペプチド、ハイブリダイゼーション誘引架橋剤、輸送因子、ハイブリダイゼーション誘引切断剤など）に結合され得る。

## 【0340】

PNA 化学および適用は、特に、R a y ら、F A S E B J . 14 (9) : 1041-60 (2000) ; N i e l s e n ら、P h a r m a c o l T o x i c o l . 86 (1) : 3-7 (2000) ; L a r s e n ら、B i o c h i m B i o p h y s A c t a 1489 (1) : 159-66 (1999) ; N i e l s e n , C u r r . O p i n . S t r u c t . B i o l . 9 (3) : 353-7 (1999)、および N i e l s e n , C u r r . O p i n . B i o t e c h n o l . 10 (1) : 71-5 (1999)（これらの開示は、その全体が本明細書中において参考として援用される）において概説される。

## 【0341】

（診断方法）

（A. 核酸診断）

上記のように、本発明の単離されたポリヌクレオチドは、核酸プローブとして使用され、このような異常な組織の誤発現が疑われる場合、g p 354 m R N A が通常発現される組織（例えば、膵臓および C N S）、およびこれが通常発現されない組織、において g p 354 m R N A のレベルを評価し得る。

## 【0342】

次いで、本発明は、生物学的サンプル（例えば、患者由来の細胞抽出物、流体または組織サンプル）における g p 354 ポリヌクレオチドの存在を、このサンプルと本発明の単離されたポリヌクレオチドとを接触させることによって検出するための方法を提供し、この本発明のポリヌクレオチドは、g p 354 ポリヌクレオチド配列とハイブリダイズすることによって特異的に検出され得る。

## 【0343】

好ましくは、本方法は、高度にストリンジェントなハイブリダイゼーション条件下で、サンプルと単離された核酸と接触させる工程、およびサンプル中の、核酸に対するこの単離されたポリヌクレオチドのハイブリダイゼーションを検出する工程、を包含し、ここで、このハイブリダイゼーションの発生は、サンプルにおける g p 354 コード配列の存在を示す。

## 【0344】

本発明の単離されたポリヌクレオチドは、核酸プローブとして使用され得、これらの核酸プローブは、膵臓および中枢神経系における g p 354 の特異的な発現に基づいて、これらの組織における特定の細胞型に対して特異的である。従って、本発明は、生物学的サンプル（例えば、患者に由来する流体サンプルまたは組織サンプル）における g p 354 ポリヌクレオチドの存在を、このサンプルを本発明の単離されたポリヌクレオチドと接触させることによって検出することによって、細胞を膵臓細胞または神経細胞として同定するための方法を提供し、この本発明のポリヌクレオチドは、g p 354 ポリヌクレオチド配列とハイブリダイズすることによって特異的に検出され得る。

## 【0345】

本発明はまた、以下：（i）G P 3 5 4 タンパク質をコードする遺伝子の異常な改変または変異；（i i）G P 3 5 4 タンパク質をコードする遺伝子の誤調節；および（i i i）G P 3 5 4 タンパク質の異常な翻訳後修飾、の少なくとも1つによって特徴付けられる遺伝的損傷または変異の存在または非存在を同定するための診断アッセイを提供し、ここで、遺伝子の野生型は、G P 3 5 4 生物学的活性を有するタンパク質をコードする。

## 【0346】

本発明はさらに、ヒト g p 3 5 4 遺伝子のホモログを同定する方法を提供し、この方法は、中間または高ストリンジェンシーなハイブリダイゼーション条件下で、核酸ライブラリーを、配列番号1、3、5、6、7、9または11、あるいは少なくとも17ヌクレオチドを有するその一部を含む核酸プローブとハイブリダイズさせる工程；およびこの核酸プローブがライブラリー中の核酸配列とハイブリダイズするかどうかを決定する工程を包含する。ライブラリー中の核酸配列が、このような選択された条件下でハイブリダイズする場合、これは、ヒト g p 3 5 4 遺伝子のホモログである。

10

## 【0347】

（B．抗体診断）

本発明の抗体は、このような異常な組織誤発現が疑われる場合、G P 3 5 4 タンパク質が通常発現される組織（例えば、膵臓およびC N S）およびそれが通常発現されない組織における、G P 3 5 4 タンパク質の発現レベルを、評価するために使用され得る。

## 【0348】

次いで、本発明は、生物学的サンプル（例えば、患者に由来する細胞抽出物、流体または組織サンプル）におけるG P 3 5 4 タンパク質の存在またはその活性を、このサンプルを、G P 3 5 4 タンパク質の存在またはその活性の指標を検出し得る薬剤と接触させることによって検出する方法を提供する。好ましくは、この薬剤は、G P 3 5 4 タンパク質の少なくとも1つのエピトープに対して特異的な抗体である。

20

## 【0349】

従って、本発明は、サンプル中にG P 3 5 4 タンパク質が存在するかどうかを決定するための方法を提供し、この方法は、このサンプルを少なくとも1つのG P 3 5 4 エピトープを有する抗体と接触させる工程、およびこの抗体の抗原に対する特異的な結合を検出する工程を包含し、この結合は、サンプルにおけるG P 3 5 4 タンパク質の存在を示す。

30

## 【0350】

上記方法はまた、被験体における試験細胞を、被験体試験細胞由来の生物学的サンプル中（例えば、被験体に由来する細胞抽出物、流体または組織）に存在するG P 3 5 4 ポリペプチドの量を、非膵臓組織または非神経組織由来の比較生物学的サンプル中に存在するG P 3 5 4 ポリペプチドの量と比較することによって、膵臓細胞または神経細胞として同定するのに有用である。

## 【0351】

いくつかの実施形態において、本明細書中に記載される抗P a n C A M抗体は、初期I型糖尿病におけるランゲルハンス島損傷およびランゲルハンス島の機能の回復を意図する任意の治療を定量するために使用され得る。例えば、抗体は、M R I 適用およびC T 適用のための対照マーカー（例えば、ガドリニウム）または核医学適用のための放射性マーカー（例えば、テクネチウム）に結合される。次いで、関連する診断モダリティが、ランゲルハンス島の大きさおよび／または数の変化を定量し、詳細に記録するために使用される。抗体は、このように標識され、このような診断モダリティを使用するランゲルハンス島細胞腫瘍もまた局在化し得る。

40

## 【0352】

（疾患を診断するための方法）

本発明のg p 3 5 4 単離ポリヌクレオチド、タンパク質およびG P 3 5 4 特異的な抗体は、異常なg p 3 5 4 発現に関連する種々の障害および疾患状態を診断するための方法において有用である。

50

## 【0353】

本発明は、次いで、被験体において疾患状態を診断するための方法を提供し、この方法は、被験体由来の組織サンプル中のG P 3 5 4タンパク質の量または活性を、コントロールサンプル中のG P 3 5 4ポリペプチドの量または活性（健康な被験体由来のものと等価）と比較する工程を包含し、コントロールサンプル中のG P 3 5 4ポリペプチドの量または活性に対する、組織サンプル中のG P 3 5 4ポリペプチドの量または活性の有意な差異は、被験体が疾患状態を有することを示す。

## 【0354】

好ましい実施形態において、組織サンプル中のG P 3 5 4タンパク質の量または活性は、本発明のG P 3 5 4ポリペプチドまたはフラグメントを使用する競合的結合アッセイによって、あるいは、本発明のG P 3 5 4特異的抗体を使用する免疫アッセイによって評価される。好ましくは、本方法は、脾臓または神経系に関連する疾患状態を診断するために使用される。

10

## 【0355】

異なる組織における相対的g p 3 5 4 mRNAレベルをモニタリングすることによって、被験体における疾患状態を診断するための方法もまた提供される。好ましくは、本方法は、被験体に由来する試験組織サンプル中のg p 3 5 4 mRNAの量を、コントロールサンプル中のg p 3 5 4 mRNAの量と比較する工程を包含し、ここで、コントロールサンプル中の量に対する試験サンプル中のmRNAの量の有意な差異は、被験体が疾患状態を有することを示す。

20

## 【0356】

好ましい実施形態において、組織サンプル中のg p 3 5 4 mRNAの量は、本発明の単離されたg p 3 5 4ポリヌクレオチドまたは核酸フラグメントを使用するハイブリダイゼーションによって評価される。好ましくは、本方法は、脾臓または神経系に関連する疾患状態を診断するために使用される。

## 【0357】

（コンピューター読み取り可能手段）

本発明のさらなる局面は、本発明のg p 3 5 4核酸配列およびアミノ酸配列を保存するためのコンピューター読み取り可能手段である。好ましい実施形態において、本発明は、配列の完全なセットとしてまたは任意の組み合わせにおいて、本明細書中に記載されるような配列番号を保存するためのコンピューター読み取り可能手段を提供する。コンピューター読み取り可能手段の記録は、読み取りおよび表示ならびにプログラムの適用のためのコンピューターシステムとの接続のためにアクセスされ、ここで、このコンピューターシステムは、特定の基準に合うデータについてのクエリに対するデータの位置付け、配列の比較、一連の基準に合う配列のアライメントまたは順序付けなどを可能にする。

30

## 【0358】

本発明の核酸配列およびアミノ酸配列は、検索分析、有用なデータベース、および配列分析アルゴリズムにおいてにおける成分として特に有用である。これらの実施形態において使用される場合、用語「本発明の核酸配列」および「本発明のアミノ酸配列」は、コンピューター読み取り可能な形態に換算（reduce）または保存されるか、あるいは換算または保存され得る、本発明のポリヌクレオチドまたはポリペプチドの任意の検出可能な化学的特徴または物理的特徴を意味する。これらとしては、限定しないが、クロマトグラフィー走査データまたはピークデータ、それら由来の写真データまたは走査データ、および質量分析データが挙げられる。

40

## 【0359】

本発明は、本発明の配列を保存するコンピューター読み取り可能な媒体を提供する。コンピューター読み取り可能な媒体は、以下の1つ以上を含み得る：本発明の核酸配列の配列を含む核酸配列；本発明のアミノ酸配列を含むアミノ酸配列；配列の少なくとも1つが本発明の核酸配列の配列を含む、1セットの核酸配列；配列の少なくとも1つが、本発明のアミノ酸配列の配列を含む、1セットのアミノ酸配列；本発明の1つ以上の核酸配列の配

50

列を含む核酸配列を表すデータセット；本発明のアミノ配列の配列を含むアミノ配列をコードする核酸配列を表すデータセット；少なくとも1つの配列が、本発明の核酸配列の配列を含む、1セットの核酸配列；少なくとも1つの配列が、本発明のアミノ酸配列の配列を含む、1セットのアミノ酸配列；本発明の核酸配列の配列を含む核酸配列を表すデータセット；本発明のアミノ酸配列の配列を含むアミノ酸配列をコードする核酸配列を表すデータセット。コンピューター読み取り可能な媒体は、情報またはデータを保存するために使用される任意の組成の物体（例えば、市販のフロッピー（登録商標）ディスク、テープ、ハードドライブ、コンパクトディスク、およびビデオディスクを含む）であり得る。

#### 【0360】

従って、本発明は、ヒト g p 3 5 4 遺伝子のホモログを同定するための診断アッセイを提供し、このアッセイは、核酸データベースを、配列番号1、3、5、6、7、9または11からなる問い合わせ配列、または300以上のヌクレオチドを有するその一部を用いてスクリーニングする工程を包含し、ここで、配列番号1、3、5、6、7、9または11、またはその一部と少なくとも65%同一であるが100%未満同一なデータベース内の核酸配列が、存在する場合、ヒト g p 3 5 4 遺伝子のホモログである。 10

#### 【0361】

特徴配列（特に、本発明の遺伝子配列）の分析のための方法がまた、本発明によって提供される。配列分析の好ましい方法としては、例えば、配列相同性分析（例えば、同一性分析および類似性分析）、RNA構造分析、配列アセンブリ、分岐（c l a d i s t i c）分析、配列モチーフ分析、オープンリーディングフレーム決定、核酸塩基コーリング（c 20 a l l i n g）、および配列決定クロマトグラムピーク分析の方法が挙げられる。

#### 【0362】

コンピューターベースの方法を、核酸の相同性の同定を実施するために提供する。この方法は、本発明の核酸の配列を含む核酸配列をコンピューターで読み取り可能な媒体中に提供する工程；および相同性を同定するために、この核酸配列を少なくとも一つの核酸配列またはアミノ酸配列と比較する工程を包含する。

#### 【0363】

コンピューターベースの方法をまた、アミノ酸の相同性の同定を実施するために提供し、この方法は、本発明のポリペプチドの配列を含むアミノ酸配列をコンピューターで読み取り可能な媒体中に提供する工程；および相同性を同定するために、このアミノ酸配列を少 30 なくとも一つの核酸配列またはアミノ酸配列と比較する工程を包含する。

#### 【0364】

コンピューターベースの方法を、またさらに、オーバーラップする核酸配列の構築のために、単一の核酸配列に提供し、この方法は、本発明の核酸の配列を含む第1の核酸配列をコンピューターで読み取り可能な媒体中に提供する工程；および第1の核酸配列と第2の核酸配列との間の少なくとも1つのオーバーラップする領域をスクリーニングする工程を包含する。

#### 【実施例】

#### 【0365】

以下の実施例は、本発明の方法および材料を例証することを意味する。記載した条件およびパラメーターの適切な変更および追加は、分子生物学の分野において通常になされ、このことは、当業者に明らかであり、本発明の精神および範囲内である。 40

#### 【0366】

以下に記載される実験について、全てのR T - P C Rおよびフラグメントを、クローニングする前にゲル精製した。このフラグメントを、標準的な方法によって、アガロースゲル電気泳動によって分離した。DNAフラグメントを、アガロースゲルから切り出し、そして製造者の説明書（Q i a g e n、V a l e n c i a、C A）に従い、Q I A E X樹脂を使用してゲルから精製した。このゲル精製したフラグメントを、プラスミドベクターにクローニングし、次いで、このプラスミドを用いて、コンピテントT O P 1 0 E . c o l i 30 宿主細胞を形質転換した。この宿主細胞によって産生されたプラスミドを、標準的なア

ルカリ溶解ミニプレップ手順 (Qiagen、Valencia、CA) によって単離した。配列決定を、標準的なジデオキシ終結法 (Applied Biosystems、Foster City、CA) によって実施した。

#### 【0367】

(実施例1. インシリコ遺伝子予測および配列分析)

遺伝子予測ソフトウェアプログラム GENSCAN (Burgeら、J. Mol. Biol. 268:78-94 (1997)) ならびに GENEMARKHMM (Lukashinら、Nucleic Acids Res. 26:1107-1115 (1998)) を使用して、GenBankに寄託されたハイスループットゲノム配列において新規な遺伝子を同定した。そのように同定するために、Genbankデータ全体を、ローカルサーバーにダウンロードし、個々の配列のコンティグを、配列の入力に伴って提供されるアノテーションに従って分離した。この分析において使用されたパラメーターは、そのプログラムとともに含まれるデフォルトパラメーターであった (Burgeら、前出; および Lukashinら、前出)。

10

#### 【0368】

GENSCANおよびGENEMARKHMMから予測される少なくとも2つの同一のエキソンを有する遺伝子候補を、さらに分析した。具体的には、これらの遺伝子配列を、タンパク質配列に翻訳し、次に、これらのタンパク質配列を、Genpept、SwissprotおよびPIRを組み込む非重複タンパク質配列データベースのBLAST分析において問い合わせ配列として使用した。これらのタンパク質配列をまた、SIGNA 20 LP (Nielsenら、Int. J. Neural Syst. 8:581-99 (1997)) を使用して、シグナルペプチドについて分析した。シグナルペプチドおよびIgスーパーファミリータンパク質との類似性を有するタンパク質をコードする候補遺伝子を、実験的可視化およびクローニングに供した。

20

#### 【0369】

BLAST (「Basic Local Alignment Search Tool」) アルゴリズムは、配列の類似性を決定するのに適切である (Altschulら、J. Mol. Biol. 215:403-410 (1990))。BLAST分析を実施するためのソフトウェアは、ウェブサイトの the National Center for Biotechnology Informationを介して公共利用可能である。このアルゴリズムは、まず、問い合わせ配列中のW長の短いワードを同定することによってハイスコアリング配列ペア (high scoring sequence pair (HSP)) を最初に同定することを含み、これは、データベース配列中の同じ長さのワードでアライメントした場合、いくつかの正の値の閾値スコアTと一致するか、このスコアTを満足するかのいずれかである。Tは、隣接するワードのスコア閾値として言及される (Altschulら、前出)。これらの最初の隣接するワードのヒットは、サーチを開始するためのシードとして作用し、これらを含むHSPを見出す。ワードのヒットは、累積的アライメントスコアを増加させ得る限り、各配列に沿って両方向に伸展される。各方向でのワードのヒットの伸展は、以下の場合に停止される: (1) この累積的アライメントスコアが、最大の達成値から量Xに下落する場合; (2) この累積的スコアが 40、1以上の負のスコアを与える残基のアライメントの累積に起因して、0以下になる場合; または (3) いずれかの配列の末端に達する場合。BLASTアルゴリズムパラメーターW、TおよびXは、このアライメントの感度および速度を決定する。このBLASTプログラムは、11のワード長 (W)、50のBLOSUM62スコアマトリクス (Henikoffら、Proc. Natl. Acad. Sci. USA、89、10915-10919 (1992)) を参照のこと) アライメント (B)、10の期待値 (E)、M=5、N=4、および両方の鎖の比較をデフォルトとして使用する。

30

40

#### 【0370】

BLAST (Karlinら、Proc. Natl. Acad. Sci. USA、90、5873-5877 (1993)) およびGAPPED BLASTは、2つの配列間の

50

類似性の統計的な分析を行う。B L A S Tアルゴリズムによって提供される1つの類似性の測定は、最小の合計確率（P（N））であり、これは、2つのヌクレオチド配列間または2つのアミノ酸配列間での一致が偶然生じるかの確率の指標を提供する。例えば、テスト核酸をg p 3 5 4核酸と比較して最小の合計確率が、約1未満、好ましくは、約0.1未満、より好ましくは、約0.01未満、そして最も好ましくは、約0.001未満である場合、核酸がg p 3 5 4の遺伝子またはc D N Aと類似しているとみなされる。

#### 【0371】

このg p 3 4 5遺伝子（O R F）を、G e n B a n k登録番号A C 0 2 2 3 1 5（これは、2000年2月10日に寄託された）のB A Cのコンティグ38中において同定した。この遺伝子についてのG E N S C A N予測は、逆方向であり、そして表3に示される以下の14個のエキソンを含んだ。 10

#### 【0372】

#### 【表3】

表 3: GENSCAN 結果

| エキソン | 始点   | 終点   | 長さ  |
|------|------|------|-----|
| 14   | 1844 | 1779 | 66  |
| 13   | 3567 | 3464 | 104 |
| 12   | 4007 | 3903 | 105 |
| 11   | 4695 | 4476 | 220 |
| 10   | 4959 | 4859 | 101 |
| 09   | 5378 | 5246 | 133 |
| 08   | 5591 | 5464 | 128 |
| 07   | 5981 | 5833 | 149 |
| 06   | 6203 | 6098 | 106 |
| 05   | 7019 | 6869 | 151 |
| 04   | 7796 | 7636 | 161 |
| 03   | 8092 | 7943 | 150 |
| 02   | 9157 | 9008 | 150 |
| 01   | 9373 | 9322 | 52  |

20

30

公共利用可能なE S Tのデータベースに対するg p 3 5 4遺伝子のB L A S T分析は、予測された遺伝子と一致するE S Tを示さなかった。

#### 【0373】

（実施例2：g p 3 5 4の増幅）

g p 3 5 4 c D N Aの配列を、M A R A T H O N－R E A D Y R A C Eキット（C l o n t e c h、P a l o A l t o、C A）を使用して、c D N A末端の迅速な増幅（R A C E）を実施することによって得る。M A R A T H O N－R E A D Y c D N Aは、ヒトの組織のm R N Aから合成した二重鎖c D N Aであり、これを、標準的なアダプターのセット（C l o n t e c h）に連結されている。全てのR A C E反応は、このキットと共に提供されるアダプタープライマーA P－1、5'－C C A T C C T A A T A C G A C T C A C T A T A G G G C－3'（配列番号15）を使用する。g p 3 5 4についての3' R A C Eは、順方向プライマーG X 1－218、5'－A C T G G G G G C T A G T T C A G T G G A C T A A－3'（配列番号16）、または逆方向プライマーG X 1－219の相補体、5'－C C A A A C A G C A C A T C C A G C G C A G T A C－3'（配列番号17）と共にA P－1を使用し得る。g p 3 5 4についての5' R A C Eは、逆方向プライマーG X 1－219、または順方向プライマーG X 1－218の相補体と共にA P－1を使用し得る。A D V A N T A G E 2 D N Aポリメラーゼ（C l o n t e c h）を、 40 50

増幅反応のために使用し得る。「タッチダウン (touchdown)」PCR (Donら、Nuc. Acids Res. 19:4008 (1991)) 条件を、熱サイクルに使用したことを除いて、MARATHON-READYキットを、製造者の説明書に従って使用し得る。熱サイクル条件は、以下の通りである：94℃で1分間；94℃で15秒間、72℃で15秒間、68℃で15秒間の1サイクル；94℃で15秒間、71℃で15秒間、68℃で15秒間の1サイクル；94℃で15秒間、70℃で15秒間、68℃で15秒間の1サイクル；94℃で15秒間、69℃で15秒間、および68℃で15秒間の1サイクル；94℃で15秒間、68℃で30秒間の35サイクル；ならびに68℃で10分間。

#### 【0374】

10

(実施例3：RT-PCRによるGP354発現の確認)

インターエキソン-PCRを、この予測されたgp354遺伝子が実際に発現されることを確認するため、および(予測された遺伝子構造ではなく)本当の遺伝子構造を決定するクローニングプロセスを開始するために使用した。PCRのためのテンプレートは、多組織cDNAパネル中に含まれる一本鎖cDNA(逆転写PCR(「RT-PCR」))によって、複数のヒト組織から単離されたmRNAから生成される)を使用して、製造者の説明書(Clontech)に従って実施した。これらの組織は、脳、心臓、腎臓、肝臓、肺、膵臓、下垂体、骨格筋、結腸、卵巣、末梢血白血球、前立腺、小腸、脾臓、精巣、および胸腺であった。

#### 【0375】

20

インターエキソンPCRのためのプライマーを、全てのPCR実験が、スプライスされていないメッセージまたはcDNAパネル中のゲノムの混入物から、スプライスされたメッセージを区別するための一体型コントロールを有するように、別個のエキソンから選択した。一对のこのようなプライマーは、GX1-220およびGX1-221(実施例13、前出)であった。使用した別の対は、GX1-218およびGX1-219(前出)であった。多組織cDNAパネル(Clontech)と共に含まれる説明書に従って、PCRを実施した。PCRのための代表的な熱サイクラー条件は、以下の通りであった：94℃で1分間、続いて、94℃で20秒間、68℃で2分間の35サイクル、続いて、68℃で5分間(最終サイクル後)。PCR産物を、1.0% アガロースゲルで分離し、エチジウムブロミド染色によって可視化した。

30

#### 【0376】

これらの結果は、膵臓中に約785bpのフラグメントが存在し、他の組織には存在しないことを示した。同じ観察が、GX1-218およびGX1-219の場合得られた。このフラグメントを単離し、そしてPCR2.1プラスミドベクター(Invitrogen, Carlsbad, CA)にクローニングした。得られたプラスミド構築物CS0026を、TOP10 E. coli 宿主細胞中で増殖させた。標準的なアルカリ溶解ミニプレップ(miniprep)手順(Qiagen, Valencia, CA)によって、プラスミドを精製した。標準的なジデオキシ終結方法(Applied Biosystems, Foster City, CA)によって、配列決定を完了した。PCRフラグメントの配列(配列番号3)を、推定gp354遺伝子の配列と比較して、この遺伝子予測の精度を確認した。

40

#### 【0377】

(実施例4：RACEによる全長gp354cDNAの同定)

遺伝子予測プログラムGENSCANおよびGENEMARKは、予測可能なエラー比率を有するので(Burgeら、前出；Lukashinら、前出)、実施例3に記載されるPCRフラグメントを、RACE反応を介して、gp354cDNA配列の残りを得るための種配列として使用する。この3'RACE反応について、このプライマーは、GX1-218またはGX1-219の相補体であり、そしてこのテンプレートは、ヒトの膵臓組織由来のcDNAである(実施例3を参照のこと)。この5'RACEについて、このプライマーは、GX1-219またはGX1-218の相補体であり、そしてこのテン

50

プレートはまた、ヒトの膵臓組織由来の cDNA である。このように得られた 5' RACE フラグメントおよび 3' RACE フラグメントを、ゲル精製し、クローニングし、そして配列決定した。全長 gp354 cDNA 配列を構築するために、初期 PCR 産物、5' RACE 産物および 3' RACE 産物を、GCG コンピューターパッケージ中の ASSEMBLE プログラム (Genetics Computer Group, Madison, Wisconsin) を使用して、単一の連続配列中に構築する。

#### 【0378】

(実施例 5: ノーザンブロット分析による GP354 発現の確認)

GP354 発現を確認するために、ノーザンブロット分析を、2  $\mu$ g のポリ A RNA を含むブロット (Clontech カタログ番号 7760-1) の各レーンで実施した。このブロットに示される組織は、心臓、脳、胎盤、肺、肝臓、骨格筋、腎臓、および膵臓を含んだ。このノーザンブロットのためのプローブは、実施例 3 に記載される PCR フラグメント (配列番号 3) であった。50 ng のプローブを、Feinberg および Vogelstein のランダムプライム化方法によって標識した (Anal. Biochem. 132: 6-13 (1983))。ハイブリダイゼーションを、EXPRESSHYB 溶液 (Clontech カタログ番号 8015-1) 中において 68°C、1 時間で実施した。オートラジオグラフィの前に、ノーザンブロットを、室温にて 2 $\times$ SSC/0.05% SDS で洗浄し、続いて、50°C の 0.1 $\times$ SSC/0.1% SDS で 2 回洗浄した。

#### 【0379】

約 3.2 kb の RNA のハイブリダイゼーションが、膵臓において観察されたが、試験した他の組織のいずれにおいても、観察されなかった。

#### 【0380】

(実施例 6: ゲノムライブラリーの PCR スクリーニングおよび GP354 コード領域のサブクローニング)

gp354 ゲノム遺伝子座のサブクローニングを、ゲノムライブラリーから、またはゲノム DNA から直接的に、PCR によって達成し得る。例えば、2  $\mu$ l のヒトゲノムライブラリー (約 10<sup>8</sup> PFU/ml) (Clontech) を、6 ml の K802 細胞 (一晚培養) (Clontech) に添加し、次いで、250 ml のアリコートとして、各々の 24 本のマイクロチューブに分配した。このマイクロチューブを、15 分間、37°C でインキュベートする。7 ml のアガロース (0.8%) を、各チューブに添加し、混合し、次いで、LB 寒天 + 10 mM の MgSO<sub>4</sub> プレート上に注ぎ、そして 37°C で一晚インキュベートする。5 ml の SM フェージ緩衝液 (0.1 M NaCl、8.1 mM MgSO<sub>4</sub> · 7H<sub>2</sub>O、50 mM Tris · Cl (pH 7.5)、0.01% ゼラチン) を各プレート (5 ml) に添加し、そしてトップアガロースを、顕微鏡用のスライドを用いて取り出し、そして 50 ml の遠心管に入れる。一滴のクロロホルムを添加し、そしてこのチューブを、37°C の振盪器に 15 分間配置し、次いで、20 分間、4000 rpm で遠心分離 (Sorvall RT6000 卓上用遠心分離器) し、そしてこの上清を、ストック溶液として 4°C で保存した。

#### 【0381】

次いで、PCR を、以下を含む 20 ml 中で実施した: 8.8 ml の H<sub>2</sub>O、4 ml の 5 $\times$ RAPID-LOAD BUFFER (Origene)、2 ml の 10 $\times$ PCR BUFFER II (Perkin Elmer)、2 ml の 25 mM の MgCl<sub>2</sub>、0.8 ml の 10 mM の dNTP、配列番号 1 の gp354 ポリヌクレオチドの 5' 末端の配列の少なくとも一部分を含む、0.12 ml のプライマー (1 mg/ml)、配列番号 1 の gp354 ポリヌクレオチドの 3' 末端に対して相補的である配列の少なくとも一部分を含む、0.12 ml のプライマー (1 mg/ml)、0.2 ml の AMPLITAQ GOLD ポリメラーゼ (Perkin Elmer) および 24 本のチューブの各々からの 2 ml のフェージ溶液。この PCR 反応は、以下を含む: 80°C で 20 分間、95°C で 10 分間の 1 サイクル; 次いで、95°C で 30 秒間、72°C で 4 分間 (サイクル毎に 1°C 低下させる)、68°C で 2 分間の 22 サイクル、続いて、95°C で 30 秒間、55°C で 3

10

20

30

40

50



0秒間、68℃で60秒間での30サイクル。この反応物を、2%のアガロースゲル上に充填する。

#### 【0382】

正確なサイズのPCR産物を与えるチューブから、5μlを、5個の1:10の希釈物を設定するために使用し、これを、LB寒天+10mMのMgSO<sub>4</sub>プレート上に配置し、そして一晩インキュベートした。BA85ニトロセルロースフィルター(Schleicher & Schuell)を、1時間、各プレートの上部に配置する。フィルターを取り出し、ペトリ皿にファージ側を上にして配置し、そしてファージを溶出するために、15分間、4mlのSM緩衝液で覆う。1mlのSM緩衝液を、各プレートから取り除き、そして上記のようにPCR反応を設定するために使用する。PCR産物を得るために最も希釈したプレートを細分し、フィルターを載せ、そしてPCR反応を繰り返す。このプレートの一連の希釈および細分を、ポジティブなPCRバンドを与える単一のプラークが単離されるまで、続ける。一旦、単一のプラークが単離されると、10mlのファージ上清を、1プレート当たり100mlのSMおよび200mlのK802細胞に添加し、全部で8個のプレートを設定する。このプレートを一晩37℃でインキュベートする。8mlのSMを、各プレートに添加し、そしてトップアガロースを、顕微鏡用のスライドで擦り落とし、そして遠心管中に収集する。

10

#### 【0383】

3滴のクロロホルムを、遠心管に添加する。続いて、このチューブをボルテックスにかけ、15分間、37℃でインキュベートし、そして20分間、4000rpmで遠心分離(Sorvall RT6000卓上用遠心分離器)し、ファージを回収する。回収したファージを使用して、QIAGEN LAMBDA MIDI KITを使用してゲノムファージDNAを単離する。プライマーに関する配列は、本明細書中で示す配列に由来し得る。

20

#### 【0384】

gp354遺伝子のコード領域をサブクロニングするために、PCRを、以下を含む50μlの反応物中で実施する: 33μlのH<sub>2</sub>O、5μlの10×TT緩衝液(140mMの硫酸アンモニウム、0.1%のゼラチン、0.6MのTris-トリシン pH8.4)、5μlの15mM MgSO<sub>4</sub>、2μlの10mMのdNTP、4μlのゲノムファージDNA(0.1μg/ml)、配列番号1のgp354ポリヌクレオチドの最も5'側のコード配列の少なくとも一部分を含む、0.3μlのプライマー(1μl/ml)、配列番号1のgp354ポリヌクレオチドの最も3'側のコード配列の少なくとも一部分に対して相補的である配列を含む、0.3μlのプライマー(1μg/ml)、0.4μlのHIGH FIDELITY Taqポリメラーゼ(Boehringer Mannheim)。PCR反応を、94℃で2分間の1サイクルで開始し、続いて、94℃で30秒間、55℃で60秒間の15サイクル、そして68℃で2分間実施した。

30

#### 【0385】

PCR産物を、2%のアガロースゲル上に充填する。期待されるサイズのDNAバンドを、ゲルから切り出し、GENELUTE AGAROSEスピンカラム(Supelco)に配置し、そして10分間、最大速度で遠心した。溶出したDNAを、エタノール沈殿し、そしてライゲーションのために12μlのH<sub>2</sub>O中で再懸濁した。これらのPCRプライマー配列は、本明細書中で提供される配列に由来し得る。

40

#### 【0386】

このライゲーション反応は、TOPO TAクローニングキット(Invitrogen)からの溶液を使用する。この反応は、室温で5分間かけて、4μlのPCR産物および1μlのpCRII-TOPOベクターを含む溶液中で進める。この反応を、1μlの6×TOPOクローニング停止溶液の添加によって終結させる。次いで、このライゲーション産物を、氷上に配置する。2μlのライゲーション反応物を使用して、ONE-SHOT TOP10細胞(Invitrogen)を形質転換する。簡潔には、このライゲーション反応物を、この細胞と共に混合し、そして30分間、氷上に配置する。次いで、こ

50

の細胞に、30秒間、42℃の熱ショックを与え、そして2分間、氷上に配置する。次に、250μlのSOCを、細胞に添加し、これを、1時間、37℃で振盪させながらインキュベートし、次いで、アンピシリンプレート上にプレートする。

#### 【0387】

このプレートからの単一のコロニーを使用して、5mlのLB培地の培養物を播種する。プラスミドDNAを、CONCERT RAPID PLASMID MINIPREP SYSTEM (GibcoBRL) を使用して、この培養物から精製し、次いで、プラスミドDNAの挿入物を、配列決定する。

#### 【0388】

このgp354ゲノムファージDNAを、ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems) (これは、高度なキャピラリー電気泳動技術およびABI PRISM BIGDYE Terminator Cycle Sequencing Ready Reaction Kitを使用する) を使用して配列決定し得る。サイクルー配列決定反応系は、14mlのH<sub>2</sub>O、16mlのBIGDYE Terminator混合物、7mlのゲノムファージDNA (0.1mg/ml) および3mlのプライマー (25ng/ml) を含み得る。この反応を、Perkin-Elmer 9600熱サイクラーで、以下により実施する：95℃で5分間、続いて95℃で30秒間、55℃で20秒間および60℃で4分間の99サイクル。この産物を、CENTRIFLEXゲル濾過カートリッジを使用して精製し、減圧下で乾燥し、次いで16μlのTemplate Suppression Reagent (テンプレート抑制剤) (PE Applied Biosystems) 中に溶解する。これらのサンプルを、95℃で5分間加熱し、次いで、310 Genetic Analyzer中に配置する。

#### 【0389】

pCRIIにサブクロニングしたDNAを、ABI PRISM 310 Genetic Analyzer (前出) を使用して配列決定する。各々のサイクルー配列決定反応系は、6mlのH<sub>2</sub>O、8mlのBIGDYE Terminator混合物、5mlのミニプレップDNA (0.1mg/ml)、および1mlのプライマー (25ng/ml) を含み、そしてPerkin-Elmer 9600熱サイクラーで、96℃で10秒間、50℃で10秒間、および60℃で4分間の25サイクルにより実施する。この産物を、CENTRIFLEXゲル濾過カートリッジを使用して精製し、減圧下で乾燥し、そして次いで16μlのTemplate Suppression Reagent中に溶解する。これらのサンプルを、95℃で5分間加熱し、次いで、310 Genetic Analyzer中に配置する。

#### 【0390】

(実施例7：脳におけるGP354の発現を実証するためのハイブリダイゼーション分析)

哺乳動物 (例えば、ラット) におけるgp354の発現は、インサイチュハイブリダイゼーション組織化学によって研究され得る。脾臓におけるgp354発現を研究するために、例えば、Reichert-Jungクリオスタットを使用して、冠状および矢状のラットの脾臓凍結切片 (20μm厚) を調製する。個々の切片を、シラン処理したヌクレアーゼを含まないスライド (CEL Associates, Inc., Houston, TX) に解凍マウントし、そして-80℃で保存する。切片を、固定後に、4%の冷パラホルムアルデヒド中で始めて処理し、冷リン酸緩衝化生理食塩水 (PBS) 中でリンスし、トリエタノールアミン緩衝液中の無水酢酸を使用してアセチル化し、そして室温における70%、95%および100%のアルコールでの一連のアルコール洗浄によって脱水する。続いて、切片をクロロホルム中で脱脂し、続いて、室温で100%および95%のアルコールに連続的に暴露することによって再水和する。処理した凍結切片を含む顕微鏡用スライドを、ハイブリダイゼーションの前に風乾させる。他の組織を、同様の様式でアッセイし得る。

## 【0391】

g p 3 5 4 特異的プローブを、P C R および配列番号 1 または配列番号 3 からの配列情報を使用して作製し得る。P C R 増幅後、このフラグメントを制限酵素で消化し、そして同じ酵素を用いて切断した p B l u e s c r i p t I I にクローニングする。g p 3 5 4 のセンス鎖に特異的なプローブの生成のために、p B l u e s c r i p t I I にクローニングされたクローン化 g p 3 5 4 フラグメントを、適切な制限酵素で線状化し得、これはベクターが有する T 7 プロモーターおよび市販の T 7 R N A ポリメラーゼを使用して、標識されたランオフ (r u n - o f f) 転写物 (すなわち、c R N A リボプローブ) に対する基質を提供する。g p 3 5 4 のアンチセンス鎖に特異的なプローブをまた、組換えプラスミドを適切な制限酵素で切断し、T 3 プロモーターおよび同族ポリメラーゼを使用し得る。10

## 【0392】

リボプローブを、 $[^3\text{S}]$ -U T P で標識して、アンチセンスリボプローブについて、約  $0.40 \times 10^6$  c p m / p m o l の比活性、およびセンス鎖リボプローブについて、約  $0.65 \times 10^6$  c p m / p m o l の比活性を生じ得る。続いて、各リボプローブを変性し得、そして 50 % ホルムアミド、10 % デキストラン、0.3 M N a C l、10 m M T r i s (p H 8.0)、1 m M E D T A、1 × デンハルト溶液、および 10 m M ジチオトレイオールを含むハイブリダイゼーション緩衝液に添加し得る (2 p m o l / m l)。20

## 【0393】

臍臓連続凍結切片を含む顕微鏡用スライドを、スライド 1 枚あたり 45  $\mu$  l のハイブリダイゼーション溶液に独立してそれぞれ暴露し得、そしてハイブリダイゼーション溶液に暴露されている切片の上にシラン処理されたカバースリップを置き得る。切片を 52  $^{\circ}$ C で一晚 (例えば、15 ~ 18 時間) インキュベートし、ハイブリダイゼーションを起こす。次いで、同等の一連の凍結切片を、センスまたはアンチセンスの g p 3 5 4 特異的 c R N A リボプローブに暴露する。

## 【0394】

ハイブリダイゼーション期間の後、このカバーガラスを 1 × S S C 中でスライドから拭き取り、続いてこのスライドを、37  $^{\circ}$ C で 45 分間、10 m M T r i s · H C l (p H 7.4)、0.5 M E D T A および 0.5 M N a C l を含む緩衝液中 20  $\mu$  g / m l の R N a s e A に暴露することによって、R N a s e A 処理する。次いで、この凍結切片を、それぞれ 20 分間の、0.1 × S S C 中、52  $^{\circ}$ C の 3 回の高ストリンジェンシーの洗浄に供する。一連の洗浄の後、この凍結切片を、アルコール中 70 %、95 % および 100 % の酢酸アンモニウムに連続的に曝すことによって脱水し、続いて空気乾燥し、そして K O D A K B I O M A X M R - 1 フィルムに露光する。露光の 13 日後、このフィルムを現像し、そして任意の有意なハイブリダイゼーションシグナルを検出する。30

## 【0395】

これらの結果に基づいて、フィルムのオートラジオグラムにより示されるように、ハイブリダイズした組織を含むスライドを、K O D A K N T B - 2 核トラックエマルジョンでコーティングし、そしてこのスライドを暗室で 32 日間保存する。次いで、このスライドを現像し、そしてヘマトキシリンで対比染色する。エマルジョンでコーティングされた切片を、顕微鏡で分析して、標識の特異性を決定する。オートラジオグラフィイン (アンチセンスプローブのハイブリダイゼーションによって生成する) が、クレシルバイオレットで染色された細胞体と明らかに会合している場合、このシグナルを特異的であると決定する。細胞体間に見出されるオートラジオグラフィインは、プローブの非特異的結合を示す。40

## 【0396】

臍臓および脳における G P 3 5 4 の発現 (下記) は、G P 3 5 4 活性のモデュレータが、50

神経系におけるGP354の活性を阻害または増加することによって、特定の神経障害を処置するための有用性を有するという指標を提供する。

#### 【0397】

(実施例8：gp354-RNAのノーザンブロット分析)

ノーザンブロットハイブリダイゼーションを実施して、gp354 mRNAの発現を試験し得る。配列番号1、配列番号3またはそれらの相補体の配列の少なくとも一部を含むクローンを、プローブとして使用し得る。ベクター特異的プライマーを、PCRに使用して、<sup>32</sup>P標識のためのハイブリダイゼーションプローブフラグメントを生成する。PCRを、以下のように実施する：

(1) 以下の試薬を混合する：

1 μl gp354含有プラスミド

2 μl 順方向プライマー

2 μl 逆方向プライマー

10 μl Taqポリメラーゼのメーカー（例えば、Amersham Pharmacia Biotech）により提供される10×PCR緩衝液

1 μl 10mM dNTP（例えば、Boehringer Mannheimカタログ番号1969064）

0.5 μl Taqポリメラーゼ（例えば、Amersham Pharmacia Biotechカタログ番号27-0799-62）

83.5 μl 水。

#### 【0398】

(2) 以下のプログラムを使用して、サーモサイクラーでPCRを実施する：

94℃で5分；94℃で1分、55℃で1分および72℃で1分を30サイクル；次いで72℃で10分。

#### 【0399】

PCR産物を、QIAQUICK PCR Purification Kit（Qia genカタログ番号28104）を使用して精製し得る。この精製したPCRフラグメントを、「Ready-to-go DNA Labeling Beads」（Amersham Pharmacia Biotechカタログ番号27-9240-01）を使用するランダムプライミングによって、<sup>32</sup>P-dCTP（Amersham Pharmacia Biotechカタログ番号AA0005/250）で標識する。メーカーのプロトコルに従って、Clontechからのヒト多組織ノーザンブロットにおいてハイブリダイゼーションを行う。Molecular Dynamics PHOSPHORIMAGERスクリーン（カタログ番号MD146-814）で一晩露光した後、約1.35 kbのバンドを可視化する。

#### 【0400】

(実施例9：真核生物宿主細胞におけるGP354の組換え発現)

(A. 哺乳動物細胞におけるgp354の発現)

GP354タンパク質を生成するために、GP354コードポリヌクレオチドを、組換え技術を使用して発現させる。例えば、実施例1に記載のGP354コード配列を、市販の発現ベクターのpzeoSV2（Invitrogen）にサブクローニングする。得られた発現構築物を、トランスフェクション試薬のFUGENE6（Boehringer-Mannheim）および製品挿入物に提供されるトランスフェクションプロトコルを使用してチャイニーズハムスター卵巣（CHO）細胞にトランスフェクトする。他の真核生物細胞株（ヒト胚性腎臓（HEK 293）およびCOS細胞を含む）も同様に適切である。

#### 【0401】

GP354を安定に発現する細胞を、100 μg/mlのゼオシン（zeocin）（Stratagene, La Jolla, CA）の存在下における増殖によって選択する。必要に応じて、GP354を、標準的なクロマトグラフィー技術を使用して、細胞から精

10

20

30

40

50

製し得る。精製を容易にするために、G P 3 5 4 アミノ酸配列の部分に対応する 1 つ以上の合成ペプチド配列に対して抗血清を惹起し、そしてこの抗血清を使用して G P 3 5 4 を親和性精製する。この G P 3 5 4 タンパク質をまた、タグ配列（例えば、ポリヒスチジン、赤血球凝集素または F L A G）とインフレイムで発現して、精製を促進し得る。さらに、G P 3 5 4 ポリペプチドの多くの用途（例えば、以下に記載されるアッセイ）は、宿主細胞からの G P 3 5 4 の精製を必要としないことが理解される。

#### 【0402】

（B. 293 細胞における G P 3 5 4 の発現）

哺乳動物細胞 293（形質転換されたヒトまたは霊長類の胚性腎臓細胞）における G P 3 5 4 の発現のために、関連の g p 3 5 4 コード配列を有するプラスミドを、ベクター p S e c T a g 2 A（I n v i t r o g e n）を使用して調製する。ベクター p S e c T a g 2 A は、分泌のためのマウス I g K 鎖リーダー配列、抗 m y c 抗体を用いて組み替えタンパク質を検出するための c - m y c エピトープ、ニッケルキレートクロマトグラフィーを用いて精製するための C 末端ポリヒスチジン、および安定なトランスフェクト体の選択のためのゼオシン耐性遺伝子を含む。この g p 3 5 4 c D N A の増幅のための順方向プライマーを、慣用的な手順によって決定し、そしてこの順方向プライマーは、好ましくは、H i n d I I I クローニング部位を導入するためのヌクレオチドの 5' 伸長部分および g p 3 5 4 配列と一致するヌクレオチドを含む。逆方向プライマーをまた、慣用的な手順によって決定し、そしてこの逆方向プライマーは、好ましくは、クローニングのための X h o I 制限部位を導入するためのヌクレオチドの 5' 伸長部分および g p 3 5 4 配列の逆相 20 補体に対応するヌクレオチドを含む。P C R 条件は、アニーリング温度として 55℃である。P C R 産物をゲル精製し、そしてベクターの H i n d I I I - X h o I 部位にクローニングする。Q I A G E N クロマトグラフィーカラムを使用して、この D N A を精製し、そして D O T A P トランスフェクション培地（B o e h r i n g e r M a n n h e i m）を使用して、293 細胞にトランスフェクトする。一過性にトランスフェクトされた細胞を、トランスフェクションの 24 時間後に、抗 H i s ペプチド抗体および抗 G P 3 5 4 ペプチド抗体を使用してプローブするウエスタンブロットを使用して、発現について試験する。

#### 【0403】

永久的にトランスフェクトされた細胞を、Z e o c i n を用いて選択し、そして増殖させる。組換えタンパク質の産生を、抗 H i s ペプチド抗体、抗 M y c ペプチド抗体または抗 G P 3 5 4 ペプチド抗体でプローブするウエスタンブロットを使用して、細胞および培地の両方から検出する。 30

#### 【0404】

（C. COS 細胞における G P 3 5 4 の発現）

C O S 7 細胞における G P 3 5 4 の発現について、例えば、配列番号 1 の配列を有するポリヌクレオチドを、ベクター p 3 - C I にクローニングし得る。このベクターは、b G H（ウシ成長ホルモン）ポリアデニル化配列から上流に位置する H C M V（ヒトサイトメガロウイルス）プロモーターイントロン、および多クローニング部位を含む、p U C 1 8 由来プラスミドである。さらに、このプラスミドは、安定な形質転換体の選択のための薬物 40 メトトレキセート（M T X）の存在下における選択を提供する d h r f（ジヒドロホレートダクターゼ）遺伝子を含む。

#### 【0405】

順方向プライマーを、慣用的な手順で決定し、そしてこの順方向プライマーは、好ましくは、クローニングのための X b a I 制限部位を導入する 5' 伸長部位、続いて配列番号 1 のヌクレオチド配列に対応するヌクレオチドを含む。逆方向プライマーをまた、慣用的な手順で決定し、そしてこの逆方向プライマーは、好ましくは、S a l I クローニング部位を導入するヌクレオチドの 5' 伸長部分、続いて配列番号 1 のヌクレオチド配列の逆相補体に対応するヌクレオチドを含む。

#### 【0406】

このPCRは、95℃で5分の最初の変性工程；95℃で30秒の変性、58℃で30秒のアニーリングおよび72℃で30秒の伸長の30サイクル；続いて72℃で5分の伸長からなる。このPCR産物をゲル精製し、そしてベクターp3-CIのXbaIおよびSalI部位に連結する。この構築物を使用して、コンピテントE. coli細胞を形質転換する。次いで、QIAGENクロマトグラフィーカラムを用いて、E. coli培養物からプラスミドDNAを精製し、そしてメーカーの仕様書に従って、BRL製のLIPOFECTAMINE試薬を使用して、COS7細胞にトランスフェクトする。トランスフェクションの48時間後および72時間後に、この培地および細胞を、組換えタンパク質の発現について試験する。

#### 【0407】

10

COS細胞培養物から発現したGP354を、最初に、細胞増殖培地を約10mgタンパク質/mlまで濃縮することによって精製し得る。この精製を、例えば、クロマトグラフィーによって達成し得る。

#### 【0408】

精製したGP354を、YM-10膜を備えたAMICON濃縮機で、0.5mg/mlまで濃縮し、そして-80℃で保存する。

#### 【0409】

(D. 昆虫細胞におけるGP354の発現)

バキュロウイルス系におけるGP354の発現について、配列番号1の配列を有するポリヌクレオチドを、PCRによって増幅する。順方向プライマーを、慣用的な手順によって決定し、そしてこの順方向プライマーは、好ましくは、NdeIクローニング部位を付加する5'伸長部分、続いて配列番号1のヌクレオチド配列に対応するヌクレオチドを含む。逆方向プライマーをまた、慣用的な手順によって決定し、そしてこの逆方向プライマーは、好ましくは、KpnIクローニング部位を導入する5'伸長部分、続いて配列番号1のヌクレオチド配列の逆相補体に対応するヌクレオチドを含む。

20

#### 【0410】

PCR産物をゲル精製し、NdeIおよびKpnIで消化し、そして発現ベクターpAcHTL-A (Pharming, San Diego, CA)の対応する部位にクローニングする。このpAcHTLベクターは、Autographa californica核多角体病ウイルス (AcMNPV)の強力な多角体プロモーター、および多クローニング部位から上流の6×Hisタグを含む。リン酸化のためのタンパク質キナーゼ部位および組換えタンパク質の切除のためのトロンビン部位をコードする核酸配列は、多クローニング部位の前にある。

30

#### 【0411】

もちろん、多くの他のバキュロウイルスベクター (例えば、pAc373、pVL941およびpAcIM1)を、pAcHTL-Aの代わりに使用し得る。GP354ポリペプチドの発現のために適切な他のベクターをまた使用し得るが、ただし、ベクター構築物は、転写、翻訳および輸送のための適切に配置されたシグナル (例えば、インフレームのAUGおよびシグナルペプチド)を必要に応じて含む。このようなベクターは、例えば、Luckowら、Virology 170:31-39 (1989)に記載される。

40

#### 【0412】

このウイルスを、標準的なバキュロウイルス発現方法 (例えば、Summersら、A MANUAL OF METHODS FOR BACULOVIRUS VECTOR S AND INSECT CELL CULTURE PROCEDURES, Texas Agricultural Experimental Station Bulletin No. 1555 (1987)に記載される方法)を使用して、増殖および単離する。好ましい実施形態において、gp354遺伝子を含むpAcHTL-Aを、BACULOGOLDトランスフェクションキット (Pharming)を使用して、バキュロウイルスを導入する。感染の24時間後に、感染細胞を<sup>35</sup>S-メチオニンで放射標識することによって、個々のウイルス単離体をタンパク質産生について分析する。感染

50

細胞を、感染の48時間後に収集し、そして標識タンパク質をSDS-PAGEによって可視化する。高い発現レベルを示すウイルスを単離し得、そしてスケールアップした発現のために使用する。

#### 【0413】

Sf9細胞におけるGP354ポリペプチドの発現について、配列番号1の配列を有するポリヌクレオチドを、バキュロウイルス発現について上に記載された方法を使用して、PCRによって増幅し得る。gp354 cDNAを、Sf9昆虫細胞における発現のためのベクターpAcHLT-A (Pharmingen) にクローニングする。(バキュロウイルスにおける発現のための上記の同じプライマーを使用して) 内部NdeI部位を除去した後に、この挿入物をNdeI部位およびKpnI部位にクローニングする。DNAを、QIAGENクロマトグラフィーカラムを用いて精製し、そしてSf9細胞において発現させる。GP354特異的抗体を使用して、精製していないプラークからの予備的ウェスタンブロット実験を、推定サイズの組換えタンパク質の存在について試験する。さらなる精製およびHiG5細胞における発現の最適化の後に、この結果を確認する。

#### 【0414】

(実施例10：相互作用捕捉／ツーハイブリッドシステム)

GP354相互作用タンパク質についてアッセイするために、相互作用捕捉／ツーハイブリッドライブラリースクリーニング方法を使用し得る。このアッセイは、最初に、Fieldsら、Nature 340:245 (1989) において記載された。プロトコルは、CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY (1999) およびAusubel, F. M. ら、SHORT PROTOCOLS IN MOLECULAR BIOLOGY、第4版、Greene and Wiley-Interscience, NY (1992) に公開される。キットは、例えば、Clontech (MATCHMAKER Two-Hybrid System 3) から市販される。

#### 【0415】

GP354の全てまたは一部をコードする核酸配列と、酵母転写因子GAL4のDNA結合ドメイン (DNA-BD) との融合物を、適切なベクター (すなわち、pGBKT7) を使用して構築する。同様に、GAL4活性ドメイン (AD) 融合ライブラリーを、潜在的なGP354結合タンパク質のcDNA由来の第2のプラスミド (すなわち、pGADT7) において構築する。cDNAライブラリーの作製のプロトコルについては、例えば、Sambrookら、MOLECULAR CLONING: A LABORATORY MANUAL, 第2版、Cold Spring Harbor Press, Cold Spring Harbor, NY (1989) を参照のこと。

#### 【0416】

DNA-BD/GP354融合構築物を、配列決定によって確認し、そして自立的なレポーター遺伝子の活性化および細胞毒性 (この両方は、首尾良いツーハイブリッド分析を妨害する) について試験する。同様のコントロールを、AD/ライブラリー融合構築物を用いて実施して、宿主細胞における発現および転写活性の欠失を証明する。標準的な手順 (Ausubelら、前出) に従って、酵母細胞を、GP354プラスミドおよびライブラリー融合プラスミドの両方を用いて形質転換する (約10<sup>5</sup>形質転換体/mg DNA)。DNA-BD/GP354とAD/ライブラリータンパク質とのインビボ結合により、特定の酵母プラスミドレポーター遺伝子 (すなわち、lacZ、HIS3、ADE2、LEU2) の転写が生じる。酵母細胞を栄養欠乏培地にプレートし、レポーター遺伝子の発現についてスクリーニングする。Xgal (5-ブロモ-4-クロロ-3-インドリル-b-D-ガラクトシド) を補充した培地中での増殖において、コロニーをβ-ガラクトシダーゼ活性について二連でアッセイする (β-ガラクトシダーゼ活性のフィルターアッセイは、Breedenら、Cold Spring Harb. Symp. Quant. Biol., 50:643 (1985) に記載される)。陽性のAD-ライブラリープラスミドを、形質転換体からレスキューし、そしてもとの酵母株および無関係のDNA-B

D融合タンパク質を含む他の株に再導入して、特異的G P 3 5 4 /ライブラリータンパク質の相互作用を確認する。挿入DNAを配列決定して、G A L 4 A Dに融合したオープンリーディングフレームの存在を証明し、そしてG P 3 5 4結合タンパク質の同一性を決定する。

#### 【0417】

(実施例11：G P 3 5 4ポリペプチドに対する抗体)

G P 3 5 4に対するポリクローナル抗体またはモノクローナル抗体を生成し、そしてこれらの有用な抗原結合フラグメントまたはそれらの改変体(「ヒト化」改変体を含む)を生成するために、標準的な技術を用いる。このようなプロトコルは、例えば、S a m b r o o kら(前出)、およびH a r l o wら編(前出)に見出され得る。いくつかの実施形態において、組換えG P 3 5 4ポリペプチド(またはこのようなポリペプチドを含む細胞もしくは細胞膜)を、抗体を生成するための抗原として使用する。他の実施形態において、G P 3 5 4の免疫原性部分に対応するアミノ酸配列を有する1つ以上のペプチド(例えば、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20またはそれ以上のアミノ酸)を、抗原として使用する。G P 3 5 4の細胞外部分に対応するペプチド、特に、親水性細胞外部分が好ましい。抗原をアジュバントと混合するか、またはハプテンに連結して、抗体産生を増加し得る。

10

#### 【0418】

(A. ポリクローナル抗体またはモノクローナル抗体)

1つの代表的なプロトコルにおいて、組換えG P 3 5 4またはその合成フラグメントを使用して、マウスを免疫してモノクローナル抗体を生成するか、またはポリクローナル抗体については、より大きな哺乳動物(例えば、ウサギ)を免疫する。抗原性を増大するために、ペプチドを、例えば、P i e r c eから市販されるキーホールリンペットヘモシアニンに結合し得る。最初の注射について、この抗原をフロイント完全アジュバントと共に乳化し、そして皮下注射する。2~3週間の間隔で、G P 3 5 4抗原のさらなるアリコート、フロイントの不完全アジュバントと共に乳化し、そして皮下注射する。最後のブースト注射の前に、血清サンプルを、この免疫したマウスから採取し、そしてウエスタンブロットによってアッセイして、G P 3 5 4と免疫反応性の抗体の存在を確認する。この免疫した動物由来の血清は、ポリクローナル抗血清として使用され得るか、またはG P 3 5 4を認識するポリクローナル抗体を単離するために使用され得る。

20

30

#### 【0419】

あるいは、マウスを屠殺し、モノクローナル抗体の生成のために、それらの脾臓を取り出す。モノクローナル抗体を生成するために、この脾臓を、10mlの血清を含まないR P M I 1640に入れ、そして2mM L-グルタミン、1mM ピルビン酸ナトリウム、100単位/ml ペニシリン、および100μg/ml ストレプトマイシン(R P M I)(G i b c o, C a n a d a)を補充した血清を含まないR P M I 1640中でこの脾臓を粉砕することによって、単一の細胞の懸濁液を形成する。この細胞懸濁液を濾過し、そして遠心分離によって洗浄し、そして血清を含まないR P M Iに再懸濁する。3匹のネイティブのB a l b / cマウスから採取した胸腺細胞を、類似の様式で調製し、そして支持細胞層として使用する。10%ウシ胎児血清(F B S)(H y c l o n e L a b o r a t o r i e s, I n c., L o g a n, U t a h)を含むR P M I中で、融合前の3日間対数期に保たれたN S - 1骨髄腫細胞を、同様に遠心分離し、そして洗浄する。

40

#### 【0420】

ハイブリドーマ融合物を生成するために、免疫したマウス由来の脾臓細胞をN S - 1細胞と合わせ、そして遠心分離し、そして上清を吸引する。この細胞ペレットを、チューブを叩くことによって取り出し、そして2mlの37℃のP E G 1500(75mM H E P E S (pH 8.0)中50%)をこのペレットに加えて攪拌し、続いて血清を含まないR P M Iを添加する。その後、この細胞を遠心分離し、15% F B S、100μM ナトリウムヒポキサンチン、0.4μM アミノプテリン、16μM チミジン(H A T)(G i b c o)、25単位/ml I L - 6(B o e h r i n g e r - M a n n h e i m

50



）および  $1.5 \times 10^6$  胸腺細胞／ml を含む RPMI に再懸濁し、そして 10 の平底 96 ウェル組織培養プレートに入れる。

#### 【0421】

融合の 2、4 および 6 日後に、 $100 \mu\text{l}$  の培地をこれらの組織培養プレートのウェルから取り出し、そして新鮮な培地と交換する。8 日目に、融合物を ELISA によってスクリーニングし、GP354 に結合するマウス IgG の存在について試験する。抗 GP354 抗体を産生するモノクローナル培養物が得られるまで、希釈によって、選択したウェルの細胞を、さらにクローニングする。

#### 【0422】

(B. 抗 GP354 モノクローナル抗体のヒト化)

本明細書中に報告される GP354 の発現パターン、および治療的介入のための標的としての GP354 の可能性は、GP354 インヒビター（アンタゴニスト）に対する治療的指標を示唆する。GP354 中和抗体は、GP354 アンタゴニストとして有用な治療剤の 1 つのクラスを含む。以下は、抗 GP354 モノクローナル抗体を「ヒト化」することによって、抗 GP354 モノクローナル抗体の、ヒトにおける治療剤としての有用性を改善するためのプロトコルである。ヒト化抗体は、改善された血清半減期を有し、そしてヒトにおける免疫原性が低い。抗体のヒト化の原理は、文献に記載されている。例えば、相補体への潜在的結合を最小にするためには、ヒト化抗体は、IgG<sub>4</sub> サブタイプのものであることが好ましい。

#### 【0423】

ヒト化の 1 つのレベルは、目的の非ヒト抗体の可変ドメインおよびヒト抗体の定常ドメインを含むキメラ抗体を作製することによって、達成され得る。例えば、Morrisson ら、Adv. Immunol., 44: 65-92 (1989) を参照のこと。抗 GP354 抗体の可変ドメインを、適切な B 細胞ハイブリドーマのゲノム DNA からか、またはこのハイブリドーマ由来の cDNA からクローニングし得る。V 領域の遺伝子フラグメントは、ヒト抗体定常ドメインをコードするエキソンに連結される。得られる構築物は、適切な哺乳動物宿主細胞（例えば、ミエローマ細胞または CHO 細胞）中で発現される。

#### 【0424】

ヒト化のさらに高いレベルを達成するためには、非ヒトモノクローナル抗体の抗原結合相補性決定領域 (CDR) をコードする可変領域の遺伝子フラグメントの部分のみが、ヒト抗体配列にクローニングされる。例えば、Jones ら、Nature 321: 522-525 (1986)；Riechmann ら、Nature 332: 323-327 (1988)；Verhoeven ら、Science 239: 1534-36 (1988)；および Tempest ら、Bio/Technology 9: 266-71 (1991) を参照のこと。必要であれば、CDR3 領域を囲むヒト抗体の  $\beta$ -シートフレームワークもまた、もとのモノクローナル抗体の抗原結合部位の三次元構造をより綿密に反映するために、改変（すなわち、「復帰突然変異」）される。Kettlborough ら、Protein Engin. 4: 773-783 (1991)；および Foote ら、J. Mol. Biol. 224: 487-499 (1992) を参照のこと。

#### 【0425】

代替のアプローチにおいて、目的の非ヒトモノクローナル抗体の表面が、非ヒト抗体の選択された表面残基を、例えば部位特異的変異誘発によって変更し、一方でこの非ヒト抗体の内部残基および接触残基の全てを維持することによって、ヒト化される。Padlan, Mol. Immunol., 28 (4/5): 489-98 (1991) を参照のこと。

#### 【0426】

上述のアプローチは、抗 GP354 モノクローナル抗体およびこれらの抗体を産生するハイブリドーマを使用して、利用される。ヒト化抗 GP354 抗体は、GP354 の発現またはリガンドにより媒介される GP354 シグナル伝達が所望でない状態を処置または軽減するための治療剤として、有用である。

10

20

30

40

50

## 【0427】

(C. ファージディスプレイ由来のヒトGP354中和抗体)

抗GP354抗体はまた、Aujameら, Hum. Antibodies 8(4): 155-168(1997); Hoogenboom(前出); および Raderら, Curr. Opin. Biotechnol. 8: 503-508(1997)に記載されるもののような、ファージディスプレイ技術によって生成され得る。例えば、Fabフラグメントまたは連結した短鎖Fvフラグメントの形態での、抗体可変領域は、線維状ファージマイナーコートタンパク質pIIIのアミノ末端に融合される。この融合タンパク質の発現およびこの融合タンパク質の成熟ファージコートへの組み込みは、その表面に抗体を提示し、そしてこの抗体をコードする遺伝物質を含むファージ粒子を生じる。このよう

10

## 【0428】

(D. トランスジェニックマウス由来のヒトGP354特異的抗体)

ヒトGP354特異的抗体は、本質的にBrueggemannら, Immunol. Today 17(8): 391-97(1996)およびBrueggemannら, Curr. Opin. Biotechnol. 8: 455-58(1997)に記載されるように、トランスジェニックマウスにおいて生成される。ヒトV遺伝子セグメントを生殖細胞系配置で保有し、そしてこれらの導入遺伝子をそのリンパ様組織において発現するトランスジェニックマウスを、従来の免疫プロトコルを使用して、GP354組成物で免疫する。この免疫されたマウス由来のB細胞を使用し、従来のプロトコルを使用して、ハイブリドーマを生成し、そしてスクリーニングして、抗GP354ヒト抗体を分泌するハイブリドーマを同定する(例えば、上記のように)。

20

## 【0429】

(実施例12: GP354活性の調節因子を同定するためのアッセイ)

GP354活性の調節因子(アゴニストおよびアンタゴニスト)を同定するためのいくつかの非限定的なアッセイが、以下に記載される。これらのアッセイによって同定され得る調節因子の中には、レセプターの天然リガンド;天然リガンドの合成アナログおよび誘導体;天然の抗体または抗体様コンビナトリアルライブラリー由来の抗体および/または抗体様化合物;ならびに/あるいはライブラリーのハイスループットのスクリーニングによって同定される合成化合物などがある。

30

## 【0430】

GP354を結合する全ての調節因子が、組織サンプルにおいてGP354を同定するために有用である(例えば、診断目的または治療目的で)。アゴニスト調節因子およびアンタゴニスト調節因子は、GP354媒介疾患を処置するために、GP354活性を、それぞれアップレギュレートおよびダウンレギュレートするために有用である。これらのアッセイは、単一の推定調節因子を使用して実施され得、そして/または候補アンタゴニストと組み合わせた既知のアゴニスト(またはその逆)を使用して実施され得る。

## 【0431】

(A. cAMPアッセイ)

1つの型のアッセイにおいて、環状アデノシンリン酸(cAMP)のレベルが、候補調節因子化合物に曝露されたgp354トランスフェクト細胞において測定される。cAMPアッセイのためのプロトコルは、文献に記載されている。例えば、Sutherlandら, Circulation 37: 279(1968); Frandsenら, Life Sci. 18: 529-541(1976); Doolleyら, J. of Pharmacol. Exp. Therap. 283(2): 735-41(1997); およびGeorgeら, J. Biomol. Screening 2(4): 235-40(1997)を参照のこと。NEN Life Science Productsのアデニル(Adenylyl)シクラーゼ活性化FLASHPLATEアッセイを使用する、

40

50

このようなアッセイのための例示的なプロトコルを、以下に記載する。

#### 【0432】

簡単に言えば、GP354コード配列を、pzeoSV2 (Invitrogen) のような発現ベクターにサブクローニングする。CHO細胞を、得られる発現構築物で、公知の方法（例えば、FUGENE6トランスフェクション試薬を供給する場合には、Boehringer-Mannheimによって提供されるトランスフェクションプロトコル）を使用して一時的にトランスフェクトする。トランスフェクトされたCHO細胞を、FLASHPLATEアッセイキットから96ウェルマイクロプレートに播種し、これを、cAMPに対する抗血清を結合させた固体シンチレーション剤でコーティングする。コントロールとして、いくつかのウェルに、トランスフェクトされていないCHO細胞を播種する。このプレートにおける他のウェルに、標準曲線を作成する際に使用するのための種々の量のcAMP標準溶液を与える。1つ以上の試験化合物を各ウェルの細胞に添加し、化合物を含まない培地または緩衝液をコントロールとして用いる。処理後、室温でちょうど15分間、cAMPを細胞中に蓄積させる。このアッセイを、 $[^{125}\text{I}]$ -cAMPを含む溶解緩衝液の添加によって終了し、そしてプレートを、Packard TOPCOUNT 96ウェルマイクロプレートシンチレーションカウンターを使用して計数する。溶解細胞または標準に由来する非標識cAMPと、固定量の $[^{125}\text{I}]$ -cAMPとは、このプレートに結合した抗体に関して競合する。標準曲線を構築し、そして未知のものに対するcAMP値を、内挿によって得る。試験化合物への曝露に応答しての、細胞の細胞内cAMPレベルの変化は、GP354の活性の調節の指標である。Gタンパク質のGsサブタイプに結合するレセプターのアゴニストとして作用する調節因子は、cAMPの産生を刺激し、cAMPレベルの適度な（例えば、3～10倍の）上昇を導く。Gタンパク質のGi/oサブタイプに結合するレセプターのアゴニストは、フォルスコリンにより刺激されるcAMP産生を阻害し、cAMPレベルの適度な減少（例えば、50～100%）を導く。インバースアゴニストとして作用する調節因子は、構成的に活性であるかまたは既知のアゴニストによって活性化されるかのいずれかであるレセプターにおいて、これらの効果を逆にする。

#### 【0433】

(B. エクオリンアッセイ)

別のアッセイにおいて、細胞（例えば、CHO細胞）は、gp354発現構築物および発光タンパク質アポアクオリンをコードする構築物で、一時的に同時トランスフェクトされる。補因子セレンテラジンの存在下で、アポアクオリンは、測定可能な発光を生じ、この発光は、細胞質の遊離カルシウムの量に比例する。一般的に、Cobboldら「Aequorin measurements of cytoplasmic free calcium」: McCormack J. G. および Cobbold P. H. 編, CELLULAR CALCIUM: A PRACTICAL APPROACH. Oxford: IRL Press (1991); Stablesら, Anal. Biochem. 252: 115-26 (1997); ならびに Haugland, HANDBOOK OF FLUORESCENT PROBES AND RESEARCH CHEMICALS, 第6版, Eugene OR (1996)を参照のこと。

#### 【0434】

1つの例示的なアッセイにおいて、gp354コード配列を、pzeoSV2 (Invitrogen) にサブクローニングする。CHO細胞を、得られる発現構築物および発光タンパク質アポアクオリンをコードする構築物 (Molecular Probes) で、トランスフェクション試薬FUGENE6 (Boehringer-Mannheim) および製品の挿入物で提供されるトランスフェクションプロトコルを使用して、一時的に同時トランスフェクトする。

#### 【0435】

これらの細胞を、10%ウシ胎仔血清、2mMグルタミン、10U/mlペニシリンおよび10μg/mlストレプトマイシンを補充したMEM (Gibco/BRL, Gait

10

## 20

30

## 40

50

50

## 50

50

## 50

トランスフェクションの24時間後、これらの細胞を、37℃に予熱したPBSで1回洗浄する。次いで、無血清MEMを、これらの細胞に、単独で（コントロール）かまたは1つ以上の候補調節因子と共にかのいずれかで、添加する。次いで、これらの細胞を37℃で5時間インキュベートする。その後、これらの細胞を氷冷PBSで1回洗浄し、そして100μlの溶解緩衝液を、Promegaによって供給されるルシフェラーゼアッセイキットのウェルごとに添加することによって、溶解する。室温で15分間のインキュベ-

ションの後に、 $15\mu\text{l}$ の溶解物を、不透明白色の96ウェルプレートで $50\mu\text{l}$ の基質溶液 (Promega) と混合し、そして発光を、Wallaceモデル1450MICROBETAシンチレーションおよび発光カウンター (Wallace Instruments, Gaithersburg, MD) で即座に読み取る。

#### 【0440】

候補調節因子化合物の存在下対非存在下での発光の差異は、調節活性の指標である。構成的に活性であるかまたはアゴニストによって活性化されるかのいずれかであるレセプターは、レポーター遺伝子単独でトランスフェクトした細胞と比較して、代表的に3倍～20倍の発光の刺激を与える。インバースアゴニストとして作用する調節因子は、この効果を逆にする。

10

#### 【0441】

(D. FLIPRを用いる細胞内カルシウムの測定)

細胞内カルシウムレベルの変化は、レセプター活性の別の認識される指標であり、そしてこのようなアッセイを使用して、GP354活性の調節因子についてスクリーニングし得る。例えば、gp354発現ベクターで安定にトランスフェクトされたCHO細胞を、 $4\times 10^4$ 細胞/ウェルの密度で、Packard黒壁96ウェルプレート (このプレート上の種々のウェルから発する蛍光シグナルを区別するよう特別に設計されている) にプレートする。これらの細胞を、 $36\text{mg/L}$ ピルベートおよび $1\text{g/L}$ グルコースを含み、1%ウシ胎仔血清および4種のカルシウム指示色素 (FLUO-3 AM、FLUO-4 AM、CALCIUM GREEN-1 AM、またはOREGON GREEN 488 BAPTA-1 AM) のうちの1種を、各々 $4\mu\text{M}$ の濃度で添加した改変ダルベッコPBS (D-PBS) 中、 $37^\circ\text{C}$ で60分間インキュベートする。プレートを、1%ウシ胎仔血清を含まない改変D-PBSで1回洗浄し、そして $37^\circ\text{C}$ で10分間インキュベートして、残余の色素を細胞膜から除去する。さらに、1%ウシ胎仔血清を含まない改変D-PBSでの一連の洗浄を実施し、その直後にカルシウム応答を活性化させる。

20

#### 【0442】

カルシウム応答を、1種以上の候補レセプターアゴニスト化合物 (カルシウムイオノフォアA23187 ( $10\mu\text{M}$ ; 陽性コントロール)、またはATP ( $4\mu\text{M}$ ; 陽性コントロール)) を添加することによって、開始する。蛍光を、アルゴンレーザーを備えるMolecular DeviceのFLIPRによって測定する ( $488\text{nm}$ で励起)。例えば、Kuntzweilerら, Drug Dev. Res. 44 (1): 14-20 (1998) を参照のこと。検出器カメラのためのFストップを2.5に設定し、そして露光の長さは0.4ミリ秒である。細胞の基準蛍光を20秒間測定し、その後、候補アゴニスト、ATPまたはA23187を添加する。基準蛍光のレベルを応答シグナルから減算する。カルシウムシグナルを2秒ごとに読み取りながら約200秒間測定する。カルシウムイオノフォアA23187およびATPは、代表的に、カルシウムシグナルをベースラインレベルの約200%上に増加させる。一般に、活性化GP354は、ベースラインシグナルの少なくとも約10～15%上に、カルシウムシグナルを増加させる。

30

#### 【0443】

(E. 有糸分裂誘発アッセイ)

有糸分裂誘発アッセイにおいて、候補調節因子がgp354媒介細胞分裂を誘導または阻害する能力を決定する。例えば、Lajinessら, J. Pharmacol. Exp. Therap. 267 (3): 1573-1581 (1993) を参照のこと。例えば、GP354を安定して発現するCHO細胞を、96ウェルプレートに、5000細胞/ウェルの密度で播種し、そして10%ウシ胎仔血清を含むMEM中 $37^\circ\text{C}$ で48時間増殖させ、この時点で、これらの細胞を無血清MEMで2回リンスする。リンス後、 $80\mu\text{l}$ の新鮮なMEM、または既知のマイトジェンを含むMEMを、無血清培地中に希釈した種々の濃度の1種以上の試験化合物を含む $20\mu\text{l}$ のMEMと共に添加する。コントロールとして、各プレート上のいくつかのウェルに、無血清培地単独を入れ、そしていくつかには10%ウシ胎仔血清を含む培地を入れる。トランスフェクトされていない細胞またはベ

40

50

クター単独でトランスフェクトされた細胞もまた、コントロールとして働き得る。

#### 【0444】

16～18時間の培養の後に、1  $\mu$ Ciの [ $^3$ H]ーチミジン (2 Ci/mm ol) をこれらのウェルに添加し、そして細胞を37℃でさらに2時間インキュベートする。これらの細胞をトリプシン処理し、そして細胞採取器 (cell harvester) (Tomtec) を用いてフィルタマット上に収集する；次いで、このフィルタをBeta plate カウンターで計数する。無血清試験ウェルへの [ $^3$ H]ーチミジンの取込みを、血清で刺激した細胞 (陽性コントロール) において達成された結果と比較する。複数の濃度の試験化合物の使用により、非線形の、最小自乗当てはめ式  $A = B \times [C / (D + C)] + G$  を使用する、用量応答曲線の作成および分析が可能となる。この式において、Aは、10  
血清刺激のパーセントであり；Bは、最大の効果からベースラインを減算したものであり；Cは、EC50であり；Dは、化合物の濃度であり；そしてGは、最大の効果である。パラメータB、CおよびGは、単純な最適化 (Simplex optimization) によって決定される。

#### 【0445】

レセプターに結合するアゴニストは、細胞への [ $^3$ H]ーチミジンの取込みを増加させ、血清への80%までの応答を示すと予測される。レセプターに結合するアンタゴニストは、既知のアゴニストを用いて見られる刺激を、100%まで阻害する。

#### 【0446】

(F. [ $^{35}$ S] GTP  $\gamma$ S 結合アッセイ)

GP354がGタンパク質媒介経路を通してシグナル伝達するか否かを評価することが可能である。Gタンパク質結合レセプターは、細胞内Gタンパク質 (この活性は、GTPの結合、および結合GDPを与える加水分解を含む) を通してシグナル伝達する。従って、候補調節因子の存在下および非存在下での、加水分解不可能なGTPアナログである [ $^{35}$ S] GTP  $\gamma$ S の結合の測定は、調節因子活性のための別のアッセイを提供する。例えば、Kowalら, Neuropharmacology 37:179-187 (1998) を参照のこと。

#### 【0447】

1つの例示的なアッセイにおいて、gp354発現ベクターで安定にトランスフェクトされた細胞を、10cmの組織培養ディッシュ中でサブコンフルエントまで増殖させ、Ca $^{2+}$ /Mg $^{2+}$  を含まない5mlの氷冷リン酸緩衝化生理食塩水で1回リンスし、そして5mlの同じ緩衝液中に掻き取った。細胞を、遠心分離 (500  $\times$  g、5分) によってペレット化し、TEE緩衝液 (25mM Tris (pH7.5)、5mM EDTA、5mM EGTA) 中に再懸濁させ、そして液体窒素中で凍結させる。解凍後、これらの細胞をDounceホモジナイザー (細胞のプレートあたり1ml TEE) を使用してホモジナイズし、そして1,000  $\times$  gで5分間遠心分離して、核および破壊されていない細胞を除去する。

#### 【0448】

ホモジネートの上清を20,000  $\times$  gで20分間遠心分離して、膜の画分を単離し、そして膜のペレットをTEEで1回洗浄し、そして結合緩衝液 (20mM HEPES (pH7.5)、150mM NaCl、10mM MgCl $_2$ 、1mM EDTA) 中に再懸濁させる。再懸濁させた膜を、液体窒素中で凍結させ得、そして使用まで-70℃で貯蔵し得る。

#### 【0449】

上記のように調製され、そして-70℃で保存された細胞膜のアリコート、を、解凍し、ホモジナイズし、そして20mM HEPES、10mM MgCl $_2$ 、1mM EDTA、120mM NaCl、10  $\mu$ M GDP、および0.2mMアスコルビン酸を含む緩衝液中に、10～50  $\mu$ g/mlの濃度で希釈する。90  $\mu$ lの最終容量で、ホモジネートを種々の濃度の候補調節因子化合物または100  $\mu$ M GTPと共に、30℃で30分間インキュベートし、次いで氷上に置く。各サンプルに、10  $\mu$ l グアノシン5'-O-

10

20

30

40

50

(3  $[^3\ ^5\text{S}]$  チオ)三リン酸 (NEN、1200 Ci/mmole;  $[^3\ ^5\text{S}]$  -GTP gS) を、100~200 pM の最終濃度まで添加した。サンプルを 30℃ でさらに 30 分間インキュベートし、1 ml の 10 mM HEPES (pH 7.4)、10 mM MgCl<sub>2</sub> (4℃) を添加し、そして濾過によって反応を停止させる。

#### 【0450】

サンプルを、Whatman GF/B フィルタで濾過し、そしてこのフィルタを 20 ml の氷冷 10 mM HEPES (pH 7.4)、10 mM MgCl<sub>2</sub> で洗浄する。フィルタを、液体シンチレーション分光法によって計数する。 $[^3\ ^5\text{S}]$  -GTP gS の非特異的結合を、100  $\mu$ M GTP の存在下で測定し、そして全体から減算する。細胞において  $[^3\ ^5\text{S}]$  -GTP gS 結合の量を調節する化合物を選択し、トランスフェクトされていないコントロール細胞と比較する。アゴニストによるレセプターの活性化は、 $[^3\ ^5\text{S}]$  -GTP gS 結合の 5 倍までの増加を与える。この応答は、アンタゴニストによってブロックされる。 10

#### 【0451】

(G. MAP キナーゼ活性アッセイ)

GP354 を発現する細胞における、MAP キナーゼ活性の評価は、GP354 活性の調節因子を同定するための別のアッセイを提供する。例えば、Lajiness ら、上記および Boulton ら、Cell 65:663-675 (1991) を参照のこと。1 の実施形態において、gp354 で安定にトランスフェクトされた CHO 細胞を、アッセイの 48 時間前に、6 ウェルプレートに  $7 \times 10^4$  細胞/ウェルの密度で播種する。この 48 時間の間に、これらの細胞を 10% ウシ胎仔血清、2 mM グルタミン、10 U/ml ペニシリンおよび 10  $\mu$ g/ml ストレプトマイシンを補充した MEM 培地中 37℃ で培養する。これらの細胞を 1~2 時間血清飢餓させ、その後、刺激因子を添加する。 20

#### 【0452】

アッセイのために、これらの細胞を、培地単独でかまたは候補アゴニストもしくは 200 nM Phorbol エステル-酢酸ミリストイル (すなわち、PMA、陽性コントロール) のいずれかを含む培地で処理し、そしてこれらの細胞を、37℃ で種々の時間にわたってインキュベートする。この反応を停止させるために、これらのプレートを氷上に置き、培地を吸引し、そして細胞を、1 mM EDTA を含む 1 ml の氷冷 PBS でリンスする。その後、200  $\mu$ l の細胞溶解緩衝液 (12.5 mM MOPS (pH 7.3)、1 2.5 mM グリセロホスフェート、7.5 mM MgCl<sub>2</sub>、0.5 mM EGTA、0.5 mM パナジン酸ナトリウム、1 mM ベンズアミジン、1 mM ジチオトレイトール、10  $\mu$ g/ml ロイペプチン、10  $\mu$ g/ml アプロチニン、2  $\mu$ g/ml ペプスタチン A、および 1  $\mu$ M オカダ酸) を、これらの細胞に添加する。これらの細胞をプレートから掻き取り、そして 23  $\frac{3}{4}$  G 針を 10 回通過させることによってホモジナイズし、そして細胞質ゾル画分を、20,000  $\times$  g で 15 分間の遠心分離によって、調製する。 30

#### 【0453】

細胞質ゾルのアリコート (1~5  $\mu$ g のタンパク質を含む 5~10  $\mu$ l) を、1 mM MAPK 基質ペプチド (APRTPGGR (配列番号 18)、Upstate Biotechnology, Inc., NY) および 50  $\mu$ M  $[^3\ ^2\text{P}]$  ATP (NEN、3000 Ci/mmole) と混合し、約 2000 cpm/pmol の最終比活性に、総容量 25  $\mu$ l に希釈する。これらのサンプルを 30℃ で 5 分間インキュベートし、そして反応を、2 cm<sup>2</sup> の正方形の Whatman P81 ホスホセルロース紙上に 20  $\mu$ l をスポットすることによって、停止させる。これらの正方形のフィルタを、1% H<sub>3</sub>PO<sub>4</sub> を 4 回交換して洗浄し、そしてこれらの正方形を液体シンチレーション分光法に供して、結合した標識を定量する。等価な細胞質ゾル抽出物を、MAPK 基質ペプチドなしでインキュベートし、そしてこれらのサンプル由来の結合した標識を、基質ペプチドを用いた一致するサンプルから減算する。各ウェルからの細胞質抽出物を、分離点として使用する。タンパク質濃度を、色素結合タンパク質アッセイ (Bio-Rad Laboratories) によって決定する。このレセプターのアゴニスト活性化は、MAPK 酵素活性の 5 40 50

倍までの増加を生じると予測される。この増加は、アンタゴニストによってブロックされる。

#### 【0454】

(H.  $[^3\text{H}]$  アラキドン酸放出)

GP354の活性化はまた、細胞におけるアラキドン酸の放出を増強し得、これはまた、GP354活性の修飾因子についてのさらに別の有用なアッセイを提供する。例えば、Kantermanら、Mol. Pharmacol. 39:364~369 (1991)を参照のこと。例えば、GP354発現ベクターで安定にトランスフェクトされるCHO細胞を、 $1.5 \times 10^4$  細胞/ウェルの密度で24ウェル中にプレートし、そして10%ウシ胎仔血清、2mM グルタミン、10U/ml ペニシリンおよび10 $\mu\text{g}/\text{ml}$  ストレプトマイシンを補充したMEM培地中、使用前に37℃で48時間増殖させる。各ウェルの細胞を、10mM HEPES、pH7.5、および0.5%脂肪酸非含有ウシ胎仔血清アルブミンを補充した1ml MEM中0.5 $\mu\text{Ci}/\text{ml}$ の $[^3\text{H}]$ アラキドン酸 (Amersham Corp.、210Ci/mmol)とともに、2時間37℃でインキュベーションすることによって標識する。次いで細胞を2回、1mlの同じ緩衝液を用いて洗浄する。候補の化合物を、1mlの同じ緩衝液に、単独で、または10 $\mu\text{M}$  ATPと共にのいずれかで添加し、そして細胞を37℃で30分間インキュベートする。緩衝液単独と、偽トランスフェクト細胞を、コントロールとして使用する。各ウェルからのサンプル(0.5ml)を、液体シンチレーション分光法によって計数する。レセプターを活性化するアゴニストは、 $[^3\text{H}]$ アラキドン酸のATP刺激放出の増強を誘導する。この増強は、アンタゴニストによってブロックされる。

10

20

#### 【0455】

(I. 細胞外酸性化速度)

さらに別のアッセイにおいて、GP354活性の候補修飾因子の効果を、試験化合物により誘導されるpHの細胞外変化をモニタリングすることでアッセイする。例えば、Dunlopら、J. Pharmacol. Toxicol. Meth. 40(1):47~55 (1998)を参照のこと。1つの実施形態において、GP354発現ベクターでトランスフェクトしたCHO細胞を、12mmカプセルカップ(Molecular Devices Corp.)に、 $4 \times 10^5$  細胞/カップで、10%ウシ胎仔血清、2mM L-グルタミン、10U/ml ペニシリンおよび10 $\mu\text{g}/\text{ml}$  ストレプトマイシンを補充したMEM培地中に播種する。この細胞を、この培地中で、5%CO<sub>2</sub>中37℃で、24時間インキュベートする。

30

#### 【0456】

細胞外酸性化速度を、CYTOSENSOR MICROPHYSIOMETER (Molecular Devices Corp.)を使用して測定する。このカプセルカップを、MICROPHYSIOMETERのセンサーチャンバーに取り付け、そしてこのチャンバーを、泳動緩衝液(4mM L-グルタミン、10単位/ml ペニシリン、10 $\mu\text{g}/\text{ml}$  ストレプトマイシン、26mM NaClを補充した、ビカルボネート非含有MEM)で、100 $\mu\text{l}/\text{分}$ の流動速度で灌流する。候補のアゴニストまたは他の薬剤を、泳動緩衝液中に希釈し、そして第2の流動経路を通して灌流する。各60秒のポンプサイクルの間に、このポンプを38秒間作動させ、そして残りの22秒間消す。センサーチャンバーにおける泳動緩衝液のpHを、サイクル中の43~58秒から記録し、そしてこのポンプを60秒で再スタートさせて、次のサイクルを開始させる。記録時間の間の泳動緩衝液の酸性化速度は、Cytosoftプログラムによって計算される。酸性化の速度の変化は、修飾因子候補の添加後に得られた、最高の速度測定値から、基底値(修飾因子候補の添加の直前の4回の速度測定の平均値)を減算することによって計算される。この選択された機器は、61mV/pH単位を検出する。レセプターのアゴニストとして作用する修飾因子は、アゴニストの非存在下における速度と比較して、細胞外酸性化の速度において上昇をもたらす。この応答は、レセプターのアントゴニストとして作用する修飾因子によってブロックされる。

40

50



## 【0457】

(実施例13：他の g p 3 5 4 遺伝子またはフラグメントのクローニング)

他の g p 3 5 4 遺伝子またはフラグメントを得るために、本質的に実施例3（前出）に記載されるようにPCRを実施した。最初の g p 3 5 4 遺伝子モデルのインターエクソンPCR検査に対して、オリゴヌクレオチドGX1-220およびGX1-221を用いた。GX1-386およびGX1-387を用いて、予測される遺伝子部分の全範囲を増幅した。GX1-218およびGX1-221を用いて、マウスPanCAMオルソログを増幅およびクローニングした。GX3-238およびGX3-232を用いて、豚臓cDNAプールから全長PanCAM cDNAを増幅した。それらの配列は、以下のようである：

GX1-218：ACTGGGGGGCTAGTTTCAGTGGACTAA（配列番号16）

GX1-220：GCCAGCATGACCTCCACATTAG（配列番号19）

GX1-221：CAGCCTTCAGCTCTGCAACACATTAGTC（配列番号20）

GX1-386：ATGCGGGGTCCCCGCCCCCTCCTCGTCCCT（配列番号21）

GX1-387：CTTGGTCTCCAGAGTCCCTTCTCTCTCTC（配列番号22）

GX3-232：GGAGGAGTAAGGTGCGCCAACTTGTC（配列番号23）

GX3-238：ACCTTGGGGGGACGAATGCTTC（配列番号24）

このPCR産物を、実施例3（前出）に記載されるように単離し、そしてクローニングした。

## 【0458】

クローニングしたGX1-220/GX1-221 PCR産物を、ランダムヘキサマーキット（Invitrogen, Carlsbad, CA）に提供されるプロトコルを用いてランダムヘキサマー伸長により標識化した。この標識化フラグメントを、ノーザンブロットおよび多組織発現（MTE）ブロットにおいて、プローブとして用いた。ハイブリダイゼーションを、多組織発現ブロット（Clontech）に提供されるプロトコルに従って実施した。

## 【0459】

ノーザンブロットは、2.2～3.5 Kbの範囲のサイズの、豚臓においてのみ検出される拡散シグナルを示す。76のヒト正常組織のアレイを用いて行われるハイブリダイゼーションは、豚臓RNAサンプルの位置に、陽性のスポットを示した。これらのデータは、g p 3 5 4 遺伝子が、豚臓において優性発現していることを確証した。

## 【0460】

(実施例14：全長ヒト g p 3 5 4 遺伝子のクローニング)

遺伝子部分と予測される全長1,779（ストップコドンを含む）塩基対を、ヒト豚臓cDNAプールから、予測されるオープンリーディングフレーム（図7）の全長を含むプライマーGX1-386およびプライマーGX1-387（前出）を用いて、PCRによって増幅した。全長cDNAを増幅するために、発明者らは、AL136654配列の推定の5'非翻訳領域配列および3'非翻訳領域配列および内に、1対のオリゴヌクレオチド（GX1-238およびGX1-232、前出）を設計し、そしてこれらをヒト豚臓cDNAプールのPCRのために用いた。異なる2つのcDNAクローンを得た。いずれもAL136654より小さく、そしてノーザンブロットで決定したサイズの範囲内であった。それらは、元の遺伝子予測から得たオープンリーディングフレームに類似した、オープンリーディングフレームが続いており、そして4Igドメインまたは5Igドメインを有するタンパク質をコードする（図9A）。

## 【0461】

10

20

30

40

50

この遺伝子は、膵臓内で優性発現し、細胞接着分子を特徴付ける、複数のC2型Igドメインを有するタンパク質をコードしている(Williamsら, *Annu. Rev. Immunol.* 6:381-405(1988))ので、発明者らはまた、この遺伝子をPanCAM(pancreatic cell adhesion molecule (膵臓細胞培養分子))と名付けた。

#### 【0462】

最も長いPanCAMの潜在的ORFによってコードされる推定タンパク質は、708アミノ酸残基から構成され、そしてシグナルペプチド、5つのIgドメイン(これらのうち3つはC2型)、膜貫通ドメイン、および174アミノ酸細胞内ドメインを含む。アミノ末端およびカルボキシ末端の両方について発明者らの元の遺伝子予測によりコードされるタンパク質よりも長い。 10

#### 【0463】

公開データベースのBLASTP検索により、PanCAMとの有意な類似性を有するいくつかのタンパク質を同定した。特徴付けられた3つの最も類似しているタンパク質は、ヒトおよびマウスのNEPH1、ネフリン、ショウジョウバエ不規則交叉Cラフェスト(roughest)(「Irre」)タンパク質(例えば、図2を参照のこと)である。これらすべてのケースにおいて、相同性は、タンパク質の細胞外部分に集中しており、細胞内部分に有意な相同性はない。

#### 【0464】

ヒトNEPH1(NP\_060710)およびマウスNEPH1タンパク質(NP\_570937)は、腎臓の発達で重要な働きを担う潜在的細胞接着分子である(Donovielら, *Mol. Cell. Biol.* 21(14):4829-36(2001))。PanCAMのように、NEPH1は、5つのIgドメインを有し、そしてそのヒトタンパク質は、細胞外ドメインにおいてPanCAMに対して44%の同一性があり、一方そのマウスタンパク質は、細胞外ドメインにおいてPanCAMに対して47%の同一性がある。 20

#### 【0465】

ヒトネフリン(NPHS1, NP\_004637)は、2番目に最も類似しているBLASTPスコアであり、GP354の細胞外ドメインにおいて、33%の同一性を共有する。このタンパク質は、8つのIgドメインを含み、糸球体スリット隔膜に局在している。ネフリン(nepgrin)は、細胞接着において役割を担っており、腎臓の糸球体の発達に関与している(Ruotsalainenら, *Proc. Natl. Acad. Sci. USA* 96:7962-7967(1999); Kestilaら, *Mol. Cell* 1(4):575-82(1998))と考えられている。ネフリン遺伝子における変異は、ヒトにおいて先天的な腎炎を引き起こす(Kestilaら, 前出)。 30

#### 【0466】

ショウジョウバエ不規則交叉Cラフェスト前駆体(irre, NP\_525058)は、別の相同なタンパク質である。Irreは、複眼(Ramosら, *Genes Dev.* 7(12B):2533-47(1993); Schneiderら, *Neuron* 15(2):259-71(1995))、および触覚(Reddyら, *Dev. Genes Evol.* 209(10):581-91(1999))の発達に必要な細胞接着タンパク質である。このタンパク質は、5つのC2型Igドメインを含み、PanCAMと31%の相同性を共有する。ショウジョウバエにおけるIrre対立遺伝子座での変異は、ハエの感覚器官の発達に影響を及ぼし、明らかに、少なくとも一部の異常なアポトーシス活性に起因する(Ramosら, 前出)。 40

#### 【0467】

NEPH1、ネフリンおよびIrreは、発達のパターン付けおよび細胞-細胞コミュニケーションに関与する。GP354とこれらのタンパク質との間の類似性は、GP354がまた、類似発達経路および特に、膵臓および神経系において、器官および組織の発達および/または維持に関与するシグナル伝達経路の引き金を引く細胞-細胞相互作用におい 50

て役割を担っていることを示唆する。

#### 【0468】

(実施例15: PanCAMの選択的スプライシング)

2つの膵臓cDNAクローンの配列と2つの公開EST配列(AL136654およびBC007312)との配列比較は、少なくとも4つのPanCAMの選択的スプライシング産物(図9B)が存在することを示した。AL136654(胎児の脳cDNAライブラリーにおいて同定された)は、予測される全てのエクソン包含し、そしておそらく最も長い全長cDNAを表す。BC007312(網膜芽細胞腫cDNA由来のESTクローンは、今までに同定された最も短いPanCAM転写産物である。BC007312は、エクソン2およびエクソン11内の潜在的スプライシング部位(図9B)を用いること

10

#### 【0469】

膵臓cDNAクローンは、細胞外ドメインにおいて互いに異なるが、同一の膜貫通ドメインおよび細胞内ドメインを有する。膵臓由来のより短いcDNAクローン(PanCAM-2)は、エクソン2外でのスプライシング(図9B)に起因して、Igドメイン1の欠失部分であるタンパク質をコードする。これらの2つの膵臓cDNAクローンの細胞内ドメインは、最後のエクソン部分外でのスプライシングに起因する他の推測のタンパク質配列(図9A)より短い。

#### 【0470】

(実施例16: マウスPanCAMの一部のcDNAのクローニング)

PanCAMのマウスオルソログ検出およびクローニングするために、ヒト遺伝子配列に基いて、複数のインター-エクソンオリゴヌクレオチド対を設計し、複数のマウス組織由来のcDNAプールについて、PCRによって試験した。1オリゴヌクレオチド対は、マウスの脳のcDNAプールから弱いバンドを生成した。このPCR産物をクローニングし、配列決定した。マウスcDNAクローン(配列番号32)は、934ヌクレオチド長であり、コードするタンパク質は、ヒト対照物と85%の同一性である(図10A-B)。

20

#### 【0471】

(実施例17: PanCAMは、膵臓においてβ細胞特異的である)

PanCAMタンパク質についてさらなる特徴付けをするために、ファージ抗体についてのスクリーニングにおいて細胞外ドメインの可溶性形態を用いた。このタンパク質の細胞外部分に対して、4つのFabフラグメントを、ファージ抗体ライブラリーからスクリーニングした(Knappikら, J. Mol. Biol. 296(1): 57-86(2000))。これらFabは全て、ウェスタンブロットにおける組換えタンパク質を認識した。それらのうち3つは、パラフィン固定組織スライド上のPanCAMを認識した。

30

#### 【0472】

免疫組織化学(IHC染色)を、ヒトの心臓組織、脳組織、腎臓組織、肝臓組織、肺組織、膵臓組織、脾臓および筋肉組織を含むスライド上、ならびにサル(Cynomolous)およびマウス由来の膵臓組織上で実施した。試験したヒト組織の中で、膵臓のみが染色を示した。この結果は、この遺伝子が、ヒト、サルおよびマウスの膵臓ランゲルハンス島で特異的に発現していることを明らかに示した。

40

#### 【0473】

膵臓染色に加えて、マウスの脳下垂体前葉もまた、IHCに対して陽性であったが、より弱くかつ散乱した染色であった。この脳下垂体染色は、脳ライブラリーにおけるESTクローンの存在、およびマウスの脳cDNA由来のオルソログをクローニングする能力を示した。

細胞レベルでのPanCAMの発現パターンを同定するために、発明者らは、複数の細胞株由来のRNAにおけるRT-PCRを実施した。(293E、Cos7、マウス膵臓内皮細胞(CRL2279)(Arbiserら, Proc. Natl. Acad. Sci. USA 94(3): 861-6(1997))、α細胞(CRL2350)(Po

50

wersら, Diabetes 39 (4) : 406-14 (1990))、肥満／肥満マウス由来のβ細胞 (CRL 2364) (カルボキシペプチダーゼE変異を有し、プロインスリンプロセッシングおよびインスリン分泌が不完全なマウス株由来) ; Naggerら, Nat Genet 10 (2) : 135-42 (1995) ; Varlamovら, Endocrinology 138 (11) : 4883-92 (1997))、および野生型マウス由来のβ細胞 (CRL, 2055) (Hamaguchiら, Diabetes 40 (7) : 842-9 (1991))。この結果は、PanCAMの発現が、β細胞特異的であることを確認した。

#### 【0474】

特に、肥満／肥満マウス由来β細胞は、2つのRT-PCR産物を与えた。各RT-PCRバンドは、野生型β細胞から増幅した単一のバンドよりもずっと弱かった。変異細胞株由来の2つのRT-PCR産物の配列は、より大きな分子量のバンドが、野生型コントロールのバンドと同じであり、その一方で、より小さな分子量のバンドは、選択的スプライシング産物であることを確証した。潜在的スプライシング部位を用いて、より短い転写産物を産生し、このエクソンの3'側半分および4つの下流側エクソンを生じた。結果的に、オープンリーディングフレーム内で、Igドメイン2およびIgドメイン3が欠失し、その一方で、下流側リーディングフレームは、そのままである。

#### 【0475】

(実施例18：NODマウスにおいて糖尿病を発症している間のPanCAM-陽性細胞の減少)

I型糖尿病の発症においてPanCAMが役割を担っているか否かを試験するために、第1工程として、4週齢、8週齢、および12週齢のNODマウス由来の膵臓、およびリタイアした繁殖動物(糖尿病を完全に発症している動物)を、IHCにより実験した。4週齢で、染色パターンは野生型マウスと同じであり、膵臓ランゲルハンス島において細胞の大部分は、PanCAMに対して陽性であった。8週齢で、マウスの膵臓が発達し始めたとき、ランゲルハンス島のT細胞侵襲がみられ、そしてPanCAM-陽性細胞の数が減少し始めた。12週齢で、膵臓ランゲルハンス島において、より多くのT細胞と、より少ないPanCAM-陽性細胞が存在した。リタイアした繁殖動物では、大部分のPanCAM-陽性細胞が失われ、そしてT細胞が、以前ランゲルハンス島であった場所において観察された。この結果は、NODマウスにおいて、PanCAMの全体的な発現が、糖尿病の発症に関連していることを示した。

#### 【0476】

(実施例19：抗PanCAM抗体を用いたβ細胞の単離)

PanCAM(例えば、ヒトPanCAM)の細胞外部分に特異的なモノクローナル抗体またはポリクローナル抗体に特異的なを固体基質上または半固体基質上に固定し、次いで分散した膵臓細胞と接触させた。非β細胞を、適切な緩衝液で基質を洗浄することによって除去した。次いで、結合したβ細胞を回収する。あるいは、抗PanCAM抗体を用いたフローサイトメトリーによって、β細胞を単離し得る。次いで精製したβ細胞を、マクロポーラス(macroporous)ゼラチンマイクロキャリア(例えば、Cult hi Spher-S)に包含し、そして糖尿病を患っている患者または糖尿病を患う危険性のある患者に移植する(例えば、Guerraら, Biotechnol. Bioeng. 75 (6) : 741-4 (2001)を参照のこと)。

#### 【0477】

別に規定しない限り、本明細書中で用いた全ての技術用語および科学用語は、本発明が属する分野の当業者の技術者によって一般に理解されているのと同じ意味を有する。例示的な方法および例示的な材料を以下に記載したが、本明細書中に記載したのと類似または同等の方法および材料もまた、本発明の実行または試験に使用し得る。本明細書中で言及した全ての刊行物および他の参考は、その全体が本明細書中に援用される。矛盾する場合、定義を含む本明細書が調整する。物質、方法、および実施例は、実例のみであり、いかなる制限も意図しない。本明細書および特許請求の範囲を通して、単語「含む」、またはそ

の変形（例えば、「含む（c o m p r i s e s）」または「含む（c o m p r i s i n g）」）は、規定の完全体または完全体の群の包含を意味するが、任意の他の完全体または完全体の群の排除は意味しないと理解される。本明細書中に記載される全てのA C T T寄託は、B u d a p e s t 条約の規定に従ってなされた。上記A C T T寄託の公開に対する有効性に関する全ての制限は、本出願に基く特許の承認に基づいて、変更不能に除去される。

【0 4 7 8】

【数 1】

|                                                                                               |                |
|-----------------------------------------------------------------------------------------------|----------------|
| 出願人または代理人の参照番号<br>A097 PCT2                                                                   | 国際出願番号<br>新規出願 |
| 寄託された微生物または他の生物学的材料に関する表示<br>(PCT規則13の2)                                                      |                |
| A. 表示は明細書中下記の箇所で言及された寄託された微生物または他の生物学的材料に関してなされた。                                             |                |
| 6 頁、                                                                                          | 1 行～1.0 行      |
| B. 寄託物の表示 <input checked="" type="checkbox"/> さらなる寄託物が追加の用紙に記載されている。 <input type="checkbox"/> |                |
| 寄託機関の名称<br>アメリカン タイプ カルチャー コレクション(ATCC)                                                       |                |
| 寄託機関の住所（郵便番号および国を含む）<br>アメリカ合衆国<br>バージニア 20110-2209、マナサス<br>ユニバーシティ プールバード 10801              |                |
| 寄託者による識別参照番号：FL86-3                                                                           |                |
| 寄託日<br>2002年6月11日                                                                             | 受託番号           |
| C. 追加の表示（なければ空白のまま） <input type="checkbox"/> この情報は追加の用紙に続く <input type="checkbox"/>           |                |
| D. 表示がなされた指定締結国（全ての指定締結国に対する表示ではない場合）                                                         |                |

10

20

30

### E. 別に添付する表示物（なければ空白のまま）

下記の表示物は、追って国際事務局に提出する（表示物の一般的性質、例えば「寄託物の受託番号」の特定）

寄託物の受託番号

| 受理官庁記入欄                                                                                                                                                           | 国際事務局記入欄                                          |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|
| <input checked="" type="checkbox"/> この用紙は国際出願とともに受理された                                                                                                            | <input type="checkbox"/> この用紙は以下の日付で国際事務局により受理された |
| 認定官 <br>Cynthia Barnes<br>PCT/Internat'l Appl Processing Div<br>(703) 305-3665 | 認定官                                               |
| PCT/RO/134様式（1998年7月）                                                                                                                                             |                                                   |

40

出願人：バイオジェン，インコーポレイテッドら。

提出：本明細書と同時に

表題：P a n C A M 核酸およびポリペプチド

50

代理人参照番号：A 0 9 7 P C T 2

本明細書において言及される寄託された生物学的材料に関する専門的解決についての表示  
寄託された生物学的材料に関する表示は、全て、本明細書中に含まれる。以下のさらなる  
表示は、明細書の一部であることを必要とせず、「別の表示」として扱われるべきである  
。これらは、専門的な解決のためだけに関連する。

【0 4 7 9】

以下になされるさらなる表示は、明細書の6頁、1～10行目において、F L 8 6 - 3と  
して言及される寄託された生物学的材料に関する。

【0 4 8 0】

寄託は、寄託番号\_\_\_\_\_のもとで、2 0 0 2 年 6 月 1 1 日に、以下においてなされた：

10

アメリカン タイプカルチャー コレクション (A T C C)

アメリカ合衆国

バージニア 2 0 1 1 0 - 2 2 0 9、マナサス

ユニバーシティー プールバード 1 0 8 0 1

さらなる表示は、以下である：

C A (カナダ) 指定について：

カナダの指定に関して、寄託された生物学的材料のサンプルは、カナダ国特許法のもと、  
特許規則の規則107および108において提供されるように、カナダ国特許の付与まで  
、あるいは同出願が拒絶または放棄されて回復され得なくなるかもしくは取り下げられる  
日までは、長官により指名された独立の専門家に対するサンプルの支給によってのみ、利  
用可能にされる（規則104（4））。

20

【0 4 8 1】

E P (欧州特許) 指定について：

E P Oの指定に関して、寄託された生物学的材料のサンプルは、E P Cのもと、履行規則  
の規則28（3）において提供されるように、欧州特許の付与の言及の公開まで、あるい  
は同出願が拒絶または取り下げられるか、もしくは取り下げられたとみなされる場合は出  
願の日から20年まで、要求者により指名された専門家に対するサンプルの支給によって  
のみ、利用可能にされる（規則28（4）E P C）。

【0 4 8 2】

F I (フィンランド) 指定について：

30

フィンランドの指定に関して、特許および統制委員会 (N a t i o n a l B o a r d  
o f P a t e n t s a n d R e g u l a t i o n s) による付与の言及の公開まで  
、あるいは同出願が特許および統制委員会による特許の付与を生じないと最終的に決定さ  
れる場合は出願の日から20年まで、寄託された生物学的材料にのサンプルの供給は、当  
該分野の専門家に対してのみ利用可能にされる。

【0 4 8 3】

G B (英国) 指定について：

英国の指定について、出願人はここにおいて、寄託された生物学的材料のサンプルの供給  
が当該分野の専門家に対してのみ利用可能にされることを、認める。

【0 4 8 4】

40

I S (アイスランド) 指定について：

アイスランドの指定について、特許がアイスランド特許庁によって付与されるか、あるい  
は同出願が特許の付与を生じないと最終的に決定される場合はアイスランド特許庁による  
最終的な決定まで、寄託された生物学的材料のサンプルの供給は、当該分野の専門家にの  
みもたされる。

【0 4 8 5】

S E (スウェーデン) 指定について：

スウェーデンの指定について、出願が（スウェーデン特許庁により）公開に付されるかあ  
るいは公開を経ずにスウェーデン特許庁による最終的な決定を受けるまでは、寄託された  
生物学的材料のサンプルの供与は当該分野の専門家に対してのみ行われる。

50

【0486】

S G (シンガポール) 指定について:

出願人は、上記で同定された培養物のサンプルが、特許規則1995に対する第4のスケジュールのパラグラフ3に従って、専門家にのみ利用可能であることを認める。

【0487】

【数2】

Application No. 502/20275

出願人または代理人の参照番号  
A097 PCT2

国際出願番号  
新規出願

10

寄託された微生物または他の生物学的材料に関する表示  
(PCT規則13の2)

A. 表示は明細書中下記の箇所で言及された寄託された微生物または他の生物学的材料に関してなされた。

それぞれ10頁および101頁、それぞれ、25~27行および10~11行

B. 寄託物の表示

さらなる寄託物が追加の用紙に記載されている。 ☒

20

寄託機関の名称

アメリカン タイプ カルチャー コレクション(ATCC)

寄託機関の住所(郵便番号および国を含む)

アメリカ合衆国

バージニア 20110-2209、マナサス

ユニバーシティ ブールバード 10801

寄託者による識別参照番号: CS0026

30

寄託日

2002年6月11日

受託番号

C. 追加の表示(なければ空白のまま)

この情報は追付の用紙に続く ☐

D. 表示がなされた指定締結国(全ての指定締結国に対する表示ではない場合)

40

50

**E. 別に添付する表示物（なければ空白のまま）**

下記の表示物は、追って国際事務局に提出する（表示物の一般的性質、例えば「寄託物の受託番号」の特定）

寄託物の受託番号

**受理官庁記入欄****国際事務局記入欄**

☒ この用紙は国際出願とともに受理された

☐ この用紙は以下の日付で国際事務局により受理された

認定官 **Bonita D Barnes**  
PCT/Internat'l Appl Processing Div  
(703) 305-3663

認定官

PCT/RO/134様式（1998年7月）

10

出願人：バイオジェン，インコーポレイテッドら。

提出：本明細書と同時

表題：PanCAM核酸およびポリペプチド

代理人参照番号：A097 PCT2

本明細書において言及される寄託された生物学的材料に関する専門的解決についての表示  
寄託された生物学的材料に関する表示は、全て、本明細書中に含まれる。以下のさらなる  
表示は、明細書の一部であることを必要とせず、「別の表示」として扱われるべきである  
。これらは、専門的な解決のためだけに関連する。

20

【0488】

以下になされるさらなる表示は、明細書の10頁、25～27行および101頁、10～  
11行において、C S 0 0 2 6として言及される寄託された生物学的材料に関する。

【0489】

寄託は、寄託番号\_\_\_\_\_のもとで、2002年6月11日に、以下においてなされた：

アメリカン タイプカルチャー コレクション（ATCC）

アメリカ合衆国

バージニア 20110-2209、マナサス

ユニバーシティー ブールバード 10801

さらなる表示は、以下である：

CA（カナダ）指定について：

カナダの指定に関して、寄託された生物学的材料のサンプルは、カナダ国特許法のもと、  
特許規則の規則107および108において提供されるように、カナダ国特許の付与まで  
、あるいは同出願が拒絶または放棄されて回復され得なくなるかもしくは取り下げられる  
日までは、長官により指名された独立の専門家に対するサンプルの支給によってのみ、利  
用可能にされる（規則104（4））。

30

【0490】

EP（欧州特許）指定について：

EPOの指定に関して、寄託された生物学的材料のサンプルは、EPCのもと、履行規則  
の規則28（3）において提供されるように、欧州特許の付与の言及の公開まで、あるい  
は同出願が拒絶または取り下げられるか、もしくは取り下げられたとみなされる場合は出  
願の日から20年まで、要求者により指名された専門家に対するサンプルの支給によって  
のみ、利用可能にされる（規則28（4）EPC）。

40

【0491】

FI（フィンランド）指定について：

フィンランドの指定に関して、特許および統制委員会（National Board  
of Patents and Regulations）による付与の言及の公開まで

50



、あるいは同出願が特許および統制委員会による特許の付与を生じないと最終的に決定される場合は出願の日から20年まで、寄託された生物学的材料にのサンプルの供給は、当該分野の専門家に対してのみ利用可能にされる。

【0492】

GB（英国）指定について：

英国の指定について、出願人はここにおいて、寄託された生物学的材料のサンプルの供給が当該分野の専門家に対してのみ利用可能にされることを、認める。

【0493】

IS（アイスランド）指定について：

アイスランドの指定について、特許がアイスランド特許庁によって付与されるか、あるいは同出願が特許の付与を生じないと最終的に決定される場合はアイスランド特許庁による最終的な決定まで、寄託された生物学的材料のサンプルの供給は、当該分野の専門家にもたされる。 10

【0494】

SE（スウェーデン）指定について：

スウェーデンの指定について、出願が（スウェーデン特許庁により）公開に付されるかあるいは公開を経ずにスウェーデン特許庁による最終的な決定を受けるまでは、寄託された生物学的材料のサンプルの供与は当該分野の専門家に対してのみ行われる。

【0495】

SG（シンガポール）指定について：

出願人は、上記で同定された培養物のサンプルが、特許規則1995に対する第4のスケジュールのパラグラフ3に従って、専門家によりのみ利用可能であることを認める。 20

【0496】

【数3】

~~Application No.~~ 1502/20275

出願人または代理人の参照番号  
A097 PCT2

国際出願番号  
新規出願

寄託された微生物または他の生物学的材料に関する表示  
(PCT規則13の2)

A. 表示は明細書中下記の箇所で言及された寄託された微生物または他の生物学的材料に関してなされた。

10

11 頁、1～7行および10～11行

B. 寄託物の表示

さらなる寄託物が追加の用紙に記載されている。 ☒

寄託機関の名称

アメリカン タイプ カルチャー コレクション(ATCC)

寄託機関の住所(郵便番号および国を含む)

アメリカ合衆国

バージニア 20110-2209、マナサス

ユニバーシティ ブールバード 10801

寄託者による識別参照番号: CS0067-1

寄託日

2002年6月11日

受託番号

C. 追加の表示(なければ空白のまま)

この情報は追付の用紙に続く ☐

30

D. 表示がなされた指定締結国(全ての指定締結国に対する表示ではない場合)

40

**E. 別に添付する表示物（なければ空白のまま）**

下記の表示物は、追って国際事務局に提出する（表示物の一般的性質、例えば「寄託物の受託番号」の特定）

寄託物の受託番号

**受理官庁記入欄****国際事務局記入欄**

☒ この用紙は国際出願とともに受理された

☐ この用紙は以下の日付で国際事務局により受理された

認定官 Sonya D. Barnes  
PCT/Internat'l Appl Processing Div  
(703) 305-3665

認定官

PCT/RO/134様式（1998年7月）

10

出願人：バイオジェン，インコーポレイテッドら。

提出：本明細書と同時

表題：PanCAM核酸およびポリペプチド

代理人参照番号：A097 PCT2

20

本明細書において言及される寄託された生物学的材料に関する専門的解決についての表示  
寄託された生物学的材料に関する表示は、全て、本明細書中に含まれる。以下のさらなる  
表示は、明細書の一部であることを必要とせず、「別の表示」として扱われるべきである。  
これらは、専門的な解決のためだけに関連する。

【0497】

以下になされるさらなる表示は、明細書の11頁、1～7行；および11頁、10～11  
行において、CS0067-1として言及される寄託された生物学的材料に関する。

【0498】

寄託は、寄託番号 \_\_\_\_\_ のもとで、2002年6月11日に、以下においてなされた：

アメリカン タイプカルチャー コレクション（ATCC）

30

アメリカ合衆国

バージニア 20110-2209、マナサス

ユニバーシティー プールバード 10801

さらなる表示は、以下である：

CA（カナダ）指定について：

カナダの指定に関して、寄託された生物学的材料のサンプルは、カナダ国特許法のもと、  
特許規則の規則107および108において提供されるように、カナダ国特許の付与まで  
、あるいは同出願が拒絶または放棄されて回復され得なくなるかもしくは取り下げられる  
日までは、長官により指名された独立の専門家に対するサンプルの支給によってのみ、利  
用可能にされる（規則104（4））。

40

【0499】

EP（欧州特許）指定について：

EPOの指定に関して、寄託された生物学的材料のサンプルは、EPCのもと、履行規則  
の規則28（3）において提供されるように、欧州特許の付与の言及の公開まで、あるい  
は同出願が拒絶または取り下げられるか、もしくは取り下げられたとみなされる場合は出  
願の日から20年まで、要求者により指名された専門家に対するサンプルの支給によって  
のみ、利用可能にされる（規則28（4）EPC）。

【0500】

FI（フィンランド）指定について：

フィンランドの指定に関して、特許および統制委員会（National Board

50

of Patents and Regulations) による付与の言及の公開まで、あるいは同出願が特許および統制委員会による特許の付与を生じないと最終的に決定される場合は出願の日から20年まで、寄託された生物学的材料にのサンプルの供給は、当該分野の専門家に対してのみ利用可能にされる。

【0501】

GB (英国) 指定について：

英国の指定について、出願人はここにおいて、寄託された生物学的材料のサンプル供給が当該分野の専門家に対してのみ利用可能にされることを、認める。

【0502】

IS (アイスランド) 指定について：

アイスランドの指定について、特許がアイスランド特許庁によって付与されるか、あるいは同出願が特許の付与を生じないと最終的に決定される場合はアイスランド特許庁による最終的な決定まで、寄託された生物学的材料のサンプルの供給は、当該分野の専門家にのみもたされる。

【0503】

SE (スウェーデン) 指定について：

スウェーデンの指定について、出願が (スウェーデン特許庁により) 公開に付されるかあるいは公開を経ずにスウェーデン特許庁による最終的な決定を受けるまでは、寄託された生物学的材料のサンプルの供与は当該分野の専門家に対してのみ行われる。

【0504】

SG (シンガポール) 指定について：

出願人は、上記で同定された培養物のサンプルが、特許規則1995に対する第4のスケジュールのパラグラフ3に従って、専門家によりのみ利用可能であることを認める。

【0505】

【数4】

出願人または代理人の参照番号  
A097 PCT2

国際出願番号  
新規出願

寄託された微生物または他の生物学的材料に関する表示  
(PCT規則13の2)

A. 表示は明細書中下記の箇所で言及された寄託された微生物または他の生物学的材料に関してなされた。

11 頁、 1~7行

B. 寄託物の表示 ☐ さらなる寄託物が追加の用紙に記載されている。

寄託機関の名称

アメリカン タイプ カルチャー コレクション(ATCC)

寄託機関の住所 (郵便番号および国を含む)

アメリカ合衆国

バージニア 20110-2209、マナサス

ユニバーシティ プールバード 10801

寄託者による識別参照番号：CS0067-3

寄託日

2002年6月11日

受託番号

C. 追加の表示 (なければ空白のまま)

☐ この情報は追加の用紙に続く

D. 表示がなされた指定締結国 (全ての指定締結国に対する表示ではない場合)

10

20

30

40

**E. 別に添付する表示物（なければ空白のまま）**

下記の表示物は、追って国際事務局に提出する（表示物の一般的性質、例えば「寄託物の受託番号」の特定）

寄託物の受託番号

| 受理官庁記入欄                                                                                                                                                    | 国際事務局記入欄                                                       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|
| <input checked="" type="checkbox"/> この用紙は国際出願とともに受理された<br>認定官 <u>Sonya D. Barnes</u><br><u>PCT/Internat'l Appl Processing Div</u><br><u>(703) 305-3665</u> | <input type="checkbox"/> この用紙は以下の日付で国際事務局により受理された<br>認定官 _____ |
| PCT/RO/134様式（1998年7月）                                                                                                                                      |                                                                |

10

出願人：バイオジェン，インコーポレイテッドら。

提出：本明細書と同時

20

表題：P a n C A M 核酸およびポリペプチド

代理人参照番号：A 0 9 7 P C T 2

本明細書において言及される寄託された生物学的材料に関する専門的解決についての表示  
寄託された生物学的材料に関する表示は、全て、本明細書中に含まれる。以下のさらなる  
表示は、明細書の一部であることを必要とせず、「別の表示」として扱われるべきである。  
これらは、専門的な解決のためだけに関連する。

【0506】

以下になされるさらなる表示は、明細書の11頁、1～7行において、C S 0 0 6 7 - 3  
として言及される寄託された生物学的材料に関する。

【0507】

30

寄託は、寄託番号 \_\_\_\_\_ のもとで、2002年6月11日に、以下においてなされた：

アメリカン タイプカルチャー コレクション（A T C C）

アメリカ合衆国

バージニア 20110-2209、マナサス

ユニバーシティー プールバード 10801

さらなる表示は、以下である：

C A（カナダ）指定について：

カナダの指定に関して、寄託された生物学的材料のサンプルは、カナダ国特許法のもと、  
特許規則の規則107および108において提供されるように、カナダ国特許の付与まで  
、あるいは同出願が拒絶または放棄されて回復され得なくなるかもしくは取り下げられる  
日までは、長官により指名された独立の専門家に対するサンプルの支給によってのみ、利  
用可能にされる（規則104（4））。

40

【0508】

E P（欧州特許）指定について：

E P Oの指定に関して、寄託された生物学的材料のサンプルは、E P Cのもと、履行規則  
の規則28（3）において提供されるように、欧州特許の付与の言及の公開まで、あるい  
は同出願が拒絶または取り下げられるか、もしくは取り下げられたとみなされる場合は出  
願の日から20年まで、要求者により指名された専門家に対するサンプルの支給によって  
のみ、利用可能にされる（規則28（4）E P C）。

【0509】

50

FI（フィンランド）指定について：

フィンランドの指定に関して、特許および統制委員会（National Board of Patents and Regulations）による付与の言及の公開まで、あるいは同出願が特許および統制委員会による特許の付与を生じないと最終的に決定される場合は出願の日から20年まで、寄託された生物学的材料にのサンプルの供給は、当該分野の専門家に対してのみ利用可能にされる。

【0510】

GB（英国）指定について：

英国の指定について、出願人はここにおいて、寄託された生物学的材料のサンプル供給が当該分野の専門家に対してのみ利用可能にされることを、認める。

10

【0511】

IS（アイスランド）指定について：

アイスランドの指定について、特許がアイスランド特許庁によって付与されるか、あるいは同出願が特許の付与を生じないと最終的に決定される場合はアイスランド特許庁による最終的な決定まで、寄託された生物学的材料のサンプルの供給は、当該分野の専門家にもたされる。

【0512】

SE（スウェーデン）指定について：

スウェーデンの指定について、出願が（スウェーデン特許庁により）公開に付されるかあるいは公開を経ずにスウェーデン特許庁による最終的な決定を受けるまでは、寄託された生物学的材料のサンプルの供与は当該分野の専門家に対してのみ行われる。

20

【0513】

SG（シンガポール）指定について：

出願人は、上記で同定された培養物のサンプルが、特許規則1995に対する第4のスケジュールのパラグラフ3に従って、専門家によりのみ利用可能であることを認める。

【図面の簡単な説明】

【0514】

【図1-1】図1は、ヒトGP354（PanCAM）ポリペプチドのコードアミノ酸配列（配列番号1）および推定アミノ酸配列（配列番号2）である。細胞外部分の免疫グロブリン（Ig）ドメインおよび膜貫通型ドメインは、下線が付されている。

30

【図1-2】図1-1のつづき。

【図1-3】図1-2のつづき。

【図2-1】図2は、図1に示されるヒトGP354アミノ酸配列（配列番号2）と、Drosophilaイレギュラーキアスマ（ICCR）タンパク質の配列（配列番号13）およびヒトネフリン（neph rin）タンパク質の配列（配列番号14）とのアライメントである。ダッシュは、配列の任意のギャップである。アスタリスクは、3つの配列において同一であるアミノ酸を示す。

【図2-2】図2-1のつづき。

【図2-3】図2-2のつづき。

【図3】図3は、プライマーGX1-218（配列番号16）およびGX1-219（配列番号17）を使用して得られるRT-PCRフラグメントの配列である。このフラグメント（配列番号3）は、プラスミドCS0026の挿入物であり、このプラスミドは、2002年6月11日にATCCに寄託された。

40

【図4-1】図4は、ヒトゲノムgp354配列（配列番号5）である。エキソンには、下線を付してある。

【図4-2】図4-1のつづき。

【図4-3】図4-2のつづき。

【図4-4】図4-3のつづき。

【図4-5】図4-4のつづき。

【図4-6】図4-5のつづき。

50

【図 4-7】図 4-6 のつづき。

【図 4-8】図 4-7 のつづき。

【図 4-9】図 4-8 のつづき。

【図 4-10】図 4-9 のつづき。

【図 4-11】図 4-10 のつづき。

【図 4-12】図 4-11 のつづき。

【図 4-13】図 4-12 のつづき。

【図 4-14】図 4-13 のつづき。

【図 5-1】図 5 は、ヒト P a n C A M の c D N A 配列（配列番号 7）および誘導されたアミノ酸配列（配列番号 8）である。配列番号 7 は、配列番号 12 ならびに P a n C A M - 1（2002 年 6 月 11 日に A T C C に寄託されたプラスミド C S 0067-1 の挿入物（配列番号 25））についてのコード配列および P a n C A M - 2（2002 年 6 月 11 日に A T C C に受託されたプラスミド C S 0067-3 の挿入物（配列番号 27））についてのコード配列に基づいて作製される。配列番号 7 は、5' 側非翻訳領域および 3' 側非翻訳領域を含む。 10

【図 5-2】図 5-1 のつづき。

【図 5-3】図 5-2 のつづき。

【図 5-4】図 5-3 のつづき。

【図 5-5】図 5-4 のつづき。

【図 6-1】図 6 は、膵臓 g p 354 c D N A のヌクレオチド配列（配列番号 11）および推定アミノ酸配列（配列番号 12）である。 20

【図 6-2】図 6-1 のつづき。

【図 6-3】図 6-2 のつづき。

【図 7-1】図 7 は、代表的な G P 354 c D N A クロンのヌクレオチド配列（配列番号 25；プラスミド C S 0067-1 の挿入物）である。この推定アミノ酸配列（配列番号 26；P a n C A M - 1 と称される）は、対応するヌクレオチド配列の下に示される。シグナルペプチドは、2 重の下線が付されており、膜貫通型ドメインは、四角で囲まれており、そして 5 つの I g ドメインは、下線が付されている。最初から起源遺伝子予測の増幅のために使用されるオリゴヌクレオチドに対応する配列は、太線で強調されており、これらは、配列番号 11 をはさんでいる。 30

【図 7-2】図 7-1 のつづき。

【図 8-1】図 8 は、P a n C A M - 2 のヌクレオチド配列（配列番号 27）および推定アミノ酸配列（配列番号 28）である。

【図 8-2】図 8-1 のつづき。

【図 9A-1】図 9A は、P a n C A M : P a n C A M - 1（配列番号 26）、P a n C A M - 2（配列番号 28）、P a n C A M（配列番号 8）、および b c 007312（配列番号 29）の選択的スプライス転写物によってコードされるタンパク質のアライメントである。保存された I g ドメインは、下線が付されている。

【図 9A-2】図 9A-1 のつづき。

【図 9B】図 9B は、P a n C A M 遺伝子座および選択的スプライスされた m R N A のダイアグラムである。エキソンは、1~15 の番号が付けられ、四角で囲まれており、そして潜在性のスプライシング部位を有するエキソンは、アスタリスクで印されている。P a n C A M 遺伝子座中で影が付けられている領域は、潜在性イントロンを示し、これは、P a n C A M - 1 転写物および P a n C A M - 2 転写物においてスプライスされたエキソンの一部である。 40

【図 10A】図 10A は、部分的マウス P a n C A M c D N A クロンのヌクレオチド配列（配列番号 32）および推定アミノ酸配列（配列番号 31）である。

【図 10B】図 10B は、ヒト P a n C A M の部分タンパク質配列（配列番号 30）とマウス P a n C A M の部分タンパク質配列（配列番号 31）とのアライメントである。免疫グロブリンドメインは、下線が付されている。マウスの配列は、下側である。 50

【图 1-2】

**FIGURE 1 (つづき)**

661 cagtaccocccagaggtgactctgtctgcttcgccacacactgtgcagaggaggagaga  
Q Y P P E V T L S A S P H T V Q E G E K

721 gtcatcttctgtgccaggccacagccagcctctgtcacaggctacaggtgggcaaaa  
V I F L C Q A T A Q P P V T G Y R W A K

781 gggggctctccggtgctggggcccgcgggccaaggttagagtgctggtgcagacgctcg  
G G S P V L G A R G P R L E V V A D A S

841 ttctgactgagcccggtgctctgcgaggtgcacacgcggtggtagcgccaacccgcagt  
F L T E P V S C E V S N A V G S A N R S

901 actgcgctggatgctgtctgttgggccgattcttcgaggdaagccggagcccggtgctcg  
T A L D V L F G P I L Q A K P E P V S V

961 gacgtgggggaagacgcttcttcagctgcgcctggcggggaacccgcttccacgggta  
D V G E D A S F S C A W R G N P L P R V

1021 acctggaccocgcgcggtggcgcgaggtgctgggctctggagccacactgcgtcttcg  
T W T R R G G A Q V V L G S G A T L R L F

1081 tcggtggggcccgaggacgcaggcgactatgtgtcagagctgagctgggctatctggc  
S V G P E D A G D Y V C R A E A G L S G

1141 ctgcggggccgcgcgagagctgcgctgactgtgaacgctccccigtatgacgcgc  
L R G G A A E A R L T V N A P P V V T A

1201 ctgcactctgcgctgcctctctgaggggccctgctgcctccagtgctgtgttttcgc  
L H S A P A F L R G P A R L Q C L V F A

1261 tctccgcgccagatgcgctggtctgtgtctgggatgagggtcttctggaggcgggtcg  
S P A P D A V V W S W D E G F L E A G S

【图 2-1】

FIGURE 2

FIGURE 2

|       |                                                         |
|-------|---------------------------------------------------------|
| GP354 | -----MRVFAALLVLLFCFRGASGSPHFHQQ-----PEDIUVLGLGEE        |
| ICCR  | -----MLETHQQLLLATITVGWVRSSPYTYSQVORFAHEPQDQTAUVGAR      |
| ネフリン  | MAIGATTILRASILLGLLLEGLAQIAIFASVPRG-----FWALPENLTVVEGAS  |
|       | *****                                                   |
| GP354 | ARLPCALGATGVLVQVTKSGLALGGORDLGRWSYMIISGNAQGHDLHI        |
| ICCR  | VTLPCRVINKQVLTOWTKDPLGLTSRDLGSGFERAYMDESLQGYLSDI        |
| ネフリン  | VELRCGVSTPGSAVORAKDGLLLGCPDFRIPGFYVIRLEGDPAEGFHLHI      |
|       | *****                                                   |
| GP354 | RPVELEDEASIEYCQAT-----QAGLSRSPAQLHLVLPPEAPQVL-----GGPS  |
| ICCR  | YPMVLDDIARYQCVQSPGPGEQPAIRSTFGAGITVLVPPPEAKPTI-----QGDV |
| ネフリン  | EACDLSDDAEYECQVGRS-EMGPSELVSRVLISILVPEKLLITLPEAGTM      |
|       | *****                                                   |
| GP354 | VSLVAGVPANLTCRCSRSGDARPTFELLWFR-----DGVLLITFTPHQTLKEGT  |
| ICCR  | YIATADRVKLEICGIVG-KGPAEITITWIDGLQVLTDNIEYTVIPLDQR       |
| ネフリン  | VTVWAGQETVNCVS-GDAKEAPDITILLS-----GQTIISDIASVNESSQOK    |
|       | *****                                                   |
| GP354 | RPGVESTLITLTFPSHDDGATFVCRARSQALPTGRDITAIKLEQYFPEVTL     |
| ICCR  | RTAKSVLRLTFKKEHHNTFSCQANTARDYSAIKTEIVKEVKAPEKVV         |
| ネフリン  | LFTVEATARVTPSSRDNRLQLVCEASSPALEAPIKASFTVNLVFPFGEFV      |
|       | *****                                                   |
| GP354 | -----SASPTHVQ-----EGEKVIFLCQNTAOPFVTG                   |
| ICCR  | NVMGSLPGGAGGSGVGGAGGVSMTSGSRIVERHSQVRLERADANPSFVR       |
| ネフリン  | IENP-----GLDEGHEVR-----AGQSLELPCVARGNPLAT               |
|       | *****                                                   |
| GP354 | YRWAKGSGSPVLGAR-GRLEIVADASFLTEP-----VSCEVSNVAGVS        |
| ICCR  | YRFWINDPEIIGGQKTEVMVIRNVTRKHFDAI-----VKCEVQNSVGA        |
| ネフリン  | LQWLKNGQFVSTAWGTHTQAVARSVILVMTVRPEDHGAQLSCERNSSVK       |
|       | *****                                                   |
| GP354 | --SANSRTALDVLPFGLIQAQPEPVSDVDGEDASFCAWRGN-PLFRVWTVT     |
| ICCR  | --SEDSETLIDISAPFSFRQPSQMEADGVGSVSLTCEVDN-SPOEIVMI       |
| ネフリン  | GTQEHGITLQVTFPSPAIITILGSASQTEKNKNTLSCVSKS SRPFVLLRW     |
|       | *****                                                   |
| GP354 | R-----RGGAQVLG-----SGATIRLSPVSGFEDAGDYVCPRAEAGLSG       |
| ICCR  | Q-----HPSDRVVG-----TSNLTIF-SVSNMAGTRYCYCRAN--VPG        |
| ネフリン  | LWRQLLPMEETVMDGLHGHSIMNLTFLARREDNGRLITLCEAFSEAT         |



【図 2 - 2】

FIGURE 2 (つづき)

GP354 LRGGAAEARLTVNAPPVYVTLH----SAPAFRLGPARLQCLVFASAPDA  
ICCR YAEISADAYVYLGSPAGSGOR---TQYGLVGTARIECFASVPRARH  
ネフリン KETFKKSLILNKYPAQKLWIEGPEGGKLRAGTRVRLVCLAIIGNPEPS  
+

GP354 VVWSWD-----FGFLE-AGSQGR--FLVETFPAP---  
ICCR VSWTFNG-----QEISSE-SGHDYS--ILVDAVPG---  
ネフリン LMMYKDSRTVTESRLPQESRRVHLGSEVKSGSTFSRELVLVTGPDNOAK  
+

GP354 -----ESRGGLGPGILISVLHISGTOESDFSRSPNCARNRLGEGG  
ICCR -----GVKSTLIIRDSQAYHYG-KYNECTVVNDYGNLV  
ネフリン FTCHAGQLSASTQLAVQFPPTNTILANASALRFGDALNLTCVSVSNNP-  
+

GP354 AQASLGRRDLLPTVRIVAG-----VAAATTTLLMVIT--GVALCC  
ICCR AETQLQAKKSVSLMTIVGGI-----SVVAFLLVLTILV--VYIIC  
ネフリン -FVHLSWDKEGERLEGVAPPPRRAPFKGSAARSVLLQVSSRDHGQVTC  
+

GP354 WRHSAKFSEQKNLMRIP-----GS-----  
ICCR KKRKLPPADVISEHQITONGG-----VSCKLEPGDRTSNYSDLKV  
ネフリン RAHSAELRETIVSSFYRLNVLRYPEFLGEGVLVVTAVEQGEALLFVSVSAN  
+

GP354 -----SDGSSSRGPEE--EETG-----SREDRG--PIVHTD  
ICCR DISGGY---VPYGDYTHYSPPQYLITC-----STKNGSSSTIMQNN  
ネフリン PAPAFAFWTFRGRLSPAGGPRHRLSSGALHLWNVTTRADDG-LYQLHCQ  
+

GP354 HSDVLLEEETLETK-----  
ICCR HQNLQLOQQQQSHHOHTTOTTLLPMTFLTN-----SSGSLTGS  
ネフリン NSEGTAEARLRLDVHYAPTIRALQDPTFVNVGSGVDIVCTVDANFILPGM  
+

GP354 IIGSRIR--QDNGLPSLQSTT--ASVSSSPNGSCSNQSTTAATTTTHV  
ICCR  
ネフリン FNWERLGEDEEDQLSDIMEKISRGPTGRLRIHAKLAQAGAYQCIVDNV  
+

GP354 VVPSSMALSVDPYSAIYGNFYLRSSNSLLFPPTAV-----  
ICCR APPARRLLRLVVRFAQVQEHPTFLTKVAAGDSTSSATLHCRGVPNVIV  
ネフリン

GP354

【図 2 - 3】

FIGURE 2 (つづき)

ICCR -----  
ネフリン FTWTKNGVPLDLQDPRTYHTYHQGGVHSLLTIANVSAQDYALFTCTA  
+

GP354 -----  
ICCR -----  
ネフリン TNALGSDQTNQLVSIISRPDPSPGLKVVSILTPHSVGLEWKPFGDGLPQR  
+

GP354 -----  
ICCR -----  
ネフリン FCIRYEALGTPGFHYVDVVPQATTFTLTGLQFSTRYRWLLASNALGDS  
+

GP354 -----  
ICCR -----  
ネフリン GLADKGTQLPITTPGLHQPSEPEQLPTPEPPSGPSGLPLFLVLFALGGL  
+

GP354 -----  
ICCR -----  
ネフリン LLLSNASCVGGLWQRRRLRLAEGISEKTEAGSEEDRVNREYESQWTGE  
+

GP354 -----  
ICCR -----  
ネフリン RDTQSSSTVSTAEAPYRSLRDFSPQLPPTQEEVSYSRGFTGEDEDMAPP  
+

GP354 -----  
ICCR -----  
ネフリン GHLYDEVERTYPPSGAWGLYDEVQMGFMDLHWPEDTYQDPERGIYQVAG  
+

GP354 -----  
ICCR -----  
ネフリン DLDITLEPDSLFPFLRGHLV  
+

【図 4 - 1】

FIGURE 4

1 TCCCGCTCT TCTCACTCC TTGCTGGGT GTACATGCA CCTATCCCT  
51 CAGCTTCTCA TCTGTCAAC AGCGCTACT CCATATATC TATCTGGGCC  
101 TCAGCCTTGT GCTGGTGTCT GCGGCGCTCG ATGTGCGCTC GCATCCACTG  
151 GGTCCGACAC TGGCGTCAGC ATCTCCGAC ACCTTCTCTT GGTCCGCAAT  
201 CCGAGGATG ACATCTTTTC TGGGCGCTCT AGAAGGATAC TGTTCAGGAA  
251 CACACACCTT TCCACTCCA GAGGCTTCAT GCTGCGCCCT GCCACCCGAT  
301 TCACCCGAC TCATCTAGGA GAATGTGTAT GTACGTCTCT GCTTCGCGT  
351 CCGCAGACAC ACAGTTGACC ACCTACTCTT GCGCAGCTAC CCAGGTGACC  
401 ATGTGCGCTG CTTCTGGGT CAGCAGGAG AGCTTGGAG GAATGTGTGA  
451 GAGAAGGATC TGGGGTAGC TTCCAGCACT GAGAGCACT TGAAGATTGG  
501 AGTTGCTTAC CCAGGCTCTG GAGAGAGGA GCTCGGGGCT TTGACTTCC  
551 GGGTTGCGG GTAGGGGAG GCTTGAGCC CAGACTCAIG GGTCTGGG  
601 GTCTCTCACC CATACCGAG ATGGAGAGA TCACTCTGG AGACACGAGC  
651 TCGGGCCCCA TCTCAGAGC GCGGACCTTG CACTCATCT CCGCTCATC  
701 GCTGAGGTCA CAGGCTCGA TGTGAGGTC GAATTCACCT CAGGGGGAG  
751 CCGGAGTCA GGGCGGAGC TTCCGCTGCT GGTGAGGCT CTCAGCTCT  
801 GATCCCTTAC CTCTAGCAG GTCCGCTTCC AGGCGGTACC TCGGAGAGC  
851 TGGGATCTG GGTGCGGGC CCGAGGAGC CCATCTTTG GCCATTCGA  
901 CCGCACTGCC AGGGGTCTC ACCTCAAC CAGCTCCAC TGAAGCCCCC  
951 TCCACACCG TCAGTTTTC AGGCAAGGCC CAGAAGCCCC GGGGAACGGA  
1001 GCGAGGAATC GCCAATCTCG CAGGCGCTGA GAGACAGCG CGGTGCAAGG  
1051 AAGGCGAGA GGTTTTCTCT AGGGAAGTA ATGGGAAT AGGGGCACT  
1101 TGGCGCTGG TACAAGCTG GGATCCCAT CACCTTCAGT CAGCAGCCCT  
1151 AGGAGCAGA GAGAAGCCT GAGCGTCTG CCGAGGCGCA TCACAGCTCC  
1201 CCTACTCTG ACCCCGACAG CGCCGCGCT CAGCCAGCTG COTCTCTG  
1251 GCTTCTCTG GGTCCCTCTC TGTGTGTCT TGGCACCCTG TTTCTTTTT  
1301 TATCTCTTC GGTACTCTC CTCCCTTCT GGTTCCTCT TCCGCTCTT  
1351 CCTCTGAGT ATCTCTCTT GTCTGTCTCT CAGTCTCAAT CTCTGAGCT  
1401 CTTCCTCTG CTCTTAAAA AAATCTTTTT TTTCTTTTT TTTTTTTT  
1451 CTTTTTTTT TTTTAGAGA CGGGGTCTCA CTATGTTGGC CAGGTTGATC

【図 4 - 2】

FIGURE 4 (つづき)

1501 TCAGACTCT TCTTCAAGC CATCTCCCA CCTTGGCTC CCCAAGTGT  
1551 GGGATTACG GCGTACGCA CTGCGGCGG TCTCTTATC TTTCACTTT  
1601 TCTCTCTTG TCAAGCGCT TCTCTCTCT TTTGCTCTG TCTCTCTCT  
1651 TCTCTCTCT TCTCTCTCT TCTCTCTCT AATCTCTAT TTTCTCTCT  
1701 CCACCGCTCA CTCTGCTCC TTCTCTCAT ACTCAGAGC TTTCAAGAG  
1751 GACCTGCAGC CCAGAGTCCA GAGGCGGAG AGCCTAGAG AGCGATGAG  
1801 CTGATGCAGC CACTGCGGA GTCAAGCGTG CTCTCTGAC CAGCTGAGC  
1851 TCAITCTCAC AGTGCACCT CCGCCAGTA CTTCCAGAG CCGCCAGCTC  
1901 TGGGCTCTGC CCAGCAGCT CCGCCAGCT GCGCCAGCT GAGCATAAA  
1951 TCCGCTGTC GCACATGCA GCGCGCTTC TGGGTGCTC CCGAGCTCTC  
2001 GTGAGCCGAG GCGCTGCTT AGGCTGGGA GATGGAGAA GTGATGAA  
2051 GGTGCTCTC CAGCTCAGC GGGGAGCAG CAGCTGGGC CCGAGCTCTT  
2101 CTTTGGCTG ATACATGAC TTGGCAAGT TCTTCTTTC TTTCTTTCT  
2151 TTTCTGAGT AGTCTTCTC TGTGCTCAG GCTGAGTGC AGTGCATCT  
2201 CGGCTCACT CACTTCCAC CTCCCATGCT TTGACCTCC CAGGCTCAG  
2251 TAATCTGCC ACCTCTGCT CCCAAGTAG TGTGCTACA GGTATATAG  
2301 ACCATGCTG GCTAATTTT GTATTTTAC TAGAGAGGG GTTTCATAT  
2351 GTTGGCCAG CTGCTCTGA ACTCTGACC TCAGGTGATC CATCTGCTC  
2401 AGCTCTCCCA AATGCTGGA TTACAGCAT GAGCCAGCG ACCTGGCTC  
2451 CTTCTCTTT TTAGTAGCA TCAGTGCTA AATGATGCA GGGATCTCT  
2501 CTGGGAGGA TGCAGAGT AGTGTGACAG GCTGGGAGG TGTGGGAGG  
2551 AGGGAAGTA TGCATCTGT TACGTGGTG TGAGAGTGG GAAGGTAGA  
2601 GTGAGTGG ATCTGTATA AGCATGTGA GAGCTGTGT GTAGCATGT  
2651 CTGAGTGG AGTGCACAG GTGTGGGCG GATGAAGG CAGAGTCCA  
2701 TTCCACCGC CCGCAGTGT GGTGAGTGT GAGCCAGAG TGGGCGCCT  
2751 TTGGCAAGA CTGATGAGC TTTCTCTCC CTCTTTTCT TGGCTCTCT  
2801 CCACTCTCT CTCTCTCT CCATGTCTCT CTCTCTCTCT CCGCTATCT  
2851 ATCTGTATT ATCTTCTTT CTTTGTGAT GGAATCTTG TCTGTGGCC  
2901 AGGCTGAGG GCAGTGGAT GATCTGGTT CATTCAGCC TCACTTCT  
2951 GGGCTCAGT GATCTCTCT CCGCAGCTC CTGAATAGC GGGACTACG

【図 4 - 3】

FIGURE 4 (つづき)

3001 GTGACACCA CCACTCCAGC TAATTTTAA AAATTTGTT GTAGAGACG  
 3051 GGTCTTCTC TATTCGCCG GCTGAGTGC AGTGGTGTG TCATGCTCA  
 3101 TTGAAGGCTC AAACCTCCTA GGTCAAGTG TTCTTTCTG CTCAGCCTC  
 3151 TGAGTAGTG GGAATCAGG CCGGATCAG CACTGTGCT ATTTTTTTT  
 3201 TTTTTTTTT TTTTGTAGA GGGAGCTTG CTCTGTACC CAGGCTGGG  
 3251 TGCAATGGT CGATGTGGC TCAGTGTAC TCCTGCTCC CAGTCCAGG  
 3301 CGATCTCTT GCTCAGGCT CCGTGTAGC TGGGAATACA GGCATTGAC  
 3351 ACCACACCA GCTAATTTT GTATTTTGA TAGAGACGG GTTGTCCAT  
 3401 GTTGGCCAG CAGGTCTCG ACTCTGAGC TCAGGTACG CACTGTGCT  
 3451 GGGCCGCCA AGTGTGGGA TTACAGTGG GAGGCTGTC ACCGCGCAC  
 3501 TTGCTTAAT TTTTAAAT GTTTTTCAG AGACAGATC TTGTATAT  
 3551 GCGCAGGCT GTCTGAGCT CCGTGTGCA AGCAATCCT CCAATCTGG  
 3601 CTCCCAAGT ACTAGATTA CAGGATGAG CCGGCGACC TGGCCCTGA  
 3651 TTTATCTTC TTTTCTTCT TTTTCTCTT TTTTCTTTT TTGAGATGA  
 3701 GTTCTACTT TGTTCGCCG ACTGTAGTG AATAGTGTG TCTGGCTCA  
 3751 GTGACACTC TGGTCCCGG GTTCAGGGA TTCTCTGCT TCAGCTGCC  
 3801 TAGTAGTGG GATTACAGG ATGCGCCAG AGGCTGTGT AATTTTTGT  
 3851 ATTTTATGA AGACGGGGT TTCTCTAGT TGATCAGGT GTCTGTGAG  
 3901 TCTGAGTGT AGGTGATCAG CCGTGTGAG CCGTCCAGG TGTGTGAT  
 3951 GCGGCGTGA GTCAATGTC CCGAGTAT TATCTTCTA TCTTCTGCA  
 4001 TCTGTGAG ACTCTCTGC TCTCTATAT AAGTGTGAA ATCTAGTCA  
 4051 ATCTTATTT CAGTGTGCT CTCATCTCT ATAGTGTCT GTTATCTGT  
 4101 TTCTGTGCT CTGTCTCAG CTCTGTGCT CCGTCCAGC CAGCTGTGT  
 4151 CTCACACCA CAGAGAGTC CATAAATAT TGTCTCAGC CACACTGTG  
 4201 CCGGCGCTT TCTCTTATG TGTCTCTCA TCTCCAGTG GCTCTGTCA  
 4251 TCACATGCT GATTTTATA ACCATATGT GGTGGGCTG CCGTCCCGC  
 4301 GTGCACATC ACTGTCTGG GATAGCTTC TTCTGCTGC TTATCTGCT  
 4351 CCGGAATGG AATGTAGT TTTCTCTCA CCGTCCAGC ACCGCGGCA  
 4401 ATATCTTCC CAGGAACAG ATGAGGATC TGGCTCAGC CCGAGCTTC  
 4451 ATCTGTGAG CTGAGTTC ATCTCTCTC ACATGCTGA CTCCTTGGC

【図 4 - 4】

FIGURE 4 (つづき)

4501 TCTCTCCAG TCCGCTCTC CAATGCTCT CTACACCCC TGGGAATGG  
 4551 GCTGGAGTC GAGTGGGGG AGTGTGCTG TCTGGGGG CCGTCCAGG  
 4601 TGTGTCTCT AGGTGCCAG GCTGTGAGT GTCTGTGCT GGGGCTTTA  
 4651 ATAGACACA AGCTTCCAG CTCACCTTG GCTTTTTC TTCTGTGAT  
 4701 GCTCTGGGA CTTTGGGCT TCAATCTTC CATGTCTTA GCGCCAGAG  
 4751 GCGAGGAGG GGAAGCGGG TCAATGCTC AGAGTCTCA GCGCATGCT  
 4801 TGATCTCTC CTAAGACTC AGAGTCTCA GCGCATGCT CATCTATTT  
 4851 CAAATCCAG GCACTAGTA AGCCACACG AGCGGGGGG TGTAGGCCC  
 4901 CAGATCTCT GCGTCTGAG CCGTGGGCT TCAATCTTC CTCTCCAGG  
 4951 ACTCGGGAAT TGGGCTCTC GATCTCTCT GCGCATGCT CCGCCCGCG  
 5001 ACCTCCCAT ACACACACG ACACACACG ACACACACG ACACACACG  
 5051 ACACACACG ATACACACG GACTTAGAG AGATGTTCAG GGTCTGATT  
 5101 CCAATCTCT CTTGGGCTCT GTGGGGGTT GGAAGATTG GCAATAGAT  
 5151 CCGCGACTC TTAAGACTTA AGACAGATA TTCTGACCC TGTACCTCT  
 5201 TTCAAGTGC ACCATGACT TGATGTGAC TTGATCTCC AGCTCTCAA  
 5251 GGAAGCGGA GATCAGGCA TCAGGCTCT AGCCAGCAA GCGCTGAGC  
 5301 ACCATGCTT TATAAGCTG TAGTGTGAG CCGCATCTC AACCCACTC  
 5351 CCGCATTTAG GACACAGAG TCTGAGCTA AGAAGCTGG AGAATCTGA  
 5401 TGTGAGCTC CAGTCTCTA CAGTCTGAG AATGTGAGT TGTGTCTCA  
 5451 GCGCCCGAG CCGTCTCTC GCTGTCTGG GTCTCTCCA CCGTCTGCG  
 5501 CAGTCTGCA GCGTGGGAG CCGCCAGCT GCGTGTGAG AGCGCTCAG  
 5551 AGAAGTGGC CCGTCTGAG CCGTCTGAG TGGCCCGGG CCGCATGCG  
 5601 GCGTCTGAG CCGTCTGAG CCGCATGAG GAGGGGGGG CCGCGCGAG  
 5651 TGGGAGAGG GTGGGGGCT GAGTGTGAG GCGTGTGAG AGGAGGGGG  
 5701 GAGGGGCTG ATCTCTGCT CTTAGGACT GGTGTAGTT GCGCAATTA  
 5751 CAGGGAACA GAGGATATT TTGAGAGGG GTTTGAGGG TGGGGAGTC  
 5801 GAGTAGGGG TCCCAACTG CCGCCAGGA TCGGTGTGG CTCTCCGGA  
 5851 CAGCAGGCG CCGGGAGCG CCGGAGCGG CATCCCGAG GCGCGGAGA  
 5901 CTGGGAGCG CCGAGGACT CCGATTTTA GCTGTGCTC CAGTGTGCG  
 5951 AGTCTCTCA GCGCTGGG TGAGCAAGT TCTCTCTCT CTCCTTCTC

【図 4 - 5】

FIGURE 4 (つづき)

6001 CTCTCTCTC CCGCTTGGT GCGCTCTCA CCGCGAGG CCGAGGTGC  
 6051 CAACCCAGC GCGAGCGCT CCGCTCTCA ACCAGGTGT TCGGCTTCC  
 6101 AGACCCCAT TGAGTGGGG GCGCCAGCC CCGGGGGAT CCGGCTTGC  
 6151 GTCCCGCAT CATCTGCTC TCAGCGGCG GAGTTTCTA CCGGAGAGG  
 6201 GCGGAGCTC CCGGGGGGG GACCGGGGG GGGGAGCGG GATCGGGCC  
 6251 TCTTGGGCT TCCAGAGC CCGGCGGCG GACTGTGAG CCGTTTACA  
 6301 GCGTCTGAG CTCTGATGA CCGAGTGG AGCATCTCG CCGAGACT  
 6351 AGCGTGGCG AGAGTCTAG CCGGAGGG CCGTGGGCG AGGTGAGAG  
 6401 GAGGCTGCT TGAGAGAGG AAGTGTGCG GAGGCGAGT GCGAGGCAA  
 6451 ATCTGTGTA CTTGGGGGA GGAATCTCA GATCGGGT CCGCGGCTC  
 6501 CTCTCTCTC TCTCTGCT CAGAGGAGA CAGGTACG CAGGAGGGA  
 6551 GCGAGGAT ATGGGGTGG GGTGGGAGT TGTGTGGG CTGCTTTC  
 6601 ACTAGCGAG AGGAGCTGG GCGTGGAG TCTGTGGTC TGAATGAGA  
 6651 GCGCGCTGA GGTCTTAGC TCAGCGGCG TCGCGGAG TCTCTAGG  
 6701 CCGTGGGCT ATTTCTGCA AGAGCGAGG GACTGTGAG TGTCTGAG  
 6751 GAGGAGGCG CCGTGGCT GTCTGTGG CCGTCTGAG GCGTGTGTC  
 6801 AGTGTACTA GATGTGGCT CCGTGGGG GCGAGGAGA CCGTGTGCT  
 6851 AAGATGTTT TCTCAGCT GCGAGGCT GCGTGGGG AGAGTGTCT  
 6901 GCGTGGGCT TACCTGAGT TCTATTTT ACATTTTCA GTTGGGAA  
 6951 TTGATGGCT CCGGTAAAC TTAGAGGCT CCGATTTTG AGTGTCTCT  
 7001 TTTGGGGTA CCGAGGAG TTGGGATTA TATGTATT GCAATATAG  
 7051 ACATAGGCA GGTGAGTGG CTCAGGCTC TATTCGAC GCTTGGAG  
 7101 GTTGGGCA GAGGATGCT TGAAGCAG AGTTGTAG CAGGCTGGG  
 7151 AACATAACA GAGCTTATC CTACACAA TATATATAT ATTATAACA  
 7201 AATAGCGCG GTATGTGCT GTGATCTAT AGTCCAGT ACTCAGAGG  
 7251 CTTAGTGT AGGATGCT GAGCTAGG GTTCAGGCT GAGTGTGCT  
 7301 ATGATCAGC CAGTGTGCT CAGGATAG TGAGAGGCT TCTCAAAA  
 7351 AAAAAAAA GAGAACATA ATGCAAAA GTACAGTAA TATAATGGA

【図 4 - 6】

FIGURE 4 (つづき)

7401 AGATTACCA AATAAATAG ACACACAG CCAATCCCA AGTCAATGC  
 7451 TAGTCTCTC GAGAGCGCG TGTCTTTC CCGTATATA GCGGCTTCC  
 7501 CCGTCTCTC GAGTGTAT CTAATTAAT TTTATGCA ATCATTTCT  
 7551 TGTCTCTCT CCGTCTTA TTACCGCTAA GTTGTAGT ACTCTGCT  
 7601 GGGAGGAGT TAGAGTCTC CTGGGCGAG TACACATT TTAATGTGT  
 7651 CTTACACCA AATGTGTGG CAGTTTCT GGTGTAGAT GTCTGGGAT  
 7701 GAGGCGTGA GCGAGAT TCAAGACAT GGTGTGTA CCGCTTCTC  
 7751 CCGTCTCTC GGTGTGCT CCGTGTGCT AGTGTGTA GCGGAGT  
 7801 CCGAGTGA CCGTCAAT AGCGGCTG AGTGTGTA TGAAGTCA  
 7851 TATGATGTC AGCTACCA AGCAGGCTC GCGTCTGAG CAGGCAACT  
 7901 GCGTGTCTC GGTAGAGC TCGGCACT GTCTGTGG AGCGGAGG  
 7951 GCGAGGCTC ACTAGTGT AGTAGAGG CCGAGGAG CCGAGGAG  
 8001 GGTCAATGG AGTGTAGG ATCAGATCC ATCTGTGAG CCAATCTCA  
 8051 CTTGAGTCT CCGAGAGC CCGGAGGT CTGGGCGCG CCGTGTGTC  
 8101 TGTGTGCT GAGTGTCT CCGAGTGT AGTGTGAG CCGTGTGTC  
 8151 CCGGCGTCA CCGTGTCT CCGTGTCT GAGTGTGCT CCGTGTGTC  
 8201 GAGGCGCTC TCAATGCT CAGTGTGTA TGTGTGCT AGCGTGTG  
 8251 CATTAGGGA AGTGTGCT ACATGTGAG CAGGCTCA TCGAGAGG  
 8301 AAGAGATAT GGAAGGAG AGGTCTAG GTTGTGCT TGAATATG  
 8351 ATGAGGATA AGATCTAG GCGAGACT CCGGCTTCA AATGTGCT  
 8401 CAGGAGTGT CAGTGTAT AGATCTAG TATGTAGT AGCTGTCT  
 8451 GCGTGTGCT CCGTGTAT AGATCTAG TATGTAGT AGCTGTCT  
 8501 TAGGATATG TGAAGTAA AATAAATAT ACCTATAAT CCGGAGAG  
 8551 GTGACATA CATGTGAG ACTATATAG TGAAGAGG TGTGTGAG  
 8601 TCGGCGCTC GCGTGTGCT CAGGCGAG TGTGTGTA TGTGTGAG  
 8651 CAGAAATCA TAAATTTT TAAACATA TGAGACTTT TGTATTTT  
 8701 TTGCTCATC AGTATCAT AGTGTAGT TGTGTGCT AAGCAATTT  
 8751 TTCTCCAT GTGGCGAG AAGCGGAG GATTGTGAG CCGTGTGAG  
 8801 TAGTGTGCT TATATCTT ATCTCTCT CCGAGGCTC CCGTGTGAG

【図 4 - 7】

FIGURE 4 (つづき)

8851 GGACCCCTGG CTGCTGGAG AGCAGCTTAA CCTGACCCC TTTCAGCCAT  
 8901 GATGATGAG CCACTTTTCT CTGCCCCGCG CGGAGCCAG CCTGCCCCAC  
 8951 AGGAGAGAC ACAGCTATCA CACTGAGCCT GCAGTGTGAG TGCAGCTGGC  
 9001 CCTGGGAAG AGGGGTGTGG GGCCCTGACT CTTGGGTATG AGGAAGGAGG  
 9051 GGACTGTGGC CCTTGGGAA TGAAGAACT GAGAGCTGGA CTCCTGGATC  
 9101 TAAATATGCA GGAGAGGGCT GGTATGCTA GCTACGCTT GACTACACAG  
 9151 AACCTTGGGA GGTGAGGCA GCGGATCAT CTAGATCAG GATTTGAGA  
 9201 CCAGTGTGGC TAAATGTGC AAACCCGCTC TCTACTAAA ATACAAJAT  
 9251 TTGCCCCGCG TGTATGACA CACTTGTAT TCCAGTATC TGGAGGCTG  
 9301 AGGCAGGGA ATCACTGTA CCGGGAGCG AGATGTTGCG GTGAGCCGAG  
 9351 ATCATGCCAC TCAGCAGCAG AGTGAAGCT CGAGCAGGAG AGGACAGACA  
 9401 GCTGGGGTCC CTGGGAJAG AGAAAGCTGG GCTTGACTC TCACATCGGG  
 9451 GAGACTAGGA GAGGCGAGAA GGTGGGACA TTGAGTAACT TGGGGAJAT  
 9501 GGGAGCTGAA AGCCAGACT CTTGGCTCAA AGGGAJAGG GATTAAGGG  
 9551 CCGAGCTCC TGGATGAGG GAACAGGGA CTGAGACCT AGGCACTGA  
 9601 CGAGGTGTT CTTGGCTCTT GCGCATCTGA CCAATGTCCT ACCCTCACAG  
 9651 ACCCGCCAGA GGTGACTCTG TCTGTGCGC CACACACTGT GCAGGAGGA  
 9701 GAGAGGTCA TTTCTGTG CAGGCGACA GCGGAGCTC CTCTCACAG  
 9751 CTACAGTGA GGAGGAGAC CCACTGTCTT CCAGCCCCAA GATGAGCTT  
 9801 GGGAGGGCT GGAAGCTGAG TAGGTGTGCC AGAGAGGCA GGAACAGTT  
 9851 AACAGGCCCA CCAATTCCTC AGTGGGCAA AAGGGGGCTC TCGGCTCTC  
 9901 GGGGCCGCG GCGCAGGTT AGAGGTCTG CAGAGGCTC CTTCTGAC  
 9951 TGAGCCGCTG TCTGCGAGG TCAGCAACGC CGTGGGTAGC GCGACCGCA  
 10001 GTACTGCTCT GATGTGCTG TGTGAGCTGG GCGGCGCTG TGGGTGTGT  
 10051 CAAGGTGCG CTTGGCTTTC AGGGCTGTTG AGGCTGGGG CTTGGAGGG  
 10101 CGGGGCGGG AGAGCGAGCG TGGGTATTA GAGGAGGAG AGTGTGAGC  
 10151 TGGGGCATAT TCTTGGCCCT TAGAGGGTGT GGTGTGTTCT TGGGCTGCG  
 10201 TGATCCGAGG TCAGTGGCTG CATTCGCGCC CGGCAATGTG ACCCTAGTC  
 10251 TCCTTGTCTC AGTTGGGCGC ATTCTGAGG CAAGCGCGGA GCGGCTGTC  
 10301 GTGACGCTGG GGAAGACGC TTCTTCAGC TGGGCTGCG GCGGAGACC

エキソン6

エキソン7

エキソン8

【図 4 - 8】

FIGURE 4 (つづき)

10351 GCTTCCAGG GTACCTGGA CCGCCGCGG TGGCGGAG GTACAGCCCT  
 10401 AAATCTGAG CGGTGCTGG AGGGGAGCA GGCCTCTTA CAATCTGGC  
 10451 TTCTGAGCC CCTTCCCTGT CGAGGTGCT GGCCTCTGGA GCGCACTGC  
 10501 GTCTTCCCTC GGTGGGGCCC GAGGAGGCA GCGCATATGT GTGAGAGCT  
 10551 GAGGCTGGG TATCGGCTT GCGGGGCGG GCGGAGGAG CTTGGCTGAC  
 10601 TGTGACGCT GAGAGGCGG GGTCTCTAG GCGAGCTGGC CCGCTCTGG  
 10651 ATAGGAGCG GACAGAGGG GCAAGGCTA ATGAGTGGG AGTGGCTGG  
 10701 AAGGAGCTT ACACCCAGG GGGGCTGAG ACCGAGCTA TTGAGGCGA  
 10751 GCTTTTAGG AGAATCGAG TTGGAGGCG GGTGGGCTG ATTGATTGAG  
 10801 GTAGCGGAG AGTGGCTGG ACAGACCGG CTTGTATCA GCTTTGGGG  
 10851 AGGCAAGAC CTCTCTCTG AGTGAGCTAG AGTCTCATC CCAGCTGCC  
 10901 CAGTAGTAC GCGCTGCGC TCTGCGCTG CTTCTGAG GGGGCTGCT  
 10951 GCGCTCCAGT GTCTGTTTT GCGCTCTCC GCGGAGTGT CCGTGTATG  
 11001 GAATGTGAC TCTTCCCGT ACCCATCCAG CCGTGTATCC TGACCTGCCA  
 11051 CCTGGCCCG GAAACTACT GTGACCTTT CTGACTGCC AGACATCCCT  
 11101 CCTGCTCTT CTCTCTCTG TCAGTCTCT CCGTGTCTCT CCGTTTTGT  
 11151 GCGCCAGGT CTGGCTTGG GATGAGGCT TCGTGGAGG GGGGTGCGAG  
 11201 GCGGGTTC TGTGGAGAG ATTCTGCTC CCAGAGAGCC GCGGGAGCT  
 11251 GGTTCGAGG CTGATCTCTG TGTACACAT TTGCGGAGC CAGAGCTCTC  
 11301 ACTTAGCAG GAGCTTAAC TGCATGCCG GAGACGCGT GCGGCGGGA  
 11351 GTGCGCCAG CAGGCTGGG CCGTAGAGT GAGACCCAG CCGGAGAGC  
 11401 CAAACTCTG AGAGCTTAA CCGCACAAG CGAGGATCC CCGAGCCGAG  
 11451 GCTGCAAAA CTGATACCC TCAATGCG AGAGAGCTC CAAAGCTGG  
 11501 GAGTCTCAA ACTGTGGCT CATGATTC CAGAGACCC CTCAGCCACA  
 11551 AATGCTTCA CATTCGAT CTAATAGTA GAGACTCTC ACACCTAGG  
 11601 GCGCCAAA GGGAACTCC AATGATGCA AAGCAATG CAAATTAAG  
 11651 GACCCCTCA ATTCTAAGC TCCCTAAGC CAGGAGTTT AAATCACTC  
 11701 TCAACTTGG GGAACCCAA ATTCAAGGG CTTGATCT TCAATGTGC  
 11751 GACCTTTGA ACCAGGAAT CCGAACTCA ATCCCTGAGC CCGGCTTCC  
 11801 TGTTCCCTC TCAGCTTCT CAGGATGTC CCGTCTCTC CTGAGACTT

エキソン9

エキソン10

エキソン11

【図 4 - 9】

FIGURE 4 (つづき)

11851 GCTGCCACT GTGCGATAG TGGCGGAGT GCGGCTGCC ACCACAATC  
 11901 TCTTATGCT CATCATGCG GTGCGCTCT GCTGCTGCG CCACAGCAG  
 11951 GCTTATGCT TGAGCCCGG CCGGCTGCC GAGGCGGAG CCGCAGAGC  
 12001 GCGCTGCTG CCGACTGAC TGAGCTGCG TTGGGCTTG GCTCAGTCC  
 12051 CAATTCCAG CTGCGACAG GCTTACTCT CTTTCACTT CCGTTCCTT  
 12101 CCTTAAGCC TAATAGGCC TTCCAGGCT CAGACTCTC GTGGGAGG  
 12151 ATTCTTATG GTTCTAACA GCGCTACCA ATCAGCTCA TTGTTCCCA  
 12201 GTCTCTCTC TTCCGCTTA TTGCTGTGA CACATTGGA CCGGCGCAT  
 12251 CGCTTAATC CAGGCTGCG TGTGTGAG CTTGAGCTC TTCTCGAG  
 12301 CAAGAGAAC TGATCGGAT CCTTGGGCG AGGAGGCT CAGTTGAG  
 12351 AGCTCTGAA GAGAGGAGA CAGGAGGCG GAGGAGGCG GTAGATGCC  
 12401 AGGGTCCCA GACTGAGCT TGCTCCAGA CTAATAAT AGCCAGTCC  
 12451 CAAGAGGCT CCAAAATCA AATAGACTC TAAGCCAGG CATGTTGCT  
 12501 GAGCTTGTG ATACACTTT GGGAGTGA GACAAAGGA TCACITAGG  
 12551 CCAGGAATC AAGGCGAGC TGACAGCAT AGCAGAGCC CATCTTACA  
 12601 AAATACAAA CTAAATPAA ATAAATAAG AACGGGTAT GGTGGCATC  
 12651 ACCTAATGC CAGCTACTC AGGACACTG GGTGGGAGA TCCCTTGGC  
 12701 ACAGAGGTA AAGCTGAG TGAGCTATGA TTGACCATG CACTGAGCC  
 12751 TGGGTACAG AGCAAGACC TGCTCCATT TTTTITTTT TTTTITATG  
 12801 AGGAGGCTC TACTTTTTT TTTTITGGA GAATTCAT CTGTACCCA  
 12851 GCTTGAATA CAGTGTGCG ATCTGCGCT ACTGCAACT CTGCTCCCT  
 12901 GTTCAAGTG ATTCTTCTG CTGAGCTCC TGAGTAGCT GATTAAGG  
 12951 GCGGACGAC CAGCGTGC TGATTTGTA TTTTATAG AGATTGGTT  
 13001 TCACCATGT GCGGAGCTG GTCTCAACT CTTGACCTCA GGTATCCAC  
 13051 CGGCTCGAC CTCCAAAGT GCTAGGATTA CAGGATGAG CCGTCCGCG  
 13101 CGGCTGAGG GCTCAAGCT TTTTITTTT TTTTITTTT TTTTITGAG  
 13151 GGATCTTGG TCTGAGCC AGGCTGGAT GCACTGGCG GAATCGACT  
 13201 CACTGCAAG TCACTGCC GGGTTCAC CATCTCTG CTTGAGCTC  
 13251 CAGATAGCT GAGACTAGC GCAGCGGCA CATGCTCAG CTAATTTTT

エキソン12

エキソン13

【図 4 - 10】

FIGURE 4 (つづき)

13301 TGATTTTGA GTAGAGAGA GGTATATCC GTGTAGCCA GGATGGTCTG  
 13351 GATCTCTGA CCGTGTGTC GCGTCTGCT GCGCTCCAA AGTGTGGGA  
 13401 TTAGAGGCT GAGCAGGCG GCGGCGGCA GGGCTTAGT CTTAAGAGT  
 13451 ACCGACACC CAAATGTC CCAATGCTT GCGGCTGAC CAATATTCC  
 13501 GTAGGCGAG TATGTCCCA CAGGCGGTT TTGTGTGT TTGTGTGT  
 13551 GTGTGTGT TTGAGTAGA GTTGTGCT GTGCTCAG CTGAATGCA  
 13601 GTGTGCAAT CAGAGTAC TGACCTCTC ACCTCCAGT TGAATGAT  
 13651 CTGTGCTT AGGCTCTGA GTAGCTGAA TTACAGTGC CTGCAACAT  
 13701 GCGTGGCTT TTTTTCATT TTTAGTAG ACAGATTC GCGATGTAG  
 13751 CCAGGCTGT CTCAACTC TGGCTCAG TGATCTGCT GCTGCGGCT  
 13801 CCAAAGTGC TGGATACA TGATAGGCC ACTGTGCTG CTTCTTAGC  
 13851 CGCTTTTAT GTCTGAGTG CAGTCCGAG CTTTGGGT CTCTACTCC  
 13901 CTCTGCTG GCGTCTGCT GTGTGCTGA GGTCTTGG GTCTGGCAT  
 13951 GGGGGCGGG GATCTCTGA CTATTCCTC TCAATAGT CCGTACCCA  
 14001 GGGGGCGAT GTGACACTG ACCAGAGTA TGTGTTCTG GAGGAGGAG  
 14051 GAGCTCTGA GACCAAGTG AGTGTGAGA GGGGTGGGC TCCCTTCACT  
 14101 GTGGGAGAG GCGGGGCTCC CTTACTGT TTTCCGCTC TCCGACGCG  
 14151 CTGTGCTCT CTTTTCCTT CTTGTGCTT CAGATGGG ACTCAGCTC  
 14201 CCAGGCACT CTTCTGCGC CTTGGGCTT CTACTGAG TCCGAGCTC  
 14251 AGTTGCTG TGTGAGACT CTTCCACAC ATCTGTCCA GCGTACGCT  
 14301 CAATCTGAG CCGGAGACA GGGCTCAGC TGCTGTGCT CTCTCATCA  
 14351 CGGTCAAGC CCGTTTCTG CAGCAGGTC TACATGGCC CCACTCTCC  
 14401 CCAGCTGCT TCTTCTCTG GGGCTCTAC TCAGGAGAG CACATGACT  
 14451 CGAGGCTAT CCGAGGCGG AGCACTTCT TCTCTCTCT TCTCTGCT  
 14501 ACTGACACA TGCACTCTC TCTCTCTCT TTTTITTTT TTTTITGAG  
 14551 ACGAGCTCT ATTCTGTTG CAGGCTGGA GTGAGTGT GCAATCTAG  
 14601 CTCATGCAA CTTGCTCTC CTGGGTCAA GCAATCTCC TGCTCAGGC  
 14651 TCCATATAG CTGGTAAT TTTCTGTAT TTTTATAGA GATGGAGTT  
 14701 CACATGTTG CCGAGGCTG TCTGGAATC CTGAGCTCA GTGATCGCT  
 14751 GCGGACGCG TCCAAAGT CTGGGATTA AGGCTGAGC CACTGTGGCC

エキソン14

## 【図 4 - 1 1】

FIGURE 4 (つづき)

14801 AGCGGACATG CCAATTCCTT GGCTGAAAC ACTCTACCT TCCTTCCAT  
 14851 GTCTACCTAA TTCTCTCTT TAGTCTCCA GTCTCAGCT AGACATCTT  
 14901 TGTTCATAGA AGCCCATGCT TCGCTCATGA CAGCTGCATC AATTGCGTG  
 14951 TGTTCACCC ATCAGAGCCA TGACCATCTT GATCTGGCT TCCTTATCC  
 15001 ACCCATATG CTGAGGCTC TACCATCACA GCGCTGTCA TTGCTATGC  
 15051 CTTTCCGAG CAGAGCCCTG ACCCTCTGG GTACTGTCTC ATGATCTGT  
 15101 ATTCTTCTT TGTGTGGGA TTCTGTGAGG ACAGGCTCCA GTTCTATCT  
 15151 AGTGACATGC CTTGTAGAG CACACAGGG TGTGACATG AATCAAGCC  
 15201 TAGAGGCTGT TGGGCGAGT AGTGTCTCTC TCTGTCTCTC TGTGACATG  
 15251 CCACACGAC ACCCTCAGC AGGCTATAT CCGCTGTCTC CTACCTTCT  
 15301 CTGCTATGT CCAATTCAT TGGCTCTTAT CACTGTCTCT GTGTCTCTC  
 15351 TTCTCTCTC CCACTCCAT GTGTCTCTCT GTGTCTCTC CCACTCTCT  
 15401 CTCTTTTCT CTCTCTCAG GTCTGTCTA TTCTAGTGTG TCTCTCTATC  
 15451 AGTGACCTTC ATCCCTCTG CAGCTCTACA GACTTACTG AGTCCATTT  
 15501 GTCCCTCAG GACCAACCA AGGCTTACTA CAGGTCCGA GAGTCTAGT  
 15551 TGAGCTGAG CTTTGGCGAA GCGCTGGAG GAGTCTCTT CCGTCCACA  
 15601 CCGTCCCTC TTGGGCGCC AGGACCTCT ACCTTCTATG ACTTCAACC  
 15651 ACACCTGGC ATGCTCTCC CCGTCAAGT TTACAGAGC AGGCTAGCT  
 15701 ATCTACCAT ACCCACTCT CAGCTCTCA CAGCTACAT CAACCCACA  
 15751 TCTTTGGGC CCGCATCTT GCGCTCTGG ACTCTCTCT TCCCATATG  
 15801 TGGCTTCCC ACACCTAGC ACCGCTCTC CAGACTCAG GTGTGACATC  
 15851 TTCCATGAG AAGATCTGT GATCTCTCA CAGTCTCTA TGGATGTTC  
 15901 TGATTTCTA GAGGCGAGA CAGTGTGGG ACTTACTCC TCCAAACTG  
 15951 AACACAGGG GAGGGAAGA TCATTACAT TTCTAGGAGC ATTGTATAC  
 16001 AGTCAGTCA GCGAAGAGG ATGCGCTAG TGGGAGCAG ATGCGACCC  
 16051 AATATGCCA CCAATTCCTC GGTGTAAAG AGATTCAAG TGGCAGTAG  
 16101 GCGCTTTCG GAGAGATGG GACAGGGAG TGGGTCTTG GAGTTGGGG  
 16151 CCGGATGGA AGTTGTCTT AGCCTAGAA AGAAGATATT TCAGATGAC  
 16201 CATCTCAT TGAAGGAAAG GTAGCATAGG ATAGATGAG ATGAGAGCA

エキソン15

## 【図 4 - 1 2】

FIGURE 4 (つづき)

16251 TACGAGGCGC CAGCTGGCT CTCTCTGAG GGAATTTTC TGGGCAATG  
 16301 GAAATGAGC CAGATGGC ATATATCCC TAGGACCCA AGATGGCAC  
 16351 CATCTTCTT TTACTTCTT TAAAGACTCA GAAGACTG GACCCAGGA  
 16401 GTGGGATAC AGTGAGAT ACCACTGTG GGGCAATAA TTGGATATA  
 16451 AATATTATG TTTATATA AAAAAAGT AAGAGGCA GTGTGTCTTA  
 16501 GGGATCTAC TGGCATATC ACTCTCAC CAGGAGGGG TCCCTTAGC  
 16551 CTGTCCAGG GTCCCTCTC TACCTAGCC TATGAGTGG CTGTAGAGT  
 16601 AAACTGTGA GCACTCTC AGCTCTTC TACCTGAA GCACTTAGG  
 16651 CTCTTTTTT TTTTCTTG AGCAAGATC TGGCTATG GCGCATATG  
 16701 GAGTCTAGT GCATATCTC AGCCATGC TACTCTGA TCTGGGCTC  
 16751 AAGCATCTT TCCACTCAG CCGCTAGT ACTGGGAT ACAGTGTCT  
 16801 GCGACACAG CAGCTAAT TTTGTATTG TTTGTATGA GCGTTTACC  
 16851 ATGTGGCCA GGTGTCTCT AAACTCTGA CCAAGTGA TCCGCCACC  
 16901 TAGGCTCCC AATGTCTG GATTACAGC ATGAGCTCT GTGCCAGCC  
 16951 ATGGCTCTT TTAATATA TCTTACACA CACACACAG CACACACAG  
 17001 CCGACACAG CACATAGT GCAACAGAA AAGACACAG CATAGCATG  
 17051 TATGACAGA CACAGCTA GATGTCCAG CAGTTGACA CAGTGCAGC  
 17101 GGTGCGCCA GGGTCTCTG GGAAGCTGA ATTCTAATC TCATTAGAG  
 17151 AGCAAAAGC GTAGGCTCTG AGTGTGAGA GGAAGGGA GACTTGGGT  
 17201 AGGAAATGG CAATCCCTG GTCTTACAG CAGGCTGGA GATCCAGAC  
 17251 CTATCTCTA GGTGTCTCT CCACTGAG CAGTGTCTG GTGCTCTCT  
 17301 GGAATCTCT TAAAGAGG TCGTGGGGG TTTGTTCTT GATCTCTTA  
 17351 GGTAGAGTT AAGGTCTG GCGCTGGA CCGTGGGA CAGAGCAAG  
 17401 AAGAGTAA CCGTGGCTT GAGGAGAT GGGTGGGG CTGTCTCT  
 17451 GGTCTGAGA GAGAGTGC CAGACTCTT GATCTGAA GAGGAGGA  
 17501 CTAGGTCTA ACTGTCTCT TCTGACTGG GCAATTTTG GAGGCTGTA  
 17551 TTCTGAGCC CTCACAGAG GAATGACTA GGGATGGGG GTCTGTGAT  
 17601 CTGCTCTCT TGAAGAGG CAAAGCTGT AGTGTCTGG TCTAAGAGG  
 17651 GTGAGATCC TGGGAGGA CTCAAATCC ACAAAGGCG GAGCCATAG  
 17701 CCGACTCTT GGTGGGCG TCTATGGGG GGAAGGCTG GAATCTGAG

## 【図 4 - 1 3】

FIGURE 4 (つづき)

17751 GCGGCGGGG CTGCGCAG CTCCGCGCG GGTTCGGA GTGTCTCAC  
 17801 TCTGCGGAA GCGCCACGC TATTCTCTG ACCAGAGG CCGGCGGGC  
 17851 GCGGCGGGG GTGCGGCG CCGGCGCAG CCGGCGGGG CCGGAGGAG  
 17901 AGGTGCGGG GTGCGGCGT GAGGCTCTG TCACTCTGG GCGCGGGCG  
 17951 GCGGCGGAG CAGGCGACT GAGGCGGCA GCGGCGGAG ATGCGGCGG  
 18001 CCGGCGGGG TGTCTGCGT TAACTGCG CCGGCGGCA GCGGCGGCG  
 18051 CCGTCTCTG TGCATATT GTGTCTCTT CTGCGGCGG AGCGGCGAT  
 18101 CCGGCGGAG GCGGCGGCA GCGGCGGCG GCGGCGGAG AGCGGCGG  
 18151 GGTCTCTGG GATGCGGGA AGGCGGGA CCGGCTCTT GAGCGGCGG  
 18201 GCGGCGGAT CAGGCGGGA AGGCGGCGT TTCCAGGCA GTGTCTGAG  
 18251 CCGGCGGCG CCGGCGGCA ATTATGAG CCGGCTCTC AGCGGCGG  
 18301 TCGGCGGAT TCGGCGGCG CAGGCGGCG CCGGCTCTT CATGAGGAT  
 18351 GGTCTATTCA GCGGCGGCG TCGGCGGAG CCGGCGGCG CTGCGGCGG  
 18401 AGGTGACTG GCGGCGGCG CCGGCGGCG GCGGCGGAG AGCGGCGG  
 18451 AGCGGCGAT TTGATGCA TATCTGAG CCGTCTATT CTCTGAGCG  
 18501 ATATGAGCG CTTGCGGCG TCGGCGGCG CAGGATCTA TCGGCTGAG  
 18551 CAGGCGGCG CCGGCGGCG TGTGCTCTC CAGTCTCTA GCGGCTGAG  
 18601 TCGGCGGAG TCGGCGGCG AAGCGGCGG TGTGCTCTC TCGGCTGAG  
 18651 AAGCGGCGG TCGGCGGAG ATGAGGCGG GAGGCGGCG AAGCTCTG  
 18701 TTCTGCGCT CTCTGCGG ATTGAGTTT AACCGGCTG GCGGCGGAG  
 18751 CCGTCTCTG GCGGCGGCG CCGGCTCTT CTGCGGCGG ACGGCGGAG  
 18801 AAGATCTCG GCGGCGGCG GTGCGGAG CAGGCGGAG CCGGCGGCG  
 18851 CCGGCGGCG CCGGCGGCG TCGGCGGAG CCGGCTCTT CATGAGGCG  
 18901 GTGTGCGGT CTTGCGGAG GCGGCTGAG CTGCGGCTT CCGGCTGAG  
 18951 GAGCAAGAA CCGGCGGAG TCGGCGGCA GCGGCGGCG GCGGCTGAG  
 19001 TCGGCTCTC TCGGCGGAG GCGGCGGCG CCGGCTGAG GCGGCTGAG  
 19051 TCGGCGGCG TCGGCTCTC CCGGCGGAG CCGGCGGCG GCGGCTGAG  
 19101 AGCGGCGAG CCGGCTCTC GCGGCTCTC GCGGCGGCG CCGGCTGAG  
 19151 AGTCTGAG CAGGCTGAG GCGGCGGAG GCGGCGGCG GCGGCTGAG  
 19201 TAGAAGCTA GACTGAGAA AGGCGGCTC CAGGCTGAG CCGGCTGAG

## 【図 4 - 1 4】

FIGURE 4 (つづき)

19251 GCGGCGGAG CTGCGGCGG TCGGCGGAG AAGCGGCA CTATCTGAG  
 19301 GCGGCGGAG CTGTTTCTG ATACTAGT AAGGAGAG GAGGCGGCG  
 19351 AGATGCGAG AGGAGGAGT GAATTAGGA ATCTGCTGT CTTGTAAAG  
 19401 AGTAGAGTG TCGGCGGAG TCGGCGGAG ATAGGCGG CTATGAGAG  
 19451 AAGAGGCTT AAGGCGGAG AGATGAGGA AAGGCGGCG CCGGCGGAG  
 19501 AGCGGCGAG GCGGCGGCG CCGGCGGAG AAGGCGGCG TTTATAGCG  
 19551 GAGGCGGAG TTTGCGGAG CCGGCGGAG GGAAGGCG GCGGCGGAG  
 19601 GAGGCGGAG CTTTGGGAG GAGGCGGAG ACTGCGGCG GCGGCGGAG  
 19651 GAGGCGGAG CTTGAGGAG CTTGAGGAG ACTGCGGCG GCGGCGGAG  
 19701 ATAGGCGGTA AAGGCGGAG GACTTTTAA GGAAGGCG TAGTGGGCT  
 19751 GCGGCGGAG CAGGCGGAG TTGATATCT AAGGCGGCG CTATGAGAG  
 19801 GCGGCGGAG GCGGCGGAG GTCTCTGA ATACTAGAG CTTGAGGAG  
 19851 GAGGCGGAG GCGGCGGAG GTCTCTGA ATACTAGAG CTTGAGGAG  
 19901 GCGGCGGAG GCGGCGGAG GTCTCTGA ATACTAGAG CTTGAGGAG  
 19951 AAGCGGCGA AACTCGGCT CTACTAGAA TACAAAGAT AGCGGCGAT  
 20001 GGTGTCTGT GCGGCGGAG CAGGCTCTC AAGGCGGAG GAGGCGGAG

【图 5-2】

1  
GGGAAGTCTGGCAGGGCTTTTCAGAGCTTCAGAGGCTTCGGGATGACAGACTTGGAGGAACT 60

61  
AGGGCAGAGACTAGGCTGGGCGAAGASTCGAGCTGAAGGGGCTCCGGGGCCAGGGTGAC 120

121  
AGGAGGCGGTGCTTCGACAGGAGAAATGATGACGGGAAGGCAGTTCGACGCGCAATCTCTG 180

181  
AAGCTTGGGGGACGAATGCTCAGGATCGGGTCCCGCCGCTCTGTGCTCTCTCTTCTGC 240  
M L R N R V P A L L V L L F C

241  
TTCAGAGGGGAGACAGGCGGCTCGCCCAATTCTTCGAACAGCAGAGACTTGTGGTG 300  
F R G R A G F S F H F L Q Q P E D L V

301  
CTGCTGGGGAGGAAGCCGGCTCGCTGTGCTCTGGGCGGCTACTGGGGCTAGTTCAG 360  
L L G E E A R L F C A L G A Y W G L V Q

361  
TGAATTAAGATCTGGGCTGGCCATAGGGGCGAAAGGGACCTACCAGGTTGGTCCCGGTAC 420  
W T K S G L A L G G Q R D L P G W S R Y

421  
TGAATATCAGGGAATCGAGCCAATGGCCAGCATGACCTCCACATTAGGCCCTGGAGCTA 480  
W I S G N A A A N G Q H D L H I R P V E L

481  
GAGGATGAAGCATCATATGAATCTCAGGCTACACAAGCAGGCTCTCGCTCCAGACCAAC 540  
E D E A S Y E C Q A T Q A G L R S R P A

541  
CAACTGCAGCTGCTGTGCTCCCCCAGAGGCCCGCCAGGTGCTGGGGGGCCCTCTGTGTCT 600  
Q L H V L V P P E A P Q V L L G G F S V S

601  
CTGGTGTCTGGAGTCTCTGCAAGCTGCATGCTCGAGCGCTGGGGATGCCCGCCCTACC 660  
L V A G V P A N L T C R S R G D A R P T

661  
CCTGAATCTCTGGTTCCGAGATGGGGTCTGTGTGGATGGAGCAACTCTCCATCAGACC 720  
P E L L L W F R D G V L L D G A T F H Q T

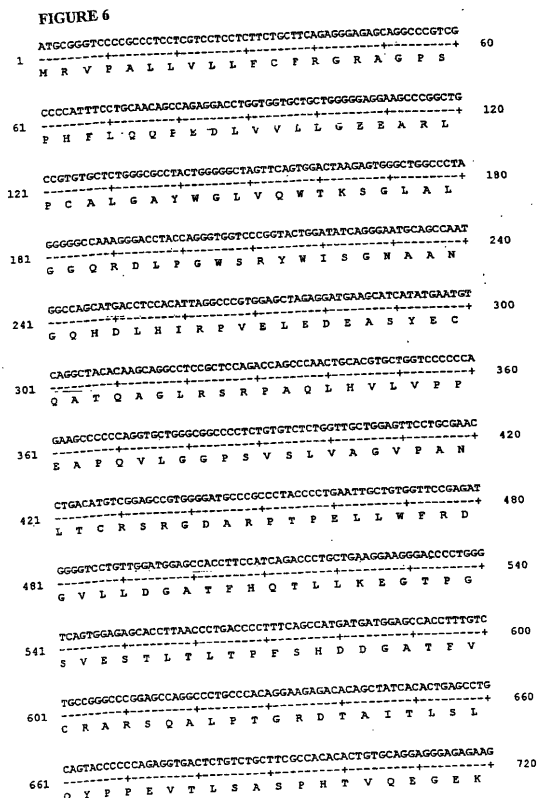
| FIGURE 3 (continued) |                                                                                                          |
|----------------------|----------------------------------------------------------------------------------------------------------|
| 721                  | TCGCTGAGGAGGAGGACCCCTGGCTTCAGAGCAGCACTTAACCGTACGCCCTTTTCAGC<br>L L K E G T P G S V E S T L T L T P F S   |
| 781                  | CATGATGATGGAGCACCCTTTGTCTGCGGGGCGGGAGCAGGCGCTGCCACAGGAAGA<br>H-D D G A-T F-V C-R-A R S Q A L P T G R     |
| 841                  | GACACAGCTATCACACTGAGCGTGGAGTACCCCGACAGGTGACTGTCTGTCTGCTCCGCA<br>D T A I T L S L Q Y P F E V T L S A S P  |
| 901                  | CACACTGTGCGAGGAGGAGAGAGGTCAITTTCTGTGCGAGGCGACGCCAGCCTCTT<br>H T V Q E G E K V I F L C Q A T A Q P P      |
| 961                  | GTCACAGGCTACAGGTGGGCAAAAGGGGGCTCTCGGTGCTCGGGGGCCCGGGGCCAAGG<br>V T G Y R W A K G G S F V L G A R G F R   |
| 1021                 | TTAGAGGTCTGTGGCAGAGCGCTCGTCTCTGACTGAGCGCGGTGCTCTCGAGGTGACAAAC<br>L E V V A D A S F L T E P V S C-E V S N |
| 1081                 | GCGCTGGGTAGCGCAACCGCACTACGCGCTGGATGTGCTGTTTGGGCGGATCTGCGAG<br>A V G S A N R S T A L D V L F G F I L Q    |
| 1141                 | GCAAGCGCGGAGCCCGTGTCTCGTGAAGCTGGGGGAGACGCTTCTTCAGCTGCGGCTGG<br>A K P E P V S V D V G E D A S C F S C A W |
| 1201                 | CGCGGGAAACCGCTTCACCGGTAACCTGGACCGCGCGGGTGGCGCTCAGGTGCTGGGCG<br>R G N P L P R V T W T R R G G A Q V L G   |
| 1261                 | TCTGAGCGACACTGCGTCTTCGTGCGTGGGGGCGGAGGACGAGGCGACTATGTGTGC<br>S G A T L R L F S V G P E D A G D Y V C     |
| 1321                 | AGAGCTCAGGGTGGGCTATCGGGCTCGCGGGCGCGCGCGAGGCTCGGCTGACTGTG<br>R A E A G L S G L R G G A A E A R L T V      |
| 1381                 |                                                                                                          |

【图 5-4】

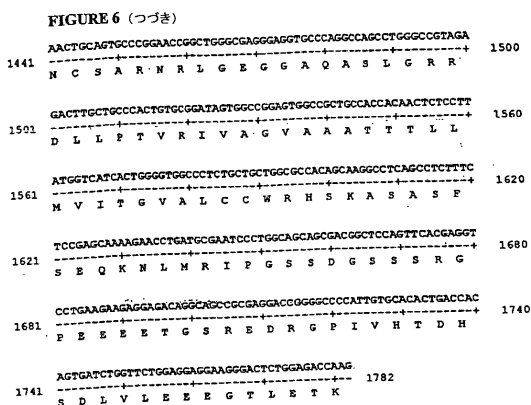
|      |      |      |
|------|------|------|
|      | 1440 | 1440 |
| 1381 | 1440 | 1440 |
| 1441 | 1500 | 1500 |
| 1501 | 1560 | 1560 |
| 1561 | 1620 | 1620 |
| 1621 | 1680 | 1680 |
| 1681 | 1740 | 1740 |
| 1741 | 1800 | 1800 |
| 1801 | 1860 | 1860 |
| 1861 | 1920 | 1920 |
| 1921 | 1980 | 1980 |
| 1981 | 2040 | 2040 |

[illegible]

【图 6-1】



【图 6-3】



【図 7-1】

**FIGURE 7**

ACTCTGGGAGCGAACTCTCAGATCTGGGCTTCCCGGCGCTCTGCGACTCTCTCTCTCGT  
M L E M R V P A T T G G V L L F C C  
 TCAGAGGAGGAGAGCGCTCTCCGCGCTTTCTTCTCAAGCAGCGCAGAGAGAGCTCTGGTGTG  
R G R A G P S P N H F L Q O F E D L V L V  
 TCGCTGGGAGGAGAGCGCTCTCGCGCTGTGCTCTGGCGGCGCTACTCGGGGCGTAGTTCAGT  
L G E A R L P C A L G A T Y N G L V C  
 GCGATCAGATGGGTGGCCCTTAGGGGCGCAAGAGGACTCTACAGAGCGTGGTCCCGGTACT  
T K S G L A L G G Q R D L P G W S R Y H  
 GCGATCAGAGGATTCGACGATCGATCGACGATCGCTCCAGATTAGCGCCCTGGAGCTAG  
I S Q N A A N G Q H D L H I R P V E L E  
 AGGATAGGACATCATGATGATCTCAGGCTACACAAAGCAGGCTCCGCTCCAGACAGACGCC  
D E A S T E C Q A T Q A G L R S R A Q  
 AACTGACATCTCTGGTCCCCCAGAGAGGCCGCCAGGTCTCTGGGCGCGGCTCTCTGTCTCT  
L H V L V F P P A A Q P V L G G S V L T  
 TGGTGTCTGGAGATCTCTCGAGACCTGACATCTCGAGAGCTCGGGAGTCCCGGCCCTACCC  
V A G Y P A N L T C S E R R G D A R P T P  
 CTGATATGCTCTGGTCCGAGATGGGCTCTGCTGTGGAGCGGACAGCTTCCATCGAGCC  
T G L T F D G V L L D G A T F H Q T L  
 TCGTAGAGAGAGAGCGCTCGGCTGGGTCACTGGAGAGCACCTTAACCGTACGCCCTTCTCAGCG  
L K E S T F P G S V S E S T L T D T P F S H  
 ATGATATGATGCGCACCTCTGTCTCGCGGGCGCGGAGCGACGAGGCCCTGCCACAGAGAG  
D D G A T L V C R A R S Q A L P T G R D  
 ACACAGCTACATACATGAGCTCGACGTACGCCCGCAGAGGTGACTCTGTCTGCTCGGCAC  
T A I T L S L Q Y P P V E T L S A S P H  
 ACACCTCTCAGAGAGGAGAGAGATTCATTCTCTCTCGCAGACGACAGCGACGCTCTGT  
T V O E G E K V I L F C Q A T A Q P P V  
 TCACAGGCTACAGGTGGGCAAGAGGGGCTCTCTGGGTGCTCGGGCGCGCGGGCGAGGT  
T G Y R N A K G S G P V L G A R G P L C  
 TAGAGGTCTGCGACAGCGCTCTGCTCTGACCTGAGCGCTGTCTGCGGTGCTGAGCAAC  
E V V A D S A F L T E P V S C E V S N A  
 CCGTGGGTAGGCGCAACCGCAGATACGTGGCTGGATGTGCTGTTGGTGGCATCTTCAGAG  
V G S A N R S T A L D V L G P I L A C  
 CAAGCGCGAGCGGCTGTCTCTGGAGCTGTGGGGAAGAAGCCCTCTCTCAGTCTGGCGTGG  
K P E P V S V D V G E D A S F S C A M

【図 8-1】

FIGURE 8

[illegible]

【図 7-2】

**FIGURE 7 (つづき)**

CGCGAATCCCGCTCCACCGGTAACTCGACCGCGCGCGCGCGCGAGGTGTGGCGT  
G N P L P R V T W T T R R G G A Q V L G S  
CTCGAGCCACATCGCGTCTTCGCGTGTGTGGCGCGCGAGACGACACTATGTGGCA  
G A T L R L P S V G P E D A D D Y V C R  
GAGCTGAGCGTGGCGTATCGCGGCTCGCGGCGCGCGCGCGAGCGTCTCGCGTGACTGTGA  
A E A G L S G L R G G A A B E A R L T V  
ACGCTCCCCGATGATGACCGCGCTGCACCTTCGCGCTGGCTGCTTCTCGAGGGCGCTGGTGC  
A P P V V T A L H S A P A F L R G F A R  
GCTTCAGGTCTGTGGTTTGTGCGCTTCGCGCGCCGAGTGGTGTGTGGTCTGTGCTTGGAGTG  
L Q C L V F A S P A P D A V V M S D D E  
AGCGTCTCTCGAGAGCGCGGTTCGACGGGTGGGTCTCTGGTGAGACATTCCTCGCCCGG  
G A G L A G S Q G R L V E T F P A P E  
AGAGCGCGCGGGACCTGGGTCCGCGCGTGTCTGTGCTACACATTCTGGGAGCCACGG  
S R G G L G P G G L I S V L H I S G T Q F  
AGTCTGACTTTAGCGAGAGCTTTAACTGCGATGTCCCGAACCGCGTGGCGCGAGCGAGGTG  
S D F S R S F N C S A R N L R L G E G A  
CCGAGCGCGACGCTGGGCGTAGAGACTTGTGCCCACTGTGTGGGATGTGGAGCGATGGTGG  
Q A S L G R R D L L P T R I V A S V V  
TGCGTGCACCAACACTCTCTTATGTGTCATCATCGGGGTGGCGGCTCTGCTGTGGCGCG  
A A T T T L L M V I T G V A L C C W R I  
ACAGCAGGCTCCAGCGCTCTTCTCCGAGCAAAAGAACCTCATGTGGAACTCTGTGCGAGT  
S K A S A S F S E Q K N L M R I P G S S  
GCGACGCGTCCAGTCTCAGCAGGTCTCTGAAGAGAGGAGACGACGACGCGCAGAGACCGG  
D G S S R G P E E E R T G S R E D R G  
GCCCATATGTGCACATCGACCAAGTGTATCTGGTTCTTGAGAGGAGAGGATCTGTGGAG  
F I V I T H D S D L V L E E E G T L E  
CGAGAGCCCAACCAACGTTTACTACAAGGTCGAGAGGAGCAGTCCACCGCGGTCTCCAG  
K D P T N G Y V K V R G V S P P A S P I  
ACTCAGTGTGACACTCTTTCCAAATGAAGAGGTCTCGGGATCTCCAACTTGCACATATGG  
S R V T S F Q W K S P G I S N L P  
TTGTCTGATTTCTGAAGAGACGAGGACAGTGTGGACAGCTTACTCTCC

【図 8-2】

FIGURE 8 (つづき)

|      |                                                                                                       |      |
|------|-------------------------------------------------------------------------------------------------------|------|
| 1321 | -----                                                                                                 | 1380 |
|      | L G E G G A Q A S L E E E D L E E T                                                                   |      |
|      | A T A G C C C G A A T G G T C T C T C C A C A A C T C T C T A T G T C A T C A C T G A G G T G G C C   |      |
| 1381 | -----                                                                                                 | 1440 |
|      | I V A G C G A A A T T T L L M V I T G V A                                                             |      |
|      | C T C T C T G T G G G C C A A G G C T C A G C T C T T T C C G A C A A M G A C T A G T G               |      |
| 1441 | -----                                                                                                 | 1500 |
|      | L C C M R H S K A S A F S E Q K N L M                                                                 |      |
|      | G A A T C T C G A C A C A G C G C T C A G T C A G A G G T C T C G A G A G G A G A C A G C             |      |
| 1501 | -----                                                                                                 | 1560 |
|      | R I F S S S S S S R G P E E E E T G                                                                   |      |
|      | A G C C G A G A G A C G G C C C A T T G T G C A C T A C C A C A C T T G T T T C G A G A G             |      |
| 1561 | -----                                                                                                 | 1620 |
|      | S D R I S E                                                                                           |      |
|      | G A G G A C T C T G A G A C C A G A D C C A C A C A C G T T A C T A C A A G T C C G A G A G T C A G T |      |
| 1621 | -----                                                                                                 | 1680 |
|      | E G T L E T K D P T N G Y Y K V R G V S                                                               |      |
|      | C C A G T C T C C A C T C A C T G T A C A C T T C A A T G A A C C T C T G A G A T C T C C             |      |
| 1681 | -----                                                                                                 | 1740 |
|      | P P A S P D S R V T S F Q W K S P G I S                                                               |      |
|      | A A C T G C C A A T A G T A T T T C A T T T C T G A G A C C C A A G T T G G C A C T A C T             |      |
| 1741 | -----                                                                                                 | 1800 |
|      | N L P * W I V L I S E E P G Q V G D L T                                                               |      |
|      | C T C C C                                                                                             |      |
| 1801 | ----- 1805                                                                                            |      |

## 【図 9 A - 1】

FIGURE 9A

1 60  
 PanCAM-1 MLAMRVALL VILFCFRGA GPSPHFLOQP EDLVVLGEE ARLPALGAY WGLVQMTKSG  
 PanCAM-2 MLAMRVALL VILFCFRGA .....  
 PanCAM MLAMRVALL VILFCFRGA GPSPHFLOQP EDLVVLGEE ARLPALGAY WGLVQMTKSG  
 bc007312 MLAMRVALL VILFCFRGA GPSPHFLOQP EDLVVL .....

61 120  
 PanCAM-1 LALQOQDLP GMSRWISGN ANQGHDLHI RPVELEAS YCOATONGL RSRPAQLVL  
 PanCAM-2 LALQOQDLP GMSRWISGN ANQGHDLHI RPVELEAS YCOATONGL RSRPAQLVL  
 PanCAM LALQOQDLP GMSRWISGN ANQGHDLHI RPVELEAS YCOATONGL RSRPAQLVL  
 bc007312 .....

121 180  
 PanCAM-1 VPPEAPQVL GPSVSLVAGV PANLTCSRSG DARPTPELIM FRDGVLLDGA TTHQTLAKEG  
 PanCAM-2 VPPEAPQVL GPSVSLVAGV PANLTCSRSG DARPTPELIM FRDGVLLDGA TTHQTLAKEG  
 PanCAM VPPEAPQVL GPSVSLVAGV PANLTCSRSG DARPTPELIM FRDGVLLDGA TTHQTLAKEG  
 bc007312 .....

181 240  
 PanCAM-1 TPQSVESTLT LTPESHDDGA TLWCARSDA LPTGRDTAIT LSLQYPPVET LSASPTVQE  
 PanCAM-2 TPQSVESTLT LTPESHDDGA TLWCARSDA LPTGRDTAIT LSLQYPPVET LSASPTVQE  
 PanCAM TPQSVESTLT LTPESHDDGA TLWCARSDA LPTGRDTAIT LSLQYPPVET LSASPTVQE  
 bc007312 .....

241 300  
 PanCAM-1 GEKVIFLQQA TAQPPVTGYR WAKGSPVLG ARGPRLEVVA DASFLTEPVS CEVSNVAGSA  
 PanCAM-2 GEKVIFLQQA TAQPPVTGYR WAKGSPVLG ARGPRLEVVA DASFLTEPVS CEVSNVAGSA  
 PanCAM GEKVIFLQQA TAQPPVTGYR WAKGSPVLG ARGPRLEVVA DASFLTEPVS CEVSNVAGSA  
 bc007312 .....

301 360  
 PanCAM-1 NRSTALDVL GPILQAKPEP VSDVGDGAS FSCAMRGNPL PRVTWTRGG AQVLSSGATL  
 PanCAM-2 NRSTALDVL GPILQAKPEP VSDVGDGAS FSCAMRGNPL PRVTWTRGG AQVLSSGATL  
 PanCAM NRSTALDVL GPILQAKPEP VSDVGDGAS FSCAMRGNPL PRVTWTRGG AQVLSSGATL  
 bc007312 .....

361 420  
 PanCAM-1 RLPSVGFEDA DQYVCAEAG LSLRGGAEE ARLTVNAPPV VTALRSAPAF LRGPARIQCL  
 PanCAM-2 RLPSVGFEDA DQYVCAEAG LSLRGGAEE ARLTVNAPPV VTALRSAPAF LRGPARIQCL  
 PanCAM RLPSVGFEDA DQYVCAEAG LSLRGGAEE ARLTVNAPPV VTALRSAPAF LRGPARIQCL  
 bc007312 .....

421 480  
 PanCAM-1 VFASAPADAV VMSWDEGFLE AGSQORFLVE TFPAPESRGG LGPGLISVIR ISQTESDFS  
 PanCAM-2 VFASAPADAV VMSWDEGFLE AGSQORFLVE TFPAPESRGG LGPGLISVIR ISQTESDFS  
 PanCAM VFASAPADAV VMSWDEGFLE AGSQORFLVE TFPAPESRGG LGPGLISVIR ISQTESDFS  
 bc007312 .....

481 540  
 PanCAM-1 RSPNCSARNR LGEQGAQASL GRDDLPTVR IVAGVVAATT TLMVITGVA LCCWRHSKAS  
 PanCAM-2 RSPNCSARNR LGEQGAQASL GRDDLPTVR IVAGVVAATT TLMVITGVA LCCWRHSKAS  
 PanCAM RSPNCSARNR LGEQGAQASL GRDDLPTVR IVAGVVAATT TLMVITGVA LCCWRHSKAS  
 bc007312 ..... LGEQGAQASL GRR ..... AS

541 600  
 PanCAM-1 ASFSSEQNLM RIPGSSDGSS SRGPEEETG SREDRGPVH TMSDLVLEE EGTLETQDPT  
 PanCAM-2 ASFSSEQNLM RIPGSSDGSS SRGPEEETG SREDRGPVH TMSDLVLEE EGTLETQDPT  
 PanCAM ASFSSEQNLM RIPGSSDGSS SRGPEEETG SREDRGPVH TMSDLVLEE EGTLETQDPT  
 bc007312 .....

601 660  
 PanCAM-1 NGYTKVRGVS PPASPDSEVT SPQWSPGIS NLP .....  
 PanCAM-2 NGYTKVRGVS PPASPDSEVT SPQWSPGIS NLP .....

## 【図 9 A - 2】

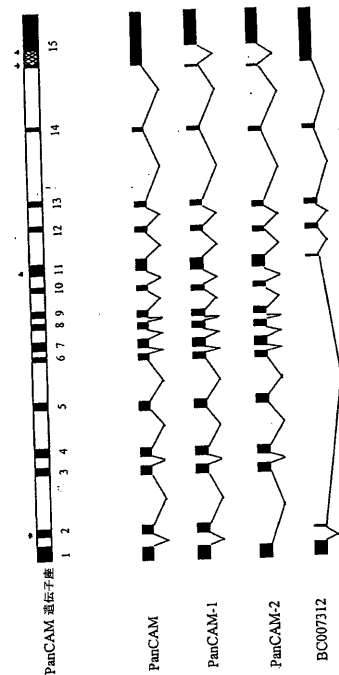
FIGURE 9A (つづき)

PanCAM NGYTKVRGVS VSLSLGEAPG GGLFLPPSP LGPRGPTTFY DFNPHLMVP PCRLYRARG  
 bc007312 NGYTKVRGVS VSLSLGEAPG GGLFLPPSP LGPRGPTTFY DFNPHLMVP PCRLYRARG

661 708  
 PanCAM-1 .....  
 PanCAM-2 .....  
 PanCAM YLITTEHPRAF TSVIKPTSPG PDLAGTTP PPYAFPTPS HPRLOTEV  
 bc007312 YLITTEHPRAF TSVIKPTSPG PDLAGTTP PPYAFPTPS HPRLOTEV

## 【図 9 B】

FIGURE 9B





## 【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  
3 January 2003 (03.01.2003)

PCT

(10) International Publication Number  
WO 03/000021 A2

- (51) International Patent Classification: Not classified (74) Agents: HALEY, James, F. et al.; Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).
- (21) International Application Number: PCT/US02/20275
- (22) International Filing Date: 24 June 2002 (24.06.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
PCT/US01/19904 22 June 2001 (22.06.2001) US
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:  
US PCT/US01/19904 (CIP)  
Filed on 22 June 2001 (22.06.2001)
- (71) Applicant (for all designated States except US): BIO-GEN, INC. [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US).
- (72) Inventors: and
- (75) Inventors/Applicants (for US only): SUN, Chao [CN/US]; 11 Clinton Road, Needham, MA 02494 (US). CARULLI, John, P. [US/US]; 9 Harris Drive, Southborough, MA 01772 (US). LUKASHIN, Alexander, V. [US/US]; 42 8th Street, #4403, Charlestown, MA 02129 (US). KILBURN, Daniel, R. [US/US]; 6 Hancock Avenue, Lexington, MA 02420 (US).
- (81) Designated States (national): AI, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GI, 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, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NI, SN, TD, TG).
- Published:**  
without international search report and to be republished upon receipt of that report  
with (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description
- 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 03/000021 A2

(54) Title: PANCAM NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: An isolated polynucleotide encoding a novel immunoglobulin superfamily member named PanCAM (also named GP354) is provided. PanCAM has a predicted single membrane spanning domain and five immunoglobulin (Ig) domains in the extracellular portion of the protein. The protein structure and tissue distribution of PanCAM indicate that it plays a role in cell-cell recognition, binding, signaling and adhesion events in the pancreas (especially the beta islets), the pituitary gland, and the central nervous system (CNS). Provided by the invention are isolated PanCAM related polynucleotides and polypeptides, vectors, and host cells comprising any of the above, antibodies directed to PanCAM, cells which produce such antibodies, and related diagnostic and therapeutic methods.

WO 03/000021

PCT/US02/20275

**PANCA M NUCLEIC ACIDS AND POLYPEPTIDES****RELATED APPLICATIONS**

This application claims priority from PCT/US01/19904, filed June 22, 2001, which claims priority from United States Provisional Application No. 60/213,611, filed June 22, 2000. The disclosures of both the PCT and provisional applications are incorporated herein by reference in their entirety.

**FIELD OF THE INVENTION**

The present invention relates generally to the field of molecular biology. More particularly, this invention relates to members of the immunoglobulin superfamily.

**BACKGROUND OF THE INVENTION**

Many proteins have been classified into superfamilies based on conserved structural motifs and biological functions. A superfamily is broadly defined as a group of proteins that share a certain degree of sequence homology, usually at least 15%. The conserved sequences shared by superfamily members often contribute to the formation of compact tertiary structures referred to as domains, and often the entire sequence of a domain characteristic of a particular superfamily is encoded by a single exon (see, e.g., Abbas *et al.*, CELLULAR AND MOLECULAR IMMUNOLOGY, W.B. Saunders Co., Philadelphia, PA. 1997). Members of a superfamily are likely derived from a common precursor gene by divergent evolution, and multidomain proteins may belong to more than one

WO 03/000021

PCT/US02/20275

-2-

superfamily. Examples of protein superfamilies include the ligand-gated ion channel receptor superfamily, the voltage-dependent ion channel receptor superfamily, the receptor tyrosine kinase superfamily, the receptor protein tyrosine phosphatase superfamily, the G protein-coupled receptor superfamily, and the immunoglobulin (Ig) superfamily.

The Ig superfamily encompasses proteins that share partial amino acid sequence homology and tertiary structural features that were originally identified in Ig heavy and light chains. The common structural motif of the Ig superfamily is the so-called "Ig domain." Ig domains are three-dimensional globular structures having about 70 to 110 amino acid residues and an internal Cys-Cys disulfide bond. These domains contain two layers of  $\beta$ -pleated sheet, each layer composed of three to five antiparallel strands of five to ten amino acid residues. Ig domains are classified as V-like or C-like on the basis of closest homology to either the Ig V or C domains. For a general review, see, e.g., Abbas et al., *supra*.

Most identified members of the Ig superfamily are integral plasma membrane proteins with Ig domains in the extracellular portions and widely divergent cytoplasmic tails, usually with no intrinsic enzymatic activity. One recurrent characteristic of the Ig superfamily members is that interactions between Ig domains on different polypeptide chains (of the same or different amino acid sequences) are essential for the biological activities of the molecules. Heterophilic interactions can also occur between Ig domains on entirely distinct molecules expressed on the surfaces of different cells. Such interactions provide adhesive forces that stabilize cell-cell binding.

Many members of the Ig superfamily are cell surface or soluble molecules that mediate cell recognition, adhesion and binding functions in the vertebrate immune system. Two prominent cell types that produce Ig superfamily molecules are B and T lymphocytes. Exemplary Ig superfamily member proteins of importance in the immune system include antibodies, T cell receptors, Class I and II major histo-compatibility complex (MHC) molecules, CD2, CD3, CD4, CD5, CD8, CD28, CD20 (B1), CD32 (Fc $\gamma$ RII), CD44, CD54 (ICAM-1), CD80

WO 03/000021

PCT/US02/20275

-3-

(B7-1), CD86 (B7-2), CD90 (Thy-1), CD102 (ICAM-2), CD106 (VCAM-1), CD121 (IL-1R), CD152 (CTLA-4), p-IgR, NCAM, and CD140 (PDGFR) (Abbas et al., *supra*).

Several Ig superfamily members have been identified outside the immune system, for instance, in the nervous system. Based on their conserved structural motifs and the well known functions of such motifs in the immune system, these Ig superfamily members likely perform cell recognition, binding and adhesion functions in non-immune tissues as well. Novel Ig superfamily members localized to particular cell types will be useful cell and tissue markers for diagnostic purposes. Tissue specific Ig superfamily members will also be suitable therapeutic targets for treating abnormal conditions, disorders and/or diseases related to improper cell-cell adhesion and signaling in the tissue, particularly during tissue development or during tissue regeneration, e.g., after tissue damage or trauma.

#### 15 **SUMMARY OF THE INVENTION**

The present invention is based, at least in part, on the discovery of a gene encoding a heretofore unknown Ig superfamily member, termed GP354. Unless indicated otherwise, the name in lower case, gp354, refers to the new nucleic acids of the invention, whereas the name in uppercase, GP354, refers to the new polypeptides of the present invention.

GP354 is an integral membrane protein predominantly expressed in the pancreas. This protein is also named PanCAM, for pancreas cell adhesion molecule. In the pancreas, GP354 (i.e., PanCAM) expression is localized to  $\beta$  cells. GP354 is also detected in low levels in the central nervous system (CNS), such as the brain. GP354 has a predicted single membrane spanning domain and five extracellular Ig domains. Human GP354 shares about 44% identity to a 494 amino acid fragment of human NEPH1 and 47% amino acid identity with a 368 amino acid fragment of mouse NEPH1. Both fragments are in the extracellular portion of NEPH1. GP354 is likely to play a role in cell-cell interactions (e.g.,

WO 03/000021

PCT/US02/20275

-4-

adhesion) in the pancreas and the CNS, affecting these organs' structures and functions.

The invention provides isolated polynucleotides encoding a vertebrate GP354 (*e.g.*, mammalian such as human or mouse) or a biologically active portion thereof. This invention also provides polynucleotide fragments suitable for use as primers or hybridization probes for the detection of GP354-encoding polynucleotides. Even though this disclosure exemplifies particular GP354 or gp354 sequences, all allelic variations of these sequences (*e.g.*, polymorphism) are encompassed by this invention.

In some embodiments, the invention provides an isolated human GP354 polypeptide comprising amino acid residues 68-594 of SEQ ID NO:12. Such an isolated polypeptide may comprise SEQ ID NO:12, 26, 28, or 8. The polypeptide may further comprise a heterologous sequence. The invention also provides an isolated mouse PanCAM polypeptide comprising SEQ ID NO:31.

The invention also provides an isolated polypeptide comprising an extracellular domain of a PanCAM protein, for instance, comprising amino acid residues 4-513 of SEQ ID NO:8, or a portion thereof having at least 10 amino acid (*e.g.*, 25, or 50) residues. This polypeptide can be used for raising PanCAM-specific antibodies or used as PanCAM modulators.

The invention also provides an isolated polypeptide comprising an intracellular domain of a PanCAM protein, for instance, comprising amino acid residues 535-708 of SEQ ID NO:8, or amino acid residues 535-633 of SEQ ID NO:26, or a portion thereof having at least 10 amino acid (*e.g.*, 25, or 50) residues. This polypeptide can be conjugated to a heterologous extracellular domain to form a hybrid molecule. The hybrid molecule can transmit a PanCAM-like intracellular signal via stimulation of the nonPanCAM extracellular domain.

The invention provides isolated polynucleotides comprising nucleotide sequences encoding the above-described isolated polypeptides. Exemplary isolated polynucleotides are those comprising part (*e.g.*, nucleotides 202-1782) or all of SEQ ID NO:11; part (*e.g.*, nucleotides 196-2319) or all of SEQ ID NO:7; and part or all of SEQ ID NOs:1, 3, 5, 6, 9, 25, 27, and 32. In some

WO 03/000021

PCT/US02/20275

-5-

embodiments, the polynucleotides further comprise a transcription regulatory sequence operatively linked to said nucleotide sequence.

Also included in this invention are vectors comprising these isolated polynucleotides. The vectors may be plasmid vectors, or viral vectors derived from  
5 baculoviruses, adenoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, lentiviruses and retroviruses.

The invention also includes host cells harboring the isolated polynucleotides of this invention. The host cells can be, *e.g.*, bacterial cells, insect  
10 cells, yeast cells, plant cells or mammalian cells (*e.g.*, mouse cells, or human cells). In some embodiments, these host cells can be used to produce a PanCAM polypeptide. To do this, the host cells are cultured under conditions in which the PanCAM-encoding sequence is expressed, and recovering the polypeptide from the cells or from the culture medium.

15 This invention also features antibodies that binds specifically to a PanCAM protein. The antibodies can be whole antibodies or fragments thereof (*e.g.*, Fab fragments, F(ab')<sub>2</sub> fragments). They can also be singled-chained antibodies, chimeric antibodies, humanized and fully human antibodies.

The invention also features a method of isolating beta cells from a  
20 dispersed pancreatic cell population (*e.g.*, a bovine or swine pancreatic cell population). This method comprises the steps of providing an antibody specific for a PanCAM protein; contacting the dispersed pancreatic cell population with the antibody under conditions that allow beta cells in the population to bind to the antibody; and separating cells unbound by the antibody from beta cells bound to  
25 the antibody, thereby isolating beta cells.

The beta cells so isolated can be used to treat a human displaying a sign or symptom of diabetes or a pre-diabetic condition. This treatment method comprise the steps of introducing the isolated beta cells into the human, *e.g.*, subcutaneously. The beta cells can be encapsulated in a biological acceptable carrier such as  
30 alginate spheres (see, *e.g.*, Ryu *et al.*, *J. Clin. Invest.* 2001, 108(1): 31-3).

WO 03/000021

PCT/US02/20275

-6-

This invention further features a polynucleotide that includes a nucleotide sequence which encodes a protein that comprises an amino acid sequence that is at least 80% (85%, 95% or 98%) identical to the amino acid sequence of SEQ ID NO:2 (encoded by a predicted gp354 cDNA); SEQ ID NO:4 (encoded by a partial gp354 pancreatic cDNA); SEQ ID NO:8 (encoded by a derived gp354 cDNA); SEQ ID NO:10 (encoded by a partial derived gp354 cDNA); or SEQ ID NO:12 (encoded by a gp354 pancreatic cDNA, an insert of plasmid FL86-3, which was deposited at American Type Culture Collection (ATCC) on June 11, 2002); or to at least one Ig domain of any one of SEQ ID NOs:2, 4, 8, 10 and 12.

In some embodiments, the polynucleotide comprises the sequence of SEQ ID NO:1 (a gp354 cDNA), or a fragment thereof having at least 17 nucleic acid units (e.g., nucleotides). An example of such a fragment is SEQ ID NO:3. In another embodiment, a polynucleotide comprises the sequence of SEQ ID NO:5 (genomic DNA comprising gp354), or a fragment thereof having at least 17 nucleic acid units. An exemplary fragment is that of SEQ ID NO:6 (gp354 upstream genomic DNA). In other embodiments, a polynucleotide comprises the sequence of SEQ ID NO:7 (a derived gp354 cDNA), or a fragment thereof having at least 17 nucleic acid units. An exemplary fragment is that of SEQ ID NO:9 (C-terminal fragment of a derived gp354 cDNA). In other embodiments, a polynucleotide comprises the sequence of SEQ ID NO:11 (pancreatic gp354 cDNA), or a fragment thereof having at least 17 nucleic acid units. Preferred fragments encode part or all of at least one extracellular Ig domain and/or an intracellular domain of GP354.

The invention also provides a polynucleotide which encodes a naturally occurring, allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the nucleic acid hybridizes to SEQ ID NO:1 or SEQ ID NO:11 under stringent conditions. The invention also provides a polynucleotide which encodes a naturally occurring, allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:4, 8, 10 or 12, wherein the nucleic acid hybridizes to SEQ ID NO:1 or 11 under stringent conditions.

WO 03/000021

PCT/US02/20275

-7-

Also provided by the invention is an isolated GP354 protein comprising an amino acid sequence that is at least 80% (85%, 95% or 98%) identical to the amino acid sequence of SEQ ID NO:2, 4, 8, 10 or 12; or to an Ig domain encoded by any one of those sequences.

5 The invention also provides an isolated GP354 protein encoded by a polynucleotide comprising a sequence which is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:1, 3, 5, 7, 9 or 11; or to a portion of any one of those sequences that encodes at least one Ig domain. Also provided is an isolated GP354 protein encoded by a polynucleotide having a sequence which  
10 hybridizes under stringent conditions to a nucleic acid having the sequence of SEQ ID NO:1 or 11.

The invention provides gp354 polynucleotides that specifically detect gp354 nucleic acids relative to nucleic acids encoding other members of the Ig superfamily. The invention also provides a nucleic acid construct, e.g., a  
15 recombinant vector (*e.g.*, a cloning, targeting or expression vector), comprising a gp354 polynucleotide of the invention.

Isolated or recombinant GP354 proteins and polypeptides are provided by the invention. In some embodiments, GP354 proteins and polypeptides display at least one of the following (overlapping) biological  
20 activities: (1) interaction with (*e.g.*, binding to) a ligand (*e.g.*, a protein receptor, a polysaccharide, etc.) that naturally binds to GP354 protein; (2) binding to an auto-antibody to naturally occurring human GP354 or an antibody raised against naturally occurring human GP354; (3) participation in a pancreatic function (*e.g.*, a  
25 signal transduction function in the pancreas or a step in the organ development of the pancreas); (4) participation in a neural function (*e.g.*, a signal transduction function in the nervous system or step in the development of the nervous system); and (5) mediation of cell-cell interactions such as recognition, binding and/or adhesion.

The GP354 proteins or biologically active portions thereof can be  
30 operably linked to a non-GP354 polypeptide (*e.g.*, heterologous amino acid



WO 03/000021

PCT/US02/20275

-8-

sequences, such as sequences that facilitate protein stability, detection, purification, or *in vivo* delivery to target cells) to form GP354 fusion proteins.

The invention provides pharmaceutical compositions comprising at least one of the above-described gp354-related isolated polynucleotides, GP354  
5 proteins or biologically active portions thereof, antibodies or fusion proteins; which optionally include pharmaceutically acceptable carriers. Such compositions are useful in therapeutic methods for ameliorating conditions in a subject associated with abnormal GP354 cellular localization, expression and/or activity.

As such, the present invention also provides methods of treatment  
10 comprising the step of administering a gp354-related compound or composition of the invention. Such methods will be useful, for example, for treating abnormal conditions, disorders or diseases which correlate with cell recognition, binding, signaling and adhesion functions in the developing or adult pancreas and central nervous system.

As a pancreatic enriched protein, GP354 will be a suitable  
15 therapeutic target for treating abnormal conditions, disorders and/or diseases related to improper cell-cell binding, adhesion and signaling in the developing and adult pancreas, particularly during tissue development and during tissue regeneration and/or healing, *e.g.*, after pancreatic damage, trauma or degenerative conditions. It is also envisioned that GP354 will be a suitable therapeutic target for  
20 inhibiting pancreatic cell death associated with immune, auto-immune, and degenerative conditions (including type I and type II diabetes). The neural form of GP354 will be a similarly suitable therapeutic target for treating tissue abnormalities, for tissue regeneration and repair, and for inhibiting tissue  
25 degeneration and cell death in the central nervous system. The GP354-related compositions can also be used to treat diseases in the pituitary gland such as pituitary hormone deficiency, growth hormone deficiency, and pituitary tumors.

The invention provides a method for modulating GP354 activity. In  
this method, a target cell is contacted with an agent that modulates (*e.g.*, inhibits or  
30 stimulates) GP354 activity or expression such that the GP354 activity or expression is altered. In some embodiments, the agent is an antibody that

WO 03/000021

PCT/US02/20275

-9-

specifically binds to GP354. In other embodiments, the agent modulates the GP354 activity or expression by modulating transcription of a gp354 gene, splicing of gp354 RNA, or translation of a gp354 mRNA. In yet other embodiments, the agent is a nucleic acid having a sequence that is antisense to the coding strand of the gp354 mRNA or the gp354 gene. In other embodiments, the agent can be a GP354 protein, a nucleic acid encoding a GP354 protein, or an antagonist or agonist of the GP354 protein such as a peptide, a peptidomimetic, or other small molecules.

The invention also provides a method for identifying a compound that binds to a GP354 protein. In another aspect, the invention provides a method for identifying a compound that modulates the biological activity of a GP354 protein, comprising measuring a biological activity or expression of the protein in the presence and absence of a test compound and identifying those compounds which alter the activity of the protein. Combinatorial libraries can be used as sources of candidate compounds in these methods.

The invention provides a method for detecting the presence of a gp354 polynucleotide, a GP354 protein or its activity in a biological sample (*e.g.*, a fluid or tissue sample derived from a patient) by contacting the sample with an agent capable of detecting an indicator of the presence of gp354 polynucleotide sequences, GP354 protein or its activity.

A diagnostic assay is provided for identifying the presence or absence of a gp354-related genetic lesion or mutation, characterized by at least one of the following: (i) aberrant modification or mutation of a gene encoding a GP354 protein; (ii) mis-regulation (*e.g.*, transcription, splicing or translation) of a gene encoding a GP354 protein; and (iii) aberrant post-translational modification or localization of a GP354 protein; wherein the wild-type form of the gene encodes a protein with a GP354 biological activity.

The invention provides a non-human animal (*e.g.*, a mammal such as a mouse, rat, guinea pig, sheep, goat, horse or cow) at least some cells of which comprise an isolated polynucleotide of this invention. Such an animal can be chimeric where only some of its somatic and/or germ cells carry the

WO 03/000021

PCT/US02/20275

-10-

polynucleotide. Such an animal can alternatively be transgenic where all of its somatic and germ cells carry the polynucleotide.

The invention also provides a non-human animal whose endogenous ortholog of the gp354 gene is disrupted by gene targeting (i.e., "knocked out").

- 5 Cells containing a gp354 polynucleotide, biological samples such as tissues and fluids and GP354-related products derived from these and the above-mentioned animals are also within the scope of this invention.

- 10 The invention provides a computer readable means of storing the nucleic acid and amino acid sequences of the instant invention. The records of the computer readable means can be accessed for reading and display of sequences and for comparison, alignment and ordering of the sequences of the invention to other sequences.

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

15 **BRIEF DESCRIPTION OF THE DRAWINGS**

**FIG. 1** The coding (SEQ ID NO:1) and deduced amino acid sequences (SEQ ID NO:2) of a human GP354 (PanCAM) polypeptide. The immunoglobulin (Ig) domains in the extracellular portion and the transmembrane domain are underlined.

- 20 **FIG. 2** The alignment of the human GP354 amino acid sequence shown in Fig. 1 (SEQ ID NO:2) with sequences of *Drosophila* irregular chiasm (ICCR) (SEQ ID NO:13) and human nephrin (SEQ ID NO:14) proteins. Dashes indicate gaps in any of the sequences. Asterisks denote amino acids that are identical in the three sequences.

- 25 **FIG. 3** Sequence of the RT-PCR fragment obtained using primers GX1-218 (SEQ ID NO:16) and GX1-219 (SEQ ID NO:17). This fragment (SEQ ID NO:3) is an insert of plasmid CS0026, which was deposited at ATCC on June 11, 2002.

**FIG. 4** A human genomic gp354 sequence (SEQ ID NO:5). Exons are underlined.

WO 03/000021

PCT/US02/20275

-11-

- FIG. 5** The cDNA (SEQ ID NO:7) and derived amino acid sequences (SEQ ID NO:8) of a human PanCAM. SEQ ID NO:7 is generated based on SEQ ID NO:12 and the coding sequences for PanCAM-1 (an insert of plasmid CS0067-1, which was deposited at ATCC on June 11, 2002 (SEQ ID NO:25)) and for
- 5 PanCAM-2 (an insert of plasmid CS0067-3, which was deposited at ATCC on June 11, 2002 (SEQ ID NO:27)). SEQ ID NO:7 includes 5' and 3' untranslated regions.
- FIG. 6** The nucleotide (SEQ ID NO:11) and deduced amino acid sequences (SEQ ID NO:12) of a pancreatic gp354 cDNA.
- 10 **FIG. 7** The nucleotide sequence (SEQ ID NO:25; an insert of plasmid CS0067-1) of a representative GP354 cDNA clone. The deduced amino acid sequence (SEQ ID NO:26; termed PanCAM-1) is shown below the corresponding nucleotide sequence. The signal peptide is double-underlined, the transmembrane domain is boxed, and the five Ig domains are underlined. The sequences
- 15 corresponding to oligonucleotides used for amplification of the original *ab initio* gene prediction are underscored with thick lines, and they bracket SEQ ID NO:11.
- FIG. 8** Nucleotide (SEQ ID NO:27) and deduced amino acid (SEQ ID NO:28) sequences of PanCAM-2.
- FIG. 9A** Alignment of proteins encoded by alternatively spliced transcripts of
- 20 PanCAM: PanCAM-1 (SEQ ID NO:26), PanCAM-2 (SEQ ID NO:28), PanCAM (SEQ ID NO:8), and bc007312 (SEQ ID NO:29). The conserved Ig domains are underlined.
- FIG. 9B** A diagram of the PanCAM locus and alternatively spliced mRNAs. Exons are numbered 1-15 and shown as filled boxes, and exons with cryptic
- 25 splicing sites are marked with asterisks. The shaded area in the PanCAM locus represents a cryptic intron, which is a portion of an exon spliced out in PanCAM-1 and PanCAM-2 transcripts.
- FIG. 10A** Nucleotide (SEQ ID NO:32) and deduced amino acid (SEQ ID NO:31) sequences of a partial mouse PanCAM cDNA clone.

WO 03/000021

PCT/US02/20275

-12-

**FIG. 10B** Alignment of a partial protein sequence of human PanCAM (SEQ ID NO:30) with a partial protein sequence of murine PanCAM (SEQ ID NO:31). Immunoglobulin domains are underlined. The mouse sequence is at the bottom.

#### **DETAILED DESCRIPTION OF THE INVENTION**

5 The present invention is based, at least in part, on the discovery of a novel human gene encoding a heretofore unknown protein, GP354 (i.e., PanCAM). This gene, gp354, was identified by computational analysis of ("mining") the published nucleic acid sequences of the human genome. The gp354 gene contains at least 14 exons and normally resides on human chromosome 19. An mRNA  
10 transcribed from this gene has an open reading frame of 1779 base pairs (including the stop codon) and encodes a protein predicted to be 592 amino acid residues. The novel GP354 protein is specifically expressed in the pancreas and the brain.

#### **DEFINITIONS**

As used herein, "nucleic acid" (also "polynucleotide") includes  
15 DNA molecules (e.g., cDNA or genomic or synthetic DNA) and RNA molecules (e.g., mRNA). The term also is intended to include analogs of DNA or RNA containing non-natural nucleotide analogs, non-native internucleoside bonds, or both. The nucleic acid can be in any topological conformation. For instance, the nucleic acid can be single-stranded, double-stranded, triple-stranded, quadruplexed,  
20 partially double-stranded, branched, hairpinned, circular, or in a padlocked conformation. See, e.g., Banér *et al.*, *Curr. Opin. Biotechnol.* 12:11-15 (2001); Escude *et al.*, *Proc. Natl. Acad. Sci. USA* 14:96(19):10603-7 (1999); Nilsson *et al.*, *Science* 265(5181):2085-8 (1994); Praseuth *et al.*, *Biochim. Biophys. Acta* 1489(1):181-206 (1999); Fox, *Curr. Med. Chem.* 7(1):17-37 (2000); Kochetkova *et al.*, *Methods Mol. Biol.* 130:189-201 (2000); Chan *et al.*, *J. Mol. Med.* 75(4):267-82 (1997).

As used herein, an "isolated nucleic acid" (also "isolated polynucleotide") is one which is separated from other nucleic acid molecules that are present in the natural source of the nucleic acid. Specifically excluded are

WO 03/000021

PCT/US02/20275

-13-

isolated, non-recombinant native chromosomes and fragments thereof that are larger than 500 kilobases. Preferably, an "isolated" nucleic acid is substantially free of sequences that naturally flank that nucleic acid in the genome of the organism from which the nucleic acid is derived. For example, a preferred isolated gp354  
5 nucleic acid is flanked by less than about 10 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences that naturally flank the nucleic acid in the genomic DNA of the cell from which the isolated nucleic acid is derived. Even more preferably, the isolated polynucleotides are no more than 5000 base pairs, often no more than 1000 base pairs, 500 base pairs, 100 base pairs or 50 base pairs.

10 However, "isolated" does not necessarily require that the nucleic acid or polynucleotide so described has itself been physically removed from its native environment. For instance, an endogenous nucleic acid sequence in the genome of an organism is deemed "isolated" herein if a heterologous sequence (i.e., a sequence that is not naturally adjacent to this endogenous nucleic acid  
15 sequence) is placed adjacent to the endogenous nucleic acid sequence, such that the expression of this endogenous nucleic acid sequence is altered. By way of example, a non-native promoter sequence can be substituted (e.g., by homologous recombination) for the native promoter of a gp354 gene in the genome of a human cell, such that this gene has an altered expression pattern. This gene would now  
20 become "isolated" because it is separated from at least some of the sequences that naturally flank it.

A nucleic acid is also considered "isolated" if it contains any modifications that do not naturally occur to the corresponding nucleic acid in a genome. For instance, an endogenous gp354-coding sequence is considered  
25 "isolated" if it contains an insertion, deletion or a point mutation introduced artificially, e.g., by human intervention. An "isolated nucleic acid" also includes a nucleic acid integrated into a host cell chromosome at a heterologous site, a nucleic acid construct present as an episome and a nucleic acid construct integrated into a host cell chromosome. Moreover, an "isolated nucleic acid" can be substantially  
30 free of other cellular material, or substantially free of culture medium when

WO 03/000021

PCT/US02/20275

-14-

produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A polynucleotide of the invention is considered "full-length" if it is able to encode a full-length GP354 protein.

5 As used herein, the phrase "degenerate variant" of a reference nucleic acid sequence encompasses nucleic acid sequences that can be translated, according to the standard genetic code, to provide an amino acid sequence identical to that translated from the reference nucleic acid sequence.

10 As used herein, the term "microarray" (also "nucleic acid microarray") refers to a substrate-bound plurality of nucleic acids, hybridization to each of the bound nucleic acids being separately detectable. The substrate can be solid or porous, planar or non-planar, unitary or distributed, or in any other configuration.

15 As so defined, the term "microarray" includes all the devices so called or similarly called in Schena (ed.), DNA Microarrays: A Practical Approach (Practical Approach Series), Oxford University Press (1999) (ISBN: 0199637768); *Nature Genet.* 21(1)(suppl):1-60 (1999); and Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing Company/BioTechniques Books Division (2000) (ISBN: 1881299376); Brenner *et al.*, *Proc. Natl. Acad. Sci. USA* 20 97(4):1665-1670 (2000). The disclosures of all of these references are incorporated herein by reference in their entireties.

25 As used herein with respect to nucleic acid hybridization, the term "probe" (also "nucleic acid probe" or "hybridization probe") refers to an isolated nucleic acid of known sequence that is, or is intended to be, detectably labeled. As used herein with respect to a nucleic acid microarray, the term "probe" (or equivalently "nucleic acid probe" or "hybridization probe") refers to the isolated nucleic acid that is, or is intended to be, bound to the substrate. In either such context, the term "target" refers to a nucleic acid intended to be bound to a probe by sequence complementarity.

30 Unless otherwise indicated, a "nucleic acid comprising SEQ ID NO:X" refers to a nucleic acid, at least a portion of which has either (i) the

WO 03/000021

PCT/US02/20275

-15-

sequence of SEQ ID NO:X, or (ii) a sequence complementary to SEQ ID NO:X. The choice between the two is dictated by the context. For instance, if the nucleic acid is used as a probe, the choice between the two is dictated by the requirement that the probe be complementary to the desired target.

5 For purposes herein, "high stringency conditions" are defined for solution phase hybridization as aqueous hybridization (*i.e.*, free of formamide) in 6X SSC (where 20X SSC contains 3.0 M NaCl and 0.3 M sodium citrate), 1% SDS at 65°C for 8-12 hours, followed by two washes in 0.2X SSC, 0.1% SDS at 65°C for 20 minutes. It will be appreciated by the skilled worker that hybridization at  
10 65°C will occur at different rates depending on a number of factors including the length and percent identity of the sequences which are hybridizing.

For microarray-based hybridization, standard "high stringency conditions" are defined as hybridization in 50% formamide, 5X SSC, 0.2 µg/µl poly(dA), 0.2 µg/µl human cot1 DNA, and 0.5% SDS, in a humid oven at 42°C  
15 overnight, followed by successive washes of the microarray in 1X SSC, 0.2% SDS at 55°C for 5 minutes, and then 0.1X SSC, 0.2% SDS, at 55°C for 20 minutes. For microarray-based hybridization, "moderate stringency conditions", suitable for cross-hybridization to mRNA encoding structurally- and functionally-related proteins, are defined to be the same as those for high stringency conditions but with  
20 reduction in temperature for hybridization and washing to room temperature (approximately 25°C).

As used herein, the terms "protein," "polypeptide," and "peptide" are used interchangeably to refer to a naturally-occurring or synthetic polymer of amino acids, irrespective of length, where amino acids here include naturally-  
25 occurring amino acids, naturally-occurring amino acid structural variants, and synthetic non-naturally occurring analogs that are capable of participating in peptide bonds. The terms "protein", "polypeptide", and "peptide" explicitly permit post-translational and post-synthetic modifications, such as N- or C-terminal amino acid cleavage reactions and glycosylation. The term "oligopeptide" herein  
30 denotes a protein, polypeptide, or peptide having 25 or fewer amino acid residues.



WO 03/000021

PCT/US02/20275

-16-

A protein, polypeptide, peptide or oligopeptide is considered "isolated" when it is encoded by an isolated polynucleotide; when it exists in a purity not found in nature, where purity can be adjudged with respect to the presence of other cellular material; and/or when it includes amino acid analogs or derivatives not found in nature or linkages other than standard peptide bonds. As thus defined, "isolated" does not necessarily require that the protein, polypeptide, peptide or oligopeptide so described has been physically removed from its native environment.

A protein, polypeptide, peptide or oligopeptide is considered "purified" herein when it is present at a concentration of at least 65% (e.g., at least 75%, 85% or 95%), as measured on a mass basis with respect to total protein in a composition. It is considered "substantially purified" when the concentration is at least 85%.

As used herein, the term "homologs" (also "homologues") encompasses "orthologs" and "paralogs." "Orthologs" are separate occurrences of the same gene in different species of organisms. The separate occurrences have similar or identical amino acid sequences, where the degree of sequence similarity depends in part on the evolutionary distance of the species from a common ancestor having the same gene. "Paralogs" indicates separate occurrences of a gene in one species of organism. The separate occurrences have similar or identical amino acid sequences, where the degree of sequence similarity depends in part on the evolutionary distance of these separate occurrences from the gene duplication event giving rise to the occurrences.

"Homologous" amino acid sequences include those amino acid sequences which contain conservative amino acid substitutions and which polypeptides have substantially the same binding and/or activity. A homologous amino acid sequence does not, however, include the amino acid sequence encoding other known Ig superfamily members. Homology (percent identity) can be determined by, for example, the GAP program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison WI), using the default settings, which uses the algorithm of Smith and

WO 03/000021

PCT/US02/20275

-17-

Waterman (*Adv. Appl. Math.*, 2:482-489 (1981), which is incorporated herein by reference in its entirety).

As used herein, the term "antibody" refers to a full antibody (consisting of two heavy chains and two light chains) or a fragment thereof. Such fragments include, but are not limited to, those produced by digestion with various proteases, those produced by chemical cleavage and/or chemical dissociation, and those produced recombinantly, so long as the fragment remains capable of specific binding to an antigen. Among these fragments are Fab, Fab', F(ab')<sub>2</sub>, and single chain Fv (scFv) fragments.

Within the scope of the term "antibody" are also antibodies that have been modified in sequence, but remain capable of specific binding to an antigen. Example of modified antibodies are interspecies chimeric and humanized antibodies; antibody fusions; and heteromeric antibody complexes, such as diabodies (bisppecific antibodies), single-chain diabodies, and intrabodies (see, e.g., Marasco (ed.), Intracellular Antibodies: Research and Disease Applications, Springer-Verlag New York, Inc. (1998) (ISBN: 3540641513), the disclosure of which is incorporated herein by reference in its entirety).

"Specific binding" refers to the ability of two molecules to bind to each other in preference to binding to other molecules in the environment.

Typically, "specific binding" discriminates over adventitious binding in a reaction by at least two-fold, more typically by at least 10-fold, often at least 100-fold. Typically, the affinity or avidity of a specific binding reaction is at least about 10<sup>-7</sup> M (e.g., at least about 10<sup>-8</sup> M or 10<sup>-9</sup> M).

By the term "region" is meant a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein.

The term "domain" refers to a structure of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof; domains may also include distinct, non-contiguous regions of a biomolecule. Examples of GP354 protein domains

WO 03/000021

PCT/US02/20275

-18-

include, but are not limited to, an extracellular Ig domain (i.e., N-terminal), a transmembrane domain, and a cytoplasmic domain (i.e., C-terminal).

As used herein, the term "compound" means any molecule, including, but not limited to, small molecule, peptide, protein, sugar, nucleotide, 5 nucleic acid, lipid, etc., and such a compound can be natural or synthetic.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Exemplary methods and materials are described below, although methods and materials similar or equivalent to those 10 described herein can also be used in the practice of the present invention and will be apparent to those of skill in the art. All publications and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. The materials, methods, and examples are illustrative only and not intended to be limiting.

Standard reference works setting forth the general principles of recombinant DNA technology known to those of skill in the art include Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York (1998 and Supplements to 2001); Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d Ed., Cold Spring Harbor 15 Laboratory Press, Plainview, New York (1989); Kaufman *et al.*, Eds., HANDBOOK OF MOLECULAR AND CELLULAR METHODS IN BIOLOGY AND MEDICINE, CRC Press, Boca Raton (1995); McPherson, Ed., DIRECTED MUTAGENESIS: A PRACTICAL APPROACH, IRL Press, Oxford (1991).

Standard reference works setting forth the general principles of immunology 25 known to those of skill in the art include: Harlow and Lane ANTIBODIES: A LABORATORY MANUAL, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1999); and Roitt *et al.*, IMMUNOLOGY, 3d Ed., Mosby-Year Book Europe Limited, London (1993). Standard reference works setting forth the general principles of medical physiology and pharmacology known to 30 those of skill in the art include: Harrison's PRINCIPLES OF INTERNAL

WO 03/000021

PCT/US02/20275

-19-

MEDICINE, 14<sup>th</sup> Ed., (Anthony S. Fauci *et al.*, editors), McGraw-Hill Companies, Inc., 1998.

#### GP354 RELATED NUCLEIC ACIDS

The gp354 gene was identified in contig 38 of a BAC clone with the  
 5 GenBank accession number AC022315, which was deposited on February 10, 2000. That deposit has the human genomic sequence of gp354 (Fig. 4 and SEQ ID NO:5), including 5' upstream (positions 1-6278) and 3' downstream (16490-20050) non-transcribed genomic sequences.

The invention provides isolated polynucleotides that encode the  
 10 entirety of the GP354 protein. As discussed above, such "full-length" polynucleotides of the present invention can be used, *inter alia*, to express full length GP354 protein. The full-length polynucleotides can also be used as nucleic acid probes; used as probes, the isolated polynucleotides of these embodiments will hybridize to gp354 polynucleotides and related polynucleotide sequences.

15 In preferred embodiments, the invention provides an isolated polynucleotide comprising part or all of (i) the nucleotide sequence of SEQ ID NO:1, 3, 5, 6, 7, 9, 11, 25, 27, or 32; (ii) a degenerate variant of thereof; or (iii) the complement of (i) or (ii).

SEQ ID NO:1 presents a predicted gp354 cDNA sequence. SEQ ID  
 20 NO:3 presents an exemplary gp354 cDNA fragment generated from RT-PCR. SEQ ID NO:5 presents the genomic DNA sequence comprising the gp354 coding sequences, including 5' and 3' non-transcribed regions. SEQ ID NO:6 presents the genomic DNA sequence comprising the upstream noncoding region. SEQ ID NO:7 presents a derived, long form of a gp354 cDNA sequence. SEQ ID NO:11  
 25 presents another exemplary pancreatic gp354 cDNA sequence. SEQ ID NOs:25 and 27 present the coding sequence of PanCAM-1 and PanCAM-2, respectively, both being alternative splice variants of SEQ ID NO:7.

In other embodiments, the invention provides an isolated polynucleotide comprising (i) a nucleotide sequence that encodes a polypeptide

WO 03/000021

PCT/US02/20275

-20-

with the amino acid sequence of SEQ ID NO:2, 4, 8, 10, 12, 26, 28 or 31; or (ii) the complement of a nucleotide sequence that encodes the above polypeptide.

SEQ ID NOs:2, 4, 8, 10, 12, 26, 28 and 31 present the GP354 amino acid sequences encoded by SEQ ID NOs:1, 3, 7, 9, 11, 25, 27 and 32, respectively.

5 See, e.g., Figs. 1-10.

In other embodiments, the invention provides an isolated polynucleotide having a nucleotide sequence that (i) encodes a polypeptide having the sequence of SEQ ID NO:2, 8 or 12, (ii) encodes a polypeptide having the sequence of SEQ ID NO:2, 8 or 12 with conservative amino acid substitutions, or  
10 (iii) that is the complement of (i) or (ii), where SEQ ID NO:2 presents the amino acid sequence of GP354 encoded by the cDNA of SEQ ID NO:1; SEQ ID NO:8 presents the amino acid sequence of GP354 encoded by sequences derived from SEQ ID NOs:5 and 11; and SEQ ID NO:12 presents the amino acid sequence of GP354 encoded by the pancreatic cDNA of SEQ ID NO:11.

#### 15 Nucleic Acids Encoding Portions Of GP354

The invention also provides isolated polynucleotides that encode select portions of GP354. As will be further discussed herein below, these "nucleic acid molecules" can be used, for example, to express specific portions of the GP354, either alone or as elements of a fusion protein. A nucleic acid fragment  
20 may also be used as a region-specific nucleic acid probe.

In preferred embodiments, the invention provides an isolated polynucleotide comprising (i) the nucleotide sequence of SEQ ID NO:3, 6 or 9, (ii) a degenerate variant of the nucleotide sequence of SEQ ID NO:3, 6 or 9, or (iii) the complement of (i) or (ii). SEQ ID NO:3 presents a 785 base pair RT-PCR  
25 fragment derived from gp354 pancreatic RNA. SEQ ID NO:6 presents genomic sequences upstream from gp354 coding sequences, and SEQ ID NO:9 presents a 1782 base pair RT-PCR fragment derived from gp354 pancreatic RNA.

In other embodiments, the isolated polynucleotide encodes, or the complement of which encodes, a polypeptide having, in at least one and preferably  
30 two, three, four or five of the Ig domains characteristic of the N-terminal

WO 03/000021

PCT/US02/20275

-21-

extracellular portion of GP354. Specifically, the five extracellular Ig domains are encoded by nucleotides 103-306, 406-609, 715-870, 967-1122 and 1228-1445, respectively, of the gp354 cDNA sequence of SEQ ID NO:1 (see Fig. 1) and by nucleotides 307-510, 610-813, 919-1074, 1171-1326 and 1432-1659, respectively, of the gp354 cDNA sequence of SEQ ID NO:8 (see Fig. 5). In preferred embodiments, the isolated polynucleotide encodes at least two, preferably three, more preferably four and most preferably all five domains in at least one copy.

For some uses, such as protein production, the nucleic acid fragments (or their complements) comprise sequences which encode a signal secretion sequence that will mediate transport of the encoded polypeptides through a membrane. Such a signal sequence is typically cleaved from the polypeptides as transport through the membrane occurs. The GP354 signal secretion sequence is encoded by nucleotides 1-54 of the gp354 cDNA sequence of SEQ ID NO:1 (see Fig. 1) and by nucleotides 1-57 of the gp354 cDNA of SEQ ID NO:8 (see Fig. 5). More preferably, the signal secretion sequence of the isolated polynucleotide of the invention is from gp354. Assuming that the signal sequence of GP354 is also cleaved during secretion, the mature GP354 polypeptide sequence has an N-terminal proline residue encoded by nucleotides 55-57 of SEQ ID NO:1 (see Fig. 1) and by nucleotides 259-261 of the gp354 cDNA of SEQ ID NO:8 (see Fig. 5).

Other preferred embodiments of the polynucleotides of the invention are those that encode, or the complements of which encode, a polypeptide having the transmembrane domain of GP354. The above preferred isolated polynucleotides, for example, may optionally encode a transmembrane domain, if insertion of the encoded polypeptides into a membrane is so-desired. The transmembrane domain may be encoded by gp354 sequences or may be encoded by a heterologous gene encoding a transmembrane domain of a heterologous membrane-associated protein. The gp354 transmembrane domain is encoded by nucleotides 1522-1590 of the gp354 cDNA sequence of SEQ ID NO:1 (see Fig. 1) and by nucleotides 1726-1794 of the gp354 cDNA of SEQ ID NO:8 (see Fig. 5).

WO 03/000021

PCT/US02/20275

-22-

If so-desired, the isolated polynucleotides of the invention may comprise sequences which encode (or their complements encode) an intracellular C-terminal domain, e.g., if specific signaling reactions are desired in response to GP354 binding interactions. The intracellular domain may be encoded by gp354  
 5 (see below) or may be encoded by a heterologous gene encoding an intracellular domain of a heterologous membrane-associated protein. Preferred polynucleotides of the invention are those that encode, or the complements of which encode, a polypeptide having a (C-terminal) intracellular domain of GP354. Specifically, one intracellular domain of GP354 is encoded by nucleotides 1591-1776 of the  
 10 gp354 cDNA sequence of SEQ ID NO:1 (see Fig. 1). A longer form of an intracellular domain of GP354 is encoded by nucleotides 1795-2319 of the gp354 cDNA sequence of SEQ ID NO:8 (see Fig. 5).

One preferred isolated polynucleotide of the invention is shown in Fig. 3 (see SEQ ID NO:3) and comprises nucleotides 139-923 of the gp354 cDNA  
 15 sequence of SEQ ID NO:1 (see Fig. 1). It comprises the sequence of an RT-PCR fragment amplified from pancreatic RNA using primers GX1-218 (SEQ ID NO:16) and GX1-219 (SEQ ID NO:17). See Example 2. This preferred isolated polynucleotide encodes amino acids 47-307 of SEQ ID NO:2, i.e., it encodes amino acids 13-68 of the first N-terminal Ig domain (i.e., it is missing the first 12  
 20 N-terminal amino acids of the Ig domain), and encodes the second and third Ig domains of GP354.

#### Cross-Hybridizing Nucleic Acids

In another series of nucleic acid embodiments, the invention provides isolated polynucleotides that hybridize to various of the gp354 nucleic  
 25 acids of the present invention. These "cross-hybridizing nucleic acids" can be used, *inter alia*, as probes for, and to drive expression of, proteins that are related to gp354 of the present invention as further isoforms, homologs, paralogs, or orthologs.

In some such embodiments, the invention provides an isolated  
 30 polynucleotide comprising a sequence that hybridizes under high stringency

WO 03/000021

PCT/US02/20275

-23-

conditions to a probe the nucleotide sequence of which comprises SEQ ID NO:1, 5, 7, 9, or 11; the complement of SEQ ID NO:1, 5, 7, 9, or 11; or a fragment thereof having at least 17 nucleic acid units.

#### 5 Preferred Nucleic Acids

Particularly preferred among the above-described nucleic acids are those that are expressed, or the complements of which are expressed, in pancreatic or neural tissues. Also particularly preferred among the above-described nucleic acids are those that encode, or the complements of which encode, a polypeptide

10 having a gp354 biological activity, as described *supra*.

#### Nucleic Acid Fragments

In another series of nucleic acid embodiments, the invention provides fragments of various of the isolated polynucleotides of the present invention which prove useful, *inter alia*, as region-specific nucleic acid probes, as

15 amplification primers, and to direct expression or synthesis of epitopic or immunogenic protein fragments.

In some embodiments, the invention provides an isolated polynucleotide comprising at least 17 nucleotides, 18 nucleotides, 20 nucleotides, 24 nucleotides, or 25 nucleotides of contiguous nucleic acid sequence selected

20 from SEQ ID NO:1, 5, 7, 9, or 11.

In other embodiments, the invention provides an isolated nucleic acid comprising a nucleotide sequence that (i) encodes a polypeptide having the sequence of at least eight contiguous amino acids of SEQ ID NO:2, 4, 8, 10 or 12 (ii) encodes a polypeptide having the sequence of at least eight contiguous amino

25 acids of SEQ ID NO:2, 4, 8, 10 or 12 with conservative amino acid substitutions, or (iii) is the complement of (i) or (ii).

#### Single Exon Probes

The invention further provides genome-derived single exon probes having portions of no more than one exon of the gp354 gene. Such single exon



WO 03/000021

PCT/US02/20275

-24-

probes have particular utility in identifying and characterizing splice variants. In particular, such single exon probes are useful for identifying and discriminating the expression of distinct isoforms of gp354.

- In some embodiments, the invention provides an isolated nucleic acid comprising a nucleotide sequence selected from one of the following exon-specific portions of SEQ ID NO:1, 5, 7, 9, or 11 or the complement of SEQ ID NO:1, 5, 7, 9, or 11, wherein the portion comprises at least 17 contiguous nucleotides, 18 contiguous nucleotides, 20 contiguous nucleotides, 24 contiguous nucleotides, 25 contiguous nucleotides, or 50 contiguous nucleotides of any one of the portions of SEQ ID NO:1, 5, 7, 9, or 11, or their complement:

TABLE 1: Exon coordinates of gp354 cDNA (SEQ ID NO:1 or 2) and genomic (SEQ ID NO:5) sequences

|            | <u>cDNA-1</u> | <u>cDNA-2</u> | <u>genomic</u> |
|------------|---------------|---------------|----------------|
| exon 1     | 1-52          | 1-52          | 6483-6534      |
| 15 exon 2  | 53-202        | 53-202        | 6699-6848      |
| exon 3     | 203-352       | 203-352       | 7762-7911      |
| exon 4     | 353-513       | 353-513       | 8058-8218      |
| exon 5     | 514-664       | 514-664       | 8835-8985      |
| exon 6     | 665-770       | 665-770       | 9651-9756      |
| 20 exon 7  | 771-919       | 771-919       | 9873-10021     |
| exon 8     | 920-1047      | 920-1041      | 10263-10390    |
| exon 9     | 1048-1180     | 1042-1180     | 10476-10608    |
| exon 10    | 1181-1281     | 1181-1281     | 10895-10995    |
| exon 11    | 1282-1501     | 1282-1501     | 11159-11378    |
| 25 exon 12 | 1502-1606     | 1502-1606     | 11847-11951    |
| exon 13    | 1607-1710     | 1607-1716     | 12287-12390    |
| exon 14    | 1711-1779     | 1717-1782     | 14002-14067    |

TABLE 2: Exon coordinates of gp354 cDNA-4 (SEQ ID NO:11) and genomic (SEQ ID NO:5) sequences

|  | <u>cDNA</u> | <u>genomic</u> |
|--|-------------|----------------|
|--|-------------|----------------|

WO 03/000021

PCT/US02/20275

-25-

|    |         |           |             |
|----|---------|-----------|-------------|
|    | Exon 1  | 1-256     | 6278-6534   |
|    | Exon 2  | 257-406   | 6699-6848   |
|    | Exon 3  | 407-556   | 7762-7911   |
|    | Exon 4  | 557-717   | 8058-8218   |
| 5  | Exon 5  | 718-868   | 8835-8985   |
|    | Exon 6  | 869-974   | 9651-9756   |
|    | Exon 7  | 975-1123  | 9873-10021  |
|    | Exon 8  | 1124-1245 | 10263-10390 |
|    | Exon 9  | 1246-1384 | 10476-10608 |
| 10 | Exon 10 | 1385-1485 | 10895-10995 |
|    | Exon 11 | 1486-1705 | 11159-11378 |
|    | Exon 12 | 1706-1810 | 11847-11951 |
|    | Exon 13 | 1811-1920 | 12281-12390 |
|    | Exon 14 | 1921-1986 | 14002-14067 |
| 15 | Exon 15 | 1987-2959 | 15511-16483 |

#### Transcription Control Nucleic Acids

In another aspect, the present invention provides genome-derived isolated polynucleotides which include nucleic acid sequence elements that control transcription of the gp354 gene. These nucleic acids can be used, *inter alia*, to

20 drive expression of heterologous coding regions in recombinant constructs, thus conferring upon such heterologous coding regions the expression pattern of the native gp354 gene. These nucleic acids can also be used, conversely, to target heterologous transcription control elements to the gp354 genomic locus, altering the expression pattern of the gp354 gene itself.

25 In a first series of such embodiments, the invention provides an isolated polynucleotide comprising nucleotides 1-6483 of SEQ ID NO:5; nucleotides 1483-6482 of SEQ ID NO:5; nucleotides 2483-6482 of SEQ ID NO:5; nucleotides 3483-6482 of SEQ ID NO:5; nucleotides 4483-6482 of SEQ ID NO:5; nucleotides 5483-6482 of SEQ ID NO:5; or nucleotides 5983-6482 of SEQ ID

30 NO:5; or the complements of such sequences.

WO 03/000021

PCT/US02/20275

-26-

In other embodiments, the invention provides an isolated polynucleotide comprising at least 17, 18, 20, 24, or 25 nucleotides of nucleotides 1-6483 of SEQ ID NO:5; nucleotides 1483-6482 of SEQ ID NO:5; nucleotides 2483-6482 of SEQ ID NO:5; nucleotides 3483-6482 of SEQ ID NO:5; nucleotides 4483-6482 of SEQ ID NO:5; nucleotides 5483-6482 of SEQ ID NO:5; or nucleotides 5983-6482 of SEQ ID NO:5; or the complements of such sequences.

Each of the isolated polynucleotides comprising nucleotides 1-6483 of SEQ ID NO:5; nucleotides 1483-6482 of SEQ ID NO:5; nucleotides 2483-6482 of SEQ ID NO:5; nucleotides 3483-6482 of SEQ ID NO:5; nucleotides 4483-6482 of SEQ ID NO:5; nucleotides 5483-6482 of SEQ ID NO:5; or nucleotides 5983-6482 of SEQ ID NO:5; or the complements of such sequences has transcription control sequences that mediate developmental and tissue specific expression and regulation of the gp354 gene. Such transcription control sequences will be useful for conferring such developmental and tissue specific expression patterns on heterologous nucleic acid sequences operatively linked thereto.

#### Other Defining Features of gp354 Nucleic Acid Molecules

All the nucleic acid sequences specifically given herein are set forth as sequences of deoxyribonucleotides. It is intended, however, that the given sequences be interpreted as would be appropriate to the polynucleotide composition: for example, if the isolated nucleic acid is composed of RNA, the given sequence intends ribonucleotides, with uridine substituted for thymidine.

Polymorphisms such as single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes. More than 1.4 million SNPs have already identified in the human genome, International Human Genome Sequencing Consortium, *Nature* 409:860-921 (2001) – and the sequence determined from one individual of a species may differ from other allelic forms present within the population. Additionally, small deletions and insertions, rather than single nucleotide polymorphisms, are not uncommon in the general population, and often do not alter the function of the protein.

-27-

Accordingly, it is particularly emphasized that the present invention not only provides isolated polynucleotides identical in sequence to those described with particularity herein (e.g., SEQ ID NOs:1, 3, 5, 6, 7, 9 and 11), but also to provide isolated polynucleotides that are allelic variants of those particularly  
 5 described nucleic acid sequences. Further, the invention provides homologs (e.g., paralogs and orthologs) of gp354 that are at least about 65% identical in sequence to SEQ ID NOs:1, 3, 5, 6, 7, 9 and 11, or to a portion of any one of those sequences that encodes at least one Ig domain, typically at least about 70%, 75%, 80%, 85%, or 90% identical in sequence, usefully at least about 91%, 92%, 93%, 94%, or 95%  
 10 identical in sequence, more usefully at least about 96%, 97%, 98%, or 99% identical in sequence, and, most conservatively, at least about 99.5%, 99.6%, 99.7%, 99.8% and 99.9% identical in sequence to those described with particularity herein. These sequence variants can be naturally occurring or can result from human intervention, as by random or directed mutagenesis.

15 Nucleic acid sequence variants have been found to occur, e.g., at positions 252, 703, 770, 1249 and 1811-1816 of the sequence presented in SEQ ID NO:7.

For purposes herein, percent identity of two nucleic acid sequences is determined using the procedure of Tatiana *et al.*, "Blast 2 sequences - a new tool  
 20 for comparing protein and nucleotide sequences", *FEMS Microbiol Lett.* 174:247-250 (1999), which procedure is effectuated by the computer program BLAST 2 SEQUENCES, available online at the National Center for Biotechnology Information.

To assess percent identity of nucleic acid sequences, the BLASTN  
 25 module of BLAST 2 SEQUENCES is used with default values of (i) reward for a match: 1; (ii) penalty for a mismatch: -2; (iii) open gap 5 and extension gap 2 penalties; (iv) gap X\_dropoff 50 expect 10 word size 11 filter, and both sequences are entered in their entirety.

The isolated polynucleotides of the present invention being useful  
 30 for expression of GP354 proteins and protein fragments, the present invention thus provide isolated polynucleotides that encode GP354 proteins and portions thereof

WO 03/000021

PCT/US02/20275

-28-

not only identical in sequence to those described with particularity herein, but degenerate variants thereof as well. As is well known, the genetic code is degenerate and codon choice for optimal expression varies from species to species.

As is also well known, amino acid substitutions occur frequently among natural allelic variants, with conservative substitutions often occasioning only *de minimis* change in protein function.

Accordingly, the present invention provides polynucleotides not only identical in sequence to those described with particularity herein, but also those that encode GP354 and portions thereof, having conservative amino acid substitutions or moderately conservative amino acid substitutions.

Although there are a variety of metrics for calling conservative amino acid substitutions, based primarily on either observed changes among evolutionarily related proteins or on predicted chemical similarity, for purposes herein a conservative replacement is any change having a positive value in the PAM250 log-likelihood matrix reproduced herein below (*see* Gonnet *et al.*, *Science* 256(5062):1443-5 (1992)):

|      | A  | R  | N  | D  | C  | Q  | E  | G  | H  | I  | L  | K  | M  | F  | P  | S  | T  | W  | Y  | V  |
|------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| A    | 2  | -1 | 0  | 0  | 0  | 0  | 0  | 0  | -1 | -1 | -1 | 0  | -1 | -2 | 0  | 1  | 1  | -4 | -2 | 0  |
| R    | -1 | 5  | 0  | 0  | -2 | 2  | 0  | -1 | 1  | -2 | -2 | 3  | -2 | -3 | -1 | 0  | 0  | -2 | -2 | -2 |
| 20 N | 0  | 0  | 4  | 2  | -2 | 1  | 1  | 0  | 1  | -3 | -3 | 1  | -2 | -3 | -1 | 1  | 0  | -4 | -1 | -2 |
| D    | 0  | 0  | 2  | 5  | -3 | 1  | 3  | 0  | 0  | -4 | -4 | 0  | -3 | -4 | -1 | 0  | 0  | -5 | -3 | -3 |
| C    | 0  | -2 | -2 | -3 | 12 | -2 | -3 | -2 | -1 | -1 | -2 | -3 | -1 | -1 | -3 | 0  | 0  | -1 | 0  | 0  |
| Q    | 0  | 2  | 1  | 1  | -2 | 3  | 2  | -1 | 1  | -2 | -2 | 2  | -1 | -3 | 0  | 0  | 0  | -3 | -2 | -2 |
| E    | 0  | 0  | 1  | 3  | -3 | 2  | 4  | -1 | 0  | -3 | -3 | 1  | -2 | -4 | 0  | 0  | 0  | -4 | -3 | -2 |
| 25 G | 0  | -1 | 0  | 0  | -2 | -1 | -1 | 7  | -1 | -4 | -4 | -1 | -4 | -5 | -2 | 0  | -1 | -4 | -4 | -3 |
| H    | -1 | 1  | 1  | 0  | -1 | 1  | 0  | -1 | 6  | -2 | -2 | 1  | -1 | 0  | -1 | 0  | 0  | -1 | 2  | -2 |
| I    | -1 | -2 | -3 | -4 | -1 | -2 | -3 | -4 | -2 | 4  | 3  | -2 | 2  | 1  | -3 | -2 | -1 | -2 | -1 | 3  |
| L    | -1 | -2 | -3 | -4 | -2 | -2 | -3 | -4 | -2 | 3  | 4  | -2 | 3  | 2  | -2 | -2 | -1 | -1 | 0  | 2  |
| K    | 0  | 3  | 1  | 0  | -3 | 2  | 1  | -1 | 1  | -2 | -2 | 3  | -1 | -3 | -1 | 0  | 0  | -4 | -2 | -2 |
| 30 M | -1 | -2 | -2 | -3 | -1 | -1 | -2 | -4 | -1 | 2  | 3  | -1 | 4  | 2  | -2 | -1 | -1 | -1 | 0  | 2  |
| F    | -2 | -3 | -3 | -4 | -1 | -3 | -4 | -5 | 0  | 1  | 2  | -3 | 2  | 7  | -4 | -3 | -2 | 4  | 5  | 0  |
| P    | 0  | -1 | -1 | -1 | -3 | 0  | 0  | -2 | -1 | -3 | -2 | -1 | -2 | -4 | 8  | 0  | 0  | -5 | -3 | -2 |

WO 03/000021

PCT/US02/20275

-29-

|          |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| <b>S</b> | 1  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | -2 | -2 | 0  | -1 | -3 | 0  | 2  | 2  | -3 | -2 | -1 |
| <b>T</b> | 1  | 0  | 0  | 0  | 0  | 0  | -1 | 0  | -1 | -1 | 0  | -1 | -2 | 0  | 2  | 2  | -4 | -2 | 0  | 0  |
| <b>W</b> | -4 | -2 | -4 | -5 | -1 | -3 | -4 | -4 | -1 | -2 | -1 | -4 | -1 | 4  | -5 | -3 | -4 | 14 | 4  | -3 |
| <b>Y</b> | -2 | -2 | -1 | -3 | 0  | -2 | -3 | -4 | 2  | -1 | 0  | -2 | 0  | 5  | -3 | -2 | -2 | 4  | 8  | -1 |
| <b>V</b> | 0  | -2 | -2 | -3 | 0  | -2 | -2 | -3 | -2 | 3  | 2  | -2 | 2  | 0  | -2 | -1 | 0  | -3 | -1 | 3  |

For purposes herein, a "moderately conservative" replacement is any change having a nonnegative value in the PAM250 log-likelihood matrix reproduced herein above.

To avoid severely reducing or eliminating biological activity, amino acid residues that are conserved among the GP354 proteins of various species or among the Ig family members are not altered (except by conservative substitution) during genetic engineering. For instance, the cysteine residues for maintaining an Ig domain of GP354 should be conserved.

Relatedness of polynucleotides can also be characterized using a functional test, the ability of the two polynucleotides to base-pair to one another at defined hybridization stringencies. The invention thus provides isolated polynucleotides not only identical in sequence to those described with particularity herein, but also to provide isolated polynucleotides ("cross-hybridizing nucleic acids") that hybridize under high stringency conditions (as defined herein) to all or to a portion of various of the isolated gp354 polynucleotides of the present invention ("reference nucleic acids").

Such cross-hybridizing nucleic acids are useful, *inter alia*, as probes for, and to drive expression of, proteins related to the proteins of the present invention such as alternative splice variants and homologs (e.g., orthologs and paralog). Particularly useful orthologs are those from other primate species, such as chimpanzee, rhesus macaque monkey, baboon, orangutan, and gorilla; from rodents, such as rats, mice, guinea pigs; from lagomorphs, such as rabbits, and from domestic livestock, such as cow, pig, sheep, horse, goat.

The hybridizing portion of the reference nucleic acid is typically at least 15 nucleotides in length, and often at least 17, 20, 25, 30, 35, 40 or 50 nucleotides (nt) in length. Cross-hybridizing nucleic acids that hybridize to a

WO 03/000021

PCT/US02/20275

-30-

larger portion of the reference nucleic acid – for example, to a portion of at least 50 nt, 100 nt, 150 nt, 200 nt, 250 nt, 300 nt, 350 nt, 400 nt, 450 nt, 500 nt or more, up to and including the entire length of the reference nucleic acid, are also useful.

The hybridizing portion of the cross-hybridizing nucleic acid is at least 75% identical in sequence to at least a portion of the reference nucleic acid. Typically, the hybridizing portion of the cross-hybridizing nucleic acid is at least 80%, often at least 85%, 86%, 87%, 88%, 89% or even at least 90% identical in sequence to at least a portion of the reference nucleic acid. Often, the hybridizing portion of the cross-hybridizing nucleic acid will be at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical in sequence to at least a portion of the reference nucleic acid sequence. At times, the hybridizing portion of the cross-hybridizing nucleic acid will be at least 99.5% identical in sequence to at least a portion of the reference nucleic acid.

The invention also provides fragments of various of the isolated polynucleotides or nucleic acids of the present invention. By “fragments” of a reference nucleic acid is here intended isolated polynucleotides or nucleic acids, however obtained, that have a nucleotide sequence identical to a portion of the reference nucleic acid sequence, which portion is at least 17 nucleotides and less than the entirety of the reference nucleic acid.

In theory, an oligonucleotide of 17 nucleotides is of sufficient length as to occur at random less frequently than once in the three gigabases of the human genome, and thus to provide a nucleic acid probe that can uniquely identify the reference sequence in a nucleic acid mixture of mammalian genomic complexity. Further specificity can be obtained by probing nucleic acid samples of subgenomic complexity, and/or by using plural fragments as short as 17 nucleotides in length collectively to prime amplification of nucleic acids, as, e.g., by polymerase chain reaction (PCR).

The nucleic acid probes of the invention can be used to detect RNA transcripts or genomic sequences encoding homologs or identical proteins. The probe may comprise a label group attached thereto, e.g., a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be

WO 03/000021

PCT/US02/20275

-31-

used as a part of diagnostic kit for identifying cells or tissues (i) that mis-express a GP354 protein (e.g., aberrant splicing, abnormal mRNA levels), or (ii) that harbor a mutation in the gp354 gene, such as a deletion, an insertion, or a point mutation. Such diagnostic kits preferably include labeled reagents and instructional inserts for their use.

The isolated polynucleotides of the invention can also be used as primers in PCR, primer extension and the like. To be useful as primers, the polynucleotides can be, e.g., at least 6 nucleotides (e.g., at least 7, 8, 9, or 10) in length. The primers can hybridize to an exonic sequence of a gp354 gene, for, e.g., amplification of a gp354 mRNA or cDNA. Alternatively, the primers can hybridize to an intronic sequence or an upstream or downstream regulatory sequence of a gp354 gene, to utilize non-transcribed, e.g., regulatory portions of the genomic structure of a gp354 gene.

The nucleic acid primers of the present invention can also be used, for example, to prime single base extension (SBE) for SNP detection (*see, e.g.*, U.S. Pat. No. 6,004,744, the disclosure of which is incorporated herein by reference in its entirety). Isothermal amplification approaches, such as rolling circle amplification, are also now well-described. *See, e.g.*, Schweitzer *et al.*, *Curr. Opin. Biotechnol.* 12(1):21-7 (2001); U.S. Patent Nos. 5,854,033 and 5,714,320 and international patent publications WO 97/19193 and WO 00/15779, the disclosures of which are incorporated herein by reference in their entireties. Rolling circle amplification can be combined with other techniques to facilitate SNP detection. *See, e.g.*, Lizardi *et al.*, *Nature Genet.* 19(3):225-32 (1998).

As described below, nucleic acid fragments that encode at least 6 contiguous amino acids (*i.e.*, fragments of 18 nucleotides or more) are useful in directing the expression or the synthesis of peptides that have utility in mapping the epitopes of the protein encoded by the reference nucleic acid. *See, e.g.*, Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1984); and U.S. Pat. Nos. 4,708,871 and 5,595,915.

And, as described below, nucleic acid fragments that encode at least 8 contiguous amino acids (*i.e.*, fragments of 24 nucleotides or more) are useful in



WO 03/000021

PCT/US02/20275

-32-

directing the expression or the synthesis of peptides that have utility as immunogens. See, e.g., Lerner, "Tapping the immunological repertoire to produce antibodies of predetermined specificity," *Nature* 299:592-596 (1982); Shinnick *et al.*, *Annu. Rev. Microbiol.* 37:425-46 (1983); Sutcliffe *et al.*, *Science* 219:660-6 (1983).

The nucleic acid fragment of the present invention is thus at least 17 nucleotides in length, typically at least 18 nucleotides in length, and often at least 24, 25, 30, 35, 40, or 45 nucleotides (nt) in length. Of course, larger fragments having at least 50 nt, 100 nt, 150 nt, 200 nt, 250 nt, 300 nt, 350 nt, 400 nt, 450 nt, 500 nt or more are also useful, and at times preferred, as will be appreciated by the skilled worker.

Having been based upon the mining of genomic sequence, rather than upon surveillance of expressed message, the present invention further provides isolated genome-derived polynucleotides or nucleic acids that include portions of the gp354 gene. The invention particularly provides genome-derived single exon probes, which comprise at least part of an exon ("reference exon") and can hybridize detectably under high stringency conditions to transcript-derived nucleic acids that include the reference exon. The single exon probe will not, however, hybridize detectably under high stringency conditions to nucleic acids that lack the reference exon but include one or more exons that are found adjacent to the reference exon in the genome.

The present invention also provides isolated genome-derived polynucleotides or nucleic acids which include nucleic acid sequence elements that control transcription of the gp354 gene. Transcription control sequences include, e.g., promoters, enhancers, operators, terminators, silencers, and the like.

When desired for use in antisense inhibition of transcription or translation, or for antisense-mediated targeting of enzymatic nucleic acid molecules such as ribozymes, the isolated polynucleotides and nucleic acids of the present invention can usefully include one or more modified bases (see below) and/or one or more modified or altered internucleoside bonds, which often provide nuclease-resistance. See Hartmann *et al.* (eds.), Manual of Antisense Methodology

WO 03/000021

PCT/US02/20275

-33-

- (Perspectives in Antisense Science), Kluwer Law International (1999) (ISBN:079238539X); Stein et al. (eds.), Applied Antisense Oligonucleotide Technology, Wiley-Liss (cover (1998) (ISBN: 0471172790); Chadwick et al. (eds.), Oligonucleotides as Therapeutic Agents - Symposium No. 209, John Wiley & Son Ltd (1997) (ISBN: 0471972797). Such altered bases and internucleoside bonds are often desired also when the isolated nucleic acid of the present invention is to be used for targeted gene correction, as described in Gamper *et al.*, *Nucl. Acids Res.* 28(21):4332-9 (2000), the disclosure of which is incorporated herein by reference in its entirety.
- 10           The antisense nucleic acid molecules (and enzymatic nucleic acids targeted by antisense) of the invention can be used in a therapeutic setting. These molecules can be expressed from an expression vector that contains an operably linked transcription regulatory sequence, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the
- 15           regulation of gene expression using antisense genes, see Weintraub *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Trends. Genet.* 1(1) (1986).
- An antisense nucleic acid of the invention may be a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable
- 20           of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes can be used to catalytically cleave gp354 mRNA transcripts to thereby inhibit translation of gp354 mRNA. A ribozyme having specificity for a gp354-encoding nucleic acid can be designed based upon the nucleotide sequence of a gp354 polynucleotide disclosed herein (i.e., SEQ ID
- 25           NO:1 or 3).
- Oligonucleotide mimetics of gp354, such as peptide nucleic acids (PNA), can be used in therapeutic and diagnostic applications. See, e.g., Hyrup *et al. Bioorg. Med. Chem. Lett.* 4:5-23 (1996). In PNA compounds, the phosphodiester backbone of the nucleic acid is replaced with an amide-containing
- 30           backbone, in particular by repeating N-(2-aminoethyl) glycine units linked by amide bonds. PNAs For example, PNAs can be used as antisense or antigene

WO 03/000021

PCT/US02/20275

-34-

agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of gp354 can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination  
 5 with other enzymes, e.g., S1 nucleases; or as probes or primers for DNA sequence and hybridization (Hyrup et al., *supra*; and Perry-O'Keefe, *supra*). PNAs of gp354 can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art  
 10 (see *infra*).

Oligonucleotide of the invention may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane or the blood-brain barrier. In addition, oligonucleotides can be modified with hybridization triggered cleavage  
 15 agents or intercalating agents. To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc. (see *infra*).

Differences from nucleic acid compositions found in nature — e.g., non-native bases, altered internucleoside linkages, post-synthesis modification —  
 20 can be present throughout the length of the gp354 polynucleotide or can usefully be localized to discrete portions thereof. As an example of the latter, chimeric nucleic acids can be synthesized that have discrete DNA and RNA domains and demonstrated utility for targeted gene repair, as further described in U.S. Pat. Nos. 5,760,012 and 5,731,181, the disclosures of which are incorporated herein by  
 25 reference in their entireties. Chimeric nucleic acids comprising both DNA and PNA have been demonstrated to have utility in modified PCR reactions. See Misra et al., *Biochem.* 37: 1917-1925 (1998); see also Finn et al., *Nucl. Acids Res.* 24: 3357-3363 (1996), incorporated herein by reference.

Polynucleotides and nucleic acids of the present invention can also  
 30 usefully be bound to a substrate. The substrate can porous or solid, planar or non-planar, unitary or distributed; the bond can be covalent or noncovalent. Bound to a

WO 03/000021

PCT/US02/20275

-35-

substrate, nucleic acids of the present invention can be used as probes in their unlabeled state. For example, the nucleic acids of the present invention can usefully be bound to a porous substrate, commonly a membrane, typically comprising nitrocellulose, nylon, or positively-charged derivatized nylon; so  
5 attached, the nucleic acids of the present invention can be used to detect gp354 nucleic acids present within a labeled nucleic acid sample, either a sample of genomic nucleic acids or a sample of transcript-derived nucleic acids, *e.g.* by reverse dot blot.

The nucleic acids of the present invention can also usefully be  
10 bound to a solid substrate, such as glass, although other solid materials, such as amorphous silicon, crystalline silicon, or plastics, can also be used. The nucleic acids of the present invention can be attached covalently to a surface of the support substrate or applied to a derivatized surface in a chaotropic agent that facilitates denaturation and adherence by presumed noncovalent interactions, or some  
15 combination thereof.

The nucleic acids of the present invention can be bound to a substrate to which a plurality of other nucleic acids are concurrently bound, hybridization to each of the plurality of bound nucleic acids being separately detectable. At low density, *e.g.* on a porous membrane, these substrate-bound  
20 collections are typically denominated macroarrays; at higher density, typically on a solid support, such as glass, these substrate bound collections of plural nucleic acids are colloquially termed microarrays. As used herein, the term microarray includes arrays of all densities. The invention thus provides microarrays that include the nucleic acids of the present invention.

25 The isolated nucleic acids of the present invention can be used as hybridization probes to detect, characterize, and quantify gp354 nucleic acids in, and isolate gp354 nucleic acids from, both genomic and transcript-derived nucleic acid samples. When free in solution, such probes are typically, but not invariably, detectably labeled; bound to a substrate, as in a microarray, such probes are  
30 typically, but not invariably unlabeled.

WO 03/000021

PCT/US02/20275

-36-

For example, the isolated nucleic acids of the present invention can be used as probes to detect and characterize gross alterations in the gp354 genomic locus, such as deletions, insertions, translocations, and duplications of the gp354 genomic locus through fluorescence *in situ* hybridization (FISH) to chromosome spreads. See, e.g., Andreeff *et al.* (eds.), Introduction to Fluorescence In Situ Hybridization: Principles and Clinical Applications, John Wiley & Sons (1999) (ISBN: 0471013455), the disclosure of which is incorporated herein by reference in its entirety. The isolated nucleic acids of the present invention can be used as probes to assess smaller genomic alterations using, e.g., Southern blot detection of restriction fragment length polymorphisms. The isolated nucleic acids of the present invention can be used as probes to isolate genomic clones that include the nucleic acids of the present invention, which thereafter can be restriction mapped and sequenced to identify deletions, insertions, translocations, and substitutions (single nucleotide polymorphisms, SNPs) at the sequence level.

The isolated nucleic acids of the present invention can be also be used as probes to detect, characterize, and quantify gp354 nucleic acids in, and isolate gp354 nucleic acids from, transcript-derived nucleic acid samples. For example, the isolated nucleic acids of the present invention can be used as hybridization probes to detect, characterize by length, and quantify gp354 mRNA by northern blot of total or poly-A<sup>+</sup>-selected RNA samples. The isolated nucleic acids of the present invention can also be used as hybridization probes to detect, characterize by location, and quantify gp354 message by *in situ* hybridization to tissue sections (see, e.g., Schwarczacher *et al.*, In Situ Hybridization, Springer-Verlag New York (2000) (ISBN: 0387915966), the disclosure of which is incorporated herein by reference in its entirety).

Further, the isolated nucleic acids of the present invention can be used as hybridization probes to measure the representation of gp354 clones in a cDNA library. For example, the isolated nucleic acids of the present invention can be used as hybridization probes to isolate gp354 nucleic acids from cDNA libraries, permitting sequence level characterization of gp354 RNA messages, including identification of deletions, insertions, truncations — including deletions,

WO 03/000021

PCT/US02/20275

-37-

insertions, and truncations of exons in alternatively spliced forms — and single nucleotide polymorphisms.

As described in the Examples herein below, the nucleic acids of the present invention can also be used to detect and quantify gp354 nucleic acids in transcript-derived samples to measure expression of the gp354 gene. Measurement of gp354 expression has particular utility in diagnostic assays for conditions, disorders and diseases associated with abnormal gp354 expression, either in pancreatic and neural tissues where and in a manner in which it is normally expressed, as well as in tissues where it may be mis-expressed, as further described in the Examples herein below.

As would be readily apparent to one of skill in the art, each gp354 nucleic acid probe — whether labeled, substrate-bound, or both — is thus currently available for use as a tool for measuring the level of gp354 expression in pancreatic and neural tissues, in which expression has already been confirmed.

As for tissues not yet demonstrated to express gp354, the gp354 nucleic acid probes of the present invention are currently available as tools for surveying such tissues to detect the presence of gp354 nucleic acids, for example, to detect gp354 RNA expression in tissues of patients who present with a condition, disorder or disease associated with abnormal gp354 cellular expression in the pancreas or nervous system or abnormal tissue distribution in other tissues.

As noted above, the nucleic acid probes of the present invention are useful in constructing microarrays; the microarrays, in turn, are products of manufacture that are useful for measuring and for surveying gene expression in, for example, drug discovery and target validation programs. When included on a microarray, each gp354 nucleic acid probe makes the microarray specifically useful for detecting that portion of the gp354 gene included within the probe, thus imparting upon the microarray device the ability to detect a signal where, absent such probe, it would have reported no signal.

Changes in the level of gp354 expression need not be observed for the measurement of expression to have utility. Where gene expression analysis is used to assess toxicity of chemical agents on cells, for example, the failure of the

WO 03/000021

PCT/US02/20275

-38-

agent to change a gene's expression level is evidence that the drug likely does not affect the pathway of which the gene's expressed protein is a part. Analogously, where gene expression analysis is used to assess side effects of pharmacologic agents — whether in lead compound discovery or in subsequent screening of lead compound derivatives — the inability of the agent to alter a gene's expression level is evidence that the drug does not affect the pathway of which the gene's expressed protein is a part. WO 99/58720, incorporated herein by reference in its entirety, provides methods for quantifying the relatedness of a first and second gene expression profile and for ordering the relatedness of a plurality of gene expression profiles, without regard to the identity or function of the genes whose expression is used in the calculation.

The genome-derived single exon probes and genome-derived single exon probe microarrays of the invention have the additional utility of permitting high-throughput detection of splice variants of the nucleic acids of the present invention.

Polynucleotides of the present invention, inserted into nucleic acid constructs such as vectors which flank the polynucleotide insert with a promoter can be used to drive *in vitro* expression of RNA complementary to either strand of the nucleic acid of the present invention. The RNA can be used as a single-stranded probe, in cDNA-mRNA subtraction, or for *in vitro* translation. Those polynucleotides which encode GP354 protein or portions thereof can further be used to express the GP354 proteins or protein fragments, either alone, or as part of fusion proteins. Expression can be from genomic or transcript-derived polynucleotides of the present invention.

Where protein expression is effected from genomic DNA, expression will typically be effected in eukaryotic, typically mammalian, cells capable of splicing introns from the initial RNA transcript. Expression can be driven from episomal vectors or from genomic DNA integrated into a host cell chromosome. As described below, where expression is from transcript-derived (or otherwise intron-less) polynucleotides of the invention, expression can be effected in a wide variety of prokaryotic or eukaryotic cells.

WO 03/000021

PCT/US02/20275

-39-

Expressed *in vitro*, the protein, protein fragment, or protein fusion can thereafter be isolated, to be used as a standard in immunoassays specific for the proteins, or protein isoforms, of the present invention; to be used as a therapeutic agent, *e.g.*, to be administered as passive replacement therapy in individuals  
 5 deficient in the proteins of the present invention; to be administered as a vaccine; to be used for *in vitro* production of specific antibody, the antibody thereafter to be used, *e.g.*, as an analytical reagent for detection and quantitation of the proteins of the present invention or to be used as an immunotherapeutic agent.

The isolated polynucleotides and nucleic acids of the present  
 10 invention can also be used to drive *in vivo* expression of the proteins of the present invention. *In vivo* expression can be driven from a vector — typically a viral vector, often a vector based upon a replication incompetent lentivirus, retrovirus, adenovirus, or adeno-associated virus (AAV) — for purpose of gene therapy. *In vivo* expression can be driven from expression control signals endogenous or  
 15 exogenous (*e.g.*, from a vector) to the nucleic acid. Other viral vectors of the invention include vectors derived, *e.g.*, from baculoviruses, adenoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, and retroviruses.

Various forms of the isolated gp354 polynucleotides of the  
 20 invention (*e.g.*, genomic or cDNA) can be microinjected into male or female pronuclei, or can be integrated into embryonic stem (ES) cells to create transgenic non-human animals capable of producing the proteins of the present invention.

Genomic nucleic acids of the present invention can also be used to target homologous recombination to a gp354 locus in a subject. *See, e.g.*, U.S.  
 25 Patent Nos. 6,187,305; 6,204,061; 5,631,153; 5,627,059; 5,487,992; 5,464,764; 5,614,396; 5,527,695 and 6,063,630; and Kmiec *et al.* (eds.), Gene Targeting Protocols, Vol. 133, Humana Press (2000) (ISBN: 0896033600); Joyner (ed.), Gene Targeting: A Practical Approach, Oxford University Press, Inc. (2000) (ISBN: 0199637938); Sedivy *et al.*, Gene Targeting, Oxford University Press  
 30 (1998) (ISBN: 071677013X); Tymms *et al.* (eds.), Gene Knockout Protocols, Humana Press (2000) (ISBN: 0896035727); Mak *et al.* (eds.), The Gene Knockout



WO 03/000021

PCT/US02/20275

-40-

FactsBook, Vol. 2, Academic Press, Inc. (1998) (ISBN: 0124660444); Torres *et al.*, Laboratory Protocols for Conditional Gene Targeting, Oxford University Press (1997) (ISBN: 019963677X); Vega (ed.), Gene Targeting, CRC Press, LLC (1994) (ISBN: 084938950X), the disclosures of which are incorporated herein by reference in their entireties.

Where the genomic region includes transcription regulatory elements, homologous recombination can be used to alter the expression of GP354, both for purpose of *in vitro* production of GP354 protein from human cells, and for purpose of gene therapy. See, e.g., U.S. Pat. Nos. 5,981,214, 6,048,524; 5,272,071; the disclosures of which are incorporated herein by reference in their entireties.

Fragments of the polynucleotides of the present invention smaller than those typically used for homologous recombination can also be used for targeted gene correction or alteration, possibly by cellular mechanisms different from those engaged during homologous recombination. See, e.g., U.S. Pat. Nos. 5,945,339, 5,888,983, 5,871,984, 5,795,972, 5,780,296, 5,760,012, 5,756,325, 5,731,181; and Culver *et al.*, "Correction of chromosomal point mutations in human cells with bifunctional oligonucleotides," *Nature Biotechnol.* 17(10):989-93 (1999); Gamper *et al.*, *Nucl. Acids Res.* 28(21):4332-9 (2000), the disclosures of which are incorporated herein by reference.

Polynucleotides of the present invention can be obtained by using the labeled probes of the present invention to probe nucleic acid samples, such as genomic libraries, cDNA libraries, and mRNA samples, by standard techniques. Polynucleotides of the present invention can also be obtained by amplification, using the nucleic acid primers of the present invention, as further demonstrated in Example 1, herein below. Polynucleotides of the present invention, especially if fewer than about 100 nucleotide, can also be synthesized chemically, typically by solid phase synthesis using commercially available automated synthesizers.

#### VECTORS AND HOST CELLS

##### A. NUCLEIC ACID CONSTRUCTS

WO 03/000021

PCT/US02/20275

-41-

The present invention provides nucleic acid constructs, such as vectors, that comprise one or more of the isolated polynucleotides of the invention, and host cells into which such vectors have been introduced.

The vectors can be used for propagating the polynucleotides of the present invention in host cells (cloning vectors), for shuttling the polynucleotides of the present invention between host cells derived from disparate organisms (shuttle vectors), for inserting the polynucleotides of the present invention into host cell chromosomes (insertion vectors), for expressing sense or antisense RNA transcripts of the polynucleotides of the present invention *in vitro* or within a host cell, and for expressing polypeptides encoded by the polynucleotides of the present invention, alone or as fusions to heterologous polypeptides (expression vectors). Vectors of the present invention will often be suitable for several such uses.

Vectors are by now well-known in the art, and are described, *inter alia*, in Jones *et al.* (eds.), Vectors: Cloning Applications : Essential Techniques (Essential Techniques Series), John Wiley & Son Ltd 1998 (ISBN: 047196266X); Jones *et al.* (eds.), Vectors: Expression Systems : Essential Techniques (Essential Techniques Series), John Wiley & Son Ltd, 1998 (ISBN:0471962678); Gacsa *et al.*, Vectors: Essential Data, John Wiley & Sons, 1995 (ISBN: 0471948411); Cid-Arregui (eds.), Viral Vectors: Basic Science and Gene Therapy, Eaton Publishing Co., 2000 (ISBN: 188129935X); Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (3<sup>rd</sup> ed.), Cold Spring Harbor Laboratory Press, 2001 (ISBN: 0879695773); Ausubel *et al.* (eds.), Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology (4<sup>th</sup> ed.), John Wiley & Sons, 1999 (ISBN: 047132938X), the disclosures of which are incorporated herein by reference in their entirety. An enormous variety of vectors are available commercially. Use of existing vectors and modifications are well within the skill in the art.

Typically, vectors are derived from virus, plasmid, prokaryotic or eukaryotic chromosomal elements, or some combination thereof, and include at least one origin of replication, at least one site for insertion of heterologous nucleic acid, typically in the form of a polylinker with multiple, tightly clustered, single

WO 03/000021

PCT/US02/20275

-42-

cutting restriction sites, and at least one selectable marker, although some integrative vectors will lack an origin that is functional in the host to be chromosomally modified, and some vectors will lack selectable markers. Vectors of the invention will further include at least one isolated polynucleotide nucleic acid of the invention inserted into the vector in at least one location. Where present, the origin of replication and selectable markers are chosen based upon the desired host cell or host cells; the host cells, in turn, are selected based upon the desired application.

For example, prokaryotic cells, typically *E. coli*, are typically chosen for cloning, i.e., for amplification of polynucleotide sequences in a host cell. In such case, vector replication is predicated on the replication strategies of coliform-infecting phage — such as phage lambda, M13, T7, T3 and P1 — or on the replication origin of autonomously replicating episomes, notably the ColE1 plasmid and later derivatives, including pBR322 and the pUC series plasmids. Where *E. coli* is used as host, selectable markers are, analogously, chosen for selectivity in gram negative bacteria: e.g., typical markers confer resistance to antibiotics, such as ampicillin, tetracycline, chloramphenicol, kanamycin, streptomycin, zeocin; auxotrophic markers can also be used.

As another example, yeast cells, typically *S. cerevisiae*, are chosen, *inter alia*, for eukaryotic genetic studies, for identification of interacting protein components, e.g. through use of a two-hybrid system, and for protein expression. Vectors of the present invention for use in yeast will typically, but not invariably, contain an origin of replication suitable for use in yeast and a selectable marker that is functional in yeast.

Examples of suitable yeast vectors include integrative Yip vectors, replicating episomal YE<sub>p</sub> vectors containing centromere sequences, CEN, and autonomously replicating sequences, ARS. YACs are based on yeast linear plasmids, denoted YL<sub>p</sub>, containing homologous or heterologous DNA sequences that function as telomeres (TEL) *in vivo*, as well as containing yeast ARS (origins of replication) and CEN (centromeres) segments.

WO 03/000021

PCT/US02/20275

-43-

Selectable markers in yeast vectors include a variety of auxotrophic markers, the most common of which are (in *Saccharomyces cerevisiae*) URA3, HIS3, LEU2, TRP1 and LYS2, which complement specific auxotrophic mutations, such as *ura3-52*, *his3-D1*, *leu2-D1*, *trp1-D1* and *lys2-201*. The URA3 and LYS2 yeast genes further permit negative selection based on specific inhibitors, 5-fluoro-orotic acid (FOA) and  $\alpha$ -aminoadipic acid ( $\alpha$ AA), respectively, that prevent growth of the prototrophic strains but allows growth of the *ura3* and *lys2* mutants, respectively. Other selectable markers confer resistance to, e.g., zeocin.

Insect cells are often chosen for high efficiency protein expression. Where the host cells are from *Spodoptera frugiperda* — e.g., Sf9 and Sf21 cell lines, and expresSF™ cells (Protein Sciences Corp., Meriden, CT, USA) — the vector replicative strategy is typically based upon the baculovirus life cycle. Typically, baculovirus transfer vectors are used to replace the wild-type AcMNPV polyhedrin gene with a heterologous gene of interest. Sequences that flank the polyhedrin gene in the wild-type genome are positioned 5' and 3' of the expression cassette on the transfer vectors. Following cotransfection with AcMNPV DNA, a homologous recombination event occurs between these sequences resulting in a recombinant virus carrying the gene of interest and the polyhedrin or p10 promoter. Selection can be based upon visual screening for lacZ fusion activity.

Mammalian cells are often chosen for expression of proteins intended as pharmaceutical agents, and are also chosen as host cells for screening of potential agonist and antagonists of a protein or a physiological pathway. Vectors intended for autonomous extrachromosomal replication in mammalian cells will typically include a viral origin, such as the SV40 origin (for replication in cell lines expressing the large T-antigen, such as COS1 and COS7 cells), the papillomavirus origin, or the EBV origin for long term episomal replication (for use, e.g., in 293-EBNA cells, which constitutively express the EBV EBNA-1 gene product and adenovirus E1A). Vectors intended for integration, and thus replication as part of the mammalian chromosome, can, but need not, include an origin of replication functional in mammalian cells, such as the SV40 origin. Vectors based upon viruses, such as lentiviruses, adenovirus, adeno-associated

WO 03/000021

PCT/US02/20275

-44-

virus, vaccinia virus, and various mammalian retroviruses, will typically replicate according to the viral replicative strategy.

Selectable markers for use in mammalian cells include resistance to neomycin (G418), blasticidin, hygromycin and to zeocin, and selection based upon  
5 the purine salvage pathway using HAT medium.

Plant cells can also be used for expression, with the vector replicon typically derived from a plant virus (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) and selectable markers chosen for suitability in plants.

For propagation of polynucleotides of the present invention that are  
10 larger than can readily be accommodated in vectors derived from plasmids or virus, the invention further provides artificial chromosomes — BACs, YACs, and HACs — that comprise gp354 nucleic acids, often genomic nucleic acids.

For propagation of polynucleotides of the present invention that are larger than can readily be accommodated in vectors derived from plasmids or  
15 viruses, the invention further provides artificial chromosomes — BACs, YACs, and HACs — that comprise gp354 nucleic acids, often genomic nucleic acids. *See, e.g., Shizuya et al., Keio J. Med.* 50(1):26-30 (2001); *Shizuya et al., Proc. Natl. Acad. Sci. USA* 89(18):8794-7 (1992); *Kuroiwa et al., Nature Biotechnol.* 18(10):1086-90 (2000); *Henning et al., Proc. Natl. Acad. Sci. USA* 96(2):592-7  
20 (1999); *Harrington et al., Nature Genet.* 15(4):345-55 (1997), the disclosures of which are incorporated herein by reference.

Vectors of the invention will also often include elements that permit *in vitro* transcription of RNA from the inserted heterologous nucleic acid. Such vectors typically include a phage promoter, such as that from T7, T3, or SP6,  
25 flanking the nucleic acid insert. Often two different such promoters flank the inserted nucleic acid, permitting separate *in vitro* production of both sense and antisense strands.

Expression vectors of the invention which will drive expression of polypeptides from the inserted heterologous nucleic acid will often include a  
30 variety of other genetic elements operatively linked to the protein-encoding heterologous nucleic acid insert, typically genetic elements that drive and regulate

WO 03/000021

PCT/US02/20275

-45-

transcription, such as promoters and enhancer elements, those that facilitate RNA processing, such as transcription termination, splicing signals and/or polyadenylation signals, and those that facilitate translation, such as ribosomal consensus sequences. Other transcription control sequences include, e.g., operators, silencers, and the like. Use of such expression control elements, including those that confer inducible expression, and developmental or tissue-regulated expression are well-known in the art.

Tissue-specific regulatory elements capable of expressing GP354 in the pancreas, nervous system or mammary glands may be particularly useful and are known in the art, e.g., the neuron-specific neurofilament promoter (Byrne and Ruddle *Proc. Natl. Acad. Sci. USA* 86:5473-5477 (1989)), a pancreas-specific promoter (Edlund *et al. Science* 230:912-916 (1985)), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters may also be selected, including but not limited to the murine hox promoters (Kessel and Gruss *Science* 249:374-379 (1990)) and the  $\alpha$ -fetoprotein promoter (Campes and Tilghman *Genes Dev.* 3:537-546 (1989)). A huge variety of inducible promoters are known and may be selected based on the particular application.

Expression vectors can be designed to fuse the expressed polypeptide to small protein tags that facilitate purification and/or visualization. Many such tags are known and available. Expression vectors can also be designed to fuse proteins encoded by the heterologous nucleic acid insert to polypeptides larger than purification and/or identification tags. Useful protein fusions include those that permit display of the encoded protein on the surface of a phage or cell, fusions to intrinsically fluorescent proteins, such as luciferase or those that have a green fluorescent protein (GFP)-like chromophore, fusions to the IgG Fc region or other immunoglobulin type constant domains, and fusions for use in two hybrid selection systems.

For secretion of expressed proteins, a wide variety of vectors are available which include appropriate sequences that encode secretion signals, such as leader peptides. Vectors designed for phage display, yeast display, and

WO 03/00021

PCT/US02/20275

-46-

mammalian display, for example, target recombinant proteins using an N-terminal cell surface targeting signal and a C-terminal transmembrane anchoring domain.

A wide variety of vectors now exist that fuse proteins encoded by heterologous nucleic acids to the chromophore of the substrate-independent, intrinsically fluorescent green fluorescent protein from *Aequorea victoria* ("GFP") and its many color-shifted and/or stabilized variants.

Vectors which allow fusions of heterologous sequences to the IgG Fc region to increase serum half-life of protein pharmaceutical products through interaction with the FcRn receptor (also denominated the FcRp receptor and the Brambell receptor, FcRb), are also widely available.

For long-term, high-yield recombinant production of the proteins, protein fusions, and protein fragments of the present invention, stable expression is preferred. Stable expression is readily achieved by integration into the host cell genome of vectors (preferably having selectable markers), followed by selection for integrants.

#### B. HOST CELLS

The present invention further includes host cells -- either prokaryotic (bacteria) or eukaryotic (e.g., yeast, insect, plant and animal cells) -- comprising the nucleic acid constructs such as vectors of the present invention, either present episomally within the cell or integrated, in whole or in part, into the host cell chromosome.

Among other considerations, some of which are described above, a host cell strain may be chosen for its ability to process the expressed protein in the desired fashion. Such post-translational modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation, and it is an aspect of the present invention to provide GP354 proteins with such post-translational modifications.

Representative, non-limiting examples of appropriate host cells include bacterial cells, such as *E. coli*, *Caulobacter crescentus*, *Streptomyces* species, and *Salmonella typhimurium*; yeast cells, such as *Saccharomyces*

WO 03/000021

PCT/US02/20275

-47-

*cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, *Pichia methanolica*; insect cell lines, such as those from *Spodoptera frugiperda* — *e.g.*, Sf9 and Sf21 cell lines, and expresSF™ cells (Protein Sciences Corp., Meriden, CT, USA) — *Drosophila* S2 cells, and *Trichoplusia ni* High Five® Cells (Invitrogen, Carlsbad, CA, USA); and mammalian cells. Typical mammalian cells include COS1 and COS7 cells, chinese hamster ovary (CHO) cells, NIH 3T3 cells, 293 cells, HEPG2 cells, HeLa cells, L cells, HeLa, MDCK, HEK293, WI38, murine ES cell lines (*e.g.*, from strains 129/SV, C57/BL6, DBA-1, 129/SVJ), K562, Jurkat cells, and BW5147. Other useful mammalian cell lines are well known and readily available from the American Type Culture Collection (ATCC) (Manassas, VA, USA) and the National Institute of General medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Cell Repositories (Camden, NJ, USA).

Methods for introducing the vectors and nucleic acids of the present invention into the host cells are well known in the art; the choice of technique will depend primarily upon the specific vector to be introduced and the host cell chosen.

#### GP354 PROTEINS, POLYPEPTIDES AND FRAGMENTS

The present invention provides GP354 proteins and various fragments thereof suitable for use as antigens (*e.g.*, for epitope mapping), for use as immunogens (*e.g.*, for raising antibodies or as vaccines), and for use in therapeutic compositions. Also provided are fusions of GP354 polypeptides and fragments to heterologous polypeptides, and conjugates of the proteins, fragments, and fusions of the present invention to other moieties (*e.g.*, to carrier proteins, to fluorophores).

In some embodiments, the invention provides an isolated GP354 polypeptide comprising the amino acid sequence encoded by a full-length gp354 cDNA (SEQ ID NO:1, 7 or 11), or a degenerate variant. The invention also provides an isolated GP354 polypeptide having the amino acid sequence encoded by a full-length gp354 cDNA (SEQ ID NO:1, 7 or 11), optionally having one or more conservative amino acid substitutions.

The invention also provides an isolated GP354 polypeptide comprising the amino acid sequence encoded by a polynucleotide sequence that



WO 03/000021

PCT/US02/20275

-48-

hybridizes under high stringency conditions to a probe having part or all of the nucleotide sequence of a gp354 cDNA (SEQ ID NO:1, 7 or 11). Preferably, an isolated GP354 polypeptide encoded by a stringently or moderately stringent cross-hybridizing polynucleotide of the invention will have at least one biological activity of GP354.

In another series of embodiments, the invention provides an isolated GP354 polypeptide comprising the GP354 amino acid sequence of SEQ ID NO:2, 8 or 12, optionally having one or more conservative amino acid substitutions. Also provided is an isolated GP354 polypeptide having the amino acid sequence encoded by the GP354 polypeptide sequence of SEQ ID NO:2, 8 or 12, optionally having one or more conservative amino acid substitutions. The invention further provides fragments of each of the above-described isolated polypeptides, particularly fragments having at least 6 amino acids, 8 amino acids, 15 amino acids up to the entirety of the sequence given in SEQ ID NO:2, 8 or 12.

Each of the above isolated polypeptides includes an N-terminal 18 or 21 amino acid signal sequence which is typically removed upon insertion of the protein through a membrane. Accordingly, the invention provides the above isolated GP354 polypeptides from which the N-terminal signal sequence has been removed. Cleavage is predicted to occur between the G and P residues at positions 18-19 of SEQ ID NO:2 or at positions 21-22 of SEQ ID NO:8.

The invention thus provides an isolated GP354 polypeptide comprising all or a portion of the predicted mature N-terminal extracellular domain of GP354. (See Figs. 1 and 5; SEQ ID NOs:2 and 8 for GP354 domains and sequences). The predicted mature extracellular domain of GP354 (i.e., lacking the secretion signal sequence), consists of amino acids 19-507 of SEQ ID NO:2, or of amino acids 22-510 of SEQ ID NO:8. Also included are fragments of the above sequences having at least 6 amino acids, 8 amino acids, 15 amino acids up to the entirety of the specified sequence.

The invention also provides an isolated GP354 polypeptide comprising or having all or a portion of the N-terminal extracellular domain of GP354. (See Figs. 1 and 5; SEQ ID NOs:2 and 8 for GP354 domains and

WO 03/000021

PCT/US02/20275

-49-

sequences). The N-terminal extracellular domain of GP354 consists of amino acids 1-507 of SEQ ID NO:2, or of amino acids 1-510 of SEQ ID NO:8. Also included are fragments of the above sequences having at least 6 amino acids, 8 amino acids, 15 amino acids up to the entirety of the specified sequence.

5 In preferred embodiments, the isolated GP354 polypeptide has or comprises the entire extracellular domain of GP354 and lacks a functional GP354 transmembrane domain. The transmembrane domain may either be excluded, deleted or mutated to render it non-functional. The transmembrane domain of GP354 consists of amino acids 508-530 of SEQ ID NO:2, or of amino acids 511-  
10 533 of SEQ ID NO:8.

In other preferred embodiments, the isolated GP354 polypeptide consists of part or all of the GP354 N-terminal extracellular domain fused to a heterologous protein domain. Preferably, the isolated GP354 polypeptide comprises at least one extracellular Ig domain, more preferably comprises two  
15 GP354 extracellular Ig domains, and most preferably comprises three, four or five GP354 extracellular Ig domains.

Also preferred is an isolated GP354 polypeptide comprising a GP354 fragment selected from the group consisting of the transmembrane domain of GP354 and the C-terminal cytoplasmic region of GP354. In other preferred  
20 embodiments, the isolated GP354 polypeptide consists of part or all of the GP354 cytoplasmic or transmembrane domains fused to a heterologous protein domain.

The GP354 fragments of the invention may be continuous portions of the native GP354 protein. However, it will be appreciated that knowledge of the GP354 gene and protein sequences as provided herein permits recombining of  
25 various domains that are not contiguous in the native GP354 protein.

The invention also provides polypeptides comprising select portions of GP354 and related proteins. As will be further discussed herein below, these protein fragments, especially when coupled to heterologous protein fragments, can be used, for example, to target agents to particular cell types through protein-  
30 protein interaction; to inhibit protein-protein interactions between Ig domain

WO 03/000021

PCT/US02/20275

-50-

containing proteins; for competitive binding assays; and to raise fragment-specific GP354 antibodies.

In a first series of such embodiments, the protein fragment comprises, in at least one copy, one, two, three, four or five of the Ig domains characteristic of the N-terminal extracellular portion of GP354. Specifically, the five extracellular Ig domains are encoded by amino acids 35-102, 136-203, 239-290, 323-374 and 410-485, respectively, of the GP354 amino acid sequence of SEQ ID NO:2 (see Fig. 1), and are encoded by amino acids 38-109, 139-206, 242-293, 326-377 and 413-488, respectively, of the GP354 amino acid sequence of SEQ ID NO:8 (see Fig. 5). In preferred embodiments, the protein fragment encodes at least two, preferably three, more preferably four and most preferably all five domains in at least one copy.

Preferably, the protein fragment contains an N-terminal signal secretion sequence that will mediate transport of the polypeptide through a membrane. The GP354 signal secretion sequence is encoded by amino acids 1-18 of the GP354 amino acid sequence of SEQ ID NO:2 (see Fig. 1) and by amino acids 1-21 of SEQ ID NO:8 (see Fig. 5). More preferably, the signal secretion sequence of the protein fragment is from GP354.

The above preferred protein fragments may optionally include a transmembrane domain, if insertion of the polypeptide into a membrane is so-desired. The transmembrane domain may be a GP354 domain (see below) or may be encoded by a heterologous gene encoding a transmembrane domain of a heterologous membrane-associated protein.

If so-desired, the above preferred protein fragments may further comprise an intracellular C-terminal domain if specific signaling reactions are desired in response to GP354 binding interactions. The intracellular domain may be derived from GP354 (see below) or may be encoded by a heterologous gene encoding an intracellular domain of a heterologous membrane-associated protein.

Other preferred embodiments of the protein fragments of the invention are those that comprise the transmembrane domain of GP354.

WO 03/000021

PCT/US02/20275

-51-

Specifically, the GP354 transmembrane domain is encoded by amino acids 508-530 of the GP354 amino acid sequence of SEQ ID NO:2 (see Fig. 1).

Yet other preferred embodiments of the above-described protein fragments have a C-terminal intracellular domain of GP354. Specifically, one intracellular domain of GP354 is encoded by amino acids 531-592 of the GP354 amino acid sequence of SEQ ID NO:2 (see Fig. 1). Another form of an intracellular domain of GP354 is encoded by amino acids 534-708 of the GP354 amino acid sequence of SEQ ID NO:8 (see Fig. 5). It is believed that these different intracellular domain forms may be produced by alternative splicing.

A preferred protein fragment of the invention is encoded by nucleotides 139-923 of the gp354 cDNA sequence of SEQ ID NO:1 (see Fig. 1). It is encoded by an RT-PCR fragment amplified from pancreatic RNA using primers GX1-218 (SEQ ID NO:16) and GX1-219 (SEQ ID NO:17; see Example 2) and consists of amino acids 47-307 of SEQ ID NO:2, i.e., it encodes most of the first N-terminal Ig domain (missing the first 12 of 68 amino acids), and the second and third Ig domains of GP354.

As described above, the invention further provides proteins that differ in sequence from those described with particularity in the above-referenced SEQ ID NOs, whether by way of insertion or deletion, by way of conservative or moderately conservative substitutions, as hybridization related proteins, or as cross-hybridizing proteins, with those that substantially retain a GP354 activity preferred. As also discussed above, the invention further provides fusions of the polypeptides, proteins and protein fragments herein described to heterologous polypeptides.

When used as immunogens, the various protein embodiments of the present invention can be used, *inter alia*, to elicit antibodies that bind to a variety of epitopes of the GP354 protein.

#### Other Defining Characteristics of GP354 Proteins

FIG. 1 presents the deduced amino acid sequences (SEQ ID NO:2) encoded by the gp354 cDNA clone (SEQ ID NO:1). Similarly, the amino acid

WO 03/000021

PCT/US02/20275

-52-

sequences presented in SEQ ID NOs:4, 8, 10 and 12 are deduced from the nucleotide sequences presented in SEQ ID NOs:3, 7, 9 and 11, respectively.

Unless otherwise indicated, amino acid sequences of the proteins of the present invention were determined as a predicted translation from a nucleic acid sequence.

- 5 Accordingly, any amino acid sequence presented herein may contain errors due to errors in the nucleic acid sequence, as described in detail above. Furthermore, single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes—more than 1.4 million SNPs have already identified in the human genome, International Human Genome Sequencing Consortium, *Nature* 409:860–921 (2001)—and the sequence determined from one individual of a species may  
10 differ from other allelic forms present within the population. Small deletions and insertions can often be found that do not alter the function of the protein.

- Accordingly, the present invention provides GP354 polypeptides not only identical in sequence to those described with particularity herein, but also  
15 isolated proteins at least about 80% identical in sequence to those described with particularity herein, typically at least about 85%, 90%, 91%, 92%, 93%, 94%, or 95% identical in sequence to those described with particularity herein, usefully at least about 96%, 97%, 98%, or 99% identical in sequence to those described with particularity herein, and, most conservatively, at least about 99.5%, 99.6%, 99.7%,  
20 99.8% and 99.9% identical in sequence to those described with particularity herein. These sequence variants can be naturally occurring or can result from human intervention by way of random or directed mutagenesis.

- For purposes herein, percent identity of two amino acid sequences is determined using the procedure of Tatiana *et al.*, “Blast 2 sequences—a new tool  
25 for comparing protein and nucleotide sequences”, *FEMS Microbiol Lett.* 174:247–250 (1999), which procedure is effectuated by the computer program Blast 2 SEQUENCES, available online at the National Center for Biotechnology Information.

- To assess percent identity of amino acid sequences, the BlastP  
30 module of Blast 2 SEQUENCES is used with default values of (i) BLOSUM62 matrix, Henikoff *et al.*, *Proc. Natl. Acad. Sci USA* 89(22):10915–9 (1992); (ii) open

WO 03/000021

PCT/US02/20275

-53-

gap 11 and extension gap 1 penalties; and (iii) gap x\_dropoff 50 expect 10 word size 3 filter, and both sequences are entered in their entireties.

As is well known, amino acid substitutions occur frequently among natural allelic variants, with conservative substitutions often occasioning only *de minimis* change in protein function. Accordingly, the present invention provides proteins not only identical in sequence to those described with particularity herein, but also isolated proteins having the sequence of GP354 proteins, or portions thereof, with conservative amino acid substitutions. Also provided are isolated proteins having the sequence of GP354 proteins, and portions thereof, with moderately conservative amino acid substitutions. These conservatively-substituted or moderately conservatively-substituted variants can be naturally occurring or can result from human intervention.

Allelic variation may account for differences in amino acid sequence between SEQ ID NO:2 and SEQ ID NO:8 at positions 195, 196, 539 and 540, for example. Splice variants (*e.g.*, differential 5' or 3' splice site selection) may also account for the differences between the C-terminal amino acid sequences of SEQ ID NO:2 and SEQ ID NO:8.

As is also well known in the art, relatedness of proteins can also be characterized using a functional test, the ability of the encoding nucleic acids to base-pair to one another at defined hybridization stringencies. It is, therefore, another aspect of the invention to provide isolated proteins not only identical in sequence to those described with particularity herein, but also to provide isolated proteins ("hybridization related proteins") that are encoded by nucleic acids that hybridize under high stringency conditions (as defined herein above) to all or to a portion of various of the isolated polynucleotides of the present invention ("reference nucleic acids").

The hybridization related proteins can be alternative isoforms, homologs, paralogs, and orthologs of the GP354 protein of the present invention. Particularly useful orthologs are those from other primate species, such as chimpanzee, rhesus macaque monkey, baboon, orangutan, and gorilla; from

WO 03/000021

PCT/US02/20275

-54-

rodents, such as rats, mice, guinea pigs; from lagomorphs, such as rabbits, and from domestic livestock, such as cow, pig, sheep, horse, goat.

Relatedness of proteins can also be characterized using a second functional test, the ability of a first protein to inhibit competitively the binding of a second protein to an antibody. It is, therefore, another aspect of the present invention to provide isolated proteins not only identical in sequence to those described with particularity herein, but also to provide isolated proteins ("cross-reactive proteins") that competitively inhibit the binding of antibodies to all or to a portion of various of the isolated GP354 proteins of the present invention ("reference proteins"). Such competitive inhibition can readily be determined using immunoassays well known in the art.

Among the proteins of the present invention that differ in amino acid sequence from those described with particularity herein—including those that have deletions and insertions causing up to 10% non-identity, those having conservative or moderately conservative substitutions, hybridization related proteins, and cross-reactive proteins—those that substantially retain one or more GP354 activities are preferred (see *supra*).

Residues that are tolerant of change while retaining function can be identified by altering the protein at known residues using methods known in the art, such as alanine scanning mutagenesis, Cunningham *et al.*, *Science* 244(4908): 1081-5 (1989); transposon linker scanning mutagenesis, Chen *et al.*, *Gene* 263(1-2):39-48 (2001); combinations of homolog- and alanine-scanning mutagenesis, Jin *et al.*, *J. Mol. Biol.* 226(3):851-65 (1992); combinatorial alanine scanning, Weiss *et al.*, *Proc. Natl. Acad. Sci. USA* 97(16):8950-4 (2000), followed by functional assay. Transposon linker scanning kits are available commercially (New England Biolabs, Beverly, MA, USA, catalog no. E7-102S; EZ::TN™ In-Frame Linker Insertion Kit, catalog no. EZI04KN, Epicentre Technologies Corporation, Madison, WI, USA).

As further described below, the isolated proteins of the present invention can readily be used as specific immunogens to raise antibodies that specifically recognize GP354 proteins, their isoforms, homologs, paralogs, and/or

WO 03/000021

PCT/US02/20275

-55-

orthologs. The antibodies, in turn, can be used, *inter alia*, specifically to assay for the GP354 proteins of the present invention — *e.g.* by ELISA for detection of protein fluid samples, such as serum, by immunohistochemistry or laser scanning cytometry, for detection of protein in tissue samples, or by flow cytometry, for  
5 detection of intracellular protein in cell suspensions — for specific antibody-mediated isolation and/or purification of GP354 proteins, as for example by immunoprecipitation, and for use as specific agonists or antagonists of GP354 action.

The isolated proteins of the present invention are also immediately  
10 available for use as specific standards in assays used to determine the concentration and/or amount specifically of the GP354 proteins of the present invention. As is well known, ELISA kits for detection and quantitation of protein analytes typically include isolated and purified protein of known concentration for use as a measurement standard (*e.g.*, the human interferon- $\gamma$  OptEIA kit, catalog no.  
15 555142, Pharmingen, San Diego, CA, USA includes human recombinant gamma interferon, baculovirus produced).

The isolated proteins of the present invention are also immediately available for use as specific biomolecule capture probes for surface-enhanced laser desorption ionization (SELDI) detection of protein-protein interactions,  
20 WO 98/59362; WO 98/59360; WO 98/59361; and Merchant *et al.*, *Electrophoresis* 21(6):1164-77 (2000), the disclosures of which are incorporated herein by reference in their entireties. Analogously, the isolated proteins of the present invention are also immediately available for use as specific biomolecule capture probes on BIACORE surface plasmon resonance probes. See Weinberger *et al.*,  
25 *Pharmacogenomics* 1(4):395-416 (2000); Malmqvist, *Biochem. Soc. Trans.* 27(2):335-40 (1999).

The isolated proteins of the present invention are also useful as a therapeutic supplement in patients diagnosed to have a specific deficiency in GP354 production or activity.

30 The invention also provides fragments of various of the proteins of the present invention. The protein fragments are useful as antigenic and



WO 03/000021

PCT/US02/20275

-56-

immunogenic fragments of GP354. By "fragments" of a protein is here intended isolated proteins (equally, polypeptides, peptides, oligopeptides), however obtained, that have an amino acid sequence identical to a portion of the reference amino acid sequence, which portion is at least 6 amino acids and less than the entirety of the reference nucleic acid. As so defined, "fragments" need not be obtained by physical fragmentation of the reference protein, although such provenance is not thereby precluded.

Fragments of at least 6 contiguous amino acids are useful in mapping B cell and T cell epitopes of the reference protein. *See, e.g., Geysen et al., "Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid," Proc. Natl. Acad. Sci. USA 81:3998-4002 (1984) and U.S. Pat. Nos. 4,708,871 and 5,595,915, the disclosures of which are incorporated herein by reference in their entirety. Because the fragment need not itself be immunogenic, part of an immunodominant epitope, nor even recognized by native antibody, to be useful in such epitope mapping, all fragments of at least 6 amino acids of the proteins of the present invention have utility in such a study.*

Fragments of at least eight contiguous amino acids, often at least fifteen contiguous amino acids, have utility as immunogens for raising antibodies that recognize the proteins of the present invention. *See, e.g., Lerner, "Tapping the immunological repertoire to produce antibodies of predetermined specificity," Nature 299:592-596 (1982); Shinnick et al., "Synthetic peptide immunogens as vaccines," Annu. Rev. Microbiol. 37:425-46 (1983); Sutcliffe et al., "Antibodies that react with predetermined sites on proteins," Science 219:660-6 (1983), the disclosures of which are incorporated herein by reference in their entirety. As further described in the above-cited references, virtually all 8-mers, conjugated to a carrier, such as a protein, prove immunogenic — that is, prove capable of eliciting antibody for the conjugated peptide; accordingly, all fragments of at least 8 amino acids of the proteins of the present invention have utility as immunogens.*

Fragments of at least 8, 9, 10 or 12 contiguous amino acids are also useful as competitive inhibitors of binding of the entire protein, or a portion thereof, to antibodies (as in epitope mapping), and to natural binding partners, such

WO 03/000021

PCT/US02/20275

-57-

as subunits in a multimeric complex or to receptors or ligands of the subject protein; this competitive inhibition permits identification and separation of molecules that bind specifically to the protein of interest, U.S. Pat. Nos. 5,539,084 and 5,783,674, incorporated herein by reference in their entireties.

5       The protein, or protein fragment, of the present invention is thus at least 6 amino acids in length, typically at least 8, 9, 10 or 12 amino acids in length, and often at least 15 amino acids in length. Often, the protein or the present invention, or fragment thereof, is at least 20, 25, 30, 35, or 50 amino acids or more in length. Larger fragments having at least 75, 100, 150 or more amino acids are  
10 also useful, and at times preferred.

      The present invention further provides fusions of each of the GP354 proteins and protein fragments of the present invention to heterologous polypeptides. By fusion is here intended that the protein or protein fragment of the present invention is linearly contiguous to the heterologous polypeptide in a  
15 peptide-bonded polymer of amino acids or amino acid analogues; by "heterologous polypeptide" is here intended a polypeptide that does not naturally occur in contiguity with the protein or protein fragment of the present invention. As so defined, the fusion can consist entirely of a plurality of fragments of the GP354 protein in altered arrangement; in such case, any of the GP354 fragments can be  
20 considered heterologous to the other GP354 fragments in the fusion protein. More typically, however, the heterologous polypeptide is not drawn from the GP354 protein itself.

      The fusion proteins of the present invention will include at least one fragment of the protein of the present invention, which fragment is at least 6,  
25 typically at least 8, often at least 15, and usefully at least 16, 17, 18, 19, or 20 amino acids long. The fragment of the protein of the present to be included in the fusion can usefully be at least 25, 50, 75, 100, or 150 amino acids long. Fusions that include the entirety of the GP354 proteins of the invention, or functional domains, such as the N-terminal GP354 Ig domains and the C-terminal  
30 intracellular domain have particular utility. Fusions comprising GP354 Ig domains will be useful in engineering fusion proteins that will recognize other Ig domain-

WO 03/000021

PCT/US02/20275

-58-

containing molecules and cells that displaying them on their surface. This, in turn, may be useful for targeting a heterologous sequence, such as a toxin or a therapeutic, to a pancreatic cell or a CNS-derived cell that expressed GP354 or a binding partner; or to all or a portion of a cell surface molecule derived from a pancreatic cell or a CNS-derived cell that expresses GP354 or a binding partner.

The heterologous polypeptide included within the fusion protein of the present invention is at least 6 amino acids in length, often at least 8 amino acids in length, and preferably, at least 15, 20, and 25 amino acids in length. Fusions that include larger polypeptides, such as the IgG Fc region, and even entire proteins (such as luciferase or GFP chromophore-containing proteins), have particular utility.

As described above in the description of vectors and expression vectors of the present invention, heterologous polypeptides included in the fusion proteins of the present invention usefully include those designed to facilitate purification and/or visualization of recombinantly-expressed proteins. Although purification tags can also be incorporated into fusions that are chemically synthesized, chemical synthesis typically provides sufficient purity that further purification by HPLC suffices; however, visualization tags as above described retain their utility even when the protein is produced by chemical synthesis, and when so included render the fusion proteins of the present invention useful as directly detectable markers of GP354 presence.

As also discussed above, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those that facilitate secretion of recombinantly expressed proteins — into the periplasmic space or extracellular milieu for prokaryotic hosts, into the culture medium for eukaryotic cells — through incorporation of secretion signals and/or leader sequences.

Other useful protein fusions of the present invention include those that permit use of the protein of the present invention as bait in a yeast two-hybrid system. See Bartel *et al.* (eds.), The Yeast Two-Hybrid System, Oxford University Press (1997) (ISBN: 0195109384); Zhu *et al.*, Yeast Hybrid Technologies, Eaton

WO 03/000021

PCT/US02/20275

-59-

- Publishing, (2000) (ISBN 1-881299-15-5); Fields *et al.*, *Trends Genet.* 10(8):286-92 (1994); Mendelsohn *et al.*, *Curr. Opin. Biotechnol.* 5(5):482-6 (1994); Luban *et al.*, *Curr. Opin. Biotechnol.* 6(1):59-64 (1995); Allen *et al.*, *Trends Biochem. Sci.* 20(12):511-6 (1995); Drees, *Curr. Opin. Chem. Biol.* 3(1):64-70 (1999); Topcu *et al.*, *Pharm. Res.* 17(9):1049-55 (2000); Fashena *et al.*, *Gene* 250(1-2):1-14 (2000), the disclosures of which are incorporated herein by reference in their entireties. Typically, such fusion is to either *E. coli* LexA or yeast GAL4 DNA binding domains. Related bait plasmids are available that express the bait fused to a nuclear localization signal.
- 10 Other useful protein fusions include those that permit display of the encoded protein on the surface of a phage or cell, fusions to intrinsically detectable proteins, such as fluorescent or light-emitting proteins, and fusions to stable protein domains such as an immunoglobulin heavy chain domain like the IgG Fc region, as described above.
- 15 The proteins and protein fragments of the present invention can also usefully be fused to protein toxins, such as *Pseudomonas* exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, ricin, or other biologically deleterious moieties in order to effect specific ablation of cells that bind or take up the proteins of the present invention.
- 20 The isolated proteins, protein fragments, and protein fusions of the present invention can be composed of natural amino acids linked by native peptide bonds, or can contain any or all of nonnatural amino acid analogues, nonnative bonds, and post-synthetic (post-translational) modifications, either throughout the length of the protein or localized to one or more portions thereof.
- 25 As is well known in the art, when the isolated protein is used, *e.g.*, for epitope mapping, the range of such nonnatural analogues, nonnative inter-residue bonds, or post-synthesis modifications will be limited to those that permit binding of the peptide to antibodies. When used as an immunogen for the preparation of antibodies in a non-human host, such as a mouse, the range of such
- 30 nonnatural analogues, nonnative inter-residue bonds, or post-synthesis modifications will be limited to those that do not interfere with the immunogenicity

WO 03/000021

PCT/US02/20275

-60-

of the protein. When the isolated protein is used as a therapeutic agent, such as a vaccine or for replacement therapy, the range of such changes will be limited to those that do not confer toxicity upon the isolated protein.

Techniques for incorporating non-natural amino acids during solid phase chemical synthesis or by recombinant methods are well established in the art. Procedures are described, inter alia, in Chan *et al.* (eds.), Fmoc Solid Phase Peptide Synthesis: A Practical Approach (Practical Approach Series), Oxford Univ. Press (March 2000) (ISBN: 0199637245); Jones, Amino Acid and Peptide Synthesis (Oxford Chemistry Primers, No 7), Oxford Univ. Press (August 1992) (ISBN: 0198556683); and Bodanszky, Principles of Peptide Synthesis (Springer Laboratory), Springer Verlag (December 1993) (ISBN: 0387564314), the disclosures of which are incorporated herein by reference in their entireties.

D-enantiomers of natural amino acids can readily be incorporated during chemical peptide synthesis: peptides assembled from D-amino acids are more resistant to proteolytic attack; incorporation of D-enantiomers can also be used to confer specific three dimensional conformations on the peptide. Other amino acid analogues commonly added during chemical synthesis include ornithine, norleucine, phosphorylated amino acids (typically phosphoserine, phosphothreonine, phosphotyrosine), L-malonyltyrosine, a non-hydrolyzable analog of phosphotyrosine (Kole *et al.*, *Biochem. Biophys. Res. Com.* 209:817-821 (1995)), and various halogenated phenylalanine derivatives.

Amino acid analogues having detectable labels are also usefully incorporated during synthesis to provide a labeled polypeptide. Biotin, for example can be added using biotinyl--(9-fluorenylmethoxycarbonyl)-L-lysine (Fmoc biocytin) (Molecular Probes, Eugene, OR, USA). (Biotin can also be added enzymatically by incorporation into a fusion protein of a *E. coli* BirA substrate peptide.) The Fmoc and tBOC derivatives of dabcyL-L-lysine (Molecular Probes, Inc., Eugene, OR, USA) can be used to incorporate the dabcyL chromophore at selected sites in the peptide sequence during synthesis. The aminonaphthalene derivative EDANS, the most common fluorophore for pairing with the dabcyL quencher in fluorescence resonance energy transfer (FRET)

WO 03/000021

PCT/US02/20275

-61-

systems, can be introduced during automated synthesis of peptides by using EDANS--Fmoc-L-glutamic acid or the corresponding *t*BOC derivative (both from Molecular Probes, Inc., Eugene, OR, USA). Tetramethylrhodamine fluorophores can be incorporated during automated Fmoc synthesis of peptides using

5 (Fmoc)--TMR-L-lysine (Molecular Probes, Inc. Eugene, OR, USA).

Other useful amino acid analogues that can be incorporated during chemical synthesis include aspartic acid, glutamic acid, lysine, and tyrosine analogues having allyl side-chain protection (Applied Biosystems, Inc., Foster City, CA, USA); the allyl side chain permits synthesis of cyclic, branched-chain, sulfonated, glycosylated, and phosphorylated peptides. A large number of other Fmoc-protected non-natural amino acid analogues capable of incorporation during chemical synthesis are available commercially, e.g., from The Peptide Laboratory (Richmond, CA, USA).

15 Non-natural amino acid residues can also be added biosynthetically by engineering a suppressor tRNA, typically one that recognizes the UAG stop codon, by chemical aminoacylation with the desired unnatural amino acid and. Conventional site-directed mutagenesis is used to introduce the chosen stop codon UAG at the site of interest in the protein gene. When the acylated suppressor tRNA and the mutant gene are combined in an *in vitro* transcription/translation system, the unnatural amino acid is incorporated in response to the UAG codon to give a protein containing that amino acid at the specified position. Liu *et al.*, *Proc. Natl Acad. Sci. USA* 96(9):4780-5 (1999); Wang *et al.*, *Science* 292(5516):498-500 (2001).

25 The isolated GP354 proteins, protein fragments and fusion proteins of the present invention can also include non-native inter-residue bonds, including bonds that lead to circular and branched forms. The isolated GP354 proteins and protein fragments of the present invention can also include post-translational and post-synthetic modifications, either throughout the length of the protein or localized to one or more portions thereof.

30 For example, when produced by recombinant expression in eukaryotic cells, the isolated proteins, fragments, and fusion proteins of the present

WO 03/000021

PCT/US02/20275

-62-

invention will typically include N-linked and/or O-linked glycosylation, the pattern of which will reflect both the availability of glycosylation sites on the protein sequence and the identity of the host cell. Further modification of glycosylation pattern can be performed enzymatically. As another example, recombinant polypeptides of the invention may also include an initial modified methionine residue, in some cases resulting from host-mediated processes.

When the proteins, protein fragments, and protein fusions of the present invention are produced by chemical synthesis, post-synthetic modification can be performed before deprotection and cleavage from the resin or after deprotection and cleavage. Modification before deprotection and cleavage of the synthesized protein often allows greater control, *e.g.* by allowing targeting of the modifying moiety to the N-terminus of a resin-bound synthetic peptide. Useful post-synthetic (and post-translational) modifications include conjugation to detectable labels, such as fluorophores. A wide variety of amine-reactive and thiol-reactive fluorophore derivatives have been synthesized that react under nondenaturing conditions with N-terminal amino groups and epsilon amino groups of lysine residues, on the one hand, and with free thiol groups of cysteine residues, on the other.

Kits are available commercially that permit conjugation of proteins to a variety of amine-reactive or thiol-reactive fluorophores: Molecular Probes, Inc. (Eugene, OR, USA), *e.g.*, offers kits for conjugating proteins to Alexa Fluor 350, Alexa Fluor 430, Fluorescein-EX, Alexa Fluor 488, Oregon Green 488, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 568, Alexa Fluor 594, and Texas Red-X. A wide variety of other amine-reactive and thiol-reactive fluorophores are available commercially (Molecular Probes, Inc., Eugene, OR, USA), including Alexa Fluor® 350, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits), BODIPY dyes, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red.

The polypeptides of the present invention can also be conjugated to fluorophores, other proteins, and other macromolecules, using bifunctional linking reagents. Common homobifunctional reagents include, *e.g.*, APG, AEDP, BASED, BMB, BMDB, BMH, BMOE, BM[PEO]3, BM[PEO]4, BS3, BSOCEs, 5 DFDNB, DMA, DMP, DMS, DPDPB, DSG, DSP (Lomant's Reagent), DSS, DST, DTBP, DTME, DTSSP, EGS, HBVS, Sulfo-BSOCEs, Sulfo-DST, Sulfo-EGS (all available from Pierce, Rockford, IL, USA); common heterobifunctional cross-linkers include ABH, AMAS, ANB-NOS, APDP, ASBA, BMPA, BMPH, BMPS, EDC, EMCA, EMCH, EMCS, KMUA, KMH, GMBs, LC-SMCC, 10 LC-SPDP, MBS, M2C2H, MPBH, MSA, NHS-ASA, PDPH, PMPL, SADP, SAED, SAND, SANPAH, SASD, SATP, SBAP, SFAD, SIA, SIAB, SMCC, SMPB, SMPH, SMPT, SPDP, Sulfo-EMCS, Sulfo-GMBs, Sulfo-HSAB, Sulfo-KMUS, Sulfo-LC-SPDP, Sulfo-MBS, Sulfo-NHS-LC-ASA, Sulfo-SADP, Sulfo-SANPAH, Sulfo-SIAB, Sulfo-SMCC, Sulfo-SMPB, Sulfo-LC-SMPT, 15 SVSB, TFCS (all available Pierce, Rockford, IL, USA).

The proteins, protein fragments, and protein fusions of the present invention can be conjugated, using such cross-linking reagents, to fluorophores that are not amine- or thiol-reactive. Other labels that usefully can be conjugated to the proteins, protein fragments, and fusion proteins of the present invention include 20 radioactive labels, echosonographic contrast reagents, and MRI contrast agents.

The proteins, protein fragments, and protein fusions of the present invention can also usefully be conjugated using cross-linking agents to carrier proteins, such as KLH, bovine thyroglobulin, and even bovine serum albumin (BSA), to increase immunogenicity for raising anti-GP354 antibodies.

25 The GP354 proteins, protein fragments, and protein fusions of the present invention can also usefully be conjugated to polyethylene glycol (PEG); PEGylation increases the serum half life of proteins administered intravenously for replacement therapy. Delgado *et al.*, *Crit. Rev. Ther. Drug Carrier Syst.* 9(3-4):249-304 (1992); Scott *et al.*, *Curr. Pharm. Des.* 4(6):423-38 (1998); 30 DeSantis *et al.*, *Curr. Opin. Biotechnol.* 10(4):324-30 (1999), incorporated herein by reference in their entireties. PEG monomers can be attached to the protein



-64-

directly or through a linker, with PEGylation using PEG monomers activated with  
tresyl chloride (2,2,2-trifluoroethanesulphonyl chloride) permitting direct  
attachment under mild conditions.

The isolated GP354 proteins of the present invention, including  
5 fusions thereof, can be produced by recombinant expression, typically using the  
expression vectors of the present invention as above-described or, especially if  
fewer than about 100 amino acids, optionally by chemical synthesis (typically,  
solid phase synthesis), and, on occasion, by *in vitro* translation.

Production of the isolated proteins of the present invention can  
10 optionally be followed by purification. Purification of recombinantly expressed  
proteins is now well within the skill in the art. *See, e.g.,* Thorner *et al.* (eds.),  
Applications of Chimeric Genes and Hybrid Proteins, Part A: Gene Expression and  
Protein Purification (Methods in Enzymology, Volume 326), Academic Press  
(2000), (ISBN: 0121822273); Harbin (ed.), Cloning, Gene Expression and Protein  
15 Purification: Experimental Procedures and Process Rationale, Oxford Univ. Press  
(2001) (ISBN: 0195132947); Marshak *et al.*, Strategies for Protein Purification  
and Characterization: A Laboratory Course Manual, Cold Spring Harbor  
Laboratory Press (1996) (ISBN: 0-87969-385-1); and Roe (ed.), Protein  
Purification Applications, Oxford University Press (2001), the disclosures of which  
20 are incorporated herein by reference in their entireties, and thus need not be  
detailed here.

Briefly, however, if purification tags have been fused through use of  
an expression vector that appends such tag, purification can be effected, at least in  
part, by means appropriate to the tag, such as use of immobilized metal affinity  
25 chromatography for polyhistidine tags. Other techniques common in the art  
include ammonium sulfate fractionation, immuno-precipitation, fast protein liquid  
chromatography (FPLC), high performance liquid chromatography (HPLC), and  
preparative gel electrophoresis. Purification of chemically-synthesized peptides  
can readily be effected, *e.g.*, by HPLC.

30 Accordingly, it is an aspect of the present invention to provide the  
isolated GP354 proteins of the present invention in pure or substantially pure form.

WO 03/000021

PCT/US02/20275

-65-

A purified protein of the present invention is an isolated protein, as above described, that is present at a concentration of at least 95%, as measured on a mass basis (w/w) with respect to total protein in a composition. Such purities can often be obtained during chemical synthesis without further purification, as, *e.g.*, by

5 HPLC. Purified proteins of the present invention can be present at a concentration (measured on a mass basis with respect to total protein in a composition) of 96%, 97%, 98%, and even 99%. The proteins of the present invention can even be present at levels of 99.5%, 99.6%, and even 99.7%, 99.8%, or even 99.9% following purification, as by HPLC.

10 Although high levels of purity are preferred when the isolated proteins of the present invention are used as therapeutic agents — such as vaccines, or for replacement therapy — the isolated proteins of the present invention are also useful at lower purity. For example, partially purified proteins of the present invention can be used as immunogens to raise antibodies in laboratory animals.

15 Thus, the present invention provides the isolated proteins of the present invention in substantially purified form. A “substantially purified protein” of the present invention is an isolated protein, as above described, present at a concentration of at least 70%, measured on a mass basis with respect to total protein in a composition. Usefully, the substantially purified protein is present at a  
20 concentration, measured on a mass basis with respect to total protein in a composition, of at least 75%, 80%, or even at least 85%, 90%, 91%, 92%, 93%, 94%, 94.5% or even at least 94.9%.

In preferred embodiments, the purified and substantially purified proteins of the present invention are in compositions that lack detectable  
25 ampholytes, acrylamide monomers, bis-acrylamide monomers, and polyacrylamide.

The GP354 proteins, fragments, and fusions of the present invention can usefully be attached to a substrate. The substrate can porous, substantially nonporous (such as plastic), or solid; planar or non-planar; the bond can be  
30 covalent or noncovalent. Porous substrates, commonly membranes, typically comprise nitrocellulose, polyvinylidene fluoride (PVDF), or cationically

WO 03/000021

PCT/US02/20275

-66-

derivatized, hydrophilic PVDF; so bound, the proteins, fragments, and fusions of the present invention can be used to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized protein of the present invention. Proteins, fragments, and fusions of the present invention when bound to substantially

5 nonporous substrates, such as plastics, may be used to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized protein of the present invention.

The proteins, fragments, and fusions of the present invention can also be attached to a substrate suitable for use as a surface enhanced laser

10 desorption ionization source; so attached, the protein, fragment, or fusion of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound protein to indicate biologic interaction therebetween.

The proteins, fragments, and fusions of the present invention can

15 also be attached to a substrate suitable for use in surface plasmon resonance detection. So attached, the protein, fragment, or fusion of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound protein to indicate significant biological interaction between the two.

20 ANTIBODIES AND ANTIBODY-PRODUCING CELLS

The invention provides antibodies, including fragments and derivatives thereof, that bind specifically to GP354 proteins and protein fragments of the invention, or that bind to one or more of the proteins and protein fragments encoded by the isolated GP354 nucleic acids of the invention. The antibodies can

25 be specific for linear epitopes, discontinuous epitopes, or conformational epitopes of such proteins or protein fragments, either as present on the protein in its native conformation or, in some cases, as present on the proteins as denatured, *as, e.g.,* by solubilization in SDS.

The invention also provides antibodies, including fragments and

30 derivatives thereof, the binding of which can be competitively inhibited by one or

WO 03/000021

PCT/US02/20275

-67-

more of the GP354 proteins and protein fragments of the present invention, or by one or more of the proteins and protein fragments encoded by the isolated gp354 polynucleotides of the present invention.

5 In a first series of antibody embodiments, the invention provides antibodies, both polyclonal and monoclonal, and fragments and derivatives thereof, that bind specifically to a polypeptide having an amino acid sequence presented in SEQ ID NO:2, 4, 8, 10 or 12.

Such antibodies are useful in a variety of *in vitro* immunoassays, such as Western blotting and ELISA. Such antibodies are also useful in isolating and purifying GP354 proteins, including related cross-reactive proteins, by  
10 immuno-precipitation, immunoaffinity chromatography, or magnetic bead-mediated purification. Such methods are well-known in the art.

In a second series of antibody embodiments, the invention provides antibodies, both polyclonal and monoclonal, and fragments and derivatives thereof,  
15 the specific binding of which can be competitively inhibited by the isolated proteins and polypeptides of the present invention.

In other embodiments, the invention further provides the above-described antibodies detectably labeled, and in yet other embodiments, provides the above-described antibodies attached to a substrate.

20 Anti-PanCAM antibodies (Examples 11 and 19, *infra*) of this invention can be used to isolate pancreatic beta cells (e.g., from large animals such as cows, pigs, and horses; or from primates such as monkeys and chimpanzees). Beta cells so obtained are useful in making limited-lifetime encapsulated (and therefore immune-protected) xenografts in diabetic patients (such as human  
25 patients).

In some embodiments, the antibodies of this invention are used to target therapeutic agents to beta islets. For instance, type I diabetes is marked by T cell infiltration into beta islets. Thus, cytotoxic agents (i.e., agents that impair the viability and/or the functions of a cell) specific for these T cells (e.g., anti-CD3, -  
30 CD4, or -CD8 antibodies) can be conjugated to anti-PanCAM antibodies and thereby directed to the diseased site. In another example, antisense therapeutics

WO 03/000021

PCT/US02/20275

-68-

(such as antisense oligos to MHC class II molecules, which are highly expressed in T lymphocytes attacking beta cells; Chaturvedi et al., *J Immunol* 164(12):6610-20 (2000)) can be linked to anti-PanCAM antibodies and be directed to the beta islets.

The antibodies of this invention can also be conjugated to cytotoxic agents (e.g., nonspecific agents such as radioactive agents; targetable toxins such as ricin and cholera toxin; and chemotherapy agents such as doxorubicin) so as to direct the cytotoxins to the site of insulinomas, which are difficult to treat because they tend to be multi-focal. The antibodies may also be used to deliver DNA to the insulinoma as gene therapy to inhibit or otherwise modify the behavior of the tumor (e.g., to deliver an antisense oligonucleotide or DNA construct to either the responsible oncogene or the insulin gene itself, or to a gene in the signal transduction machinery regulating insulin release).

As used herein, the term "antibody" refers to a polypeptide, at least a portion of which is encoded by at least one immunoglobulin gene, which can bind specifically to a first molecular species, and to fragments or derivatives thereof that remain capable of such specific binding.

By "bind specifically" and "specific binding" is here intended the ability of the antibody to bind to a first molecular species in preference to binding to other molecular species with which the antibody and first molecular species are admixed. An antibody is said specifically to "recognize" a first molecular species when it can bind specifically to that first molecular species.

As is well known in the art, the degree to which an antibody can discriminate as among molecular species in a mixture will depend, in part, upon the conformational relatedness of the species in the mixture; typically, the antibodies of the present invention will discriminate over adventitious binding to non-GP354 proteins by at least two-fold, more typically by at least 5-fold, typically by more than 10-fold, 25-fold, 50-fold, 75-fold, and often by more than 100-fold, and on occasion by more than 500-fold or 1000-fold. When used to detect the proteins or protein fragments of the present invention, the antibody of the present invention is sufficiently specific when it can be used to determine the presence of

WO 03/000021

PCT/US02/20275

-69-

the protein of the present invention in samples derived from human pancreatic and neural tissues.

Typically, the affinity or avidity of an antibody (or antibody multimer, as in the case of an IgM pentamer) of the present invention for a GP354 protein or protein fragment of the present invention will be at least about  $1 \times 10^{-6}$  molar (M), typically at least about  $5 \times 10^{-7}$  M, usefully at least about  $1 \times 10^{-7}$  M, with affinities and avidities of at least  $1 \times 10^{-8}$  M,  $5 \times 10^{-9}$  M, and  $1 \times 10^{-10}$  M proving especially useful.

The antibodies of the present invention can be naturally-occurring forms, such as IgG, IgM, IgD, IgE, and IgA, from any mammalian species.

Human antibodies can, but will infrequently, be drawn directly from human donors or human cells. In such case, antibodies to the proteins of the present invention will typically have resulted from fortuitous immunization, such as autoimmune immunization, with the protein or protein fragments of the present invention. Such antibodies will typically, but will not invariably, be polyclonal.

Human antibodies are more frequently obtained using transgenic animals that express human immunoglobulin genes, which transgenic animals can be affirmatively immunized with a GP354 protein immunogen of the present invention. Human Ig-transgenic mice capable of producing human antibodies and methods of producing human antibodies therefrom upon specific immunization are well known in the art. See, e.g., in U.S. Patent Nos. 6,162,963; 6,150,584; 6,114,598; 6,075,181; 5,939,598; 5,877,397; 5,874,299; 5,814,318; 5,789,650; 5,770,429; 5,661,016; 5,633,425; 5,625,126; 5,569,825; 5,545,807; 5,545,806, and 5,591,669, the disclosures of which are incorporated herein by reference in their entireties. Such antibodies are typically monoclonal, and are typically produced using techniques developed for production of murine antibodies.

Human antibodies are particularly useful, and often preferred, when the antibodies of the present invention are to be administered to human beings as *in vivo* diagnostic or therapeutic agents, since recipient immune response to the administered antibody will often be substantially less than that occasioned by administration of an antibody derived from another species, such as mouse.

WO 03/000021

PCT/US02/20275

-70-

IgG, IgM, IgD, IgE and IgA antibodies of the present invention are also usefully obtained from other mammalian species, including rodents — typically mouse, but also rat, guinea pig, and hamster — lagomorphs, typically rabbits, and also larger mammals, such as sheep, goats, cows, and horses. In such cases, as with the transgenic human-antibody-producing non-human mammals, fortuitous immunization is not required, and the non-human mammal is typically affirmatively immunized, according to standard immunization protocols, with the protein or protein fragment of the present invention.

As discussed above, virtually all fragments of eight or more contiguous amino acids of the proteins of the present invention can be used effectively as immunogens when conjugated to a carrier, typically a protein such as bovine thyroglobulin, keyhole limpet hemocyanin, or bovine serum albumin, conveniently using a bifunctional linker such as those described elsewhere above, which discussion is incorporated by reference here.

Immunogenicity can also be conferred by fusion of the proteins and protein fragments of the present invention to other moieties. Peptides of the present invention can, for example, be produced by solid phase synthesis on a branched polylysine core matrix; these multiple antigenic peptides (MAPs) provide high purity, increased avidity, accurate chemical definition and improved safety in vaccine development. Tam *et al.*, *Proc. Natl. Acad. Sci. USA* 85:5409-5413 (1988); Posnett *et al.*, *J. Biol. Chem.* 263, 1719-1725 (1988).

Protocols for immunizing non-human mammals are well-established in the art, Harlow *et al.* (eds.), Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory (1998) (ISBN: 0879693142); Coligan *et al.* (eds.), Current Protocols in Immunology, John Wiley & Sons, Inc. (2001) (ISBN: 0-471-52276-7); Zola, Monoclonal Antibodies : Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Background to Bench), Springer Verlag (2000) (ISBN: 0387915907), the disclosures of which are incorporated herein by reference.

Antibodies from nonhuman mammals can be polyclonal or monoclonal, with polyclonal antibodies having certain advantages in immuno-

-71-

histochemical detection of the proteins of the present invention and monoclonal antibodies having advantages in identifying and distinguishing particular epitopes of the proteins of the present invention.

Following immunization, the antibodies of the present invention can be produced using any art-accepted technique. Such techniques are well known in the art, Coligan *et al.* (eds.), Current Protocols in Immunology, John Wiley & Sons, Inc. (2001) (ISBN: 0-471-52276-7); Zola, Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Background to Bench), Springer Verlag (2000) (ISBN: 0387915907); Howard *et al.* (eds.), Basic Methods in Antibody Production and Characterization, CRC Press (2000) (ISBN: 0849394457); Harlow *et al.* (eds.), Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory (1998) (ISBN: 0879693142); Davis (ed.), Monoclonal Antibody Protocols, Vol. 45, Humana Press (1995) (ISBN: 0896033082); Delves (ed.), Antibody Production: Essential Techniques, John Wiley & Son Ltd (1997) (ISBN: 0471970107); Kenney, Antibody Solution: An Antibody Methods Manual, Chapman & Hall (1997) (ISBN: 0412141914), incorporated herein by reference in their entireties, and thus need not be detailed here.

Recombinant expression in host cells is particularly useful when fragments or derivatives of the antibodies of the present invention are desired. Host cells for recombinant antibody production—either whole antibodies, antibody fragments, or antibody derivatives—can be prokaryotic or eukaryotic.

Prokaryotic hosts are particularly useful for producing phage displayed antibodies of the present invention. The technology of phage-displayed antibodies, in which antibody variable region fragments are fused, for example, to the gene III protein (pIII) or gene VIII protein (pVIII) for display on the surface of filamentous phage, such as M13, is by now well-established, Sidhu, *Curr. Opin. Biotechnol.* 11(6):610-6 (2000); Griffiths *et al.*, *Curr. Opin. Biotechnol.* 9(1):102-8 (1998); Hoogenboom *et al.*, *Immunotechnology*, 4(1):1-20 (1998); Rader *et al.*, *Curr. Opin. Biotechnol.* 8:503-508 (1997); Aujame *et al.*, *Hum. Antibodies* 8:155-168 (1997); Hoogenboom, *Trends Biotechnol.* 15:62-70 (1997); de Kruif *et al.*,



WO 03/000021

PCT/US02/20275

-72-

17:453-455 (1996); Barbas *et al.*, *Trends Biotechnol.* 14:230-234 (1996); Winter *et al.*, *Annu. Rev. Immunol.* 433-455 (1994), and techniques and protocols required to generate, propagate, screen (pan), and use the antibody fragments from such libraries have recently been compiled, Barbas *et al.*, Phage Display: A Laboratory  
 5 Manual, Cold Spring Harbor Laboratory Press (2001) (ISBN 0-87969-546-3); Kay *et al.* (eds.), Phage Display of Peptides and Proteins: A Laboratory Manual, Academic Press, Inc. (1996); Abelson *et al.* (eds.), Combinatorial Chemistry, Methods in Enzymology vol. 267, Academic Press (May 1996), the disclosures of which are incorporated herein by reference in their entireties. Typically, phage-  
 10 displayed antibody fragments are scFv fragments or Fab fragments; when desired, full length antibodies can be produced by cloning the variable regions from the displaying phage into a complete antibody and expressing the full length antibody in a further prokaryotic or a eukaryotic host cell.

Eukaryotic cells are also useful for expression of the antibodies,  
 15 antibody fragments, and antibody derivatives of the present invention. For example, antibody fragments of the present invention can be produced in *Pichia pastoris*, Takahashi *et al.*, *Biosci. Biotechnol. Biochem.* 64(10):2138-44 (2000); Freyre *et al.*, *J. Biotechnol.* 76(2-3):157-63 (2000); Fischer *et al.*, *Biotechnol. Appl. Biochem.* 30 (Pt 2):117-20 (1999); Pennell *et al.*, *Res. Immunol.* 149(6):599-603  
 20 (1998); Eldin *et al.*, *J. Immunol. Methods.* 201(1):67-75 (1997); and in *Saccharomyces cerevisiae*, Frenken *et al.*, *Res. Immunol.* 149(6):589-99 (1998); Shusta *et al.*, *Nature Biotechnol.* 16(8):773-7 (1998), the disclosures of which are incorporated herein by reference in their entireties.

Antibodies, including antibody fragments and derivatives, of the  
 25 invention can also be produced in insect cells, Li *et al.*, *Protein Expr. Purif.* 21(1):121-8 (2001); Ailor *et al.*, *Biotechnol. Bioeng.* 58(2-3):196-203 (1998); Hsu *et al.*, *Biotechnol. Prog.* 13(1):96-104 (1997); Edelman *et al.*, *Immunology* 91(1):13-9 (1997); and Nesbit *et al.*, *J. Immunol. Methods.* 151(1-2):201-8 (1992), the disclosures of which are incorporated herein by reference in their entireties.

30 Antibodies and fragments and derivatives thereof of the present invention may also be produced in plant cells, Giddings *et al.*, *Nature Biotechnol.*

WO 03/000021

PCT/US02/20275

-73-

18(11):1151-5 (2000); Gavilondo *et al.*, *Biotechniques* 29(1):128-38 (2000); Fischer *et al.*, *J. Biol. Regul. Homeost. Agents* 14(2):83-92 (2000); Fischer *et al.*, *Biotechnol. Appl. Biochem.* 30 (Pt 2):113-6 (1999); Fischer *et al.*, *Biol. Chem.* 380(7-8):825-39 (1999); Russell, *Curr. Top. Microbiol. Immunol.* 240:119-38  
5 (1999); and Ma *et al.*, *Plant Physiol.* 109(2):341-6 (1995), the disclosures of which are incorporated herein by reference in their entireties.

Mammalian cells useful for recombinant expression of antibodies, antibody fragments, and antibody derivatives of the present invention include CHO cells, COS cells, 293 cells, and myeloma cells. Verma *et al.*, *J. Immunol. Methods*  
10 216(1-2):165-81 (1998), review and compare bacterial, yeast, insect and mammalian expression systems for expression of antibodies.

Antibodies of the present invention may also be prepared by cell free translation, as further described in Merk *et al.*, *J. Biochem. (Tokyo)*. 125(2):328-33 (1999) and Ryabova *et al.*, *Nature Biotechnol.* 15(1):79-84 (1997),  
15 and in the milk of transgenic animals, as further described in Pollock *et al.*, *J. Immunol. Methods* 231(1-2):147-57 (1999), the disclosures of which are incorporated herein by reference in their entireties.

The invention further provides antibody fragments that bind specifically to one or more of the GP354 proteins and protein fragments of the present invention, to one or more of the proteins and protein fragments encoded by  
20 the isolated gp354 polynucleotides of the present invention, or the binding of which can be competitively inhibited by one or more of the proteins and protein fragments of the present invention or one or more of the proteins and protein fragments encoded by the isolated polynucleotides of the present invention.

Among such useful fragments are Fab, Fab', Fv, F(ab)'<sub>2</sub>, and single chain Fv (scFv) fragments. Other useful fragments are described in Hudson, *Curr. Opin. Biotechnol.* 9(4):395-402 (1998). The present invention thus provides  
25 antibody derivatives that bind specifically to one or more of the GP354 proteins and protein fragments of the present invention, to one or more of the proteins and protein fragments encoded by the isolated nucleic acids of the present invention, or  
30 the binding of which can be competitively inhibited by one or more of the proteins

WO 03/00021

PCT/US02/20275

-74-

and protein fragments of the present invention or one or more of the proteins and protein fragments encoded by the isolated polynucleotides of the present invention.

Among such useful derivatives are chimeric, primatized, and humanized antibodies; such derivatives are less immunogenic in human beings, and thus more suitable for *in vivo* administration, than are unmodified antibodies from non-human mammalian species.

Chimeric antibodies typically include heavy and/or light chain variable regions (including both CDR and framework residues) of immunoglobulins of one species, typically mouse, fused to constant regions of another species, typically human. See, e.g., U.S. Pat. No. 5,807,715; Morrison *et al.*, *Proc. Natl. Acad. Sci USA* 81(21):6851-5 (1984); Sharon *et al.*, *Nature* 309(5966):364-7 (1984); Takeda *et al.*, *Nature* 314(6010):452-4 (1985), the disclosures of which are incorporated herein by reference in their entireties.

Primatized and humanized antibodies typically include heavy and/or light chain CDRs from a murine antibody grafted into a non-human primate or human antibody V region framework, usually further comprising a human constant region, Riechmann *et al.*, *Nature* 332(6162):323-7 (1988); Co *et al.*, *Nature* 351(6326):501-2 (1991); U.S. Pat. Nos. 6,054,297; 5,821,337; 5,770,196; 5,766,886; 5,821,123; 5,869,619; 6,180,377; 6,013,256; 5,693,761; and 6,180,370, the disclosures of which are incorporated herein by reference in their entireties.

Other useful antibody derivatives of the invention include heteromeric antibody complexes and antibody fusions, such as diabodies (bispecific antibodies), single-chain diabodies, and intrabodies.

The antibodies of the present invention, including fragments and derivatives thereof, can usefully be labeled. It is, therefore, another aspect of the present invention to provide labeled antibodies that bind specifically to one or more of the proteins and protein fragments of the present invention, to one or more of the GP354 proteins and protein fragments encoded by the isolated polynucleotides of the present invention, or the binding of which can be competitively inhibited by one or more of the proteins and protein fragments of the

WO 03/000021

PCT/US02/20275

-75-

present invention or one or more of the proteins and protein fragments encoded by the isolated polynucleotides of the present invention.

The choice of label depends, in part, upon the desired use. When the antibodies of the present invention are used for immunohistochemical staining of tissue samples, the label can usefully be an enzyme that catalyzes production and local deposition of a detectable product. Enzymes typically conjugated to antibodies to permit their immunohistochemical visualization are well known, and include alkaline phosphatase,  $\beta$ -galactosidase, glucose oxidase, horseradish peroxidase (HRP), and urease. The antibodies of the invention can also be labeled using colloidal gold.

A multitude of typical substrates for production and deposition of visually detectable products, luminescent and fluorescent labels, are also well known and need not be further described. See, e.g., Thorpe *et al.*, *Methods Enzymol.* 133:331-53 (1986); Kricka *et al.*, *J. Immunoassay* 17(1):67-83 (1996); and Lundqvist *et al.*, *J. Biolumin. Chemilumin.* 10(6):353-9 (1995), the disclosures of which are incorporated herein by reference in their entireties. Kits for enhanced chemiluminescent detection (ECL) are available commercially.

When the antibodies of the present invention are used, e.g., for flow cytometric detection, for scanning laser cytometric detection, or for fluorescent immunoassay, they can usefully be labeled with fluorophores. There are a wide variety of fluorophore labels that can usefully be attached to the antibodies of the present invention. Many are available, e.g., from Molecular Probes, Inc., Eugene, OR, USA.

For secondary detection using labeled avidin, streptavidin, captavidin or neutravidin, the antibodies of the present invention can usefully be labeled with biotin.

When the antibodies of the present invention are used, e.g., for Western blotting applications, they can usefully be labeled with radioisotopes, such as  $^{33}\text{P}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^3\text{H}$ , and  $^{125}\text{I}$ . As another example, when the antibodies of the present invention are used for radioimmunotherapy, the label can usefully be  $^{228}\text{Th}$ ,  $^{227}\text{Ac}$ ,  $^{225}\text{Ac}$ ,  $^{223}\text{Ra}$ ,  $^{213}\text{Bi}$ ,  $^{212}\text{Pb}$ ,  $^{212}\text{Bi}$ ,  $^{211}\text{At}$ ,  $^{203}\text{Pb}$ ,  $^{194}\text{Os}$ ,  $^{188}\text{Re}$ ,  $^{186}\text{Re}$ ,  $^{153}\text{Sm}$ ,  $^{149}\text{Tb}$ ,  $^{131}\text{I}$ ,  $^{125}\text{I}$ ,

WO 03/000021

PCT/US02/20275

-76-

<sup>111</sup>In, <sup>105</sup>Rh, <sup>99m</sup>Tc, <sup>97</sup>Ru, <sup>90</sup>Y, <sup>90</sup>Sr, <sup>88</sup>Y, <sup>72</sup>Se, <sup>67</sup>Cu, or <sup>47</sup>Sc. As another example, when the antibodies of the present invention are to be used for *in vivo* diagnostic use, they can be rendered detectable by conjugation to MRI contrast agents, such as gadolinium diethylenetriaminepentaacetic acid (DTPA), Lauffer *et al.*, *Radiology* 5 207(2):529-38 (1998), or by radioisotopic labeling. As would be understood by the skilled artisan, use of any of the labels described above is not restricted to the application as for which they were mentioned.

The antibodies of the present invention, including fragments and derivatives thereof, can also be conjugated to biologically deleterious moieties, such as toxins, in order to target the toxin's ablative action to cells that display 10 and/or express the proteins of the present invention. Commonly, the antibody in such immunotoxins is conjugated to *Pseudomonas* exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, or ricin. See Hall (ed.), Immunotoxin Methods and Protocols (Methods in Molecular Biology, Vol 166), Humana Press 15 (2000) (ISBN:0896037754); and Frankel *et al.* (eds.), Clinical Applications of Immunotoxins, Springer-Verlag New York, Incorporated (1998) (ISBN:3540640975), the disclosures of which are incorporated herein by reference in their entireties, for review.

The antibodies of the present invention can usefully be attached to a 20 substrate. The invention thus provides antibodies that bind specifically to one or more of the GP354 proteins and protein fragments of the present invention, to one or more of the proteins and protein fragments encoded by the isolated polynucleotides of the present invention, or the binding of which can be competitively inhibited by one or more of the proteins and protein fragments of the 25 present invention or one or more of the proteins and protein fragments encoded by the isolated polynucleotides of the present invention, attached to a substrate. Substrates can be porous or nonporous, planar or nonplanar.

For example, the antibodies of the present invention can usefully be conjugated to filtration media, such as NHS-activated Sepharose or CNBr- 30 activated Sepharose for purposes of immunoaffinity chromatography.

WO 03/000021

PCT/US02/20275

-77-

The antibodies of the present invention can also usefully be attached to paramagnetic microspheres, typically by biotin-streptavidin interaction, which microsphere can then be used for isolation of cells that express or display the proteins of the present invention. As another example, the antibodies of the present invention can usefully be attached to the surface of a microtiter plate for ELISA.

As noted above, the antibodies of the present invention can be produced in prokaryotic and eukaryotic cells. The invention thus also provides cells that express the antibodies of the present invention, including hybridoma cells, B cells, plasma cells, and host cells recombinantly modified to express the antibodies of the present invention.

The present invention also provides aptamers evolved to bind specifically to one or more of the GP354 proteins and protein fragments of the present invention, to one or more of the proteins and protein fragments encoded by the isolated polynucleotides of the present invention, or the binding of which can be competitively inhibited by one or more of the proteins and protein fragments of the present invention or one or more of the proteins and protein fragments encoded by the isolated polynucleotides of the present invention.

#### PHARMACEUTICAL COMPOSITIONS AND THERAPEUTIC METHODS

GP354 is a new member of the immunoglobulin (Ig) superfamily expressed predominantly in the pancreas and in lower amounts in neural tissue, e.g., the CNS. GP354, an integral cell surface membrane protein, has five signature Ig domains in its extracellular portion which are known in other family members to mediate cell-cell recognition and adhesion reactions. As a member of the Ig superfamily, GP354 is likely important for mediating cell-cell recognition, binding and adhesion functions in the pancreatic, neural and potentially other tissues in which it is expressed.

As a pancreatic enriched protein, GP354 will be a suitable therapeutic target for treating abnormal conditions, disorders and/or diseases related to improper cell-cell binding, adhesion and signaling in the pancreas,

WO 03/000021

PCT/US02/20275

-78-

particularly during tissue development and during tissue regeneration and/or healing, e.g., after pancreatic damage, trauma or degenerative conditions. It is also envisioned that GP354 will be useful for inhibiting pancreatic cell death associated with immune, auto-immune, and degenerative conditions. It is envisioned that the

5 neural form of GP354 will be a similarly suitable therapeutic target for tissue regeneration and repair and for inhibiting degeneration and cell death in CNS tissue.

The invention accordingly provides pharmaceutical compositions comprising nucleic acids, proteins, and antibodies of the present invention, as well

10 as mimetics, agonists, antagonists, or modulators of GP354 activity, may be administered as pharmaceutical agents for the treatment (i.e., the amelioration of) of disorders, conditions or diseases associated with mis-expression of GP354 or to overcome abnormal expression or activities of other components which participate in GP354 related molecular and cellular recognition pathways. As GP354

15 expression is relatively concentrated in the pancreas, it is anticipated that GP354 mis-expression may be associated with pancreatic disorder or disease, and/or with congenital defects in pancreatic development of function.

Disorders and diseases of the pancreas, for which administration of a composition of the invention may be useful, include acute pancreatitis (often but

20 not always manifesting in abnormal pancreatic exocrine functions, such as elevated serum, ascitic and/or pleural fluid amylase levels, or abnormal lipase or trypsinogen levels. Pancreatic inflammation and necrosis are also associated with acute as well as with chronic pancreatitis and exocrine insufficiency. A variety of pancreatic endocrine tumors have been characterized, and auto-immune disorders

25 which affect the pancreas have also been described. For a more detailed description of diagnoses and treatments of pancreatic disorders and diseases, see Harrison's PRINCIPLES OF INTERNAL MEDICINE, 14<sup>th</sup> Ed., (Anthony S. Fauci et al., editors), McGraw-Hill Companies, Inc., 1998, Part Eleven, Section 3, the disclosure of which is incorporated by reference in its entirety.

30 GP354 expression is also detected in neural CNS tissue, albeit at lower levels than is detected in the pancreas. It is therefore envisioned that GP354

WO 03/000021

PCT/US02/20275

-79-

mis-expression may be associated with neural dysfunction, disorder or disease, or abnormal development of the CNS. Examples of neural disorders which may be ameliorated by treatment with a composition of the invention include, without limitation, Alzheimer's disease, Parkinson's disease, senile dementia, migraine, epilepsy, neuritis, neurasthenia, neuropathy, and any other diseases involving GP354-mediated neural migration, neural degeneration (e.g., GP354-mediated autoimmune diseases such as certain forms of multiple sclerosis), and neural tumors (e.g., glioma, astroblastoma, and astrocytoma).

Some other diseases for which compositions of the invention may have utility include endocrine and hormonal problems (e.g., diabetes), pancreatic diseases, cancers (particularly pancreatic cancer), and the like. The use of GP354 modulators, including GP354 antisense reagents, GP354 ligands and anti-GP354 antibodies, to treat individuals having or at risk of developing such diseases is an aspect of the invention.

A composition of the invention typically contains from about 0.1 to 90% by weight (such as 1 to 20% or 1 to 10%) of a therapeutic agent of the invention in a pharmaceutically accepted carrier. Solid formulations of the compositions for oral administration can contain suitable carriers or excipients, such as corn starch, gelatin, lactose, acacia, sucrose, microcrystalline cellulose, kaolin, mannitol, dicalcium phosphate, calcium carbonate, sodium chloride, or alginic acid. Disintegrators that can be used include, without limitation, microcrystalline cellulose, corn starch, sodium starch glycolate, and alginic acid. Tablet binders that can be used include acacia, methylcellulose, sodium carboxymethylcellulose, polyvinylpyrrolidone(Povidone™), hydroxypropyl methylcellulose, sucrose, starch, and ethylcellulose. Lubricants that can be used include magnesium stearates, stearic acid, silicone fluid, talc, waxes, oils, and colloidal silica.

Liquid formulations of the compositions for oral administration prepared in water or other aqueous vehicles can contain various suspending agents such as methylcellulose, alginates, tragacanth, pectin, kelgin, carrageenan, acacia, polyvinylpyrrolidone, and polyvinyl alcohol. The liquid formulations can also



WO 03/000021

PCT/US02/20275

-80-

include solutions, emulsions, syrups and elixirs containing, together with the active compound(s), wetting agents, sweeteners, and coloring and flavoring agents. Various liquid and powder formulations can be prepared by conventional methods for inhalation into the lungs of the mammal to be treated.

- 5           Injectable formulations of the compositions can contain various carriers such as vegetable oils, dimethylacetamide, dimethylformamide, ethyl lactate, ethyl carbonate, isopropyl myristate, ethanol, polyols (glycerol, propylene glycol, liquid polyethylene glycol, and the like). For intravenous injections, water soluble versions of the compounds can be administered by the drip method,
- 10   whereby a pharmaceutical formulation containing the antifungal agent and a physiologically acceptable excipient is infused. Physiologically acceptable excipients can include, for example, 5% dextrose, 0.9% saline, Ringer's solution or other suitable excipients. Intramuscular preparations, *e.g.*, a sterile formulation of a suitable soluble salt form of the compounds, can be dissolved and administered
- 15   in a pharmaceutical excipient such as Water-for-Injection, 0.9% saline, or 5% glucose solution. A suitable insoluble form of the compound can be prepared and administered as a suspension in an aqueous base or a pharmaceutically acceptable oil base, such as an ester of a long chain fatty acid (*e.g.*, ethyl oleate).

- A topical semi-solid ointment formulation typically contains a
- 20   concentration of the active ingredient from about 1 to 20%, *e.g.*, 5 to 10%, in a carrier such as a pharmaceutical cream base. Various formulations for topical use include drops, tinctures, lotions, creams, solutions, and ointments containing the active ingredient and various supports and vehicles. The optimal percentage of the therapeutic agent in each pharmaceutical formulation varies according to the
- 25   formulation itself and the therapeutic effect desired in the specific pathologies and correlated therapeutic regimens.

Inhalation and transdermal formulations can also readily be prepared.

- Pharmaceutical formulation is a well-established art, and is further
- 30   described in Gennaro (ed.), Remington: The Science and Practice of Pharmacy, 20<sup>th</sup> ed., Lippincott, Williams & Wilkins (2000) (ISBN: 0683306472); Ansel *et al.*,

WO 03/000021

PCT/US02/20275

-81-

Pharmaceutical Dosage Forms and Drug Delivery Systems, 7<sup>th</sup> ed., Lippincott Williams & Wilkins Publishers (1999) (ISBN: 0683305727); and Kibbe (ed.), Handbook of Pharmaceutical Excipients American Pharmaceutical Association, 3<sup>d</sup> ed. (2000) (ISBN: 091733096X), the disclosures of which are incorporated herein by reference in their entireties. Conventional methods, known to those of ordinary skill in the art of medicine, can be used to administer the pharmaceutical formulation(s) to the patient.

Typically, the pharmaceutical formulation will be administered to the patient by applying to the skin of the patient a transdermal patch containing the pharmaceutical formulation, and leaving the patch in contact with the patient's skin (generally for 1 to 5 hours per patch). Other transdermal routes of administration (e.g., through use of a topically applied cream, ointment, or the like) can be used by applying conventional techniques. The pharmaceutical formulation(s) can also be administered via other conventional routes (e.g., enteral, subcutaneous, intrapulmonary, transmucosal, intraperitoneal, intrauterine, sublingual, intrathecal, or intramuscular routes) by using standard methods. In addition, the pharmaceutical formulations can be administered to the patient via injectable depot routes of administration such as by using 1-, 3-, or 6-month depot injectable or biodegradable materials and methods.

Regardless of the route of administration, the therapeutic protein or antibody agent typically is administered at a daily dosage of 0.01 mg to 30 mg/kg of body weight of the patient (e.g., 1mg/kg to 5 mg/kg). The pharmaceutical formulation can be administered in multiple doses per day, if desired, to achieve the total desired daily dose. The effectiveness of the method of treatment can be assessed by monitoring the patient for known signs or symptoms of a disorder.

The pharmaceutical compositions of the invention may be included in a container, package or dispenser alone or as part of a kit with labels and instructions for administration.

TRANSGENIC ANIMALS AND CELLS

WO 03/000021

PCT/US02/20275

-82-

In another aspect, the invention provides transgenic cells and non-human organisms comprising gp354 isoform nucleic acids, and transgenic cells and non-human organisms with targeted disruption of the endogenous ortholog of the human gp354 gene. The cells can be embryonic stem cells or somatic cells. The transgenic non-human organisms can be chimeric, non-chimeric heterozygotes, and non-chimeric homozygotes.

Host cells of the invention may be used to produce non-human transgenic animals. For example, in some embodiments, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which gp354 nucleotide sequences have been introduced. Such a host cell may be used to create non-human transgenic animals in which exogenous gp354 sequences have been introduced into their genome or used to alter or replace related endogenous gp354 sequences in the animal.

As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a cow, goat, sheep, or rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, dogs, chickens, amphibians, etc.

As used herein, a "transgene" is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal.

As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous gp354 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

The non-human transgenic animals of the invention will be useful for studying the function and/or activity of gp354 and for identifying and/or evaluating modulators of gp354 activity. They will also be useful in methods for

WO 03/00021

PCT/US02/20275

-83-

producing a GP354 protein or polypeptides fragment, i.e., in which the protein is produced in the mammary gland of a non-human mammal.

A transgenic animal of the invention can be created by introducing gp354-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. A polynucleotide comprising or having human gp354 DNA sequences of SEQ ID NO:1, 3, 5, 6, 7, 9, or 11, may be introduced as a transgene into the genome of a non-human animal. Alternatively, a non-human homolog of the human gp354 gene, such as a mouse gp354 gene, isolated by hybridization to an isolated polynucleotide of the invention, may be used as a transgene. Heterologous transcription control sequence sequences, intronic sequences, polyadenylation signals and the like may also be operatively linked with the transgene to increase the efficiency or otherwise regulate the expression (*e.g.*, in a developmental or tissue specific manner) the transgene in the recipient host animal.

Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the gp354 transgene in its genome and/or expression of gp354 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding gp354 can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a gp354 gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the gp354 gene. The gp354 gene can be a human gene (*e.g.*, SEQ ID NO:1, 5, 9 or 11), but more preferably, is a non-human homolog of a human gp354 gene. For

WO 03/000021

PCT/US02/20275

-84-

example, a mouse homolog of the human gp354 gene of SEQ ID NO:1, 5, 9 or 11 or can be used to construct a homologous recombination vector suitable for altering an endogenous gp354 gene in the mouse genome.

In some embodiments, the vector is designed such that, upon  
5 homologous recombination, the endogenous gp354 gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous gp354 gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to  
10 thereby alter the expression of the endogenous GP354 protein). In the homologous recombination vector, the altered portion of the gp354 gene is flanked at its 5' and 3' ends by additional nucleic acid of the gp354 gene to allow for homologous recombination to occur between the exogenous gp354 gene carried by the vector and an endogenous gp354 gene in an embryonic stem cell. The additional flanking  
15 gp354 nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See e.g., Thomas et al. *Cell* 51:503 (1987) for an exemplary description of homologous recombination vectors.

The vector is introduced into an embryonic stem cell line (e.g., by  
20 electroporation) and cells in which the introduced gp354 gene has homologously recombined with the endogenous gp354 gene are selected (see e.g., Li et al. *Cell* 69:915 (1992)). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras. See e.g., Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL  
25 APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the  
30 transgene.

WO 03/000021

PCT/US02/20275

-85-

Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley, *Curr. Opin. Biotechnol.* 2:823-829 (1991); PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

5 Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al. Nature* 385:810-813 (1997). In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G0 phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

15 Regulated expression of transgenes *in vivo* may be accomplished using controllable recombination systems, such as the cre/loxP recombinase system (see, *e.g.*, Lakso *et al. Proc. Natl. Acad. Sci. USA* 89:6232-6236 (1992)) and the FLP recombinase system (O'Gorman *et al. Science* 251:1351-1355 (1991)). If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Transgenic animals containing both elements of the system can be obtained, *e.g.*, by mating two transgenic animals, each containing either the transgene encoding the selected protein or the transgene encoding a recombinase.

#### ANTISENSE REAGENTS AND METHODS

##### 25 A. Antisense

Many of the isolated polynucleotides of the invention are antisense polynucleotides that recognize and hybridize to gp354 polynucleotides. Full-length and fragment antisense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that specifically recognize and hybridize to gp354 RNA (as determined by sequence comparison of DNA

WO 03/000021

PCT/US02/20275

-86-

encoding GP354 to DNA encoding other known molecules). Identification of sequences unique to GP354 encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be performed. Antisense polynucleotides are particularly relevant to regulating expression of GP354 by those cells expressing gp354 mRNA.

Antisense oligonucleotides, or fragments of a nucleotide sequence set forth in SEQ ID NO:1, 3, 5, 6, 7, 9 or 11, or sequences complementary or homologous thereto, derived from the nucleotide sequences encoding GP354 are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed in situ with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this enzyme or pathological conditions relating thereto. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire gp354 coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a GP354 protein of SEQ ID NO:2, 4, 8, 10 or 12, antisense nucleic acids complementary to a GP354 nucleic acid sequence of SEQ ID NO:1, 3, 5, 6, 7, 9 or 11 are additionally provided.

Antisense nucleic acid molecules of the invention may be antisense to a "coding region" or non-coding regions of the coding strand of a nucleotide sequence encoding GP354. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (e.g., a protein coding region of human GP354 corresponds to the coding region presented in SEQ ID NO:1, 7 or 11).

Antisense oligonucleotides are preferably directed to a regulatory region of a nucleotide sequence of SEQ ID NO:1, 7 or 11, or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like. The antisense nucleic acid molecule can be

-87-

complementary to the entire coding or non-coding region of gp354, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or non-coding region of gp354 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of gp354 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length.

Antisense nucleic acids of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention (preferably oligonucleotides of 10 to 20 nucleotides in length) are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a GP354 protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. Suppression of gp354 expression at either the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by aberrant gp354 expression.

The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. Phosphorothioate and



WO 03/000021

PCT/US02/20275

-88-

methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin, polylysine, or cholesterol moieties at their 5' end.

5           An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens  
10           expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule  
15           is placed under the control of a strong pol II or pol III promoter are preferred.

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

#### B. Ribozymes and Catalytic Nucleic Acids

25           In still another series of embodiments, an antisense nucleic acid of the invention is part of a gp354 specific ribozyme (or, as modified, a "nucleozyme"). Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (such as  
30           hammerhead, hairpin, Group I intron ribozymes, and the like) can be used to

WO 03/000021

PCT/US02/20275

-89-

catalytically cleave gp354 mRNA transcripts to thereby inhibit translation of gp354 mRNA. A ribozyme having specificity for a gp354-encoding nucleic acid can be designed based upon the nucleotide sequence of a gp354 polynucleotide disclosed herein (SEQ ID NO:1, 3, 5, 6, 7, 9, or 11). See, e.g., U.S. Patent Nos. 5,116,742; 5,334,711; 5,652,094; and 6,204,027, incorporated herein by reference in their entireties.

For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a GP354-encoding mRNA. See, e.g., Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, gp354 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel *et al.*, *Science* 261:1411-1418 (1993).

Expression of the gp354 gene may be inhibited by targeting nucleotide sequences complementary to the regulatory region of the gp354 (*e.g.*, the gp354 promoter and/or enhancers) to form triple helical structures that prevent transcription of the gp354 gene in target cells. See generally, Helene, *Anticancer Drug Des.* 6: 569-84 (1991); Helene *et al.* *Ann. N.Y. Acad. Sci.* 660:27-36 (1992); and Maher *Bioassays* 14: 807-15 (1992).

#### 20 C. Peptide Nucleic Acids (PNA)

In other preferred oligonucleotide mimetics, especially useful for *in vivo* administration, both the sugar and the internucleoside linkage are replaced with novel groups, such as peptide nucleic acids (PNA). See, e.g., Hyrup *et al.* *Bioorg. Med. Chem. Lett.* 4:5-23 (1996). In PNA compounds, the phosphodiester backbone of the nucleic acid is replaced with an amide-containing backbone, in particular by repeating N-(2-aminoethyl) glycine units linked by amide bonds. Nucleobases are bound directly or indirectly to aza-nitrogen atoms of the amide portion of the backbone, typically by methylene carbonyl linkages. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis

WO 03/000021

PCT/US02/20275

-90-

protocols as described in Hyrup et al., *supra*; and Perry-O'Keefe et al., *Proc. Natl. Acad. Sci. USA* 93:14670-675 (1996).

PNAs of gp354 can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of gp354 can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases; or as probes or primers for DNA sequence and hybridization (Hyrup et al., *supra*; and Perry-O'Keefe, *supra*).

In other embodiments, PNAs of gp354 can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of gp354 can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup et al., *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup, *supra* and Finn et al., *Nucl. Acids Res.* 24:3357-63 (1996).

For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl) amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al., *Nucl. Acids Res.* 17:5973-88 (1989)). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al., *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al., *Bioorg. Med. Chem. Lett.* 5:1119-11124 (1975).

WO 03/000021

PCT/US02/20275

-91-

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, *Proc. Natl. Acad. Sci. USA* 86:6553-6556 (1989); Lemaitre *et al.*, *Proc. Natl. Acad. Sci. USA* 84:648-652 (1987); PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, *BioTechniques* 6:958-976 (1988)), or intercalating agents (See, *e.g.*, Zon, *Pharm. Res.* 5: 539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

PNA chemistry and applications are reviewed, *inter alia*, in Ray *et al.*, *FASEB J.* 14(9):1041-60 (2000); Nielsen *et al.*, *Pharmacol Toxicol.* 86(1):3-7 (2000); Larsen *et al.*, *Biochim Biophys Acta.* 1489(1):159-66 (1999); Nielsen, *Curr. Opin. Struct. Biol.* 9(3):353-7 (1999), and Nielsen, *Curr. Opin. Biotechnol.* 10(1):71-5 (1999), the disclosures of which are incorporated herein by reference in their entireties.

#### DIAGNOSTIC METHODS

##### A. Nucleic Acid Diagnostics

As described above, the isolated polynucleotides of the invention can be used as nucleic acid probes to assess the levels of gp354 mRNA in tissues in which it is normally expressed (*e.g.*, pancreas and CNS), and in tissues in which it is not normally expressed, if such abnormal tissue mis-expression is suspected.

The invention thus provides a method for detecting the presence of a gp354 polynucleotide in a biological sample (*e.g.*, a cell extract, fluid or tissue sample derived from a patient) by contacting the sample with an isolated polynucleotide of the invention which is capable of specifically detecting by hybridization gp354 polynucleotide sequences.

Preferably, the method comprises the steps of contacting the sample with an the isolated nucleic acid under high stringency hybridization conditions

WO 03/000021

PCT/US02/20275

-92-

and detecting hybridization of the isolated polynucleotide to a nucleic acid in the sample, wherein the occurrence of said hybridization indicates the presence of a gp354-encoding sequence in the sample.

The isolated polynucleotides of the invention can be used as nucleic acid probes that are specific to particular cell types in the pancreas and central nervous system based on the specific expression of gp354 in these tissues. Accordingly, the present invention provides a method for identifying a cell as a pancreatic or a neural cell by detecting the presence of a gp354 polynucleotide in a biological sample (*e.g.*, a cell extract, fluid or tissue sample derived from a patient) by contacting the sample with an isolated polynucleotide of the invention which is capable of specifically detecting by hybridization gp354 polynucleotide sequences.

The present invention also provides a diagnostic assay for identifying the presence or absence of a genetic lesion or mutation characterized by at least one of: (i) aberrant modification or mutation of a gene encoding a GP354 protein; (ii) mis-regulation of a gene encoding a GP354 protein; and (iii) aberrant post-translational modification of a GP354 protein, wherein a wild-type form of the gene encodes a protein with a GP354 biological activity.

The present invention further provides a method of identifying a homolog of a human gp354 gene, comprising the step of hybridizing a nucleic acid library with a nucleic acid probe comprising SEQ ID NO:1, 3, 5, 6, 7, 9 or 11, or a portion thereof having at least 17 nucleotides, under medium or high stringency hybridization conditions; and determining whether the nucleic acid probe hybridizes to a nucleic acid sequence in the library. If the nucleic acid sequence in the library hybridizes under such selected conditions, it is a homolog of a human gp354 gene.

#### B. Antibody Diagnostics

Antibodies of the present invention can be used to assess the expression levels of GP354 proteins in tissues in which it is normally expressed (*e.g.*, pancreas and CNS), and in tissues in which it is not normally expressed, if such abnormal tissue mis-expression is suspected.

WO 03/000021

PCT/US02/20275

-93-

The invention thus provides a method for detecting the presence of a GP354 protein or its activity in a biological sample (*e.g.*, a cell extract, fluid or tissue sample derived from a patient) by contacting the sample with an agent capable of detecting an indicator of the presence of GP354 protein or its activity.

- 5 Preferably, the agent is an antibody specific for at least one epitope of GP354 protein.

Accordingly, the invention provides a method for determining whether a GP354 protein is present in a sample, comprising the step of contacting the sample with an antibody having at least one GP354 epitope and  
10 detecting specific binding of the antibody to an antigen, which indicates the presence of a GP354 protein in the sample.

- The above method will also be useful for identifying a test cell in a subject as a pancreatic or a neural cell by comparing the amount of GP354 polypeptides present in a biological sample (*e.g.*, a cell extract, fluid or tissue sample derived from the subject) from the subject test cell to the amount of GP354  
15 polypeptides present in a parallel biological sample from non-pancreatic or non-neural tissue.

- In some embodiments, anti-PanCAM antibodies described herein can be used to quantify pancreatic islet injury in early type I diabetes, as well as any  
20 therapy aimed at restoring islet function. For instance, the antibodies are attached to a contrast marker (*e.g.*, Gadolinium) for MRI and CT applications, or to a radioactive marker (*e.g.*, Technecium) for nuclear medicine applications. The relevant diagnostic modality is then used to quantify and document changes in islet sizes and/or numbers. The antibodies so labeled may also localize islet cell tumors  
25 using such a diagnostic modality.

C. Methods for Diagnosing Disease

The gp354 isolated polynucleotides, proteins and GP354 specific antibodies of the invention will be useful in methods for diagnosing a variety of disorders and disease conditions associated with aberrant gp354 expression.

- 30 The invention thus provides a method for diagnosing a disease condition in a subject, comprising the steps of comparing the amount or activity of

WO 03/000021

PCT/US02/20275

-94-

a GP354 protein in a tissue sample from the subject to the amount or activity of the GP354 polypeptide in a control sample (e.g., an equivalent one derived from a healthy subject), wherein a significant difference in the amount or activity of the GP354 polypeptide in the tissue sample relative to the amount or activity of the GP354 polypeptide in the control sample indicates that the subject has a disease condition.

In preferred embodiments, the amount or activity of a GP354 protein in a tissue sample is assessed by competitive binding assays using a GP354 polypeptides or fragment of the invention, or by an immunoassay using a GP354 specific antibody of the invention. Preferably, the method is used to diagnose a disease condition relating to the pancreas or to the nervous system.

Also provided are methods for diagnosing a disease condition in a subject by monitoring relative gp354 mRNA levels in difference tissues. Preferably, the methods comprise the step of comparing the amount of a gp354 mRNA in a test tissue sample from the subject to the amount of gp354 mRNA in a control sample, wherein a significant difference in the amount of the mRNA in the test sample relative to the amount in the control sample indicates that the subject has a disease condition.

In preferred embodiments, the amount of gp354 mRNA in a tissue sample is assessed by hybridization using an isolated gp354 polynucleotide or nucleic acid fragment of the invention. Preferably, the method is used to diagnose a disease condition relating to the pancreas or to the nervous system.

#### COMPUTER READABLE MEANS

A further aspect of the invention is a computer readable means for storing the gp354 nucleic acid and amino acid sequences of the instant invention. In preferred embodiments, the invention provides a computer readable means for storing SEQ ID NOs as described herein, as the complete set of sequences or in any combination. The records of the computer readable means can be accessed for reading and display and for interface with a computer system for the application of programs allowing for the location of data upon a query for data meeting certain

WO 03/000021

PCT/US02/20275

-95-

criteria, the comparison of sequences, the alignment or ordering of sequences meeting a set of criteria, and the like.

The nucleic acid and amino acid sequences of the invention are particularly useful as components in databases useful for search analyses as well as in sequence analysis algorithms. As used in these embodiments, the terms "nucleic acid sequences of the invention" and "amino acid sequences of the invention" mean any detectable chemical or physical characteristic of a polynucleotide or polypeptide of the invention that is or may be reduced to or stored in a computer readable form. These include, without limitation, chromatographic scan data or peak data, photographic data or scan data therefrom, and mass spectrographic data.

This invention provides computer readable media having stored thereon sequences of the invention. A computer readable medium may comprise one or more of the following: a nucleic acid sequence comprising a sequence of a nucleic acid sequence of the invention; an amino acid sequence comprising an amino acid sequence of the invention; a set of nucleic acid sequences wherein at least one of said sequences comprises the sequence of a nucleic acid sequence of the invention; a set of amino acid sequences wherein at least one of said sequences comprises the sequence of an amino acid sequence of the invention; a data set representing a nucleic acid sequence comprising the sequence of one or more nucleic acid sequences of the invention; a data set representing a nucleic acid sequence encoding an amino acid sequence comprising the sequence of an amino acid sequence of the invention; a set of nucleic acid sequences wherein at least one of said sequences comprises the sequence of a nucleic acid sequence of the invention; a set of amino acid sequences wherein at least one of said sequences comprises the sequence of an amino acid sequence of the invention; a data set representing a nucleic acid sequence comprising the sequence of a nucleic acid sequence of the invention; a data set representing a nucleic acid sequence encoding an amino acid sequence comprising the sequence of an amino acid sequence of the invention. The computer readable medium can be any composition of matter used to store information or data, including, for example, commercially available floppy disks, tapes, hard drives, compact disks, and video disks.



WO 03/000021

PCT/US02/20275

-96-

Accordingly, the invention provides a diagnostic assay for identifying a homolog of a human gp354 gene, comprising the step of screening a nucleic acid database with a query sequence consisting of SEQ ID NO:1, 3, 5, 6, 7, 9 or 11, or a portion thereof having 300 or more nucleotides, wherein a nucleic acid sequence in said database that is at least 65% but less than 100% identical to SEQ ID NO:1, 3, 5, 6, 7, 9 or 11, or said portion thereof, if found, is a homolog of a human gp354 gene.

Also provided by the invention are methods for the analysis of character sequences, particularly genetic sequences of the invention. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such as identity and similarity analysis, RNA structure analysis, sequence assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, and sequencing chromatogram peak analysis.

A computer-based method is provided for performing nucleic acid homology identification. This method comprises the steps of providing a nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and comparing said nucleic acid sequence to at least one nucleic acid or amino acid sequence to identify homology.

A computer-based method is also provided for performing amino acid homology identification, said method comprising the steps of: providing an amino acid sequence comprising the sequence of a polypeptide of the invention in a computer readable medium; and comparing said amino acid sequence to at least one nucleic acid or an amino acid sequence to identify homology.

A computer based method is still further provided for assembly of overlapping nucleic acid sequences into a single nucleic acid sequence, said method comprising the steps of: providing a first nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and screening for at least one overlapping region between said first nucleic acid sequence and a second nucleic acid sequence.

WO 03/000021

PCT/US02/20275

-97-

**EXAMPLES**

The following example is meant to illustrate the methods and materials of the present invention. Suitable modifications and adaptations of the described conditions and parameters normally encountered in the art of molecular biology which are apparent to those skilled in the art are within the spirit and scope of the present invention.

For the experiments described below, all RT-PCR and fragments were gel-purified prior to cloning. The fragments were separated by agarose gel electrophoresis by standard methods. DNA fragments were excised from the agarose gel and purified from the gel using QIAEX resin according to the manufacturer's specifications (Qiagen, Valencia, CA). The gel-purified fragments were cloned into plasmid vectors and then the plasmids were used to transform competent TOP10 *E. coli* host cells. Plasmids produced by the host cells were isolated by a standard alkaline lysis miniprep procedure (Qiagen, Valencia, CA). Sequencing was executed by a standard dideoxy termination method (Applied Biosystems, Foster City, CA).

**Example 1: *In Silico* Gene Prediction and Sequence Analysis**

The gene prediction software programs GENSCAN (Burge *et al.*, *J. Mol. Biol.* 268:78-94 (1997)) and GENEMARKHMM (Lukashin *et al.*, *Nucl. Acids Res.* 26:1107-1115 (1998)) were used to identify novel genes in the high throughput genomic sequences deposited in GenBank. To do so, the Genbank data entries were downloaded to a local server, and individual sequence contigs were separated according to the annotation provided with the sequence entries. The parameters used in the analyses were the default parameters included with the programs (Burge *et al.*, *supra*; and Lukashin *et al.*, *supra*).

Gene candidates with at least two identical exons predicted from GENSCAN and GENEMARKHMM were further analyzed. Specifically, the gene sequences were translated to protein sequences which were in turn used as queries in Blast analyses of a non-redundant protein sequence database incorporating Genpept, Swissprot, and PIR. The protein sequences were also analyzed for signal peptides using SIGNALP (Nielsen *et al.*, *Int J. Neural Syst* 8:581-99 (1997)).

-98-

Candidate genes encoding proteins with signal peptides and similarity to Ig superfamily proteins were subjected to experimental verification and cloning.

The BLAST ("Basic Local Alignment Search Tool") algorithm is suitable for determining sequence similarity (Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990)). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information website. This algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension for the word hits in each direction are halted when: (1) the cumulative alignment score falls off by the quantity X from its maximum achieved value; (2) the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or (3) the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:10915-10919 (1992)) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

BLAST (Karlin *et al.*, *Proc. Natl. Acad. Sci. USA*, 90:5873-5787 (1993)) and GAPPED BLAST perform a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a gp354 gene or cDNA if the smallest sum probability in comparison of the test nucleic acid to gp354 is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

WO 03/000021

PCT/US02/20275

-99-

The gp354 gene (ORF) was identified in contig 38 of a BAC with the GenBank accession number AC022315, which was deposited on February 10, 2000. The GENSCAN prediction for this gene was in the reverse orientation and included the following 14 exons, shown in TABLE 3.

5

TABLE 3: GENSCAN results

| Exon | Begin | End  | Length |
|------|-------|------|--------|
| 14   | 1844  | 1779 | 66     |
| 13   | 3567  | 3464 | 104    |
| 12   | 4007  | 3903 | 105    |
| 11   | 4695  | 4476 | 220    |
| 10   | 4959  | 4859 | 101    |
| 09   | 5378  | 5246 | 133    |
| 08   | 5591  | 5464 | 128    |
| 07   | 5981  | 5833 | 149    |
| 06   | 6203  | 6098 | 106    |
| 05   | 7019  | 6869 | 151    |
| 04   | 7796  | 7636 | 161    |
| 03   | 8092  | 7943 | 150    |
| 02   | 9157  | 9008 | 150    |
| 01   | 9373  | 9322 | 52     |

BLAST analysis of the gp354 gene against publicly available EST databases showed no ESTs that matched the predicted gene.

**Example 2: Amplification of gp354**

A sequence of gp354 cDNA is obtained by performing rapid amplification of cDNA ends (RACE) using the MARATHON-READY RACE kit (Clontech, Palo Alto, CA). A MARATHON-READY cDNA is a double-stranded cDNA synthesized from human tissue mRNA and ligated to a standard set of

WO 03/000021

PCT/US02/20275

-100-

adapters (Clontech). All RACE reactions use an adapter primer AP-1, 5'-CCATCCTAATACGACTCACTATAGGGC-3' (SEQ ID NO:15) provided with the kit. The 3' RACE for gp354 may use AP-1 together with the forward primer GX1-218, 5'-ACTGGGGGCTAGTTCAGTGGACTAA-3' (SEQ ID NO:16), or the complement of the reverse primer, GX1-219, 5'-CCAAACAGCACATCCAGCGCAGTAC-3' (SEQ ID NO:17). The 5' RACE for gp354 may use AP-1 together with the reverse primer GX1-219, or the complement of the forward primer GX1-218. ADVANTAGE 2 DNA polymerase (Clontech) may be used for the amplification reactions. The MARATHON-READY kit may be used according to the manufacturer's specifications except that "touchdown" PCR (Don et al., Nuc. Acids Res. 19:4008 (1991)) conditions are used for thermal cycling. The thermal cycling conditions are as follows: 94°C for 1 minute, one cycle of 94°C for 15 seconds, 72°C for 15 seconds, 68°C for 15 seconds; one cycle of 94°C for 15 seconds, 71°C for 15 seconds, 68°C for 15 seconds; one cycle of 94°C for 15 seconds, 70°C for 15 seconds, 68°C for 15 seconds; one cycle of 94°C for 15 seconds, 69°C for 15 seconds, 68°C for 15 seconds; 35 cycles of 94°C for 15 seconds and 68°C for 30 seconds; and 68°C for 10 minutes.

**Example 3: Confirmation of GP354 Expression by RT-PCR**

Inter-exon PCR was used to confirm that the predicted gp354 gene was indeed expressed and to initiate the cloning process that would determine the true (rather than the predicted) gene structure. Templates for the PCR were single-stranded cDNAs contained in a multi-tissue cDNA panel generated by reverse transcription PCR ("RT-PCR") from mRNA isolated from multiple human tissues according to the manufacturer's specifications (Clontech). These tissues were brain, heart, kidney, liver, lung, pancreas, pituitary, skeletal muscle, colon, ovary, peripheral blood leukocyte, prostate, small intestine, spleen, testis, and thymus.

Primers for the inter-exon PCR were selected from separate exons so that every PCR experiment had a built-in control to distinguish spliced messages from unspliced messages or from genomic contamination in the cDNA panels. One pair of such primers were GX1-220 and GX1-221 (Example 13, *infra*).

WO 03/000021

PCT/US02/20275

-101-

Another pair used were GX1-218 and GX1-219 (*supra*). The PCR was carried out according to the specification included with the multiple tissue cDNA panel (Clontech). Typical thermal cycler conditions for the PCR were: 94°C for 1 minute, followed by 35 cycles of 94°C for 20 seconds, 68°C for 2 minutes, followed by 5 minutes at 68°C after the last cycle. PCR products were separated on 1.0% agarose gel and visualized by ethidium bromide staining.

The results showed a fragment of approximately 785 bp in the pancreas and in no other tissues. The same observation was made when GX1-218 and GX1-219. This fragment was isolated and cloned into the PCR2.1 plasmid vector (Invitrogen, Carlsbad, CA). The resultant plasmid construct CS0026 was propagated in TOP10 *E. coli* host cells. Plasmids were purified by a standard alkaline lysis miniprep procedure (Qiagen, Valencia, CA). Sequencing was executed by standard di-deoxy terminator methods (Applied Biosystems, Foster City, CA). The sequence of the PCR fragment (SEQ ID NO:3) was compared with the sequences of the predicted gp354 gene to confirm the accuracy of the gene prediction.

**Example 4: Identification of Full-Length gp354 cDNA by RACE**

Because the gene prediction programs GENSCAN and GENEMARK have predictable error rates (Burge et al., *supra*; Lukashin et al., *supra*), the PCR fragment described in Example 3 are used as a seed sequence to obtain the rest of the gp354 cDNA sequence via RACE reactions. For the 3' RACE reaction, the primer is GX1-218 or the complement of GX1-219, and the template is cDNAs derived from human pancreas tissue (see Example 3). For the 5' RACE, the primer is GX1-219 or the complement of GX1-218, and the template is also cDNAs derived from human pancreas tissue. The 5' and 3' RACE fragments so obtained are gel-purified, cloned, and sequenced. To assemble the full-length gp354 cDNA sequence, the initial PCR product, the 5' RACE product and the 3'RACE product are assembled into a single contiguous sequence using the ASSEMBLE program in the GCG computer package (Genetics Computer Group, Madison, Wisconsin).

**Example 5: Confirmation of GP354 Expression by Northern Blot Analysis**

-102-

To confirm the expression of GP354, Northern blot analysis was conducted with each lane of the blot (Clontech catalogue no. 7760-1) containing 2 µg of polyA RNA. The tissues represented on the blot included heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas. The probe for the Northern blot was the PCR fragment described in Example 3 (SEQ ID NO:3). 50 ng of the probe was labeled by the random-primed method of Feinberg and Vogelstein (*Anal. Biochem.* 132:6-13 (1983)). Hybridization was carried out at 68°C for one hour in EXPRESSHYB solution (Clontech catalogue no. 8015-1). Prior to autoradiography, the Northern blot was washed with 2X SSC/0.05% SDS at room temperature, followed by two washes with 0.1X SSC/0.1% SDS at 50°C.

A hybridizing RNA of approximately 3.2 kb was observed in the pancreas but not in any other tissues tested.

**Example 6: PCR Screening of A Genomic Library And Subcloning of GP354 Coding Regions**

Subcloning of the gp354 genomic locus may be accomplished by PCR from a genomic library, or directly from genomic DNA. For example, two microliters of a human genomic library (~10<sup>8</sup> PFU/ml) (Clontech) are added to 6 ml of an overnight culture of K802 cells (Clontech), and then distributed as 250 µl aliquots into each of 24 microtubes. The microtubes are incubated at 37°C for 15 min. Seven milliliters of 0.8% agarose is added to each tube, mixed, then poured onto LB agar + 10 mM MgSO<sub>4</sub> plates and incubated overnight at 37°C. To each plate 5 ml of SM phage buffer (0.1 M NaCl, 8.1 mM MgSO<sub>4</sub>·7H<sub>2</sub>O, 50 mM Tris·Cl (pH 7.5), 0.01% gelatin) is added and the top agarose is removed with a microscope slide and placed in a 50 ml centrifuge tube. A drop of chloroform is added and the tube is placed in a 37°C shaker for 15 min, then centrifuged for 20 min at 4000 rpm (Sorvall RT6000 table top centrifuge) and the supernatant stored at 4°C as a stock solution.

PCR may be then performed in 20 µl containing 8.8 µl H<sub>2</sub>O, 4 µl 5X RAPID-LOAD BUFFER (Origene), 2 µl 10X PCR BUFFER II (Perkin Elmer), 2 µl 25 mM MgCl<sub>2</sub>, 0.8 µl 10 mM dNTP, 0.12 µl of a primer comprising at least a portion of the sequence of the 5' end of the gp354 polynucleotide of SEQ

WO 03/000021

PCT/US02/20275

-103-

ID NO:1 (1 mg/ml), 0.12 ml of a primer comprising at least a portion of the sequence that is complementary to the 3' end of the gp354 polynucleotide of SEQ ID NO:1 (1 mg/ml), 0.2 ml AMPLITAQ GOLD polymerase (Perkin Elmer) and 2 ml of phage solution from each of the 24 tubes. The PCR reaction involves 1 cycle  
 5 at 80°C for 20 min, 95°C for 10 min, then 22 cycles at 95°C for 30 sec, 72°C for 4 min decreasing 1°C each cycle, 68°C for 2 min, followed by 30 cycles at 95°C for 30 sec, 55°C for 30 sec, 68°C for 60 sec. The reaction is loaded onto a 2% agarose gel.

From the tube that gives a PCR product of the correct size, 5 µl is  
 10 used to set up five 1:10 dilutions that are plated onto LB agar + 10 mM MgSO<sub>4</sub> plates and incubated overnight. A BA85 nitrocellulose filter (Schleicher & Schuell) is placed on top of each plate for 1 hour. The filter is removed, placed with the phage side up in a petri dish, and covered with 4 ml of SM buffer for 15 min to elute the phage. One milliliter of SM buffer is removed from each plate and  
 15 used to set up a PCR reaction as described above. The plate of the lowest dilution to give a PCR product is subdivided, filter-lifted and the PCR reaction is repeated. The series of dilutions and subdivisions of the plate is continued until a single plaque is isolated that gives a positive PCR band. Once a single plaque is isolated, 10 ml phage supernatant is added to 100 ml SM and 200 ml of K802 cells per plate  
 20 with a total of 8 plates set up. The plates are incubated overnight at 37°C. Eight milliliters of SM is added to each plate, and the top agarose is scraped off with a microscope slide and collected in a centrifuge tube.

Three drops of chloroform are added to the centrifuge tube. Subsequently, the tube is vortexed, incubated at 37°C for 15 min, and centrifuged  
 25 for 20 min at 4000 rpm (Sorvall RT6000 table top centrifuge) to recover the phage. The recovered phage is used to isolate genomic phage DNA using the QIAGEN LAMBDA MIDI KIT. The sequences for primers may be derived from the sequences given herein.

To subclone the coding region of the gp354 gene, PCR is performed  
 30 in a 50 µl reaction containing 33 µl H<sub>2</sub>O, 5 µl 10X TT buffer (140 mM ammonium sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 µl 15 mM MgSO<sub>4</sub>, 2 µl 10 mM



WO 03/000021

PCT/US02/20275

-104-

dNTP, 4 µl genomic phage DNA (0.1 µg/ml), 0.3 µl of a primer comprising at least a portion of the 5' most coding sequence of the gp354 polynucleotide of SEQ ID NO:1 (1 µg/ml), 0.3 µl of a primer comprising a sequence that is complementary to at least a portion of the 3' most coding sequence of the gp354 polynucleotide of  
 5 SEQ ID NO:1 (1 µg/ml), 0.4 µl HIGH FIDELITY Taq polymerase (Boehringer Mannheim). The PCR reaction is started with 1 cycle of 94°C for 2 min followed by 15 cycles at 94°C for 30 sec, 55°C for 60 sec., and 68°C for 2 min.

The PCR product is loaded onto a 2% agarose gel. The DNA band of expected size is excised from the gel, placed in GENELUTE AGAROSE spin column (Supelco) and spun for 10 min at maximum speed. The eluted DNA is  
 10 ethanol-precipitated and resuspended in 12 µl H<sub>2</sub>O for ligation. The PCR primer sequences may be derived from the sequences provided herein.

The ligation reaction uses solutions from the TOPO TA Cloning Kit (Invitrogen). The reaction proceeds in a solution containing 4 µl of PCR product  
 15 and 1 µl of pCRII-TOPO vector at room temperature for 5 min. The reaction is terminated by the addition of 1 µl of 6X TOPO Cloning Stop Solution. The ligation product is then placed on ice. Two microliters of the ligation reaction is used to transform ONE-SHOT TOP10 cells (Invitrogen). Briefly, the ligation reaction is mixed with the cells and placed on ice for 30 min. The cells are then  
 20 heat-shocked for 30 seconds at 42°C and placed on ice for two minutes. Next, 250 µl of SOC is added to the cells, which are incubated at 37°C with shaking for one hour and then plated onto ampicillin plates.

A single colony from the plates is used to inoculate a 5 ml culture of LB medium. Plasmid DNA is purified from the culture using the CONCERT  
 25 RAPID PLASMID MINIPREP SYSTEM (GibcoBRL) and the insert of the plasmid DNA is then sequenced.

The gp354 genomic phage DNA may be sequenced using the ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems), which uses the advanced capillary electrophoresis technology and the ABI PRISM BIGDYE Terminator  
 30 Cycle Sequencing Ready Reaction Kit. The cycle-sequencing reaction may contain 14 ml of H<sub>2</sub>O, 16 ml of BIGDYE Terminator mix, 7 ml genomic phage DNA (0.1

WO 03/000021

PCT/US02/20275

-105-

mg/ml), and 3 ml primer (25 ng/ml). The reaction is performed in a Perkin-Elmer 9600 thermocycler at 95°C for 5 min, followed by 99 cycles of 95°C for 30 sec, 55°C for 20 sec, and 60°C for 4 min. The product is purified using a CENTRIFLEX gel filtration cartridge, dried under vacuum, and then dissolved in 16 µl of Template Suppression Reagent (PE Applied Biosystems). The samples are heated at 95°C for 5 min and then placed in the 310 Genetic Analyzer.

The DNA subcloned into pCRII is sequenced using the ABI PRISM 310 Genetic Analyzer, *supra*. Each cycle-sequencing reaction contains 6 ml of H<sub>2</sub>O, 8 ml of BIGDYE Terminator mix, 5 ml of miniprep DNA (0.1 mg/ml), and 1 ml of primer (25 ng/ml) and is performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product is purified using a CENTRIFLEX gel filtration cartridge, dried under vacuum, and then dissolved in 16 µl of Template Suppression Reagent. The samples are heated at 95°C for 5 min and then placed in the 310 Genetic Analyzer.

**Example 7: Hybridization Analysis To Demonstrate GP354 Expression in Brain**

The expression of gp354 in mammals, such as rat, may be investigated by *in situ* hybridization histochemistry. To investigate gp354 expression in the pancreas, for example, coronal and sagittal rat pancreas cryosections (20 µm thick) are prepared using a Reichert-Jung cryostat. Individual sections are thaw-mounted onto silanized, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. Sections are processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold phosphate-buffered saline (PBS), acetylated using acetic anhydride in triethanolamine buffer, and dehydrated through a series of alcohol washes in 70%, 95%, and 100% alcohol at room temperature. Subsequently, sections are delipidated in chloroform, followed by rehydration through successive exposure to 100% and 95% alcohol at room temperature. Microscope slides containing processed cryosections are allowed to air dry prior to hybridization. Other tissues may be assayed in a similar fashion.

A gp354-specific probe may be generated using PCR and sequence information from SEQ ID NO:1 or SEQ ID NO:3. Following PCR amplification,

WO 03/000021

PCT/US02/20275

-106-

the fragment is digested with restriction enzymes and cloned into pBluescript II cleaved with the same enzymes. For production of a probe specific for the sense strand of gp354, a cloned gp354 fragment cloned in pBluescript II may be linearized with a suitable restriction enzyme, which provides a substrate for labeled run-off transcripts (i.e., cRNA riboprobes) using the vector-borne T7 promoter and commercially available T7 RNA polymerase. A probe specific for the antisense strand of gp354 may also be readily prepared using the gp354 clone in pBluescript II by cleaving the recombinant plasmid with a suitable restriction enzyme to generate a linearized substrate for the production of labeled run-off cRNA transcripts using the T3 promoter and cognate polymerase.

The riboprobes may be labeled with [<sup>35</sup>S]-UTP to yield a specific activity of about 0.40 x 10<sup>6</sup> cpm/pmol for antisense riboprobes and about 0.65 x 10<sup>6</sup> cpm/pmol for sense-strand riboprobes. Each riboprobe may be subsequently denatured and added (2 pmol/ml) to hybridization buffer which contains 50% formamide, 10% dextran, 0.3 M NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA, 1X Denhardt's Solution, and 10 mM dithiothreitol.

Microscope slides containing sequential pancreas cryosections may be independently exposed to 45 µl of hybridization solution per slide and silanized cover slips may be placed over the sections being exposed to hybridization solution. Sections are incubated overnight (e.g., 15-18 hours) at 52°C to allow hybridization to occur. Equivalent series of cryosections are then exposed to sense or antisense gp354-specific cRNA riboprobes.

Following the hybridization period, coverslips are washed off the slides in 1X SSC, followed by RNase A treatment by exposing the slides to 20 µg/ml RNase A in a buffer containing 10 mM Tris•HCl (pH 7.4), 0.5 M EDTA, and 0.5 M NaCl for 45 minutes at 37°C. The cryosections are then subjected to three high-stringency washes in 0.1 X SSC at 52°C for 20 minutes each. Following the series of washes, cryosections are dehydrated by consecutive exposure to 70%, 95%, and 100% ammonium acetate in alcohol, followed by air drying and exposure to KODAK BIOMAX MR-1 film. After 13 days of exposure, the film is developed, and any significant hybridization signal is detected.

WO 03/000021

PCT/US02/20275

-107-

Based on these results, slides containing tissue that hybridized, as shown by film autoradiograms, are coated with KODAK NTB-2 nuclear track emulsion and the slides are stored in the dark for 32 days. The slides are then developed and counterstained with hematoxylin. Emulsion-coated sections are analyzed microscopically to determine the specificity of labeling. The signal is determined to be specific if autoradiographic grains (generated by antisense probe hybridization) are clearly associated with cresyl violet-stained cell bodies. Autoradiographic grains found between cell bodies indicate non-specific binding of the probe.

Expression of GP354 in the pancreas and the brain (*infra*) provides an indication that modulators of GP354 activity have utility for treating certain neural disorders by inhibiting or increasing the activity of GP354 in the nervous system.

**Example 8: Northern Blot Analysis of gp354-RNA**

Northern blot hybridizations may be performed to examine the expression of gp354 mRNA. A clone containing at least a portion of the sequence of SEQ ID NO:1, SEQ ID NO:3, or a complement thereto, may be used as a probe. Vector-specific primers are used in PCR to generate a hybridization probe fragment for <sup>32</sup>P-labeling. The PCR is performed as follows: (1) mix the following reagents:

|    |              |                                                                                                      |
|----|--------------|------------------------------------------------------------------------------------------------------|
| 1  | 1 $\mu$ l    | gp354-containing plasmid                                                                             |
|    | 2 $\mu$ l    | forward primer                                                                                       |
|    | 2 $\mu$ l    | reverse primer                                                                                       |
|    | 10 $\mu$ l   | 10X PCR buffer provided by the manufacturer of the Taq polymerase (e.g., Amersham Pharmacia Biotech) |
| 25 | 1 $\mu$ l    | 10mM dNTP (e.g., Boehringer Mannheim catalogue no. 1 969 064)                                        |
|    | 0.5 $\mu$ l  | Taq polymerase (such as Amersham Pharmacia Biotech catalogue no. 27-0799-62)                         |
| 30 | 83.5 $\mu$ l | water                                                                                                |

WO 03/000021

PCT/US02/20275

-108-

(2) perform PCR in a thermocycler using the following program: 94°C 5min; 30 cycles of 94°C, 1 min, 55°C, 1 min, and 72°C 1 min; and then 72°C, 10 min.

The PCR product may be purified using QIAQUICK PCR

Purification Kit (Qiagen catalogue no. 28104). The purified PCR fragment is

5 labeled with <sup>32</sup>P-dCTP (Amersham Pharmacia Biotech catalogue no. AA0005/250) by random priming using "Ready-to-go DNA Labeling Beads" (Amersham Pharmacia Biotech cat. no. 27-9240-01). Hybridization is carried out on a human multi-tissue Northern blot from Clontech according to the manufacturer's protocol. After overnight exposure on a Molecular Dynamics PHOSPHORIMAGER screen  
10 (cat. no. MD146-814), bands of about 1.35 kb are visualized.

**Example 9: Recombinant Expression of GP354 in Eukaryotic Host Cells**

**A. Expression of gp354 in Mammalian Cells**

To produce GP354 protein, a GP354-encoding polynucleotide is expressed using recombinant techniques. For example, the GP354-encoding  
15 sequence described in Example 1 is subcloned into the commercial expression vector pzeoSV2 (Invitrogen). The resultant expression construct is transfected into Chinese Hamster Ovary (CHO) cells using the transfection reagent FUGENE6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert. Other eukaryotic cell lines, including human embryonic kidney (HEK 293)  
20 and COS cells, are suitable as well.

Cells stably expressing GP354 are selected by growth in the presence of 100 µg/ml zeocin (Stratagene, LaJolla, CA). Optionally, GP354 may be purified from the cells using standard chromatographic techniques. To facilitate purification, antisera are raised against one or more synthetic peptide sequences  
25 that correspond to portions of the GP354 amino acid sequence, and the antisera are used to affinity-purify GP354. The GP354 protein also may be expressed in-frame with a tag sequence (e.g., polyhistidine, haemagglutinin, or FLAG) to facilitate purification. Moreover, it will be appreciated that many of the uses for GP354 polypeptides, such as assays described below, do not require purification of GP354  
30 from the host cell.

**B. Expression of GP354 in 293 cells**

WO 03/000021

PCT/US02/20275

-109-

- For expression of GP354 in mammalian cells 293 (transformed human or primate embryonic kidney cells), a plasmid bearing the relevant gp354 coding sequence is prepared, using vector pSecTag2A (Invitrogen). Vector pSecTag2A contains the murine IgK
- 5 chain leader sequence for secretion, the c-myc epitope for detection of the recombinant protein with the anti-myc antibody, a C-terminal polyhistidine for purification with nickel chelate chromatography, and a Zeocin-resistant gene for selection of stable transfectants. The forward primer for amplification of this gp354 cDNA is determined by routine procedures and preferably contains a 5'
- 10 extension of nucleotides to introduce the HindIII cloning site and nucleotides matching the gp354 sequence. The reverse primer is also determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce an XhoI restriction site for cloning and nucleotides corresponding to the reverse complement of the gp354 sequence. The PCR conditions are 55°C as the
- 15 annealing temperature. The PCR product is gel purified and cloned into the HindIII-XhoI sites of the vector.
- The DNA is purified using QIAGEN chromatography columns and transfected into 293 cells using the DOTAP transfection medium (Boehringer Mannheim).
- Transiently transfected cells are tested for expression at 24 hours after transfection,
- 20 using Western blots probed with anti-His and anti-GP354 peptide antibodies.
- Permanently transfected cells are selected with Zeocin and propagated. Production of the recombinant protein is detected from both cells and media by Western blots probed with anti-His, anti-Myc or anti-GP354 peptide antibodies.
- 25 **C. Expression of GP354 in COS cells**
- For expression of GP354 in COS7 cells, a polynucleotide having a sequence of SEQ ID NO:1, for example, can be cloned into vector p3-CI. This vector is a pUC18-derived plasmid that contains the HCMV (human cytomegalovirus) promoter-intron located upstream from the bGH (bovine growth
- 30 hormone) polyadenylation sequence and a multiple cloning site. In addition, the plasmid contains the dhfrf (dihydrofolate reductase) gene which provides selection

WO 03/000021

PCT/US02/20275

-110-

in the presence of the drug methotrexane (MTX) for selection of stable transformants.

The forward primer is determined by routine procedures and preferably contains a 5' extension which introduces an XbaI restriction site for cloning, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:1. The reverse primer is also determined by routine procedures and preferably contains 5'-extension of nucleotides which introduces a SalI cloning site followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO:1.

The PCR consists of an initial denaturation step of 5 min at 95°C; 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 58°C and 30 sec extension at 72°C; and followed by 5 min extension at 72°C. The PCR product is gel purified and ligated into the XbaI and SalI sites of vector p3-CL. This construct is used to transform competent *E. coli* cells. The plasmid DNA is then purified from the *E. coli* culture with QIAGEN chromatography columns and transfected into COS7 cells using the LIPOFECTAMINE reagent from BRL in accordance with the manufacturer's specification. Forty-eight and 72 hours after transfection, the media and the cells are tested for recombinant protein expression.

GP354 expressed from a COS cell culture can be purified by first concentrating the cell-growth media to about 10 mg protein/ml. The purification can be accomplished by, for example, chromatography.

Purified GP354 is concentrated to 0.5 mg/ml in an AMICON concentrator fitted with a YM-10 membrane and stored at -80°C.

#### **D. Expression of GP354 in insect cells**

For expression of GP354 in a baculovirus system, a polynucleotide having a sequence of SEQ ID NO:1 is amplified by PCR. The forward primer is determined by routine procedures and preferably contains a 5' extension which adds the NdeI cloning site, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:1. The reverse primer is also determined by routine procedures and preferably contains a 5' extension which introduces the

WO 03/000021

PCT/US02/20275

-111-

KpnI cloning site, followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO:1.

The PCR product is gel purified, digested with NdeI and KpnI, and cloned into the corresponding sites of expression vector pAcHTL-A (Pharmingen, San Diego, CA). The pAcHTL vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV), and a 6XHis tag upstream from the multiple cloning site. Nucleic acid sequences encoding a protein kinase site for phosphorylation and a thrombin site for excision of the recombinant protein precede the multiple cloning site.

Of course, many other baculovirus vectors, such as pAc373, pVL941 and pAcIM1, can be used in place of pAcHTL-A. Other suitable vectors for the expression of GP354 polypeptides can be also used, provided that the vector construct includes appropriately located signals for transcription, translation, and trafficking, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in, e.g., Luckow *et al.*, *Virology* 170:31-39 (1989).

The virus is grown and isolated using standard baculovirus expression methods, such as those described in Summers *et al.*, A MANUAL OF METHODS FOR BACULOVIRUS VECTORS AND INSECT CELL CULTURE PROCEDURES, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). In preferred embodiments, pAcHLT-A containing the gp354 gene is introduced into baculovirus using the BACULOGOLD transfection kit (Pharmingen). Individual virus isolates are analyzed for protein production by radiolabeling infected cells with <sup>35</sup>S-methionine at 24 hours post infection. Infected cells are harvested at 48 hours post infection, and the labeled proteins are visualized by SDS-PAGE. Viruses exhibiting high expression levels can be isolated and used for scaled up expression.

For expression of a GP354 polypeptide in a Sf9 cells, a polynucleotide having the sequence of SEQ ID NO:1 can be amplified by PCR using the methods described above for baculovirus expression. The gp354 cDNA is cloned into vector pAcHLT-A (Pharmingen) for expression in Sf9 insect cells. The insert is cloned into the NdeI and KpnI sites, after elimination of an internal



-112-

NdeI site (using the same primers described above for expression in baculovirus). DNA is purified with QIAGEN chromatography columns and expressed in Sf9 cells. Preliminary Western blot experiments from non-purified plaques are tested for the presence of a recombinant protein of the expected size using a

- 5 GP354-specific antibody. The results are confirmed after further purification and expression optimization in HiG5 cells.

**Example 10: Interaction trap/two-hybrid system**

- In order to assay for GP354-interacting proteins, the interaction trap/two-hybrid library screening method can be used. This assay was first described in Fields *et al.*, *Nature* 340:245 (1989). A protocol is published in  
10 CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY (1999) and Ausubel, F. M. *et al.* SHORT PROTOCOLS IN MOLECULAR BIOLOGY, fourth edition, Greene and Wiley-interscience, NY (1992). Kits are commercially available from, *e.g.*, Clontech (MATCHMAKER Two-Hybrid  
15 System 3).

- A fusion of the nucleotide sequences encoding all or partial GP354 and the DNA-binding domain (DNA-BD) of yeast transcription factor GAL4 is constructed using an appropriate vector (*i.e.*, pGBKT7). Similarly, a GAL4 active domain (AD) fusion library is constructed in a second plasmid (*i.e.*, pGADT7)  
20 from cDNA of potential GP354-binding proteins. For protocols on making cDNA libraries, see, *e.g.*, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, second edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989).

- The DNA-BD/GP354 fusion construct is verified by sequencing,  
25 and tested for autonomous reporter gene activation and cell toxicity, both of which would prevent a successful two-hybrid analysis. Similar controls are performed with the AD/library fusion construct to ensure expression in host cells and lack of transcriptional activity. Yeast cells are transformed (ca. 10<sup>5</sup> transformants/mg of DNA) with both the GP354 and library fusion plasmids according to standard  
30 procedure (Ausubel, *et al.*, *supra*). *In vivo* binding of DNA-BD/GP354 with AD/library proteins results in transcription of specific yeast plasmid reporter genes

WO 03/000021

PCT/US02/20275

-113-

(i.e., lacZ, HIS3, ADE2, LEU2). Yeast cells are plated on nutrient-deficient media to screen for expression of reporter genes. Colonies are dually assayed for b-galactosidase activity upon growth in Xgal (5-bromo-4-chloro-3-indolyl-b-D-galactoside) supplemented media (filter assay for b-galactosidase activity is described in Breeden *et al.*, *Cold Spring Harb. Symp. Quant. Biol.*, 50:643 (1985). Positive AD-library plasmids are rescued from transformants and reintroduced into the original yeast strain as well as other strains containing unrelated DNA-BD fusion proteins to confirm specific GP354/library protein interactions. Insert DNA is sequenced to verify the presence of an open reading frame fused to GAL4 AD and to determine the identity of the GP354-binding protein.

**Example 11: Antibodies To GP354 Polypeptides**

Standard techniques are employed to generate polyclonal or monoclonal antibodies to GP354, and to generate useful antigen-binding fragments thereof or variants thereof, including "humanized" variants. Such protocols can be found, for example, in Sambrook *et al.*, *supra*, and Harlow *et al.* (Eds.) *supra*. In some embodiments, recombinant GP354 polypeptides (or cells or cell membranes containing such polypeptides) are used as antigen to generate the antibodies. In other embodiments, one or more peptides having amino acid sequences corresponding to an immunogenic portion of GP354 (e.g., 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more amino acids) are used as antigen. Peptides corresponding to extracellular portions of GP354, especially hydrophilic extracellular portions, are preferred. The antigen may be mixed with an adjuvant or linked to a hapten to increase antibody production.

**A. Polyclonal or monoclonal antibodies**

In one exemplary protocol, recombinant GP354 or a synthetic fragment thereof is used to immunize a mouse to generate monoclonal antibodies, or to immunize a larger mammal, such as a rabbit, for polyclonal antibodies. To increase antigenicity, peptides can be conjugated to keyhole limpet hemocyanin commercially available from, e.g., Pierce. For an initial injection, the antigen is emulsified with Freund's Complete Adjuvant and injected subcutaneously. At intervals of two to three weeks, additional aliquots of GP354 antigen are

WO 03/000021

PCT/US02/20275

-114-

emulsified with Freund's Incomplete Adjuvant and injected subcutaneously. Prior to the final booster injection, a serum sample is taken from the immunized mice and assayed by Western blot to confirm the presence of antibodies that immunoreact with GP354. Sera from the immunized animals may be used as polyclonal antisera or used to isolate polyclonal antibodies that recognize GP354.

Alternatively, the mice are sacrificed and their spleen removed for generation of monoclonal antibodies. To generate monoclonal antibodies, the spleens are placed in 10 ml of serum-free RPMI 1640, and single cell suspensions are formed by grinding the spleens in serum-free RPMI 1640 supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 100 units/ml penicillin, and 100 µg/ml streptomycin (RPMI) (Gibco, Canada). The cell suspensions are filtered and washed by centrifugation and resuspended in serum-free RPMI. Thymocytes taken from three naive Balb/c mice are prepared in a similar manner and used as a feeder layer. NS-1 myeloma cells, kept in log phase in RPMI with 10% fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, Utah) for three days prior to fusion, are centrifuged and washed as well.

To produce hybridoma fusions, spleen cells from the immunized mice are combined with NS-1 cells and centrifuged, and the supernatant is aspirated. The cell pellet is dislodged by tapping the tube, and 2 ml of 37°C PEG 1500 (50% in 75 mM HEPES, pH 8.0) is stirred into the pellet, followed by the addition of serum-free RPMI. Thereafter, the cells are centrifuged, resuspended in RPMI containing 15% FBS, 100 µM sodium hypoxanthine, 0.4 µM aminopterin, 16 µM thymidine (HAT) (Gibco), 25 units/ml IL-6 (Boehringer-Mannheim) and 1.5 x 10<sup>6</sup> thymocytes/ml, and plated into 10 flat-bottom 96-well tissue culture plates.

On days 2, 4, and 6 after the fusion, 100 µl of medium is removed from the wells of the tissue culture plates and replaced with fresh medium. On day 8, the fusions are screened by ELISA, testing for the presence of mouse IgG that binds to GP354. Cells from selected wells are further cloned by dilution until monoclonal cultures producing anti-GP354 antibodies are obtained.

#### **B. Humanization of anti-GP354 monoclonal antibodies**

-115-

The expression pattern of GP354 as reported herein and the potential of GP354 as targets for therapeutic intervention suggest therapeutic indications for GP354 inhibitors (antagonists). GP354-neutralizing antibodies comprise one class of therapeutics useful as GP354 antagonists. The following are  
5 protocols to improve the utility of anti-GP354 monoclonal antibodies as therapeutics in humans by "humanizing" the monoclonal antibodies. Humanized antibodies have improved serum half-life and are less immunogenic in humans. The principles of antibody humanization have been described in the literature. For instance, to minimize potential binding to complement, a humanized antibody is  
10 preferred to be of the IgG<sub>4</sub> subtype.

One level of humanization can be achieved by generating chimeric antibodies comprising the variable domains of a non-human antibody of interest and the constant domains of a human antibody. See, *e.g.*, Morrison *et al.*, *Adv. Immunol.*, 44:65-92 (1989). The variable domains of anti-GP354 antibodies can be  
15 cloned from the genomic DNA of an appropriate B-cell hybridoma or from cDNA derived from the hybridoma. The V region gene fragments are linked to exons encoding human antibody constant domains. The resultant construct is expressed in suitable mammalian host cells (*e.g.*, myeloma or CHO cells).

To achieve an even greater level of humanization, only those  
20 portions of the variable region gene fragments that encode antigen-binding complementarity determining regions (CDRs) of the non-human monoclonal antibody are cloned into human antibody sequences. See, *e.g.*, Jones *et al.*, *Nature* 321:522-525 (1986); Riechmann *et al.*, *Nature* 332:323-327 (1988); Verhoeven *et al.*, *Science* 239:1534-36 (1988); and Tempest *et al.*, *Bio/Technology* 9:266-71  
25 (1991). If necessary, the  $\beta$ -sheet framework of the human antibody surrounding the CDR3 region is also modified (*i.e.*, "back-mutated") to more closely mirror the three dimensional structure of the antigen-binding site of the original monoclonal antibody. See Kettleborough *et al.*, *Protein Engin.* 4:773-783 (1991); and Foote *et al.*, *J. Mol. Biol.* 224:487-499 (1992).

30 In an alternative approach, the surface of a non-human monoclonal antibody of interest is humanized by altering selected surface residues of the

WO 03/000021

PCT/US02/20275

-116-

non-human antibody, e.g., by site-directed mutagenesis, while retaining all of the interior and contacting residues of the non-human antibody. See Padlan, *Mol. Immunol.*, 28(4/5):489-98 (1991).

The foregoing approaches are employed using anti-GP354

- 5 monoclonal antibodies and the hybridomas that produce them. The humanized anti-GP354 antibodies are useful as therapeutics to treat or palliate conditions wherein GP354 expression or ligand-mediated GP354 signaling is undesirable.

**C. Human GP354-neutralizing antibodies from phage display**

Anti-GP354 antibodies can be also generated by phage display

- 10 techniques such as those described in Aujame *et al.*, *Hum. Antibodies* 8(4):155-168 (1997); Hoogenboom, *supra*; and Rader *et al.*, *Curr. Opin. Biotechnol.* 8:503-508 (1997). For example, antibody variable regions in the form of Fab fragments or linked single chain Fv fragments are fused to the amino terminus of filamentous phage minor coat protein pIII. Expression of the fusion protein and incorporation  
15 thereof into the mature phage coat results in phage particles that present an antibody on their surface and contain the genetic material encoding the antibody. A phage library comprising such constructs is expressed in bacteria, and the library is screened for GP354-specific phage-antibodies using labeled or immobilized GP354 as antigen-probe.

20 **D. Human GP354-specific antibodies from transgenic mice**

Human GP354-specific antibodies are generated in transgenic mice essentially as described in Brüggemann *et al.*, *Immunol. Today* 17(8):391-97 (1996) and Brüggemann *et al.*, *Curr. Opin. Biotechnol.* 8:455-58 (1997).

- Transgenic mice carrying human V-gene segments in germline configuration and  
25 that express these transgenes in their lymphoid tissue are immunized with a GP354 composition using conventional immunization protocols. Hybridomas are generated using B cells from the immunized mice using conventional protocols and screened to identify hybridomas secreting anti-GP354 human antibodies (e.g., as described above).

30 **Example 12: Assays to Identify Modulators of GP354 Activity**

-117-

Set forth below are several non-limiting assays for identifying modulators (agonists and antagonists) of GP354 activity. Among the modulators that can be identified by these assays are natural ligands of the receptor; synthetic analogs and derivatives of the natural ligands; antibodies and/or antibody-like compounds derived from natural antibodies or from antibody-like combinatorial libraries; and/or synthetic compounds identified by high-throughput screening of libraries; and the like.

All modulators that bind GP354 are useful for identifying GP354 in tissue samples (e.g., for diagnostic purposes or therapeutic purposes). Agonist and antagonist modulators are useful for up-regulating and down-regulating GP354 activity, respectively, so as to treat GP354-mediated diseases. The assays may be performed using single putative modulators, and/or may be performed using a known agonist in combination with candidate antagonists (or visa versa).

#### A. cAMP Assays

In one type of assay, levels of cyclic adenosine monophosphate (cAMP) are measured in gp354-transfected cells that have been exposed to candidate modulator compounds. Protocols for cAMP assays have been described in the literature. See, e.g., Sutherland *et al.*, *Circulation* 37:279 (1968); Frandsen *et al.*, *Life Sci.* 18:529-541 (1976); Dooley *et al.*, *J. of Pharmacol. Exp. Therap.* 283(2): 735-41 (1997); and George *et al.*, *J. Biomol. Screening* 2(4):235-40 (1997). An exemplary protocol for such an assay, using an Adenylyl Cyclase Activation FLASHPLATE Assay from NEN Life Science Products, is set forth below.

Briefly, a GP354-encoding sequence is subcloned into an expression vector, such as pzeoSV2 (Invitrogen). CHO cells are transiently transfected with the resultant expression construct using known methods, such as the transfection protocol provided by Boehringer-Mannheim when supplying the FUGENE 6 transfection reagent. Transfected CHO cells are seeded into 96-well microplates from the FLASHPLATE assay kit, which are coated with solid scintillant to which antisera to cAMP have been bound. For a control, some wells are seeded with untransfected CHO cells. Other wells in the plate receive various amounts of a cAMP standard solution for use in creating a standard curve. One or more test

WO 03/000021

PCT/US02/20275

-118-

- compounds are added to the cells in each well, with compound-free medium or buffer as control. After treatment, cAMP is allowed to accumulate in the cells for exactly 15 minutes at room temperature. The assay is terminated by the addition of lysis buffer containing [<sup>125</sup>I]-cAMP, and the plate is counted using a Packard
- 5 TOPCOUNT 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells or from standards and fixed amounts of [<sup>125</sup>I]-cAMP compete for antibody bound to the plate. A standard curve is constructed, and cAMP values for the unknowns are obtained by interpolation. Changes in intracellular cAMP levels of cells in response to exposure to a test compound are indicative of GP354
- 10 modulating activity. Modulators that act as agonists of receptors which couple to the Gs subtype of G proteins will stimulate production of cAMP, leading to a measurable (e.g., 3-10) fold increase in cAMP levels. Agonists of receptors which couple to the Gi/o subtype of G proteins will inhibit forskolin-stimulated cAMP production, leading to a measurable decrease (e.g., 50-100%) in cAMP levels.
- 15 Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

#### B. Aequorin Assays

- In another assay, cells (e.g., CHO cells) are transiently co-transfected with a gp354 expression construct and a construct that encodes the
- 20 photoprotein apoaequorin. In the presence of the cofactor coelenterazine, apoaequorin will emit a measurable luminescence that is proportional to the amount of cytoplasmic free calcium. See generally, Cobbold, et al. "Aequorin measurements of cytoplasmic free calcium," In: McCormack J.G. and Cobbold P.H., eds., CELLULAR CALCIUM: A PRACTICAL APPROACH. Oxford:IRL
- 25 Press (1991); Stables *et al.*, *Anal. Biochem.* 252:115-26 (1997); and Haugland, HANDBOOK OF FLUORESCENT PROBES AND RESEARCH CHEMICALS, Sixth edition, Eugene OR (1996).

- In one exemplary assay, a gp354 coding sequence is subcloned into pzeoSV2 (Invitrogen). CHO cells are transiently co-transfected with the resultant
- 30 expression construct and a construct that encodes the photoprotein apoaequorin (Molecular Probes) using the transfection reagent FUGENE 6

-119-

(Boehringer-Mannheim) and the transfection protocol provided in the product insert.

The cells are cultured for 24 hours at 37°C in MEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin. Then the culture medium is changed to serum-free MEM containing 5 µM coelenterazine (Molecular Probes). Culturing is continued for two more hours at 37°C. Subsequently, the cells are detached from the plate using VERSEN (Gibco/BRL), washed, and resuspended at 2X10<sup>5</sup> cells/ml in serum-free MEM.

Dilutions of candidate GP354 modulator compounds are prepared in serum-free MEM and dispensed into wells of an opaque 96-well assay plate at 50 µl/well. The plate is then loaded onto an MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantilly, VA). The instrument is programmed to dispense 50 µl cell suspensions into each well, one well at a time, and immediately read luminescence for 15 seconds. Dose-response curves for the candidate modulators are constructed using the area under the curve for each light signal peak. Data are analyzed with SLIDEWRITE, using the equation for a one-site ligand, and EC50 values are obtained. Changes in luminescence caused by the compounds are considered indicative of modulatory activity. Modulators that act as agonists at receptors which couple to the Gq subtype of G proteins give an increase in luminescence of up to 100 fold. Modulators that act as inverse agonists will reverse this effect at receptors that are either constitutively active or activated by known agonists.

#### C. Luciferase Reporter Gene Assay

The photoprotein luciferase provides another useful tool for identifying GP354 modulators. Cells (*e.g.*, CHO cells or COS7 cells) are transiently co-transfected with a gp354 expression construct and a reporter construct which includes a gene for the luciferase protein downstream from a transcription factor binding site, such as the cAMP-response element (CRE), AP-1, or NF-kappa B. Expression levels of luciferase reflect the activation status of the signaling events. See generally, George *et al.*, *J. Biomol. Screening* 2(4):235-240



WO 03/000021

PCT/US02/20275

-120-

(1997); and Stratowa *et al.*, *Curr. Opin. Biotechnol.* 6:574-581 (1995). Luciferase activity may be quantitatively measured using, e.g., luciferase assay reagents that are available from Promega (Madison, WI).

In one exemplary assay, CHO cells are plated in 24-well culture plates at a density of  $10^5$  cells/well one day prior to transfection, and cultured at 37°C in MEM (Gibco/BRL) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin. Cells are transiently co-transfected with a gp354 expression construct and a reporter construct containing the luciferase gene. The reporter plasmid constructs CRE-luciferase, AP-1-luciferase and NF-kappaB-luciferase may be purchased from Stratagene (LaJolla, CA). Transfections are performed using the FUGENE 6 transfection reagent (Boehringer-Mannheim) according to the supplier's instructions. Cells transfected with the reporter construct alone are used as a control.

Twenty-four hours after transfection, the cells are washed once with PBS pre-warmed to 37°C. Serum-free MEM is then added to the cells either alone (control) or with one or more candidate modulators. The cells are then incubated at 37°C for five hours. Thereafter, the cells are washed once with ice-cold PBS and lysed by the addition of 100 µl of lysis buffer per well from the luciferase assay kit supplied by Promega. After incubation for 15 minutes at room temperature, 15 µl of the lysate is mixed with 50 µl of substrate solution (Promega) in an opaque-white, 96-well plate, and the luminescence is read immediately on a Wallace model 1450 MICROBETA scintillation and luminescence counter (Wallace Instruments, Gaithersburg, MD).

Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulatory activity. Receptors that are either constitutively active or activated by agonists typically give a 3-fold to 20-fold stimulation of luminescence compared to cells transfected with the reporter gene alone. Modulators that act as inverse agonists will reverse this effect.

#### **D. Intracellular calcium measurement using FLIPR**

Changes in intracellular calcium levels are another recognized indicator of receptor activity, and such assays can be employed to screen for

WO 03/000021

PCT/US02/20275

-121-

modulators of GP354 activity. For example, CHO cells stably transfected with a gp354 expression vector are plated at a density of  $4 \times 10^4$  cells/well in Packard black-walled, 96-well plates specially designed to discriminate fluorescence signals emanating from the various wells on the plate. The cells are incubated for 60 minutes at 37°C in modified Dulbecco's PBS (D-PBS) containing 36 mg/L pyruvate and 1 g/L glucose with the addition of 1% fetal bovine serum and one of four calcium indicator dyes (FLUO-3 AM, FLUO-4 AM, CALCIUM GREEN-1 AM, or OREGON GREEN 488 BAPTA-1 AM), each at a concentration of 4  $\mu$ M. Plates are washed once with modified D-PBS without 1% fetal bovine serum and incubated for 10 minutes at 37°C to remove residual dye from the cellular membrane. In addition, a series of washes with modified D-PBS without 1% fetal bovine serum is performed immediately prior to activation of the calcium response.

A calcium response is initiated by the addition of one or more candidate receptor agonist compounds, calcium ionophore A23187 (10  $\mu$ M; positive control), or ATP (4  $\mu$ M; positive control). Fluorescence is measured by Molecular Device's FLIPR with an argon laser (excitation at 488 nm). See, e.g., Kuntzweiler *et al.*, *Drug Dev. Res.* 44(1):14-20 (1998). The F-stop for the detector camera is set at 2.5 and the length of exposure is 0.4 milliseconds. Basal fluorescence of cells is measured for 20 seconds prior to addition of a candidate agonist, ATP, or A23187. The basal fluorescence level is subtracted from the response signal. The calcium signal is measured for approximately 200 seconds, taking readings every two seconds. Calcium ionophore A23187 and ATP typically increase the calcium signal about 200% above baseline levels. In general, activated GP354s increase the calcium signal at least about 10-15% above baseline signal.

#### 25 E. Mitogenesis Assay

In a mitogenesis assay, the ability of candidate modulators to induce or inhibit gp354-mediated cell division is determined. See, e.g., Lajiness *et al.*, *J. Pharmacol. Exp. Therap.* 267(3):1573-1581 (1993). For example, CHO cells stably expressing GP354 are seeded into 96-well plates at a density of 5000 cells/well and grown at 37°C in MEM with 10% fetal calf serum for 48 hours, at which time the cells are rinsed twice with serum-free MEM. After rinsing, 80  $\mu$ l of

-122-

fresh MEM, or MEM containing a known mitogen, is added along with 20  $\mu$ l MEM containing varying concentrations of one or more test compounds diluted in serum-free medium. As controls, some wells on each plate receive serum-free medium alone, and some receive medium containing 10% fetal bovine serum.

- 5 Untransfected cells or cells transfected with vector alone also may serve as controls.

After culture for 16-18 hours, 1  $\mu$ Ci of [ $^3$ H]-thymidine (2 Ci/mmol) is added to the wells and cells are incubated for an additional 2 hours at 37°C. The cells are trypsinized and collected on filter mats with a cell harvester (Tomtec); the  
 10 filters are then counted in a Betaplate counter. The incorporation of [ $^3$ H]-thymidine in serum-free test wells is compared to the results achieved in cells stimulated with serum (positive control). Use of multiple concentrations of test compounds permits creation and analysis of dose-response curves using the non-linear, least squares fit equation:  $A = B \times [C / (D + C)] + G$  where A is the  
 15 percent of serum stimulation; B is the maximal effect minus baseline; C is the EC50; D is the concentration of the compound; and G is the maximal effect. Parameters B, C and G are determined by Simplex optimization.

Agonists that bind to the receptor are expected to increase [ $^3$ H]-thymidine incorporation into cells, showing up to 80% of the response to  
 20 serum. Antagonists that bind to the receptor will inhibit the stimulation seen with a known agonist by up to 100%.

#### F. [ $^{35}$ S]GTP $\gamma$ S Binding Assay

It is possible to evaluate whether GP354 signals through a G protein-mediated pathway. G protein-coupled receptors signal through  
 25 intracellular G proteins whose activities involve GTP binding and hydrolysis to yield bound GDP. Thus, measurement of binding of the non-hydrolyzable GTP analog [ $^{35}$ S]GTP $\gamma$ S in the presence and absence of candidate modulators provides another assay for modulator activity. See, e.g., Kowal *et al.*, *Neuropharmacology* 37:179-187 (1998).

30 In one exemplary assay, cells stably transfected with a gp354 expression vector are grown in 10 cm tissue culture dishes to subconfluence, rinsed

WO 03/000021

PCT/US02/20275

-123-

once with 5 ml of ice-cold  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free phosphate-buffered saline, and scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x g, 5 minutes), resuspended in TEE buffer (25 mM Tris, pH 7.5, 5 mM EDTA, 5 mM EGTA), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (1 ml TEE per plate of cells), and centrifuged at 1,000 x g for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x g for 20 minutes to isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM  $\text{MgCl}_2$ , 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at  $-70^\circ\text{C}$  until use.

Aliquots of cell membranes prepared as described above and stored at  $-70^\circ\text{C}$  are thawed, homogenized, and diluted into buffer containing 20 mM HEPES, 10 mM  $\text{MgCl}_2$ , 1 mM EDTA, 120 mM NaCl, 10  $\mu\text{M}$  GDP, and 0.2 mM ascorbate, at a concentration of 10-50  $\mu\text{g}/\text{ml}$ . In a final volume of 90  $\mu\text{l}$ , homogenates are incubated with varying concentrations of candidate modulator compounds or 100  $\mu\text{M}$  GTP for 30 minutes at  $30^\circ\text{C}$  and then placed on ice. To each sample, 10  $\mu\text{l}$  guanosine 5'-O-(3 $^{35}\text{S}$ )thio triphosphate (NEN, 1200 Ci/mmol; [ $^{35}\text{S}$ ]-GTP $\gamma\text{S}$ ), was added to a final concentration of 100-200 pM. Samples are incubated at  $30^\circ\text{C}$  for an additional 30 minutes, 1 ml of 10 mM HEPES, pH 7.4, 10 mM  $\text{MgCl}_2$ , at  $4^\circ\text{C}$  is added and the reaction is stopped by filtration.

Samples are filtered over Whatman GF/B filters and the filters are washed with 20 ml ice-cold 10 mM HEPES, pH 7.4, 10 mM  $\text{MgCl}_2$ . Filters are counted by liquid scintillation spectroscopy. Nonspecific binding of [ $^{35}\text{S}$ ]-GTP $\gamma\text{S}$  is measured in the presence of 100  $\mu\text{M}$  GTP and subtracted from the total. Compounds are selected that modulate the amount of [ $^{35}\text{S}$ ]-GTP $\gamma\text{S}$  binding in the cells, compared to untransfected control cells. Activation of receptors by agonists gives up to a five-fold increase in [ $^{35}\text{S}$ ]-GTP $\gamma\text{S}$  binding. This response is blocked by antagonists.

#### G. MAP Kinase Activity Assay

WO 03/000021

PCT/US02/20275

-124-

Evaluation of MAP kinase activity in cells expressing GP354 provides another assay to identify modulators of GP354 activity. See, *e.g.*, Lajiness *et al.*, *supra*, and Boulton *et al.*, *Cell* 65:663-675 (1991). In one embodiment, CHO cells stably transfected with gp354 are seeded into 6-well plates at a density of  $7 \times 10^4$  cells/well 48 hours prior to the assay. During this 48 hour period, the cells are cultured at 37°C in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin. The cells are serum-starved for 1-2 hours prior to the addition of stimulants.

For the assay, the cells are treated with medium alone or medium containing either a candidate agonist or 200 nM Phorbol ester- myristoyl acetate (i.e., PMA, a positive control), and the cells are incubated at 37°C for various amounts of time. To stop the reaction, the plates are placed on ice, the medium is aspirated, and the cells are rinsed with 1 ml of ice-cold PBS containing 1 mM EDTA. Thereafter, 200 µl of cell lysis buffer (12.5 mM MOPS, pH 7.3, 12.5 mM glycerophosphate, 7.5 mM MgCl<sub>2</sub>, 0.5 mM EGTA, 0.5 mM sodium vanadate, 1 mM benzamide, 1 mM dithiothreitol, 10 µg/ml leupeptin, 10 µg/ml aprotinin, 2 µg/ml pepstatin A, and 1 µM okadaic acid) is added to the cells. The cells are scraped from the plates and homogenized by 10 passages through a 23 3/4 G needle, and the cytosol fraction is prepared by centrifugation at 20,000 x g for 15 minutes.

Aliquots (5-10 µl containing 1-5 µg protein) of cytosol are mixed with 1 mM MAPK Substrate Peptide (APRTPGGRR (SEQ ID NO:18), Upstate Biotechnology, Inc., NY) and 50 µM [ $\gamma$ -<sup>32</sup>P]ATP (NEN, 3000 Ci/mmol), diluted to a final specific activity of about 2000 cpm/pmol, in a total volume of 25 µl. The samples are incubated for 5 minutes at 30°C, and reactions are stopped by spotting 20 µl on 2 cm<sup>2</sup> squares of Whatman P81 phosphocellulose paper. The filter squares are washed in 4 changes of 1% H<sub>3</sub>PO<sub>4</sub>, and the squares are subjected to liquid scintillation spectroscopy to quantitate bound label. Equivalent cytosolic extracts are incubated without MAPK substrate peptide, and the bound labels from these samples are subtracted from the matched samples with the substrate peptide.

-125-

The cytosolic extract from each well is used as a separate point. Protein concentrations are determined by a dye binding protein assay (Bio-Rad Laboratories). Agonist activation of the receptor is expected to result in up to a five-fold increase in MAPK enzyme activity. This increase is blocked by antagonists.

## II. [<sup>3</sup>H]Arachidonic Acid Release

The activation of GP354s may also potentiate arachidonic acid release in cells, providing yet another useful assay for modulators of GP354 activity. See, e.g., Kanterman *et al.*, *Mol. Pharmacol.* 39:364-369 (1991). For example, CHO cells that are stably transfected with a GP354 expression vector are plated in 24-well plates at a density of  $1.5 \times 10^4$  cells/well and grown in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin for 48 hours at 37°C before use. Cells of each well are labeled by incubation with [<sup>3</sup>H]-arachidonic acid (Amersham Corp., 210 Ci/mmol) at 0.5 µCi/ml in 1 ml MEM supplemented with 10 mM HEPES, pH 7.5, and 0.5% fatty-acid-free bovine serum albumin for 2 hours at 37°C. The cells are then washed twice with 1 ml of the same buffer. Candidate compounds are added in 1 ml of the same buffer, either alone or with 10 µM ATP, and the cells are incubated at 37°C for 30 minutes. Buffer alone and mock-transfected cells are used as controls. Samples (0.5 ml) from each well are counted by liquid scintillation spectroscopy. Agonists which activate the receptor will lead to potentiation of the ATP-stimulated release of [<sup>3</sup>H]-arachidonic acid. This potentiation is blocked by antagonists.

## I. Extracellular Acidification Rate

In yet another assay, the effects of candidate modulators of GP354 activity are assayed by monitoring extracellular changes in pH induced by the test compounds. See, e.g., Dunlop *et al.*, *J. Pharmacol. Toxicol. Meth.* 40(1):47-55 (1998). In one embodiment, CHO cells transfected with a GP354 expression vector are seeded into 12 mm capsule cups (Molecular Devices Corp.) at  $4 \times 10^5$  cells/cup in MEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine,

WO 03/000021

PCT/US02/20275

-126-

10 U/ml penicillin, and 10 µg/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO<sub>2</sub> for 24 hours.

Extracellular acidification rates are measured using a CYTOSENSOR MICROPHYSIOMETER (Molecular Devices Corp.). The capsule cups are loaded into the sensor chambers of the MICROPHYSIOMETER and the chambers are perfused with running buffer (bicarbonate-free MEM supplemented with 4 mM L-glutamine, 10 units/ml penicillin, 10 µg/ml streptomycin, 26 mM NaCl) at a flow rate of 100 µl/min. Candidate agonists or other agents are diluted into the running buffer and perfused through a second fluid path. During each 60-second pump cycle, the pump is run for 38 seconds and is off for the remaining 22 seconds. The pH of the running buffer in the sensor chamber is recorded during the cycle from 43-58 seconds, and the pump is re-started at 60 seconds to start the next cycle. The rate of acidification of the running buffer during the recording time is calculated by the Cytosoft program. Changes in the rate of acidification are calculated by subtracting the baseline value (the average of 4 rate measurements immediately before addition of a modulator candidate) from the highest rate measurement obtained after addition of a modulator candidate. The selected instrument detects 61 mV/pH unit. Modulators that act as agonists of the receptor result in an increase in the rate of extracellular acidification compared to the rate in the absence of agonist. This response is blocked by modulators which act as antagonists of the receptor.

**Example 13: Cloning of Other gp354 Genes or Fragments**

To obtain other gp354 genes or fragments, PCR was performed essentially as described in Example 3, *supra*. Oligonucleotides GX1-220 and GX1-221 were used for inter-exon PCR verification of the *ab initio* gp354 gene model. GX1-386 and GX1-387 were used to amplify the full extent of the predicted portion of the gene. GX1-218 and GX1-221 were used to amplify and clone the mouse PanCAM orthologue. GX3-238 and GX3-232 were used to amplify the full length PanCAM cDNA from a pancreas cDNA pool. Their sequences are as following:

GX1-218: ACTGGGGCTAGTTCAGTGGACTAA (SEQ ID NO:16)

WO 03/000021

PCT/US02/20275

-127-

GX1-220: GCCAGCATGACCTCCACATTAG (SEQ ID NO:19)  
 GX1-221: CAGCCTCAGCTCTGCACACATAGTC (SEQ ID NO:20)  
 GX1-386: ATGCGGGTCCCCGCCCTCCTCGTCCT (SEQ ID NO:21)  
 GX1-387: CTTGGTCTCCAGAGTCCTTCCTCCTC (SEQ ID NO:22)  
 5 GX3-232: GGAGGAGTAAGGTCGCCAACTTGTC (SEQ ID NO:23)  
 GX3-238: ACCTTGGGGGACGAATGCTC (SEQ ID NO:24)

The PCR products were isolated and cloned as described in Example 3, *supra*.

The cloned GX1-220/GX1-221 PCR product was labeled by random hexamer extension using the protocol provided with the random hexamer kit (Invitrogen, Carlsbad, CA). The labeled fragment was used as a probe in Northern and multiple tissue expression (MTE) blots. Hybridizations were done according to the protocol provided with the multiple tissue blots (Clontech).

The Northern blot shows a diffused signal detected only in the pancreas, with a size in the range of 2.2-3.5 Kb. Hybridization done with an array of 76 normal human tissues revealed a positive spot in the position of the pancreas RNA sample. These data confirm that the gp354 gene is predominantly expressed in the pancreas.

#### **Example 14: Cloning of the Full Length Human gp354 Gene**

The entire 1,779 (including the stop codon) base pair predicted portion of the gene was amplified from a human pancreas cDNA pool by PCR, using primers GX1-386 and GX1-387 (*supra*) that encompassed the full length of the predicted open reading frame (Fig. 7). To amplify a full-length cDNA, we designed a pair of oligonucleotides (GX1-238 and GX1-232, *supra*) in the presumed 5' and 3' un-translated regions of the AL136654 sequence and used these for PCR of a human pancreas cDNA pool. Two different cDNA clones were obtained. Both were smaller than AL136654 and were within the size range determined on the Northern blot. They had uninterrupted open reading frames similar to that obtained from the original gene prediction, and encode proteins with 4 or 5 Ig domains (Fig. 9A).

Since this gene is predominantly expressed in the pancreas and encodes a protein with multiple C2 type Ig domains that are characteristic for cell



WO 03/000021

PCT/US02/20275

-128-

adhesion molecules (Williams et al., *Annu. Rev. Immunol.* 6:381-405 (1988)), we also named this gene PanCAM, for pancreatic cell adhesion molecule.

The predicted protein encoded by the longest potential ORF of PanCAM is composed of 708 amino acid residues and contains a signal peptide, five Ig domains (three of them are C2 type), a transmembrane domain, and a 174 amino acid intracellular domain. It is longer than the protein encoded by our original gene prediction on both the amino and carboxy termini.

BLASTP searches of public databases identified several proteins with significant similarity to PanCAM. The three most similar proteins that have been characterized are human and mouse NEPH1, nephrin, and the *Drosophila* irregular chiasm-C roughest ("Irre") protein (see, e.g., Fig. 2). In all of these cases, the homology is concentrated in the extracellular portion of the protein and there is no significant similarity in the intracellular portion.

Human NEPH1 (NP\_060710) and mouse NEPH1 proteins (NP\_570937) are potential cell adhesion molecules that play critical roles in kidney development (Donoviel et al., *Mol. Cell. Biol.* 21(14):4829-36 (2001)). Like PanCAM, NEPH1 has 5 Ig domains, and the human protein is 44% identical to PanCAM in the extracellular domain, while the mouse protein is 47% identical to PanCAM in the extracellular domain.

The second most similar BLASTP score is for human nephrin (NP\_004637), which shares 33% identity with GP354 in the extracellular domain. This protein contains eight Ig domains and is localized to the glomerular slit diaphragm. Nephrin is believed to play a role in cell adhesion and to be involved in the development of glomeruli of kidney (Ruotsalainen et al., *Proc. Natl. Acad. Sci. USA* 96:7962-7967 (1999); Kestila et al., *Mol. Cell* 1(4):575-82 (1998)). Mutations in the nephrin gene cause congenital nephritis in humans (Kestila et al., *supra*).

The *Drosophila* irregular chiasm C roughest precursor (irre, NP\_525058) is another homologous protein. Irre is a cell adhesion protein required for development of compound eyes (Ramos et al., *Genes Dev.* 7(12B):2533-47 (1993); Schneider et al., *Neuron* 15(2):259-71 (1995)) and

WO 03/000021

PCT/US02/20275

-129-

antenna (Reddy *et al.*, *Dev. Genes Evol.* 209(10):581-91 (1999)). This protein contains 5 C2 type Ig domains, and shares 31% identity with PanCAM. Mutations at the Irre locus in *Drosophila* affect sensory organ development in the fly, apparently due at least in part to abnormal apoptotic activity (Ramos *et al.*, *supra*).

5 NEPH1, nephrin and Irre are involved in developmental patterning and cell-cell communication. The similarity between GP354 and these proteins suggests that GP354 also plays a role in similar developmental pathways and, in particular, cell-cell interactions which trigger signal transduction pathways involved in organ and tissue development and/or maintenance in the pancreas and  
10 nervous system.

**Example 15: Alternative Splicing of PanCAM**

Sequence comparison of the two pancreas cDNA clones with two public EST sequences, AL136654 and BC007312 indicates that there are at least 4 alternative splice products of PanCAM (Fig. 9B). AL136654, identified in a fetal  
15 brain cDNA library, encompasses all of the predicted exons and most likely represents the longest full-length cDNA. BC007312, an EST clone derived from retinoblastoma cDNA, is the shortest transcript of PanCAM identified so far. BC007312 is generated by using cryptic splicing sites in exon 2 and exon 11 (Fig. 9B), where the encoded protein lacks all of the Ig domains, and has the signal  
20 peptide fused to the putative intracellular domain.

The pancreas cDNA clones differ from one another in the extracellular domain, but have identical transmembrane and intracellular domains. The shorter cDNA clone (PanCAM-2) from pancreas encodes a protein that is missing part of Ig domain 1, due to splicing out of exon 2 (Fig. 9B). The  
25 intracellular domains of the two pancreas cDNA clones are shorter than that of the other inferred protein sequences (Fig. 9A) due to splicing out part of the last exon.

**Example 16: Cloning a partial cDNA of mouse PanCAM**

To detect and clone the mouse ortholog of PanCAM, multiple inter-exon oligonucleotide pairs were designed based on the human gene sequence,  
30 and tested by PCR on cDNA pools from multiple mouse tissues. One oligonucleotide pair generated a faint band from a mouse brain cDNA pool. This

WO 03/000021

PCT/US02/20275

-130-

PCR product was cloned and sequenced. The mouse cDNA clone (SEQ ID NO:32) was 934 nucleotides in length, and the encoded protein is 85% identical to the human counterpart (Figs. 10A-B).

**Example 17: PanCAM is Beta Cell Specific in the Pancreas**

5 To further characterize the PanCAM protein, a soluble form of the extracellular domain was used in a screen for phage antibodies. Four human Fab fragments against extracellular portion of this protein were screened out from a phage antibody library (Knappik *et al.*, *J. Mol. Biol.* 296(1):57-86 (2000). All of these Fabs recognized the recombinant protein on Western blot. Three of them  
10 recognized PanCAM on paraffin-fixed tissue slides.

Immunohistochemical (IHC) staining was performed on slides containing human heart, brain, kidney, liver, lung, pancreas, spleen and muscle tissues, and on pancreas tissue from monkey (*Cynomolous*) and mouse. Among the human tissues tested, only the pancreas showed staining. The results clearly  
15 showed that the gene was expressed specifically in pancreatic islets of human, monkey and mouse.

In addition to the pancreas staining, the mouse anterior pituitary was also positive for IHC, with somewhat weaker and scattered staining. The pituitary staining explained the appearance of EST clones in brain libraries, as well as our  
20 ability to clone the mouse ortholog from brain cDNA.

To identify expression patterns of PanCAM at the cellular level, we performed RT-PCR on RNAs from multiple cell lines, including 293E, Cos7, mouse pancreatic endothelial cell (CRL 2279) (Arbiser *et al.*, *Proc. Natl. Acad. Sci. USA* 94(3):861-6 (1997)), alpha cell (CRL 2350) (Powers *et al.*, *Diabetes*  
25 39(4):406-14 (1990)), beta cell from fat/fat mice (CRL2364)(derived from a mouse strain with a carboxypeptidase E mutation and defective in proinsulin processing and insulin secretion; Naggert *et al.*, *Nat Genet* 10(2):135-42 (1995); Varlamov *et al.*, *Endocrinology* 138(11):4883-92 (1997)), and beta cells from wild type mice (CRL2055) (Hamaguchi *et al.*, *Diabetes* 40(7):842-9 (1991)). The results  
30 confirmed that PanCAM expression was beta cell specific.

Notably, beta cells from fat/fat mice yielded two RT-PCR products.

WO 03/000021

PCT/US02/20275

-131-

Each of the RT-PCR bands was much fainter than the single band amplified from wild type beta cells. Sequence of the two RT-PCR products from the mutant cell line confirmed that the larger molecular weight band was same as that of the wildtype control, whereas the lower molecular weight band was an alternative splicing product. A cryptic splicing site was used to generate the shorter transcript, causing deletion of 3' half of this exon and 4 downstream exons. Consequently, Ig domains 2 and 3 are missing in this open reading frame, while the downstream reading frame is maintained.

**Example 18: PanCAM-Positive Cells Diminish  
During Development of Diabetes in NOD Mice**

As a first step to test whether PanCAM plays a role in the development of type I diabetes, pancreata from NOD mice at ages of 4, 8, and 12 weeks, as well as retired breeders (animals with fully developed diabetes) were examined by IHC. At 4 weeks of age, the staining pattern was the same as in wild type mice, with the majority of cells in pancreatic islets being positive for PanCAM. At 8 weeks, when the mice started to develop pancreatitis, T cell invasion of islets was seen, and the number of PanCAM-positive cells started to diminish. At 12 weeks, there were more T cells and fewer PanCAM- positive cells in pancreas islets. In retired breeders, most of PanCAM- positive cells were lost, and T cells were observed where the islets used to be. The results demonstrated that overall expression of PanCAM was correlated with the development of diabetes in the NOD mice.

**Example 19: Isolation of Beta Cells Using Anti-PanCAM Antibodies**

Monoclonal or polyclonal antibodies specific for an extracellular portion of PanCAM (e.g., human PanCAM) are immobilized on a solid or semi-solid substrate and then contacted with dispersed pancreatic cells. Non-beta cells are removed by washing the substrate with an appropriate buffer. The bound beta cells are then recovered. Alternatively, the beta cells can be isolated by flow cytometry using anti-PanCAM antibodies. The purified beta cells are then entrapped into macroporous gelatin microcarriers (e.g., Cultispher-S) and

WO 03/000021

PCT/US02/20275

-132-

transplanted to a patient having or at risk of having diabetes (see, *e.g.*, Guerra *et al.*, *Biotechnol. Bioeng.* 75(6):741-4 (2001).

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Exemplary methods and materials are described below, although methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention. All publications and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. The materials, methods, and examples are illustrative only and not intended to be limiting. Throughout this specification and claims, the word "comprise," or variations such as "comprises" or "comprising" will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers. All ATCC deposits mentioned herein were made according to the provisions of the Budapest Treaty. All restrictions on the availability to the public of the above ATCC deposits will be irrevocably removed upon the granting of a patent on this application.

WO 03/000021

PCT/US02/20275

-133-

CLAIMS

What is claimed is:

1. An isolated polypeptide comprising amino acid residues 68-594 of SEQ ID NO:12.
- 5 2. The polypeptide of claim 1, comprising SEQ ID NO:12.
3. The polypeptide of claim 1, comprising SEQ ID NO:26.
4. The polypeptide of claim 1, comprising SEQ ID NO:28.
5. The polypeptide of claim 1, comprising SEQ ID NO:8.
6. The polypeptide of claim 1, further comprising a heterologous sequence.
- 10 7. An isolated polypeptide comprising amino acid residues 4-513 of SEQ ID NO:8.
8. An isolated polypeptide comprising amino acid residues 535-708 of SEQ ID NO:8.
9. An isolated polypeptide comprising amino acid residues 535-633 of SEQ ID NO:26.
- 15 10. An isolated polynucleotide comprising a nucleotide sequence encoding the isolated polypeptide of claim 1.
11. The polynucleotide of claim 10, comprising nucleotides 202-1782 of SEQ ID NO:11.

WO 03/000021

PCT/US02/20275

-134-

12. The polynucleotide of claim 10, comprising SEQ ID NO:11.
13. The polynucleotide of claim 10, comprising SEQ ID NO:25.
14. The polynucleotide of claim 10, comprising SEQ ID NO:27.
15. The polynucleotide of claim 10, comprising nucleotides 196-2319 of SEQ  
5 ID NO:7.
16. The polynucleotide of claim 10, comprising SEQ ID NO:7.
17. The polynucleotide of claim 10, further comprising a transcription  
regulatory sequence operatively linked to said nucleotide sequence.
18. A vector comprising the polynucleotide of claim 10.
- 10 19. The vector of claim 18, wherein the vector is a plasmid.
20. The vector of claim 18, wherein the vector is a viral vector.
21. The vector of claim 18, selected from the group consisting of baculoviruses,  
adenoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses,  
Semliki Forest viruses, vaccinia viruses, lentiviruses and retroviruses.
- 15 22. A host cell comprising the polynucleotide of claim 10.
23. The host cell of claim 22, wherein the host cell is selected from the group  
consisting of a bacterial cell, insect cell, yeast cell, plant cell and mammalian cell.
24. The host cell of claim 22, wherein the host cell is a human cell.

WO 03/000021

PCT/US02/20275

-135-

25. A method of producing a PanCAM polypeptide, comprising:  
culturing the host cell of claim 22 in a medium under conditions in which  
said nucleotide sequence is expressed, and  
recovering the polypeptide from the cell or from the culture medium.
- 5 26. An antibody that binds to the polypeptide of claim 1 or 7.
27. The antibody of claim 26, wherein the antibody is a monoclonal antibody.
28. The antibody of claim 26, wherein the antibody is a humanized or a fully  
human antibody.
29. A method of determining the presence of a PanCAM-encoding sequence in  
10 a sample, comprising:  
contacting the sample with a nucleic acid probe comprising SEQ ID NO:3,  
5, 6, 7, 9, 11, 25, or 27, or a portion thereof having at least 17 nucleotides, under  
high stringency hybridization conditions, and  
detecting hybridization of said polynucleotide to a nucleic acid in the  
15 sample, wherein the occurrence of said hybridization indicates the presence of a  
PanCAM-encoding sequence in the sample.
30. A method of determining the presence of a PanCAM protein in a sample,  
comprising:  
contacting the sample with the antibody of claim 26; and  
20 detecting specific binding of said antibody to an antigen, wherein the  
occurrence of said specific binding indicates the presence of a PanCAM protein in  
the sample.
31. A method of identifying a compound that binds a PanCAM protein,  
comprising:  
25 contacting the PanCAM protein with a test compound; and



WO 03/000021

PCT/US02/20275

-136-

detecting a complex formed by the PanCAM protein and the test compound, wherein the presence of the complex indicates that the test compound binds to the PanCAM protein.

32. A method of identifying a compound that modulates the activity of a  
5 PanCAM protein, comprising:  
    contacting the PanCAM protein with a test compound; and  
    determining the effect of the test compound on the activity of the PanCAM  
protein, whereas a change of the activity after the contacting step indicates that the  
test compound modulates the activity of the PanCAM protein.
- 10 33. A method of identifying a homolog of a human PanCAM gene, comprising:  
    hybridizing a nucleic acid library with a nucleic acid probe comprising SEQ  
ID NO:3, 5, 6, 7, 9, 11, 25, or 27, or a portion thereof consisting of at least 17  
nucleotides, under medium or high stringency hybridization conditions; and  
    determining whether the nucleic acid probe hybridizes to a nucleic acid  
15 sequence in the library,  
    wherein the nucleic acid sequence so hybridized is a homolog of a human  
PanCAM gene.
34. A method of diagnosing a disease condition in a subject, comprising:  
    comparing the amount or activity of a PanCAM protein in a tissue sample  
20 from the subject to the amount or activity of the PanCAM protein in a control  
sample,  
    wherein a significant difference in the PanCAM amount or activity of the  
tissue sample relative to the control sample indicates that the subject has a disease  
condition.
- 25 35. A method of diagnosing a disease condition in a subject, comprising:  
    comparing the amount of a PanCAM mRNA in a tissue sample from the  
subject to the amount of the PanCAM mRNA in a control sample,

WO 03/000021

PCT/US02/20275

-137-

wherein a significant difference in the PanCAM mRNA amount of the tissue sample relative to the control sample indicates that the subject has a disease condition.

36. The method of claim 34 or 35, wherein the tissue sample is a pancreatic  
5 tissue sample.
37. The method of claim 36, wherein the significance difference is a significant decrease, and the disease condition relates to diabetes.
38. The method of claim 34 or 35, wherein the disease condition relates to the central nervous system.
- 10 39. The method of claim 34 or 35, wherein the disease condition relates to the pituitary gland.
40. A method of identifying in a test cell the presence or absence of a genetic lesion or mutation characterized by at least one of: (i) aberrant modification or mutation of a gene encoding a PanCAM protein, (ii) mis-regulation of a gene  
15 encoding a PanCAM protein, and (iii) aberrant post-translational modification of a PanCAM protein, comprising:
- separately hybridizing nucleic acids from the test cell and from a reference cell that lacks the genetic lesion or mutation with a nucleic acid probe comprising SEQ ID NO:3, 5, 6, 7, 9, 11, 25, or 27, or a portion thereof having at least 17  
20 nucleotides, under high stringency hybridization conditions;
- separately washing the nucleic acid hybrids under high stringency wash conditions to allow dissociation of the hybrids; and
- determining whether the nucleic acid probe dissociates more readily from the nucleic acids of the test cell compared to the nucleic acids of the reference cell.

WO 03/000021

PCT/US02/20275

-138-

41. A method of treating a disease condition in a subject, comprising administering to the subject a composition that comprises the polynucleotide of claim 10, or the antibody of claim 26.
42. The method of claim 41, wherein the disease condition relates to the  
5 pancreas.
43. The method of claim 42, wherein the disease condition relates to the pancreatic beta cells.
44. The method of claim 43, wherein the disease condition relates to diabetes.
45. The method of claim 41, wherein the disease condition is selected from the  
10 group consisting of acute or chronic pancreatitis, pancreatic inflammation, pancreatic necrosis, exocrine insufficiency, pancreatic endocrine and hormonal imbalance, pancreatic tumors and associated cancers, and an auto-immune disorder which affects the pancreas.
46. The method of claim 41, wherein the disease condition relates to the central  
15 nervous system.
47. The method of claim 46, wherein the condition is selected from the group consisting of Alzheimer's disease, Parkinson's disease, senile dementia, migraine, epilepsy, neuritis, neurasthenia, neuropathy, neural degeneration and neural tumors.
48. The method of claim 41, wherein the disease condition is selected from the  
20 group consisting of pituitary hormone deficiency, growth hormone deficiency, and pituitary tumors.
49. A method of isolating beta cells from a dispersed pancreatic cell population, comprising:

WO 03/000021

PCT/US02/20275

-139-

- providing an antibody specific for a PanCAM protein,  
contacting the dispersed pancreatic cell population with the antibody under  
conditions that allow beta cells in the population to bind to the antibody, and  
separating cells unbound by the antibody from beta cells bound to the  
5 antibody, thereby isolating beta cells.
50. The method of claim 49, wherein the dispersed pancreatic cell population is  
a bovine or swine pancreatic cell population.
51. A method of treating a human displaying a sign or symptom of diabetes or a  
pre-diabetic condition, comprising:  
10 providing beta cells isolated by the method of claim 49, and  
introducing the beta cells into the human.
52. The method of claim 51, wherein the beta cells are introduced  
subcutaneously in the human.
53. The method of claim 51, wherein the beta cells are encapsulated in a  
15 biological acceptable carrier.
54. An isolated polypeptide comprising SEQ ID NO:31.
55. An isolated polynucleotide comprising a nucleotide sequence encoding  
SEQ ID NO:31.
56. The polynucleotide of claim 55, comprising SEQ ID NO:32.

1  
atcgccggtcccccgcctctcgtctctctctctctgcttcagaggaggcgcaggcccgctg  
M R V P A L L V L L F C F R G S A G P S  
ccccatttctgtcgaacagccagagacctggtggtgctgctggggagggaagcccgctg  
P H F L Q Q P E D L V V L L G E E A R L  
ccgtgtgctctgggcgcctactgtggggctagtctcagtgactaagagtggtggtgcctta  
P C A L L G A Y W G L V Q W T K S G L A L L  
gggggccaaagggaacctaccagggtggtcccggtactggatatcagggaatgcagccaat  
G G Q R D L P G W S R Y W I S G N A A N  
ggccacgatgacctccacattaggcccggtggagctagaggatgaagcatcatatgaatgt  
G Q H D L H I R P V E L E D E A S Y E C  
caggctacacaagcaggccctccgctccagaccagcccaactgcacgtgctggtccccca  
Q A T Q A G L R S R P A Q L H V L V P P  
gaagccccccagtgctgtgggcggccctctgtgtctctggtgtgctggagtctctgcgaac  
E A P Q V L G G P S V S L V A G V P A N  
ctgacatgtcggagccgtggggatgccgcctaccctgaattgctgtggttccagat  
L T C R S R G D A R P T P E L L W F K D  
ggggtcctgttgatggaaccaccttccatcagaccctgctgaaggaagggaaccttggtg  
G V L L D G T T F H Q T L L K E G T P G  
tcagtggaagagcaccttaacctgaccctttcagccatgatgatggaccaccttgtct  
S V E S T L T L T P F S H D D G A T F V  
tgcggggcccgagcagccgctgccacaggaagagacacagctatcacatgagcctg  
C R A R S Q A L P T G R D T A I T L S L

FIGURE 1 (cont.)

661 cagtacccccagaggtgactctgtctgcttgcacacactgtgcaggaggagagaag  
Q Y P P E V T L S A S P H T V Q E G E K  
721 gtcattttctgtgccaggccacagcccagcctcctgtcacaggctacaggtgggcaaaa  
V I F L C Q A T A Q P P V T G Y R W A K  
781 gggggctctccgtgctcggggcccgccggaaggttagaggtcgtggcagacgcctcg  
G G S P V L G A R G P R L E V V A D A S  
841 ttctgactgagcccggtgctcgtcgaggtcagcaacccgtgggtagcgaaccgcagt  
F L T E P V S C E V S N A V G S A N R S  
901 actgcgctggatgtgctgtttgggcccattCtgcaggcaaagccggagcccggtgctcgtg  
T A L D V L F G P I L Q A K P E P V S V  
961 gacgtgggggaagacgcttccttcagctgcgcctggcggggaacccgcttcacgggta  
D V G E D A S F S C A W R G N P L P R V  
1021 acctggaccggcccggtggcgccaggtgctgggctctggagccacactgcgtcttcog  
T W T R R G G A Q V L G S G A T L R L P  
1081 tcgggtggggcccgaggacgcagggcactatgtgtgcagagctgaggtgggctatcgggc  
S V G P E D A G D Y V C R A E A G L S G  
1141 ctgcggggcgccgcgcggaggtcgggtgactgtgaacgctccccigtagtgcgcgc  
L R G G A A E A R L T V N A P P V V T A  
1201 ctgcactctgcgcctgccttcctgagggccctgctgcctccagtgctggttttcgcc  
L H S A P A F L R G P A R L Q C L V F A  
1261 tctcccgccccagatgccgtggtctggtcttgggatgagggcttcctggaggggggtcg  
S P A P D A V V W S W D E G F L E A G S

FIGURE 1 (cont.)

1321 cagggccgggttcctggtggagacattccctgccccagagagccgcggggactgggtccg  
Q G R F L V E T F P A P E S R G G L G P  
ggcctgatctctgtgtacacatttcggggaccaggagtctgacttttagcaggagcttt  
1361 G L I S V L H I S G T Q E S D F S R S F  
aactgcagtgcccgaaccggctgggagggaggtgcccaggccagcctgggcccgtaga  
1441 N C S A R N R L G E G G A Q A S L G R R  
gacttgctgcccaactgtgcggatagtggccggagtggccgtgccaccacaactctcctt  
1501 D L L P T V R I V A G V A A A T T T L L  
atggtcatcactggggtggccctctctgctggcgccacagcaaggcctctttctccgag  
1561 M V I T G V A L C C W R H S K A S F S E  
caaaagaacctgatgcgaatccctggcagcagcgacggctccagttcacgaggtcctgaa  
1621 Q K N L M R I P G S S D G S S S R G P E  
gaagaggagacaggcagccgagggacggggcccccattgtgcacactgaccacagtgat  
1681 E E E T G S R E D R G P I V H T D H S D  
ctggttctgaggagggaagggaactctggagaccaag  
1741 L V L E E E G T L E T K 1776

WO 03/000021

4/38

PCT/US02/20275

## FIGURE 2

GP354 -----MRVPALLVLLFCFRGSAGFSPHFLOQ-----PEDLVVLLGEE  
 ICCR -----MLHTMQLLLLATIVGMVRSSPYTSYQNRFAFMEPDQQTAVVGAR  
 Nephtrin MALGTTLRASLLLLGLLLEGLAQLAIPASVPRG--FWALPENLTVVEGAS  
 \* \* \* \*

GP354 ARLPCALGAYWGLVQWTKSGLALGGQDRLPGWSRYRISGNAANGQHDLI  
 ICCR VTLPCRIVINKQGLQWTKDDFGLGTSRDLGSGFERIYAMVGSDEEGDYSLDI  
 Nephtrin VELRCGVSTPGSAVQWAKDGLLGFDFRIPGFPRYALEGDPARGEFLHI  
 \* \* \* \* \*

GP354 RPVELEDEASYEQAT-----QAGLRSRPAQLHVLVPPEAPQVL--GGPS  
 ICCR YFVMLDDDDARYQCQVSPGPEGQPAIRSTFAGLTVLVPPEAPKIT--QGQV  
 Nephtrin EACDLSDDAEYECQVGRS-EMGPVLVSPRVLISILVPPKLLLLTPEAGTM  
 \* \* \* \* \*

GP354 VSLVAGVVPANLTCRSRGDARPTPELLWFR--DGVLLDGTTFHQTLLEGT  
 ICCR IYATADRKVEIECVSVG-GKPAAEITWIDGLGNVLTNDIEYTVIPLPDQR  
 Nephtrin VTWVAGQEYVNCVS-GDAKPAPDITILLS--GQTISDISANVNEGSQOK  
 \* \* \* \*

GP354 PGSVESTLTLTFFSHDDGATFVCRARSQALPTGRDTAITLSLQYPFEVTL  
 ICCR RFTAKSVLRLTPKKEHHNTNFSCQAQNTADRTYRSKIRVEVKYAPKVKV  
 Nephtrin LFTVEATARVTPRSSDNRQLLVCEASSPALEAPIKASFTVNVLPFGPPV  
 \* \* \* \*

GP354 -----SASPHTVQ-----EGEKVIFLCQATAQPPVTG  
 ICCR NVMGSLPGGAGGSVGGAGGGSVHMSTGSRIVEHSQVRLECRADANPSDVR  
 Nephtrin IEWP-----GLDEGHVR-----AGQSLELFCVARGGNPLAT  
 \* \*

GP354 YRWAKGGSFVLGAR-GPRLEVVDASFLTEP-----VSCEVSNVAGS  
 ICCR YRWFINDEPIIGGQKTEMVIRNVTRKFHDAI-----VKCEVQNSVGK  
 Nephtrin LQWLKNGQPVSSTANGTEHTQAVARSVLVMTVRPEDHGAQLSCEAHNSVSA  
 \* \* \* \*

GP354 --ANRSTALDVLFGPILQAKPEPVSVVDVGEDASFSCAWRGN-PLPRVTWT  
 ICCR --SEDSETLDISYAFSFRQRQSMEDVGSVSLTCEVDSN-PQPEIVMI  
 Nephtrin GTQEHGITLQVTFPPSAIILGSASQTENKNVTLSCVSKSRPRVLLRWW  
 \* \* \* \*

GP354 R-----RGGAQVLG---SGATLRLPSVGPEDAGDYVCRAEAGLSG  
 ICCR Q-----HPSDRVVG---TSTNLTF-SVSNETAGRYYCKAN--VPG  
 Nephtrin LGWRQLLFMEETVMDGLHGGHISMSNLTFLARREDNGLTLTCEAFSEAF  
 \* \* \*



WO 03/000021

5/38

PCT/US02/20275

FIGURE 2 (cont.)

|         |                                                     |                                          |
|---------|-----------------------------------------------------|------------------------------------------|
| GP354   | LRGGAAEARLTVNAPFVVVTLH----                          | SAPAFLRGPRLQCLVFASAPDA                   |
| ICCR    | YAEISADAYVYLKGSFAIGSOR----                          | TQYGLVGDTRIECFASSVPRARH                  |
| Nephrin | KETFKKSLILNVKYPQKLNWIEGPPGQKLRAGTRVRLVCLAIGNPEPS    |                                          |
|         |                                                     | * *                                      |
| GP354   | VVWSWD-----                                         | FGFLE-AGSQGR--FLVETFPAP----              |
| ICCR    | VSWTFNG-----                                        | QEISSE-SGHDYS--ILVDAVPG----              |
| Nephrin | LMWYKDSRTVTESRLPQESRRVHLGSVEKSGSTFSRELVLVTGPGSDNQAK |                                          |
|         |                                                     | * * *                                    |
| GP354   | -----                                               | ESRGGGLPGLISVLHISGTQESDFSRSFNCSARNRLGEGG |
| ICCR    | -----                                               | GVKSTLIIRDSQAYHYG-KYNCTVVNDYGNDV         |
| Nephrin | FTCKAGQLSASTQLAVQFPPTNVITILANASALRPGDALNLTCSVSSNP-  |                                          |
|         |                                                     | *                                        |
| GP354   | AQASLGRRDLLPTVRIVAG-----                            | VAAATTTLLMVIT--GVALCC                    |
| ICCR    | AETQLQAKKSVSLMTIVGGI-----                           | SVVAFLLVLTILV--VVIK                      |
| Nephrin | -PVNLSWDKEGERLEGVAAPRRAPFKGSAARSVLLQVSSRDHGQVTC     |                                          |
|         |                                                     | * * *                                    |
| GP354   | WRHSKASFSEQKNLMRIP-----                             | GS-----                                  |
| ICCR    | KKRTKLPPADVISEHQITKNGG-----                         | VSKLEPGDRTSNYSCLKV                       |
| Nephrin | RAHSAELRETVSSFYRLNVLYRPEFLGEQVLVVTAVEQGEALLPVSVSAN  |                                          |
|         |                                                     | *                                        |
| GP354   | -----                                               | SDGSSSRCPPEE--EETG-----SREDRG--PIVHTD    |
| ICCR    | DISGGY---VPYGDYSTHYSPPFQYLTC-----                   | STKNGSSTIMQNN                            |
| Nephrin | PAFEAFNWTFRGYRLSPAGGPRHRLSSGALHLWNVTRADDG-LYQLHCQ   |                                          |
|         |                                                     | * *                                      |
| GP354   | HSDLVLEEETLETK-----                                 |                                          |
| ICCR    | HQNQLQLQQQQQSHHQHHTOTTTLPMFTLTN-----                | SSGSSLTGS                                |
| Nephrin | NSEGTAEARLRLDVHYAPTIRALQDPTEVNVGGSVDIVCTVDANPILFGM  |                                          |
| GP354   | -----                                               |                                          |
| ICCR    | IIGSREIR--QDNGLPSLQSTT-ASVVSSSPNGSCSNQSTTAATTTTHV   |                                          |
| Nephrin | FNWERLGEDEEDQSLDDMEKISRGTGRLRIHHAKLAQAGAYQCIVDNGV   |                                          |
| GP354   | -----                                               |                                          |
| ICCR    | VVPSSMALSVDPYSAIYGNPYLRSSNSLLPPPTAV-----            |                                          |
| Nephrin | APPARRLLRLVVRFAQVEHPTPLTKVAAAGDSTSSATLHCRARGVFNIV   |                                          |
| GP354   | -----                                               |                                          |

FIGURE 2 (cont.)

|         |                                                     |
|---------|-----------------------------------------------------|
| ICCR    | -----                                               |
| Nephrin | FTWTKNGVPLDLQDPRTYHTYHQQGVHSSLLTIANVSAQAQDYALFTCTA  |
| GP354   | -----                                               |
| ICCR    | -----                                               |
| Nephrin | TNALGSDQTNQLVVISRFDPPSGGLKVVS LTPHSVGLWKPGFDGGLPQR  |
| GP354   | -----                                               |
| ICCR    | -----                                               |
| Nephrin | FCIRYEALGTPGFHYVDVVPQATTFTLTGLQPSTRYRVLLASNALGDS    |
| GP354   | -----                                               |
| ICCR    | -----                                               |
| Nephrin | GLADKGTQLFITTPGLHQPSGEPEDQLPTÉPPSGPSGLPLLPLVLFALGGL |
| GP354   | -----                                               |
| ICCR    | -----                                               |
| Nephrin | LLLSNASCVGGVLMQRRLRLAEGISEKTEAGSEEDRVNNEYESQWTGE    |
| GP354   | -----                                               |
| ICCR    | -----                                               |
| Nephrin | RDTSSTVSTTEAEPYRSLRDFSPQLPPTQEEVSYSRGFTGEDEDMAPP    |
| GP354   | -----                                               |
| ICCR    | -----                                               |
| Nephrin | GHLYDEVERTYPPSGAWGPLYDEVQMGPDLDLHWPEDTYQDERGIYDQVAG |
| GP354   | -----                                               |
| ICCR    | -----                                               |
| Nephrin | DLDTLEPDSLPPFELRGHLV                                |

## FIGURE 3

1 TACTGGGGGC TAGTTCAGTC GACTAAGAGT GGGCTGGCCC TAGGGGGCCA  
51 AAGGGACCTA CCAGGGTGGT CCCGGTACTG GATATCAGGG AATGCAGCCA  
101 ATGGCCAGCA TGACCTCCAC ATTAGGCCCG TGGAGCTAGA GGATGAJGCA  
151 TCATATGAAT GTCAGGCTAC ACAAGCAGGC CTCGGCTCCA GACCAGCCCA  
201 ACTGCACGTG CTGGTCCCCC CAGAAGCCCC CCAGGTGCTG GCGGGCCCT  
251 CTGTGTCTCT GGTGCTGGA GTTCCTGCGA ACCTGACATG TCGGAGCCGT  
301 GGGGATGCCC GCCCTACCCC TGAATTGCTG TGGTTCCGAG ATGGGGTCCT  
351 GTTGATGGA ACCACCTTCC ATCAGACCCT GCTGAAGGAA GGGACCCCTG  
401 GGTCACTGGA GAGCACCTTA ACCCTGACCC CTTTCAGCCA TGATGATGGA  
451 GCCACCTTTG TCTGCCGGGC CCGGAGCCAG GCCCTGCCCA CAGGAAGAGA  
501 CACAGCTATC ACACTGAGCC TGCAGTACCC CCCAGAGGTG ACTCTGTCTG  
551 CTTGCCACA CACTGTGCGA GAGGGAGAGA AGGTCATTTT CTTGTGCCAG  
601 GCCACAGCCC AGCCTCCTGT CACAGGCTAC AGGTGGGCAA AAGGGGCTC  
651 TCCGGTGCTC GGGGCCCGCG GGCCAAGGTT AGAGGTGCTG GCAGACGCCT  
701 CGTTCTGAC TGAGCCCGTG TCCTGCGAGG TCAGCAACGC CGTGGGTAGC  
751 GCCAACCGBA ~~GACTGCGCT~~ ~~GGATGTGCTG~~ ~~TTTGG~~

## FIGURE 4

1 TCCCGGCTCT TCTCACTCC TTGCTGGGT GTACCATGCA CCTATCCCT  
51 CAGCTTCTCA TGTCTGCACC AGCGCTACTG CCCATATTTC TATCTGGGCC  
101 TCAGCCTTGT GCTGGTTGCT GCCGCCCTCG ATGTGCCCTC GCATCCACTG  
151 GGTCCACAC TGGCCTCAGC ATCTCCCCAC ACCTTCTCCT GGGTCCCAT  
201 CCCAGGGATG ACATCTTTTC TGGGGCCCTT AGAAGGTAC TGGTCAGGAA  
251 CACACACCT TCCACTCCA GAGGCTTCAT GCTGCCCTCT GCCACCCAGT  
301 TCACCCACAC TCACCTCAGGA GAATGGTGAT GTCAGGTGCT GGGTTCGGGT  
351 CCCAGACAC ACAGTTGACC AGCTACTCCT GCCAGCTAC CCAGGTGACC  
401 ATGGTGCCTG CCTCTGGGT CAGCAGGAGC AGCTTGGAG GAATGCTGA  
451 GAGAAGGGTC TGGGTAAGC TTCCAGCACT GAGAAGGACT TGAAGATTGG  
501 AGTTCGGTAC CCAGACTCTG GGAGAGAGA GGTGGGGC TTGGACTTCC  
551 GGGTTCGGG GTAGGGGAGG GCTTGAAGCC CAGACTCATG GGTCTGGGC  
601 GTCTCTACC CATACCCAGG ATGGAGAGGA TCACTCTGGG AGACAGAGC  
651 TCGGGCCCCA TCTCAGAGCG GCCGACCTGG CACTCATACT CCGCTCATC  
701 GCTGAGGTCA CAGGCTCGA TGTGCAAGTG GAATTCACCT GCAGGGGAG  
751 CCGGAAGTCA GGGCCGAGC TTCCGCTGGT GGTGAGGGT CTCAGGCTCT  
801 GATCCCTTAC CTCTAGCAGG GTCCCTTCC AGGCGGTACC TCGGGAAGCC  
851 TGGGATCTCG GGGTCGGGGC CCAGAGCAG CCCATCTTG GCCATTGCA  
901 CCGCACTGCC AGGGGTGCT ACCCCACAC GCAGCTCCAC TGAGGCCCCC  
951 TCCACCACCG TCAGGTTTTC AGGCAGGGCC CAGAAGCCCC GGGGAACGGA  
1001 GGCAGGAATC GCCAAGTGGC CCAGGCTGA GGACACAGCG CGGTGCAAGG  
1051 AAAGGGCAGA GGGTTGTCT AGGGAAGGTA AGTGGAAAT GGGGGCCACT  
1101 TGGCGCTGGG TACAAGGCTG GGATCCCACT CACCTTCAGT CAGCAGCCCC  
1151 AGGAGCAGGA GAGAAGCCCT GAGCGTCGTC CCCAGGGCCA TCACAGGTCC  
1201 CCTACTGTG ACCCCACAG CGCCCGCTGC CAGCCACCTG CGTCTGTCTG  
1251 GCTTCTCTG GGTCCCTCTC TGTGTGTCTC TGCCACCTGC TTTTCTTTTT  
1301 TATCTCTTC CGTTACTCTC CTCCCTTCTT CGTTTTCCTC TTCCCTCTCT  
1351 CCTGTGAGT ATCTCTCTCT GTCTTGCTCT CAGTCTCAAT CTCTGAGTCT  
1401 CTTTCTCTGT CTCTTAAAA AAACTTTTT TTCTTTTTTC TTTTTTTTTT  
1451 CTTTTTTTT TTTTAGAGA CGGGCTCTCA CTATGTTGGC CAGGTTGATC

## FIGURE 4 (cont.)

1501 TCAGACTCTT TCCTCAAGC CATCCTCCCA CTTGGGCTC CCCAAGTGT  
 1551 GGGATTACAG GCGTAGCCA CTGGGCCAG TCTCTTATC TTCCATCTT  
 1601 TCTCTCCTTG TCTAAGCGT TCTCTCTCT TTTGCTCTG TCTCTCTC  
 1651 TCTCTCTGTC TCTCTCTCT TCTCTCTCT AATCTCTATC TCTCTCTG  
 1701 CCACCCCTCA CTCCTGCTCC TTGTCTCACT ACTCAGAGC TTCAAGAG  
 1751 GACCTGCAGC CCAGAGTCCA GCAGGCCAGG AGCCTAGGAG AGCGATGAG  
 1801 CTGATGCAGG CACTGGCAGA GTGAGCCCTG CTCTCTGACC CAGCTTGAGC  
 1851 TCATTCTCAC AGTCAACCT CCCCAGGTA CCTTCCAGAG CCCCAGCTC  
 1901 TGGCCTCTGC CCAGCAGGCT CCTCCAGCT GGCCTAGCTG GAGCATAAA  
 1951 TCCCTCTCA GCACATGCCA GCGCGTCC TCCTGCTC CCCAGCTCC  
 2001 GTGACCCAG GGCCTGCTT AGGCTGGAA GATGGGAGAA GTCAGATCA  
 2051 GGTGCTCTCC CAGCTCAGCA GGGAGCAGC CAGCTGGGCC CCGACTCTT  
 2101 CCTGCCCCG ATACATGACC TTGGCAAGT TCTTCTTTC TTCTTTCTT  
 2151 TTCTTGAGAT AGTCTTCTC TGTGCTCAG GCTGAGTGC AGTGGCATCT  
 2201 CGGCTCACTG CAATCTCCAC CTCCTATGCC TTGAACCTCC CAGCTTCAG  
 2251 TAATCTCCC ACCTCTGCT CCCAAGTAG TGGTGTACA GGTATATAG  
 2301 ACCATGCTG GCTAATTTT GTATTTTAC TAGAGACGG GTTTCATCAT  
 2351 GTTGGCCAGC CTGGTCTCA ACTCCTGACC TCAGGTGATC CATCTGCTC  
 2401 AGCCTCCCA AATGCTGGA TTACAGACAT GAGCCACGC ACCTGGCTC  
 2451 CTTCTCTTT TTAGTAGACA TCACTGCTA AATGATGCA GGGATCTCT  
 2501 CTGGGGAGGA TGCAAGAGT AGTGTGACAG GCTGGGAGAG TGTGGGAGAG  
 2551 AGGGAAGATA TGCAATGCT TACGTGGGTG TGAGAGTGG GAAGGTTAGA  
 2601 GTGAATGCG ATCTGTATA AGCATGTGA GAGCTGTGT GTGACAGGT  
 2651 CTTACGTGG AGTGACAGG GTGTGGGCGG GAGTAAAGG CAGAGTCCA  
 2701 TTCCACCGGC CCCAGTGTG GGTGCACTGT GAGCCCAAG TGGGGGCCCT  
 2751 TTGGCAAGGA CTGATGAGC TTTCTTCTC CTCCTTTTCT TGGCCTCTCT  
 2801 CCCATCTCTT CTTTCTCTCT CCATGCTCT CTCCTCTCCT CCTCTATCT  
 2851 ATCTTGATTT ATCTTTCTT CTTTGTAGAT GGAATCTTG TCTGTTGCC  
 2901 AGGCTGGAGG GCACTGGCAT GATCTTGGTT CATTGCAGGC TCAACTTCT  
 2951 GGGCTCAGGT GATCCTCTG CCTCAGCCTC CTGAATAGCT GGGACTACAG

FIGURE 4 (cont.)

3001 GTGCACACCA CCACCTCCAGC TAATTTTTHA AAATTGTGTT GTAGAGACAG  
3051 GGTCTTTCTC TATTGCCAG GCTGGAGTGC AGTGGTGTGA TCATGGCTCA  
3101 TTGAAGCCTC AAACCTCCTA GGCTCAAGTG TCTTTCTGC CTCAGCCTCC  
3151 TGAGTAGCTG GGACTACAGG CCGCATCAC CACTCTGGCT ATTTTTTTTT  
3201 TTTTTTTTTT TTTTTGAGA GGGAGTCTTG CTCTGTACC CAGGCTGGAG  
3251 TGCATGGTG CGATGTTGGC TCACTGTAAC CTCGCCCTCC CAGTCCAG  
3301 CGATTCTCCT GCCTCAGCCT CCTGAGTAGC TGGGAATACA GGCATTGACC  
3351 ACCACACCCA GCTAATTTT GTATTTTAG TAGAGACGGG GTTTCGCCAT  
3401 GTTGGCCAGG CAGGTCTCGA ACTCCTGACC TCAGGTAACC CACCTGCCIT  
3451 GGGCCCCCAA AGTCTGGGA TTACAGGTGG GAGCGCTGC ACCCCGCCAC  
3501 TTGGCTAAT TTTTAAAT GTTTTGCAG AGACAGAGTC TTGCTATATT  
3551 GCCCAGGCTT GTCTGGAAT CCTGGGCTCA AGCAATCCTC CCATCTCGGC  
3601 CTCCAAAGT ACTAGGATTA CAGGCATGAG CCACCGCACC TGGCCCTTGA  
3651 TTTATCTTC TTTTTTCT TTTTCTCT TTTTCTTTT TTGAGATGA  
3701 GTTTCACCTT TGTGCCAG ACTGGAGTGT AATAGTGTGA TCTCGGCTCA  
3751 CTGCACCTC TGCTCCCGG GTTCAGGCGA TTCTCTGCC TCAGCCTCCC  
3801 TAGTAGCTGG GATTACAGGC ATGCCGCCACC AGGCTGGCT AATTTTTGT  
3851 ATTTTAGTA AAGACGGGGT TTCTCCATGT TGATCAGGCT GGTCTCGAAC  
3901 TCCTGCTCTG AGTGATCAG COTGACTCGG CCTCCCAAAG TGCTGGGATT  
3951 GCAGGCCGTA GTCATTGTC CCAGCTGATT TATCTTTCTA TCTTTCTCCA  
4001 TCTGTTGAG ACTCTCTGC TCTCTATATT AAGTTGTAA ATCTCAGTCA  
4051 ATCTTTATTT CACTGTGTCT CTCATCTCT ATATGTCTCT GTTATTCTGT  
4101 TTCTCTGTCT CTGTCTCAC CTCTGTCTCT CCGCTCACC CACAGTCTGT  
4151 CTCACACACA CCAGGAGCTC CATAAATATT TGTCTCAGC CACACTCTGA  
4201 CCAGCCCTCT TTCTTTATG TGTCTCTCCA TCTCCGAGTG GCTCTGCTCA  
4251 TCACATCCCT GGATTTTATA ACCATATGCT GGTGGGCTG CCTCCCGGC  
4301 GTGCACATAC ACTTGCCTGG GATAAGCTTC TTCTGCTGC TTATCTCTG  
4351 CGGGAATTGG AAATGCTAGT TTTCTCCCTA CCTCCCAAAG ACCCCGCCCA  
4401 ATATCGTTCC CAGGAACAAG ATGAGGCATC TGGCCTCAGC CCCCAGCTTC  
4451 ATCTCGATG CTGCACTCC ATCTTCCCTC ACATGCTTGA CTCCTTGGCC

FIGURE 4 (cont.)

4501 TCCTCCACCC TCCTCTCTCC CACTGCTCT CTACACCCC TGGGAATGG  
4551 GCTGGTGGC GAGCTGGGG AGTGGCTCTG TCCTGGGGC CCTGGCCAGA  
4601 TGGTGTCCCT AGGTGCCAGA GCGTGGAGCT GTCCCTTGCT GGGGCTTTA  
4651 ATAGCACAA ACCTTCCACC CTCCACCTG GCTGTTTTCC TTCTCTGCAT  
4701 GCTCTGGGA CCTTGGGCTC TCCATCTTC CATGTCCGA GCCCCAGAGA  
4751 GCCAGGAAGG GGAAGCGGG TCAAGTCCCT GGAARACAG CCCCATGACT  
4801 TGAGTTCTTC CCTAAGACTC AGGASTTCCA GCGCCATGTC CATCTATTT  
4851 CAAATCCAG GCACATAGATA AGCCACACAG AGCCCGGGAG TGTAGGCCCC  
4901 CAGATCCCTC-CCCTCTCAGA CCTGGGGTC TCAGTCCCTT CTCTCCAGG  
4951 ACTCGGAAT TTGGGCTCTT GATCCTCTG GCCACATAC CCACCCCCGC  
5001 ACCTCCCAT ACACACACAC ACACACACAC ACACACACAC ACACACACAC  
5051 ACACACACAC ATACACACAG GACTTAGGAC AGATGTTAC GGTCTGATT  
5101 CCAATCTCTC CTGGGCTGT GTGGGGGTGG GGAGAGATTG GCAGATAGAT  
5151 CCACCGACTC TTAAGACTTA AGACCAGATA TTCTGACCCC TGTACCCCTC  
5201 TTCCAAGTC ACCATGCACT TGAGTGCACC TTGASTCTCC AGCCTCTCAA  
5251 GGAACCGGGA GATCAGGCCA TCAGGCTCTC AGCCAGCAA GGCCTGAACC  
5301 ACCAGTCCCT TATAACCTG TAAGTCCAC CCCCCTCCC AACCCACTC  
5351 CCCCATTAG GGACACGGAG TGTAGCCTA AGAACAGTG AGAATCTGAA  
5401- TGTGGACCT CCACTTTTA CAGGTCCAGG AATGTCAGAT CAGGTCCCA  
5451 GCGCCCGAGC CCTCTTCAG GCTGCTGGG GTCCCTCCA CTGCTCGGC  
5501 CAGTGGCCA GCGTGGGAC GCGCCAGCTG GCGTGCATGG AGCGTCAGG  
5551 ACAAGCTGG CCGTTCCAG CTTCCCTGCC TGCCCGGGC CGGCACCGC  
5601 GCTCCGAGC CGTCCCGGG CAACCGGCC GAGGGGCCG GCGGCGGAG  
5651 TGGGGAGAG GGTGGGCTG GGAAGTGGG GTCTGGGAA AGGAGGGGCC  
5701 GAGGGCTGG ATTCTGGGT CTTAGGACGT GCTGTAGTT GCAGCAATAA  
5751 CAAGGAACA GAGGATATT TTAGGAGGG GTTTGAGGC TGGGGGAGTC  
5801 GAGTAGGGG TCCCAACTGT CCCCAGGTA TCGGTGTGCC CTCTTCCGA  
5851 CAGCGAGGC CGGGGAGCC CCGGACCCG CATCCCCAG GCGCGGAAA  
5901 CTGGCGAGC CCCAGGAGCT CCCATTTATA GCTCAGTTT CACTGAGGC  
5951 AGTCCCTCTA GGACCTGGG TGAGCAAGT TCTTCCACTC TCTCCCTTCC

WO 03/000021

12/38

PCT/US02/20275

FIGURE 4 (cont.)

6001 CTCCTCTCA CCCCTTGCCT GCCCCCAAC CCGGCGAGG CGCAGGTGTC  
 6051 CAACCCAGCC GGGACCCCTT CCTCTCTGA ACCAGGTGT TCCGGCTTCC  
 6101 AGACCCCAAT TGAGCTGGG GCGCCACCC GCGGGGGAT CCCGCCCTGC  
 6151 GTCCCCCAT CATCCGGTC TCAGCCGGG GAGTTTCTA ACGGGAAGAG  
 6201 GCGGGAGTC CCGGGGGGG GACCCGGGG GGGGAGCGG GATCGGGCCC  
 6251 TCTTGGGTC TCCAGAGAC CCAGGCCCG GAATGGCAG GCGTTTCAG  
 6301 GCGTCAGAG CTGCGGATGA GCAGACTTGG AGGACTCCAG GCCAGAGACT  
 6351 AGGCTGGCG AAGACTCGAG CCGAAGGGG GCTCCGGGG AGGGTGACAG  
 6401 GAGGCTGCT TGAGAGGAG AAGTTGACGG GAAGGCCAGT GCGACGGCAA  
 6451 ATCTCTGAA CTTGGGGGA CGAATGCTCA GGATCGGGT CCCCCTCTC  
 6501 CTCCTCTCC TCTTCTGCT CAGAGGAGA GCAAGTACCG CACGAGGGA  
 6551 GCGAGGAAT ATGGGTGGG GGTGGGAGT TGCTTGGGG CTGCTCTTC  
 6601 ACTAGCGAG AGGGAGCTGG GGGCTGGAC TCCTGGTCC TGAATAGGA  
 6651 GGGCCCTGA GGTGCTAAG TCAGCCCTG TGCCCCGAC TCTCTAGGC  
 6701 CGCTCGCCC ATTCTGCA ACAGCCAGAG GAGCTGGTG TCTCTGGG  
 6751 GGAGGAAGC CGGCTGCTT GTCTCTGGG CGCTACTGG GGGCTAGTC  
 6801 AGTGGACTTA GATGGGCTG GCGTAGGGG GCCAAGGGA CCTACAGGT  
 6851 AAGAGTGTTC TCTCCAGCT GGGAGGGGT GGCTAGGGG AGAGTTGCTG  
 6901 GGTCTGGCTG TACCTGAGT TTCTATTITG ACATTTTCAA GTTTGGGAA  
 6951 TTGATGGCT CGGGTAAACA TTAGGAGTC CTGATTTTG AGCTGCTCT  
 7001 TTGGGGTGA CCCACGGAG TTGGGAATTA TTATGTATT GCAAAATAGT  
 7051 ACATAGGCCA GGTGCACTGG CTCACGCTG TAATCCCAAC GCTTTGGAG  
 7101 GTTACGGCA GAGGATCGCT TGAACCAAG AGTTTGAGAC CAGCCTGGG  
 7151 AACATAACA GACCTTATCT CTACACAAAT GTATATATAT ATTTAAACA  
 7201 AATTAGCCGG GTATGGTGGT GTGATCTAT AGTCCAGTT ACTCAGGAG  
 7251 CTTAGTGGT AGGATTGCTT GAGCCTAGGA GTTCAAGCT GCAGTGAGCC  
 7301 ATGATCAAG CACTGCATT CAGGCAATGG TGAGACCCTG TCTCAAAAA  
 7351 AAAAAAAAA GAGAACATAA ATGCAAAAA GTACAGTAA TATAATGGA

Exon 1

Exon 2



FIGURE 4 (cont.)

7401 AGATTTCACCA AATAAAATAG ACACACACAG CCAATACCCA AGTCCATTGC  
7451 TAGCTCCCCA GAAGACCCCG TGTTCCTTC CCTATCATA GCCCCCTCCC  
7501 CCTCACTCCA GAAGTAGTAT CTAACCTAAT TTTATGGCA ATCATTITCT  
7551 TGCTTTCCTT CCTGACTTAA TTACCCCTAA GTTTCAGTG ACTCTGGGT  
7601 GGGAGGGAGT TAGAGTCTCT CTGGGCCAG TACACACTT TTAATAGTG  
7651 CTTACCACCA AATGTGTGGG CCAGTTTCT GGTGGAGAT GTCTGGGGAT  
7701 GGAGGCCCTGA GGCAGGATT TCAGAACCAT GGTGTGCTGA CTGCTTCTC  
7751 CCTGACTCCA GGTGGTCCC GGTACTGGAT ATCAGGGAAT GCAGCCAATG  
7801 GCCAGCATGA CCTCCACATT AGGCCCGTGG AGCTAGAGGA TGAAGCATCA  
7851 TATGAATGTC AGGCTACAGA AGCAGGCCTC CGCTCCAGAC CAGCCCACT  
7901 GCACGTGCTG GGTAGGACC TCGCCCACTT GTCCCTGGG AGCCCAAGAG  
7951 GGCAGCCCGT ACTAGCTGTG AGTAGCAGAG CCCAGGGAGC CCAGGGCAT  
8001 GGTCAATTGG AGCTGAGAAG ATCAGGATCC ATCTCTGACC CCAATCCAC  
8051 CTTGCACTCC CCCCAGAAGC CCCCAGGTG CTGGCCGGCC CCTCTGTCT  
8101 TCTGGTTGCT GGAGTTCCTG CGAACCTGAC ATGTGGGACC CGTGGGGATG  
8151 CCCGCCCTAC CCTGAATTG CTGTGGTCC GAGATGGGT CCTGTGGAT  
  
A  
8201 GGAGCCACCT TCCATCAGGT CAGGTCCAAA TTCCTGTGCT AGCCTTTGCC  
8251 CATTGAGGGA AACTTGGGT ACACCTGAC CACAGGCTCA TCCAGAAGAG  
8301 AAGAAGACAT GGGAGGGCAG AGGTTCTAGG GTTTGGACTC TTGAATATG  
8351 ATGCAGGGTA AAGATTCTAG GGCCAGACTA CCTGGGTCA AATTATGCT  
8401 CAGCCACTTG CTAGTTGATT GATCTGAGT AAGTTAGTAA ACCTCTCTGT  
8451 GCCTCAGTTG CCTTATCTAT ACAATCAGGA TAATAGTAGC ATGCATGTCA  
8501 TAGGGTATTG TGAGAATTAA ATAAATAAAT ACCTATAAAT GCCCAGAAGA  
8551 GTGACCRATA CATAGTGAGC ACTATATAAG TAAGGCAAGC TTGTCCACC  
8601 TCGGGCCCAT GGGCTGCAIG CAGCCAGGA TGGCTTTGAA TGTGGCCAC  
8651 CACAAATICA TAACTTTCT TAAACATTA TGAGACTTTT TTGTAATTT  
8701 TTAGCTCATC AGCTATCATT AGTGTAGTA TGTGTGGCCT AAGACAATTC  
8751 TTCTTCCAT GTGGCCAGG AAGCCAAAA GATTGACAC CCCTGATGG  
8801 TAGATGGCAT TATTATTCTT ATCCTTCCT CCAGACCCCTG CTGAAGGAAG

Exon 3

Exon 4

Exon 5

FIGURE 4 (cont.)

8851 GGACCCCTGG GTCAGTGGAG AGCACCTTAA CCCTGACCCC TTTCAGCCAT  
 8901 GATGATGGAG CCACCTTTGT CTGCCGSSCC CGAGCCAGG CCTGCCCCAC  
 8951 AGGAAGAGAC ACAGCTATCA CACTGAGCCT GCAGTGTGAG TGCAGCTGGC  
 9001 CCTGGGAAAG AGGGGTGTGG GGCCTGACT CCTGGGTATG AGGAAGGAGG  
 9051 GGACTGTGGC CTTGGGGAA TGAGGAACCT GGAGCCTGGA CTCCTGGATC  
 9101 TAAGATAGCA GGAGAGGGCT GGGTATGGTA GCTACGCCCT GTACTCACAG  
 9151 AACITTTGGGA GGTGAGGCA GCGGATCAT CTAAGATCAG GAGTTCGAGA  
 9201 CCAGTCTGGC TAACATGTCC AAACCCCGTC TCTACTAAA ATACAAAAT  
 9251 TTGCCGGGCG TGGTAGCACA CACTTGTAA TCCAGTACC TGGGAGGCTG  
 9301 AGGCAGAGA ATCACTTGTG CCGGGAGGC AGATGTTGCG GTGAGCCGAG  
 9351 ATCATGCCAC TCAGCAGCAG AGTGAAGCTC CGAGCAGAG AGGACAGACA  
 9401 GCTGGGTCC CTGGGAAAG AGAAAGCTGG GCCTTGACTC TCAGATCGGG  
 9451 GAGACTAGGA GAGGACAGAA GGCTGCACA TTGAGGTAAC TGGGAAAT  
 9501 GGAACTGAA AGCCAGACT CTTGGCTCAA AGGGAGAAGG GGATTAGGGG  
 9551 CCCAGACTCC TGGGATGGAG GAACAGGGA CTGACACCT AGGCCAGTGA  
 9601 CGGAGGTGTT CTTGGTCTTT GCCCATCTGA CCATTGTCCC ACCCTCACAG  
 9651 ACCCCCAGA GGTACTCTG TCTGCTTCC CACACACTGT GCAGGAGGA Exon 6  
 9701 GAGAAGTCA TTTCTCTG CAGGCCACA GCCAGCCTC CTGTACAGG  
 9751 CTACAGTGA GGACGAGAC CCACCTCTC CTAGCCCCAA GAGTGAGCTT  
 9801 GGGAAGGGCT GGAACCTGAG TAGGTGTGCC AGAGAGGCCA GGACAACGTT  
 9851 AACAGCGCCA CCATTTCCTC AGGTGGCAA AAGGGGGCTC TCCGGTCTC  
 9901 GGGGCCCCCG GGCAGGTT AGAGGTCTG GCAGAGCCT CGTTCTGAC Exon 7  
 9951 TGAGCCCCTG TCTGCGAGG TCAGCAACGC CTTGGGTAGC GCCAACCGCA  
 10001 GTACTGGCT GATGTGCTG TGTAGCTGG GGCAGGCTG TGGGTGTGGT  
 10051 CAAGGTGGC CTGGCTTTC AGGGCTGTT AGGTCGGGG CTTGAGGGG  
 10101 CGGGCCCGG AGAGCGAGCG TGGGTATTA GGAGAGGAG AGTGTGAGC  
 10151 TGGGGCATAT TCTTCCGCC TAGAGGTGT GGTGTTTCTG TGGGGCTGGC  
 10201 TGATCCAGG TCAGTGGCTG CATTCCGCC CGGCATGTG ACCCTAGTC  
 10251 TCTTTCTGTC AGTTGGGCG ATTCTGCAGG CAAGCCGGA GCGGTGTCC Exon 8  
 10301 GTGGACGTGG GGAAGAGCG TTCCTTCAGC TGGGCTTGGC GCGGGAACCC

FIGURE 4 (cont.)

10351 GCTTCCACGG GTACCTGGA CCCGCCCGG TGGCCGCGAG GTACAGCCCT  
10401 AAATCTGAGG CGGTGGCTGG AGGGGGACCA GGCTTCCTTA CAAATCCGGC  
10451 TTCTGACGCC CTTCCCTGT CGCAGTGCT GGGCTCTGA GCCACTGTC  
10501 GTCTTCGGTC GGTGGGGCCG GAGGACCGAG GCGACTATGT GTGCAGAGCT Exon 9  
10551 GAGGCTGGGC TATCGGGCTT GCGGGCGGC GCCCGGAGG CTCGGCTGAC  
10601 TGTGAACGGT GAGAAGGCGG GGCCTCCTAG GGAACCTGGC CCGTCTGGG  
10651 ATAGGGAGCG GACAGAGGGG GCAAGGCTA ATGCAGTGGG AGTGGCCTGG  
10701 AAGGAGCTTT ACACCCAGCG GGGGCTGGAG ACCCGACCTA TTGAAGCCGA  
10751 GGCTTTTAGG AGAATCGGAG TTGAGGGCG GCCTGGCCTG ATTGATTGAG  
10801 GTTAGCGGAG AGTGGCTGG ACAGACCCGG CTTTGTACA GCCTTGGGG  
10851 AGGGCAAGAC CTCTCCTCTG AGTGACCTAC AGTCTCATC CCACTCCCC  
10901 CAGTATGAC CCGCTGCAC TCTGGGCTG CCTTCCTGAG GGGGCTGCT Exon 10  
10951 CGCTCCAGT GTCTGGTTT GCGCTTCCC GCCCAGATG CCGTGGTAG  
11001 GAAATCTCAC TCCTCCCGTG ACCCATCCAG CCGTGATCC TGACCTCCCA  
11051 CCTGGCCCC CGAACTACT GTGACATT CTGACTCCC AGACATCCCT  
11101 CCTGCTCTT CCTCCCTCC TCAGTCTCT CCGTGCTCT CCTCTTTGT  
11151 GCCCCAGGT CTGCTCTTG GATGAGGCT TCCTGAGGC GGGGTCGAG Exon 11  
11201 GGCCGGTCC TGGTGGAGAC ATTCCCTGCC CCAGAGAGCC GCGGGGACT  
11251 GGTCCGGG CTGATCTCTG TGCTACAT TTGGGGACC CAGGATCTC  
11301 ACTTAGCAG GAGCTTTAAG TGCACTGCC GGAACCGCT GGGCGAGGA  
11351 GGTGGCCAGG CCAGCCTGG CCCTAGAGGT GAGACCCAG CCCGAAGAC  
11401 CCAAATCTGG AGAGCTAAA CCCCAAAAC GCAGGGATCC CCCAGCCGAG  
11451 GGCTGCAAAA CTTATACCC TCAATGCGAG AGGAGACCTC CAAACTCTCG  
11501 GAGTCTCAAA ACTGTGGGT CATTGATCC CAAGACACC CTCAACCACA  
11551 AATGCCTTCA CATTCTGAAT CCTAACTGA GAGACTCCTC ACACCTAGGG  
11601 GCCCCAAAA GGAAGACTCC AATGATTGCA AAGCAATTG CAAAGTAAAG  
11651 GACCCCTCAA ATTCTAAGAC TCCCTAAGC CAGGGATTT AAACCTACTC  
11701 TCAAACTGG GGAACCCAA ATTCAAGGGC CTTTGAATCT TCAATGTGC  
11751 GACCTTTGA ACCCAGGAAT CCCAACTCA ATCCCTGAGC CCCCCTTCC  
11801 TGGTTCCCC TCAGCCTTCT CAGGATGCC CCTCTGCTC CTGCAGACTT

FIGURE 4 (cont.)

|       |                                                               |         |
|-------|---------------------------------------------------------------|---------|
| 11851 | <u>GCTGCCCACT GTGGGATAG TGCCGGAGT GCGCGCTGCC ACCACAATC</u>    | Exon 12 |
| 11901 | <u>TCCTTATGCT CATCACTGGG GTGGCCCTCT GTGCTGGCG CCACAGCAG</u>   |         |
| 11951 | <u>GCTTAGTGCC TGAGCCCGCG CCGGCTCC GAGGCCCGAG CCCCACACGC</u>   |         |
| 12001 | <u>GCCCTGGCTG CCACTGACC TGACCTGGCC TTGGGCTTG GCTCCAGTCC</u>   |         |
| 12051 | <u>CATTTCACGC TGTGCACAGG GCTTAGCTCT CTTACAGTT CTGGTTCCCT</u>  |         |
| 12101 | <u>CCTTAGCCCC TAAC TAGGCC TTCCAGGGT CACACTCCTC GGTGGGATG</u>  |         |
| 12151 | <u>ATTCTTATTG GTTCCAAACA GCGCTACCCA ATCAGCCTCA TTGGTTCCCA</u> |         |
| 12201 | <u>GTCTCTCTC TTCCCGCTTA TTGGTCTGCA CACATTGTGA CCCCGCCAT</u>   |         |
| 12251 | <u>CGCTTAATC GACGGTGGC TGTGTGTCAG CCTCAGCCTC TTTCTCCGAG</u>   |         |
| 12301 | <u>CAAAAGAACC TGATCGAAT CCTGGCGAG AGCGACGGCT CCAGTTCCAG</u>   | Exon 13 |
| 12351 | <u>AGSTCTTGAA GAAGAGAGA CAGCGACCG CGAGGACCG GTAGGATGCC</u>    |         |
| 12401 | <u>AGGCTCCCA GACCTGACTG TGCTCCGGA CTTAATAT AGCCAGTCC</u>      |         |
| 12451 | <u>CAAGAGGGTC CCAATTC AATAGACTC TAAGGCCAGG CATGTCCT</u>       |         |
| 12501 | <u>GACCTGGTA ATACCACTT GGGAGTGGA GACACAGGA TCACTAAGG</u>      |         |
| 12551 | <u>CCAGGAATC AAAGCCAGCC TGGACAGCAT AGCAGGAGCC CATCTCTACA</u>  |         |
| 12601 | <u>AAATACAAA CTAATAAATA ATAAAAATG AACCGGTAT GGTGGCATA</u>     |         |
| 12651 | <u>ACCTATAGTC CCAGCTACTC AGGACACTGA GGTGGGAGGA TCCCTTGAGC</u> |         |
| 12701 | <u>ACAGGAGSTA AAGGCTGCAG TGAGCTATGA TTGCACCATG CACTCCAGCC</u> |         |
| 12751 | <u>TGGGCTACAG AGCAAGAGCC TGTCTCCATT TTTTTTTTTT TTTTATGT</u>   |         |
| 12801 | <u>AGGAGGGCTC TAGTCTTTT TTTTGTGCA GAATTTCACT CTGTACCCA</u>    |         |
| 12851 | <u>GGCTGGAGTA CAGTCTGCG ATCTCGGCTC ACTGCAACCT CTGCTCCCT</u>   |         |
| 12901 | <u>GGTCAAGTG ATTCTTTGC CTCAGCCTCC TGAGTAGCTG CGATTACAGG</u>   |         |
| 12951 | <u>CGCCACCAC CACGCTGAC TGATTTTGT TTTTAGTAG AGATTGGGT</u>      |         |
| 13001 | <u>TCACCATGTT GGCAGGCTG GTCTCAACT CCTGACCTCA GGTGATCCAC</u>   |         |
| 13051 | <u>CGGCTCGAC CTCCAAAGT GCTAGGATTA CAGGCATGAG CCTCCAGGC</u>    |         |
| 13101 | <u>CGGCTGAGG GCTCAAGTCT TTTTCTTCT TCTTTCTTT TTTTGAGAC</u>     |         |
| 13151 | <u>GGAGTCTTG TCTGAGGCC AGGCTGGAGT GCAGTGGCG GAACTCGACT</u>    |         |
| 13201 | <u>CACTGCAAGC TCCACCTCCC GGGTTACAC CATTCTCCTG CCTCAGCCTC</u>  |         |
| 13251 | <u>CAGAGTAGCT GGGACTACAG GCACCCGCCA CCATGTCCAG CTAATTTTT</u>  |         |

FIGURE 4 (cont.)

13301 TGTATTTTAA GTAGAGACGA GGTGTATACC GTGTTAGCCA GGATGGTCTG  
 13351 GATCTCCTGA CTTCTGTGATC CGCTCGTCTC GGCTTCCCAA AGTGTGGGA  
 13401 TTACAGGGCT GAGCCACCGC GCCCGGCCAA GGGCTCTAGT CTTAACAGTG  
 13451 ACCCCACACC CAAATGTCAC CCAAGTCCAT GCCCTGACC CAATTATTCC  
 13501 CTAGGCCCAG TATGTCCCA CAGCCCGTTT TTGTGTGTGT TGTGTGTGT  
 13551 GTTGTGTGT TTGAGATAGA GTCTTGCTCT GTCTGCCAAG CTGGAATGCA  
 13601 GTGTGCAAT CCAGACTCAC TGCACCTCC ACCTCCAGT TCAAGTGAT  
 13651 CTCGTCTCTT AGCCTCCTGA GTAGCTGAAA TTACAGGTGC CTGCCACCAT  
 13701 GCCTGCTTAT TTTTGCATT TTTAGTAGAG ACAGAGTTT GGCATGTTAG  
 13751 CCAGGCTGGT CTCAAACTTC TGGCCTCAAG TGATACTCCT GCTGCGGCTT  
 13801 CCCAAAGTGC TGGGATTACA TGCATGAGCC ACTGTGCTGG CTTCTTACAG  
 13851 CCGTTTATT GTCTGAGTG CAGTCCCAAG CTCTGGGTG CTCTTACTCC  
 13901 CTCCTGCTG GCCTCCACTG CTGGCTGAA GGTCTTGGG CTCTGGCAAT  
 13951 GGGGGGGGG GATCCTTGA CTATTCCCTC TCACTAAGTT CCCTACCCCA  
 14001 GGGCTTCATT GTGCACATG ACCACAGTGA TCTGTTCTG GAGGAGAAAG  
 14051 GCACTCTGGA GACCAAGTGA AGTGTGAGA GGGGTGGGG TCCCTTCACT  
 14101 GTTGGGAGAG GCGGGGCTCC CTTCATTGTG TTTCCGTCTC TCTCCAGGC  
 14151 CTGTCCCTC CTTTTCTT CTGTGTCTCT CAGAGTTGGG ACTCAGCTCC  
 14201 CCACCCCACT CCTCCTGCC CTTGGGCAAT CTCACTCAGC TCCAGGCTC  
 14251 AGTTTGCTG TCTGCAGACT CTTCCACAC ATCTGTCCCA GCGCTAGCCT  
 14301 CCATCTGGAG CCCCAGACCA GGGCTCACC TGCCTGTGCT CTCCTCATCA  
 14351 CGGTCAAGCC CCGTTTCAGC CACAGGTCC TACACTGGCC CCACATCTCC  
 14401 CCAGACTGGT TCTTCTCTG GGTCTCTACC TCAGGACAGC CACATTGACT  
 14451 CCAGGCCATC CCGAGGCCAG AGCACTTCTC TCTCTCTCTC TCTCTGCTG  
 14501 ACCTAGCACA TGCCATTCTC TCTCTCTTT TTTTTTTTT TTTTTTGAG  
 14551 ACGGAGTCTC ATTCTGTTC CCAGGCTGGA GTGCAGTGGT GCAATCTCAG  
 14601 CTCACTGCAA CCTCTGCTC CTGGGTCAA GCCATTCTCC TGCCTCAGGC  
 14651 TCCCTAATAG CTGGCTAATT TTTCTGTAT TTTTAGTAGA GATGGAGTTT  
 14701 CACCATGTTG GCCAGGTGA TCTGGAACTC CTGACCTCAA GTGATCCGCT  
 14751 CGGCCAGCC TCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGTGCCC

Exon 14

FIGURE 4 (cont.)

14801 AGCCGACATG CCATTCTCTT GGCTGAAGC ACTCCTACCT TCCTTCCCAT  
14851 GTCTACCTAA TTCCTTCCTT TAGTCTTCCA GTCTCAGTC AGACATTCT  
14901 TGTCTAGGA AGCCCATGCT TCCGTCTGA CAGCTCGATC ATTTTGGCTG  
14951 TGTTCACCC ATCAGAGCA TGACCACTCT GATCTGGCT TCCTTATCCC  
15001 ACCCACTATG CTGAGGGCTC TACCATCACA GCCCCTGTCA TTGCTATGC  
15051 CTTTCCCAGG CACAGCCCTG ACCCCTCTGG GTACTGTCTC ATGATCTGC  
15101 ATTTTTCCTT TGGTGTGGGA TTCTGTAGG ACAGGTCCA GTTCTATCCT  
15151 AGTGACATGC CTGTAGCAG CAACACAGG TGTGACATG AATCAAGCC  
15201 TAGAGCTGT TGGGAGGCTG AGTGTCTCTC TCCTGTCCC TCTGCACCT  
15251 CCACACGAG ACCCCTCAGC AGGCTATAT CCCTCGCTCT CTACCTTCT  
15301 CTGCTATGT CTTATCCATT TGCTCTTAT CACTGTCTCT CTGTCTCACT  
15351 TCTCTCTCT CCCAGTCCAT GTGTCTCTCT GTGTCTCTGC CCACCTCTGT  
15401 CTCTTTTGT CTCTCTCAG GTCTGTCTA TTTCAGTGT TCTCTCCATC  
15451 AGTGACCTC ATCCCCCTG CAGCTCACA GACTTACTG AGTCCCATTT  
15501 GTCCCTCAG GACCCACCA ACGTTACTA CAAGTCCA GAGTCACTG  
15551 TGAGCCTGAG CCTTGGCGAA GCCCCTGGAG GAGTCTCTT CCTGCCACCA  
15601 CCTCCCCC TTGGCCCCC AGGAGCCCT ACCTTCTATG ACTCAACCC  
15651 ACACCTGGC ATGCTCCCC CCTGCAGCT TTACAGAGC AGGGCAGGT  
15701 ATCTCACCAC ACCCACCCT CGAGTTTCA CCAGCTACAT CAACCCACA  
15751 TCCTTTGGG CCCCAGATCT GGGCCCCGG ACTCCCCCT TCCCATATG  
15801 TGCCTTCCC ACACCTAGC ACCCGCTCT CCAGACTCAG GTGTGACATC  
15851 TTCCAAATG AAGAGTCTG GGATCTCCA CTGCCATAA TGGATTGTT  
15901 TGATTCTGA GGAGCCAGGA CAAGTTGGG ACCTTACTCC TCCAAACTG  
15951 AACCAAGG GAGGGAAGA TCATTACATT TGTAGGAGC ATTGTATAC  
16001 AGTCAGTCA GCCAAGGAG ATGCCCCAG TGGAGCAGC ATGCCCCC  
16051 AATATGCCA CCTATCCCC GGTGTAAAG AGATTCAAG TGGCAGGTAG  
16101 GCCCTTGGG GAGAGATGGG GACAGGGCAG TGGGTGTGG GAGTTTGGG  
16151 CCGGATGGA AGTTGTTTCT AGCCACTGA AGAAGATATT TCAAGATGAC  
16201 CATCTGCATT GAGAGGAAG GTAGCATAGG ATAGATGAAG ATGAAGGCA

Exon 15

FIGURE 4 (cont.)

16251 TACCAGGCCC CACCCTGGCT CTCCCTGAGG GGAACCTTGC TCGGCCAATG  
 16301 CAAAATGAGC CAGATGGCC ATATACTCCC TAGCAACCA AGATGGCCAC  
 16351 CATCTTGATT TTACTTTCCT TAAAGACTCA GAAAGACTTG GACCCAAAGGA  
 16401 CTGGGGATAC AGTGAGAATT ACCACTGTTG GGGCAAAATA TTGGGATAAA  
 16451 AATATTATG TTTAATAATA AAAAAAGTC AAGAGGGCAA GTGTGTCTTA  
 16501 GGGAGCTAC TGGCATTATC ACTCTCCACC AAGGAAGGGG TCCCTTAGAC  
 16551 CTGTCCCAAG GTCCCTCCTC TACCCTAGCC TATGAGGTGG CTGTAGAGT  
 16601 AAACTGTGA GCCACCTCTC AGCCTCTTGC TACCTGCAAA GCACCTTAGG  
 16651 CTCTTTTTT TTTTCTCTG AGACAAGATC TGGCTCTATG GCCCACATTG  
 16701 GAGTGCAGTG GCATGATCTC AGCCCACTGC TACCTCTGCA TCCTGGGCTC  
 16751 AAGCCATCCT TCCACCTCAG CCTCCCAAGT AGCTGGGACT ACAGGTGCAT  
 16801 GCCACCACAC CCAGCTAATT TTTGTATTG TTTGTAGACA GGGTTTCACC  
 16851 ATGTTGGCCA GGCTGCTCTC AACTCCTGA CCTCAAGTA TCCGCCACCC  
 16901 TAGGCTCCC AATGTCTGG GATTACAGG ATGAGCCACT GTGCCAGCC  
 16951 ATGGGCTCTT TTAATATACA TCTTCACACA CACACACACA CACACACACA  
 17001 CGCACACACA CACATGAGTT GCAACAGAA AAGACACACA CATAGGCATG  
 17051 TATGCACAGA CACACGCATA GATGTCCACA CAGTTGCACA CAAGTGACAG  
 17101 GGCTGCCCCA GGGGTCTTGG GGAAGACTGA ATTCTAATC TCATTAGAGG  
 17151 AGACAAACAA GTGAGCCCTG AAGTGGAGCA GGGAGGGGA GACTATGGGT  
 17201 AGGAAATGG CAATCCCTTG GTCTTACAG CAGCGTGA GATCCAGCC  
 17251 CTAATCCTGA GGTGCTGCAT CCACAGTGGG CATGGTCTG GTGCCTGCTT  
 17301 GGATGATCCT TAAAGAAAGG TCCTGGGGGG TTGGTTTCT GATCCTTGA  
 17351 GCTAGGAGTT AAAGGTCCAG GCCCTGGGA CCCTTGGGA GCAGAGCAG  
 17401 AAGAGTGAAC TCCTGGGTCT GAAGGAGAAT GGGGTGGGG CTGGTCTCT  
 17451 GGTCTGAGA GAGAAGTGC CCAGACTTCT GGATCTGAAA GAGGAGGGA  
 17501 CTAGGTCTCA ACTGCTGCTT TCTGACTGG GGACATTTG GAGGCTGTA  
 17551 TTCTGAGCC CTCACAGAG GAATGTACTA GGGGATGGG GTCTCTGATG  
 17601 CTTGATCCT TGGAAAAGGA CAAACTGTG AGTGTCTGGG TCTAAGAGG  
 17651 GTGAGAGTCC TCCGGGAGGA CTCAAAATCC ACACGGGGG GAGCCCATAG  
 17701 CTGGACTCCT GGCTGGGCCC TTCATGGGGC GGGACGCTG GAATCTGAG

FIGURE 4 (cont.)

17751 GGGCGGGGG CTGGCGCAGG CTCCCGCCCG GGGTTCCCGA GCTGCTCCAC  
 17801 TCTGCGCGAA GCGGCCACGC TATTGTCTTG ACCAGGAAGG CGGGGCCGGC  
 17851 GCGGGCGGGG GCTGGCGGGC CCGGCGCAGC CCGGGGGCGG CCGGAGGAGG  
 17901 AGGTGGCGGG GGTGGCGGTG GGAGCTCCTG TCACCGCTGG GGGCGGGCGG  
 17951 GGGGGGAGTG CAGGGGAGGT GAGGGCGCAA GGGCGGGGCA ATGGGGGCGG  
 18001 CCAGCCCGCG TGCTCGCGGT CTAACTGCGC GCGCGGGCCA GCGCGCGCTG  
 18051 CCGCTGCTGG TGCCACTATT GCTGCTGCTT CTGCGCGCGC AGCCCGCCAT  
 18101 CGGGAGCCTG GCGGTGGGA GCGCGGGCGC GCGGAGGTG AGCGCGGGCG  
 18151 GGGTCCTGGG GGATGGGGGA AGGGGCGGGA CCGGGTCTCT GGACGCGGGC  
 18201 GCGGACATGT CCAGGGCAGA AAGCGCGGTG TTTCAGCCCA GGTGGTCAGC  
 18251 CCGCAGCGCG CCCCAATCAC ATTTATGAAC CAGGGGTTCG AGCGCCCGAG  
 18301 TCCCCATCA TCGGAGCTCC CAGCGCCCTC CCATCTCGAG CATAGGAAT  
 18351 GGTCTAATCA GAGCCCGTGG TCCCAAGAGT CCAGCCCGCT CTCAGAGCC  
 18401 AGGTGACTCG GCGCCCAAGC CCTCCCGCT GGACATAGGA CCCACCAAGC  
 18451 AGCGAGGCAT TTAGATCCAA TAATCCAGAC CCCTGTATT CTCTGAGCC  
 18501 ATATGGAGGC CCTTGAGCC TCCAGGAGC CAGGAGTCCA GTCCCTCAGT  
 18551 CACCACCCAC CCCAACAGA TGTAGCTCTC CAGTCTCAA GGACCTGCTG  
 18601 TCCAGGACTG TAGGCCCTG AAGCCAGGCC TTGTGAGCTT TGCATCCTGC  
 18651 AACGGAGCC TGAGCAAGGG ATGAGGGAG GAGGGGCCAG AACTCTGCTG  
 18701 TTCTGGCCTC CTCCTCCGCG ATTCAGGTTT AACCCCTTCG GGTCCAGAG  
 18751 CGGCTCGCT GGGGTGGGGG CGGAGTCTGT CTCGCGGCA ACAGGCGA  
 18801 AAGAATCCCG GGGGACCCAG GTCCCATAG CAACGGGAGC GCTGGGGCGC  
 18851 CCGCGCCCTA CGGAGCTGT TTCCAGGGA ACGGTGCTC CATGGAGCG  
 18901 GTGTGCGGTG CTGGGGGAG GGGGCTGGTG CTGGGGTCT CCGTCTTAGG  
 18951 GAGCAAGAA CCAGGGGACC CTCATGCCAA CGCCCCCGA GCGCTCACTG  
 19001 TCCTTTCCAC TTCCATCCAG GCGCGGGGT CGGCCAGGT GGTGGACTA  
 19051 TCGGGGCGCC TAACCTTCA CCGGGACCTG CGCACCGGCC GCTGGGAACC  
 19101 AGACCCACAG CGCTCTGAC GCTGTCTCG GAGCCCGCAG CGGTGCTGG  
 19151 AGTACTGCAG ACAGTGGGC GGGGCCGAAC GGGAGAGCG GGGCGGCCA  
 19201 TAGAAAGCTA GACTTGAAA AGCGTGGTC CAGGGTCTG CGGATCTAA



## FIGURE 4 (cont.)

19251 GGCCTGGAGG CTGGGGGGCG TGGCCAATAA AGAGGCGCAA CTATGCTAGG  
19301 GGCAGGGGAC CTGTTTGTAG ATACTAATC AGGAAAAGG GAGAGCCGCG  
19351 AGATAGCCAG AGAGGAAGTG GAATTTAGGA ATCTGGTGGT CTTTGTAAAG  
19401 AGTAGAGGTG TAGGGGGGAG TGGCGAAGG ATAGGCGGGG CTAAAGACAGA  
19451 AAGAGACCTT AAGGACCAGC AAGATGGGGA AAGGGGTGGA GCCCAATGAG  
19501 AGCGCGGAGA GCTGGGGGGG CGTGGCCATG AAAAGACAAA TTTATAACGG  
19551 GAAGGGAGAG TTTTGGAGAG GCGGAATAGA GGAAGAAGCG GGGCCTAAG  
19601 GAGGGTGAGA CTTTGGGGA GACGAATCTG ACTGCGGGGA GGGGTGACCA  
19651 GAGAGGTGGG CTTAGAGGGA CCTTCAGAAA GAAACAGCAC AGGAAAAGAG  
19701 ATAGGGCTTA AAGATGACGG GACTTTTAAG GGAAGACTGC TAGTGGGCGT  
19751 GGCCATGAG CACAAGGAGC TTGGATATCT AAGGCTGGTG CTAGGGAGAA  
19801 GCAGGGCCTA GGGAAAGCAT GTCTCATGA ATACTAGAGC CTTGAAAAG  
19851 GACCTGGCCG GCGCGGTGG CTCACGCTG TAATCGCAGC ACTTGGGGAG  
19901 GCCAGGCGAG GCGGATCACC TGAGGTCAGA AGTTGGAGAC CAGCCTGGCC  
19951 AACAGGCGCA AACTCCGTCT CTAATAAAA TACAAAATT AGCCTGGCAT  
20001 GGTGGTGGCT GCCTGTATTC CCAGCTACTC AGGAGGCTGA GACAGGAGAA

WO 03/000021

22/38

PCT/US02/20275

FIGURE 5

```

1  GGGAACTGGCAGGCGTTTCAGAGCGTCAGAGGCTCGGATGAGCAGACTTGGAGGACTCC 60
   -----+-----+-----+-----+-----+-----+
61  AGGCCAGAGACTAGGCTGGCGAAGAGTCGAGCGTGAAGGGGCTCCGGGCCAGGGTGAC 120
   -----+-----+-----+-----+-----+-----+
121  AGGAGGCGCTGCTTGACAGGCTGAGGTTGACGGGAAGGCGCACTGCGACGGCAATCTCTG 180
   -----+-----+-----+-----+-----+-----+
181  AACCTTGGGGGACGAATGCTCAGGATGCGGGTCCCGCCCTCCTGCTCCTCTCTCTG 240
   -----+-----+-----+-----+-----+-----+
      M L R M R V E A L L V L L F C
241  TTCAGAGGGAGAGCAGGCGCGTCCGCCATTTCCTGCAACAGCCAGAGACCTGGTGGTG 300
   -----+-----+-----+-----+-----+-----+
      F R G R A G F S P H F L Q Q P E D L V V
301  CTGCTGGGGAGGAAGCCCGGCTGCCGTGCTGCTGCGCGCTACTGGGGCTAGTTCAG 360
   -----+-----+-----+-----+-----+-----+
      L L G E E A R L P C A L G A Y W G L V Q
361  TGGACTAAGAGTGGGCTGGCCCTAGGCGGCAAGGGACCTACCAGGCTGGTCCCGGTAC 420
   -----+-----+-----+-----+-----+-----+
      W T K S G L A L G G Q R D L P G W S R Y
421  TGGATATCAGGGAATGCAGCCCAATGGCCAGCATGACCTCCACATTAGGCGCGTGGAGCTA 480
   -----+-----+-----+-----+-----+-----+
      W I S G N A A N G Q H D L H I R F V E L
481  GAGGATGAAGCATCATATGAATGTCAGGCTACACAAGCAGGCTCCGCTCCAGAGCAGCC 540
   -----+-----+-----+-----+-----+-----+
      E D E A S Y E C Q A T Q A G L R S R P A
541  CAATGCACTGCTGCTGCTCCCCCAGAAGCCCCCAGGTGCTGGGCGGCCCCCTCTGTGCT 600
   -----+-----+-----+-----+-----+-----+
      Q L H V L V P P E A F Q V L G G P S V S
601  CTGGTTGCTGAGTTCTCTGCGAACCTGACATGTCGGAGCCGTGGGGATGCCCGCCCTACC 660
   -----+-----+-----+-----+-----+-----+
      L V A G V P A N L T C R S R G D A R P T
661  CCTGAATTGCTGTGGTTCCGAGATGGGTCTCTGTTGGATGGAGCCACCTTCCATCAGACC 720
   -----+-----+-----+-----+-----+-----+
      P E L L W F R D G V L L D G A T F H Q T

```

WO 03/000021

23/38

PCT/US02/20275

FIGURE 5 (cont.)

721 CTGCTGAAGGAAGGGACCCCTGGGTCAGTGGAGAGCACCTTAACCCCTGACCCCTTTTCAGC 780  
L L K E G T P G S V E S T L T L T P F S

781 CATGATGATGAGCCACCTTTGCTGCGGGGCGGAGCCAGGCCCTGCCACAGGAAGA 840  
H D D G A T F V C R A R S Q A L P T G R

841 GACACAGCTATCACACTGAGCCTGCAGTACCCCCAGAGGTGACTCTGTCTGCTTCGCCA 900  
D T A I T L S L Q Y P P E V T L S A S P

901 CACACTGTGCGAGGGGAGAGAAGTCAITTTCTGTGCGGCGCACAGCCAGCCTCCT 960  
H T V Q E G E K V I F L C Q A T A Q P F

961 GTCACAGGCTACAGGTGGGCAAAGGGGGCTCTCCGGTGTCTGGGGCCCGGGGCCAAGG 1020  
V T G Y R W A K G G S F V L G A R G F R

1021 TTAGAGGTCGTGGCAGAGCCTCCTTCCTGACTGAGCCCGTCTCTGCGAGSTCAGCAAC 1080  
L E V V A D A S F I L T E P V S C E V S N

1081 GCCGTGGGTAGCGCCAACCGCAGTACTGCGCTGGATGTGCTTTGGGCCGATTCCTGCAG 1140  
A V G S A N R S T A L D V L F G P I L Q

1141 GCNAAGCCGAGCCCGCTGTCCGTGGACGTGGGGGAAGACGCTTCCTCAGCTGCGCCTGG 1200  
A K P E P V S V D V G E D A S F S C A W

1201 CGCGGGAAACCCGCTTCACGGGTAACTGGACCCCGCGGTGGCGCTCAGGTGCTGGGC 1260  
R G N P L F R V T W T R R G G A Q V L G

1261 TCTGGAGCCACACTGCGCTTCCTCGTGGTGGGGCCGAGGACGAGGCGACTATGTGTGC 1320  
S G A T L R L P S V G P E D A G D Y V C

1321 AGAGCTGAGGCTGGGCTATCGGGCCTGCGGGGCGGCGCGGAGGCTCGGCTGACTGTG 1380  
R A E A G L S G L R G G A A E A R L T V

FIGURE 5 (cont.)

1381 AACGCTCCCCAGTAGTGACCGCCCTGCACTCTGCGCCTGCCCTTCTGAGGGGCCCTGCT 1440  
N A P P V V T A L H S A P A F L R G P A

1441 CGCTCCAGTGTCTGTTTTTCGCTCTCCCGCCAGATGCCGTGGTCTGTCTTGGGAT 1500  
R L Q C L V F A S E A P D A V V W S W D

1501 GAGGGCTTCTGAGGCGGGGTCCGAGGCGCGTTCTGTGGAGACATTCCCTGCCCA 1560  
E G F L E A G S Q G R F L V E T F P A P

1561 GAGAGCGCGGGGACTGGGTCCGGGCTGATCTCTGTCTACACATTTGGGGACCCAG 1620  
E S R G G L G F G L I S V L H I S G T Q

1621 GAGTCTGACTTTAGCAGGAGCTTTAACTGCACTGCCGGAACCGGCTGGCGAGGGAGT 1680  
E S D F S R S F N C S A R N R L G E G G

1681 GCCCAGGCCAGCCTGGGCGCTAGAGACTTGTGCCACTGTGCGGATAGTGGCCGAGTG 1740  
A Q A S L G R R D L L P T V R I V A G V

1741 GCGCTGCCACCACTCTCTTATGGTCATCACTGGGGTGGCCCTCTGCTGCTGCC 1800  
A A A T T T L L M V I T G V A L C C W R

1801 CACAGCAAGGCCTCAGCTCTTTCTCCGAGCAAGAAGCTGATGCGAATCCCTGGCAGC 1860  
H S K A S A S F S E Q K N L M R I P G S

1861 AGCGACGGCTCCAGTTACGAGGTCCTGAAGAAGAGGAGACAGCGCGGAGGACCGG 1920  
S D G S S S R G P E E E E T G S R E D R

1921 GGGCCCATTTGCACACTGACCACAGTGATCTGGTTCTGGAGGAGGAGGACTCTGGAG 1980  
G F I V H T D H S D L V L E E E G T L E

1981 ACCAAGGACCCCAACGATTACTACAAGGTCGAGGAGTCACTGTGAGCCTGAGCCTT 2040  
T K D P T N G Y Y K V R G V S V S L S L

FIGURE 5 (cont.)

2041 GCGAAGCCCTGGAGGAGTCTCTCTCTGCCACACCTCCCCCTTGGCCCCCAGGG 2100  
G E A F G G G L F L P P P S P L G P P G

2101 ACCCTACCTTCTATGACTTCAACCCACACCTGGGCACTCCCCCTGCAGACTTTAC 2160  
T P T F Y D F N P H L G M V P P C R L Y

2161 AGAGCCAGGGCAGGCTATCTCACCACACCCACCTCGAGCTTTCAGGCTACATCAA 2220  
R A R A G Y L T T P H P R A F T S Y I K

2221 CCCACATCCTTTGGCCCCCAGACTCTGGCCCCGGGACTCCCCCTTCCATATGCTGCC 2280  
P T S F G P P D L A P G T P P P Y A A

2281 TTCCCCACACTAGCCACCCGGCTCTCCAGACTCAGGTGTGACATCTTCCAATGGAAGA 2340  
F P T P S H P R L Q T H V

2341 GTCTTGGGATCTCCAACTTGCATATGGATTGTCTGATTCTGAGGAGCCAGGACAAG 2400

2401 TTGGCGACCTTACTCTCCAAACTGAACACAAGGGGAGGGAAGATCATTACATTTGTC 2460

2461 AGGAGCAITTTATACAGTCAGCTCAGCCAAAGGAGATGCCCCAAGTGGGAGCAACATGG 2520

2521 CCACCCAATATGCCACCTATTCCCCGGGTGTAAGAGATTCAGATGGCAGGTAGGCCC 2580

2581 TTTGAGGAGAGATGGGACAGGGCACTGGGTCTTGGGAGTTTGGGCGGGATGAAAGTT 2640

2641 GTTCTAGCCACTGAAAGAGATATTCAAGATGACCATCTGCATTGAGAGGAAGGTAG 2700

2701 CATAGGATAGATGAAGATGAAGAGCATACAGGCCCCACCTGCTCTCCCTGAGGGGAA 2760

2761 CTTTGCTGGCCAAATGGAATGCAGCCAAGATGGCCATATACTCCCTAGGAACCAAGAT 2820  
-----+-----+-----+-----+-----+-----+-----  
2821 GGCCACCATCTTGTATTTACTTTTCITTAAGACACAGAAGACTTGGACCCCAAGGAGTGG 2880  
-----+-----+-----+-----+-----+-----+-----  
2881 GGATACAGTGAGAATTACCAC TGTGGGGCAAATATTGGGATAAAAAATATTATGTTTA 2940  
-----+-----+-----+-----+-----+-----+-----  
2941 ATATAAAAAAAAAAGTCAA 2959

WO 03/000021

27/38

PCT/US02/20275

FIGURE 6

1 ATGCGGGTCCCGCCCTCTCGTCTCTCTTCTGCTTCAGAGGGAGAGCCAGCCCGCTCG 60  
M R V P A L L V L L F C F R G R A G P S

61 CCCCATTTCTGCAACAGCCAGAGACCTGGTGGTGGTGGTGGGGAGAGCCCGGCTG 120  
P H F L Q Q F E D L V V L L G E E A R L

121 CCGTGTGCTCTGGGGCCCTACTGGGGCTAGTTCAAGTGGACTAAGAGTGGGCTGGCCCTA 180  
P C A L G A Y W G L V Q W T K S G L A L

181 GGGGGCCAAAGGGACCTACAGGGTGGTCCCGGTACTGGATATCAGGGAATGCAGCCAAT 240  
G G Q R D L F G W S R Y W I S G N A A N

241 GGCCAGCATGACCTCCACATTAGGCCCGTGGAGCTAGAGGATGAAGCATCATATGAATGT 300  
G Q H D L H I R P V E L E D E A S Y E C

301 CAGGCTACACAAGCAGGCTCCCGCTCCAGACCAGCCCAACTGCACGTGCTGGTCCCCCA 360  
Q A T Q A G L R S R P A Q L H V L V P F

361 GAAGCCCCCAGGTGCTGGGGCCCGCTCTGTCTCTGTTGCTGGAGTTCTGCGAAG 420  
E A P Q V L G G P S V S L V A G V P A N

421 CTGACATGTCGGAGCCGTGGGGATGCCCGCCCTACCCCTGAATTGCTGTGGTTCCGAGAT 480  
L T C R S R G D A R P T F E L L W F R D

481 GGGGTCTGTTGGATGGAGCCACCTTCATCAGACCCTGCTGAAGGAAGGACCCCTGGG 540  
G V L L D G A T F H Q T L L K E G T F G

541 TCAGTGGAGAGCACCTTAACCCCTGACCCCTTTTCAGCCATGATGATGGAGCCACCTTTGTC 600  
S V E S T L T L T P F S H D D G A T F V

601 TGCCGGGCCCCGAGCCAGCCCTGCCCCACAGGAGAGACACAGCTATCAGACTGAGCCTG 660  
C R A R S Q A L P T G R D T A I T L S L

661 CAGTACCCCCCAGAGGTGACTCTGTCTGCTTCGCCACACACTGTGCAGGAGGAGAGAG 720  
Q Y F P E V T L S A S P H T V Q E G E K

FIGURE 6 (cont.)

721 GTCAATTTCTCTGTCAGGCCACAGCCAGCCCTCTCTACAGGCTACAGGTGGGCAAAA 780  
V I F L C Q A T A Q P P V T G Y R W A K

781 GGGGGCTCTCCGGTCTCGGGCCCCGGGGCAAGGTTAGAGTCTGGCAGACGCTCG 840  
G G S P V L G A R G P R L E V V A D A S

841 TTCTGACTGAGCCCTGTCTCTGCGAGGTACGCAACGCCGTGGGTAGCGCAACCGCAGT 900  
F L T E F V S C E V S N A V G S A N R S

901 ACTGGCTGGATGTCTCTTTGGGCGATTCTGCAGGCAAGCCGGAGCCCTGTCCCTG 960  
T A L D V L F G P I L Q A K P E P V S V

961 GACGTGGGGGAAGACGCTTCTTCAGCTGCTGCTGGCGCGGAACCGCTTCCACGGTA 1020  
D V G E D A S F S C A W R G N P L F R V

1021 ACCTGACCCGCCCGGGTGGGCTCAGGTGCTGGGCTCTGGAGCCACACTGCTCTTCG 1080  
T W T R R G G A Q V L G S G A T L R L P

1081 TCGTGGGGCCCAGGACGCGAGGCGACTATGTGTCAGAGCTGAGGCTGGGCTATCGGGC 1140  
S V G P E D A G D Y V C R A E A G L S G

1141 CTGCGGGGGCGCCCGGGAGGCTCGGCTGACTGTGAACGCTCCCCAGTAGTGACCGCC 1200  
L R G G A A E A R L T V N A P P V V T A

1201 CTGCACTCTGCGCTGCTTCTGAGGGGGCCCTGCTCGCTCCAGTGTCTGTTTTGCGC 1260  
L H S A P A F L R G P A R L Q C L V F A

1261 TCTCCCGCCCGAGATGCCGTGGTCTGGTCTTGGGATGAGGGCTTCTGGAGGGGGGTG 1320  
S P A F D A V V W S W D E G F L E A G S

1321 CAGGGCGGGTTCTGTTGGAGACATTCCTGCCCCAGAGAGCCGCGGGGACTGGGCTCG 1380  
Q G R F L V E T F F A P E S R G G L G P

1381 GGCCTGATCTCTGTGCTACACATTTCGGGACCCAGGAGTCTGACTTAGCAGGAGCTTT 1440  
G L I S V L H I S G T Q E S D F S R S F



FIGURE 6 (cont.)

```
1441 AACTGCAGTSCCCGGARCCGGCTGGGCGAGGGAGTSCCCAGGCCAGCCTGGGCCGTAGA 1500
-----+-----+-----+-----+-----+
      N C S A R N R L G E G G A Q A S L G R R

1501 GACTTGCTGCCCACTGTGCGGATAGTGGCCGGAGTGGCCGCTGCCACCAACTCTCCTT 1560
-----+-----+-----+-----+-----+
      D L L F T V R I V A G V A A T T T L L

1561 ATGGTCATCAGTGGGGTGGCCCTCTGCTGCTGGCGCCACAGCAAGGCCCTCAGCCTCTTTC 1620
-----+-----+-----+-----+-----+
      M V I T G V A L C C W R H S K A S A S F

1621 TCCGAGCAAAAGAACCTGATGCGAATCCCTGGCAGCAGCGAGCGCTCCAGTTCACGAGGT 1680
-----+-----+-----+-----+-----+
      S E Q K N L M R I P G S S D G S S S R G

1681 CCTGAGAGAGAGGAGACAGGCGAGCCGCGAGGACCGGGGCCCCATTGTGCACACTGACCAC 1740
-----+-----+-----+-----+-----+
      P E E E E T G S R E D R G F I V H T D H

1741 AGTGATCTGGTTCTGGAGGAGGAAGGACTCTGGAGACCAAG 1782
-----+-----+-----+-----+-----+
      S D L V L E E E G T L E T K
```

ACCTTGGGGGACGAATGCTCAGGATGGGCTGCCCGGCTCTCTGCTCTCTCTTCTGT  
M L R M R V P A L L V L L F C F  
TCA GAGGGAGAGCAGGCGCGCTGCCCAATTCTCTCAACAGCCAGGAGACCTGGTGTGC  
E G R A G P S P H F L Q Q P E D L V L V  
TGTCTGGGAGAGGACCGGCTGCCCTGTGCTCTCTGGCGCTTCTGGGGCTAGTTCAGT  
L G E E A R L P C A L G A Y W G L V Q W  
GGACTAAGAGTGGCTGGCCCTAGGGGCGCAAAAGGGACATCACAGGCTGGTCCCGGTACT  
T K S G L A L G G Q R D L P G W S R Y W  
GGATATCAGGAAGTCAGCCAAATGGCAGCATGACCTCCACATTAGGCGCGCTGGAGCTAG  
I S G N A A T G G Q H D L H R I F V E L E  
AGGATGAAGCATCATATGAATGTGAGCTACACAAGCAGGCGCTCCGCTCCAGACGAGCC  
D E A S Y E C Q A T Q A G L R S R P A Q  
AATGCAGCGTGTGTGCCCCCAGAGACCCCCCAGGTGTGGCGCGCCCTCTGTGTCTC  
L H V L V P P E A P Q V L G G P S V S L  
TGGTTGCTGGAGTCTCTGGGAACTGACATGTGGAGCGGTGGGAGTACGCCCGCTACCC  
G A V F P A N T C R S G D A R T P T P  
CTGAATTGCTGTGTCCAGAGTGGGCTCGTGTGGATGAGCCACCTTCCATCAGACC  
E L L W F R D G V L L D G A T F H Q T L  
TGTCTAAGGAAGGAGCCCTGGGTCACTGGAGAGCACTTAACCTTGACCCCTTCAGCC  
L K E G T P G S V S T L T L T P F S H  
ATGATGATGGAGCACCTTGCTGTGCCGGCCGGAGCGAGCGCTCCGCAAGAGAGAG  
D D G A T L V C R A S G S Q A L P T G R D  
ACACAGCTATCACATGAGCTGTCAGTACCCCCAGAGTGACTGTGTCTGTCTGGCAC  
T A I T L S L Y T P P E T L T S A S P H  
ACACTGTGCAGGAGGAGAGAAGGTCAATTCTCTGTGCAGGCAAGCCAGGCTCTGCT  
T V Q E G E K V I F L C Q A T A Q A P V T  
TCACAGGCTACAGGTGGGCAAAAGGGGCTCTCGGTGTCTGGGGCGCGCGGGCAAGG  
T G Y R W A K G G S P V L G A R G P R L  
TAGAGGCTGTGGCAGACGCTCGTCTCTGACTGAGCCCGTGTCTCGGAGGTCAGCAAG  
E V V A D A S F L T E P V S C E Y S N A  
CCGTGGTAGCGCCCAACCGCACTAGTGGCTGGATGTGCTGTTTGGCGGATCTCGAGAG  
V G S A C T A L D L V L G F G P I L Q A  
CAAAGCCGAGGCGGTGTCCGTGGACGTGGGGGAGAGCGCTCTCTCTGAGTGGCGTGGC  
K P E P V S V D V G E D A S F S C A W R

FIGURE 7 (cont.)

GCAGGAATCCGCTTCCACGGGTAACTGGACCCGCGCGGTGGCGCGCAGGTGCTGGGCT  
G N P L P R V T W T E R G G A Q V L G S  
CTGGAGCCACACTGCGTCTTCCGTGGTGGGGCCCGAGGACGCGAGACGACTATGTGTGA  
G A T L R L P S V G P E D A D D Y V C R  
GAGCTGAGGCTGGGCTATCGGGCTGCGGGCGCGCGCGGAGGCTCGGCTGACTGTGA  
A E A G L S G L R G G A A E A R L T V N  
ACGCTCCCGCAGTAGTACGCGCCCTGCACTCTGGGCGTCCCTTCTGAGGGGCGCTGCTC  
A P P V V T A L H S A P A F L R G P A R  
GCCTCCAGTGTCTGGTTTTCGCTCTCCCGCCCGAGATGCCGTGGTCTGGTCTGGGATG  
L Q C L V F A S P A P D A V V W S H D E  
AGGGCTTCTCGGAGCGGGTCCGAGGTGCGGTTCTTGTGGAGACATTCCTGCGCCAG  
G F L E A G S Q G R F L V E T F F A P E  
AGAGCGCGGGGAGTGGTCCGGGCTGATCTCTGTGTACACATTCGGGGACCCAGG  
S R G G L G P G L I S V L H I S G T Q E  
AGTCTGACTTTAGCAGGAGCTTTAACTGCACTGCCCGAACCGCTGGGCGAGGGAGGTG  
S D F S R S F N C S A R N R L G E G G A  
CCGAGGCCAGCTGGGCGGTAGAGACTTGTGCCACTGTGGGATAGTGGCGGAGTGG  
Q A S L G R R D L L P T V R I V A S V V  
TCGCTGGCCACACAACCTCTCTTATGGTCACTGCGGCTGGCCCTCTGCTGCTGGCGCC  
A A T T F L L M V I T G V A L C C W R H  
ACAGCAAGGCTCAGCCTCTTCTCCGAGCAAAAGAACCTGATGCGAATCCCTGGCAGCA  
S K A S A S F S E Q K N L M R I P G S S  
GCGACGGCTCCAGTTACGAGGTCTGAAGAAGAGAGACAGCGAGCCGCGAGGACCCGG  
D G S S S R G P E E E E T G S R E D R G  
GCCCCATTGTGCACACTGACACAGTGATCTGGTTCTGGAGGAGGAGGAGTCTGGAGA  
P I V H T D H S D L V L E E E G T L E T  
CCAAGGACCCCAACCAACGTTACTACAAGGTCCGAGGAGTCACTCCACCGCGTCTCCAG  
K D P T N G Y Y K V R G V S P P A S P D  
ACTCAAGTGTGACATCTTCCAATGGAAGAGTCTGGGATCTCCAACCTGCCATAATGGA  
S R V T S F Q W K S P G I S N L P \*  
TTGTTCTGATTCTGAGGAGCCAGGACAAGTTGGCGACCTTACTCTCC

[illegible]

|                                                                                                         |      |
|---------------------------------------------------------------------------------------------------------|------|
| 1321                                                                                                    | 1380 |
| L G E G G A Q A S L G R R D L L P T V R<br>ATAGTGGCCGGAGTGGTGGCTGCCACCACAACTCTCTTATGTGTCATCACTGGGGTGGCC |      |
| 1381                                                                                                    | 1440 |
| I V A G V V A A T T T L L M V I T G V A<br>CTCTGCTGCTGGCCACAGCAAGCTCAGCTCTTTCTCCGACAGAGAGAGCTGATG       |      |
| 1441                                                                                                    | 1500 |
| L C C W R H S K A S A S F S E Q K N L M<br>CGAATCCCTGGCAGCAGCGACGGCTCCAGTTCACGAGGTCTCGAAGAAGAGGAGACAGGC |      |
| 1501                                                                                                    | 1560 |
| R I P G S S D G S S S R G P E E E E T G<br>AGCCCGAGGACCGGGGCCCATTTGGCACACTGACCAAGTGATCTGGTTCTTGAGAGAG   |      |
| 1561                                                                                                    | 1620 |
| S R E D R G P I V H T H S D L V L E E<br>GAAGGAGACTCTGGAGACCAGGAGCCCAACCAAGCGTTACTACAAGTCCGAGGAGTCAGT   |      |
| 1621                                                                                                    | 1680 |
| E G T L E T K D F T N G Y V K R G V S<br>CACCCCGCGTCTCCAGACTCAGTGTGACATTTCCAATGGAGAGACTCTGGGATCTCC      |      |
| 1681                                                                                                    | 1740 |
| P P A S P D S R V T S G W K S P G I S<br>AACTTGCCTAATGATGTTCGATTTCTGAGAGCCAGGACAAGTTGGCGACTTCTACT       |      |
| 1741                                                                                                    | 1800 |
| N L P * W I V L I S E E P G Q V G D L T<br>CTCCTC                                                       |      |
| 1801                                                                                                    |      |
| P                                                                                                       |      |

WO 03/000021

34/38

PCT/US02/20275

FIGURE 9A

|          |                                                                    |     |
|----------|--------------------------------------------------------------------|-----|
|          | 1                                                                  | 60  |
| PanCAM-1 | MLMRVPALL VLLFCFRGRA GPSPHLQOP EDLVVLLGEE ARLPCALGAY WGLVQWTKSG    |     |
| PanCAM-2 | MLMRVPALL VLLFCFRGRA                                               |     |
| PanCAM   | MLMRVPALL VLLFCFRGRA GPSPHLQOP EDLVVLLGEE ARLPCALGAY WGLVQWTKSG    |     |
| bc007312 | MLMRVPALL VLLFCFRGRA GPSPHLQOP EDLVVLL                             |     |
|          | 61                                                                 | 120 |
| PanCAM-1 | LALGGQRDLG GMSRYWISGN AANGQHDHLI RPVELEDEAS YECQATQAGL RSRPAQLHVL  |     |
| PanCAM-2 | GMSRYWISGN AANGQHDHLI RPVELEDEAS YECQATQAGL RSRPAQLHVL             |     |
| PanCAM   | LALGGQRDLG GMSRYWISGN AANGQHDHLI RPVELEDEAS YECQATQAGL RSRPAQLHVL  |     |
| bc007312 |                                                                    |     |
|          | 121                                                                | 180 |
| PanCAM-1 | VPPEAPQVLG GPSVSLVAGV PANLTCRSRG DARPTPELLW FRDGVLLDGA TFHQTLKEG   |     |
| PanCAM-2 | VPPEAPQVLG GPSVSLVAGV PANLTCRSRG DARPTPELLW FRDGVLLDGA TFHQTLKEG   |     |
| PanCAM   | VPPEAPQVLG GPSVSLVAGV PANLTCRSRG DARPTPELLW FRDGVLLDGA TFHQTLKEG   |     |
| bc007312 |                                                                    |     |
|          | 181                                                                | 240 |
| PanCAM-1 | TPGSVESTLT LTPFSDHDDGA TLVCRARSQA LPTGRDTAIT LSLQYPPVET LSASPHTVCE |     |
| PanCAM-2 | TPGSVESTLT LTPFSDHDDGA TLVCRARSQA LPTGRDTAIT LSLQYPPVET LSASPHTVCE |     |
| PanCAM   | TPGSVESTLT LTPFSDHDDGA TLVCRARSQA LPTGRDTAIT LSLQYPPVET LSASPHTVCE |     |
| bc007312 |                                                                    |     |
|          | 241                                                                | 300 |
| PanCAM-1 | GERVIFLQGA TAQPPVTGYR WAKGGSFVLG ARGPRELVA DASFLTEPVS CEVSNVAGSA   |     |
| PanCAM-2 | GERVIFLQGA TAQPPVTGYR WAKGGSFVLG ARGPRELVA DASFLTEPVS CEVSNVAGSA   |     |
| PanCAM   | GERVIFLQGA TAQPPVTGYR WAKGGSFVLG ARGPRELVA DASFLTEPVS CEVSNVAGSA   |     |
| bc007312 |                                                                    |     |
|          | 301                                                                | 360 |
| PanCAM-1 | NRSTALDVLG GPILQAKPEP VSDVVGEDAS FSCANRGNPL PRVTWTRGG AQVLGGGATL   |     |
| PanCAM-2 | NRSTALDVLG GPILQAKPEP VSDVVGEDAS FSCANRGNPL PRVTWTRGG AQVLGGGATL   |     |
| PanCAM   | NRSTALDVLG GPILQAKPEP VSDVVGEDAS FSCANRGNPL PRVTWTRGG AQVLGGGATL   |     |
| bc007312 |                                                                    |     |
|          | 361                                                                | 420 |
| PanCAM-1 | RLPSVGPEDA DDYVCRAEAG LSGLRGGAAE ARLTVNAPPV VTALHSAPAF LRGPARIQCL  |     |
| PanCAM-2 | RLPSVGPEDA DDYVCRAEAG LSGLRGGAAE ARLTVNAPPV VTALHSAPAF LRGPARIQCL  |     |
| PanCAM   | RLPSVGPEDA DDYVCRAEAG LSGLRGGAAE ARLTVNAPPV VTALHSAPAF LRGPARIQCL  |     |
| bc007312 |                                                                    |     |
|          | 421                                                                | 480 |
| PanCAM-1 | VFASPADAV VNSWDEGFLE AGSQGRFLVE TFPAPESRGG LGPGLISVLH ISGTQESDFS   |     |
| PanCAM-2 | VFASPADAV VNSWDEGFLE AGSQGRFLVE TFPAPESRGG LGPGLISVLH ISGTQESDFS   |     |
| PanCAM   | VFASPADAV VNSWDEGFLE AGSQGRFLVE TFPAPESRGG LGPGLISVLH ISGTQESDFS   |     |
| bc007312 |                                                                    |     |
|          | 481                                                                | 540 |
| PanCAM-1 | RSFNCSARNR LGEGBGAQASL GRDDLPTVR IVAGVVAATT TLLMVITGVA LCCWRHSKAS  |     |
| PanCAM-2 | RSFNCSARNR LGEGBGAQASL GRDDLPTVR IVAGVVAATT TLLMVITGVA LCCWRHSKAS  |     |
| PanCAM   | RSFNCSARNR LGEGBGAQASL GRDDLPTVR IVAGVVAATT TLLMVITGVA LCCWRHSKAS  |     |
| bc007312 | LGEGBGAQASL GRR.....AS                                             |     |
|          | 541                                                                | 600 |
| PanCAM-1 | ASFSEQKNLM RIPGSSDGSS SRGPEEEETG SREDRGPVH TDHSDLVLEE EGTLETKDPT   |     |
| PanCAM-2 | ASFSEQKNLM RIPGSSDGSS SRGPEEEETG SREDRGPVH TDHSDLVLEE EGTLETKDPT   |     |
| PanCAM   | ASFSEQKNLM RIPGSSDGSS SRGPEEEETG SREDRGPVH TDHSDLVLEE EGTLETKDPT   |     |
| bc007312 | ASFSEQKNLM RIPGSSDGSS SRGPEEEETG SREDRGPVH TDHSDLVLEE EGTLETKDPT   |     |
|          | 601                                                                | 660 |
| PanCAM-1 | NGYKVRGVS PPASPDSTVT SFQWKSFGIS NLP.....                           |     |
| PanCAM-2 | NGYKVRGVS PPASPDSTVT SFQWKSFGIS NLP.....                           |     |

FIGURE 9A (cont.)

|          |            |            |           |            |            |            |
|----------|------------|------------|-----------|------------|------------|------------|
| PanCAM   | NGYKVRGVS  | VSLSLGEAPG | GGLFLPPSP | LGPPGTPTFY | DFNPHLGMVP | PCRLYRARAG |
| bc007312 | NGYKVRGVS  | VSLSLGEAPG | GGLFLPPSP | LGPPGTPTFY | DFNPHLGMVP | PCRLYRARAG |
|          | 661        |            |           |            | 708        |            |
| PanCAM-1 | .....      | .....      | .....     | .....      | .....      | .....      |
| PanCAM-2 | .....      | .....      | .....     | .....      | .....      | .....      |
| PanCAM   | YLTTPHPRAF | TSYIKPTSPG | PDLAPGTPP | PPYAAFTPS  | HPRLQTHV   |            |
| bc007312 | YLTTPHPRAF | TSYIKPTSPG | PDLAPGTPP | PPYAAFTPS  | HPRLQTHV   |            |

FIGURE 9B

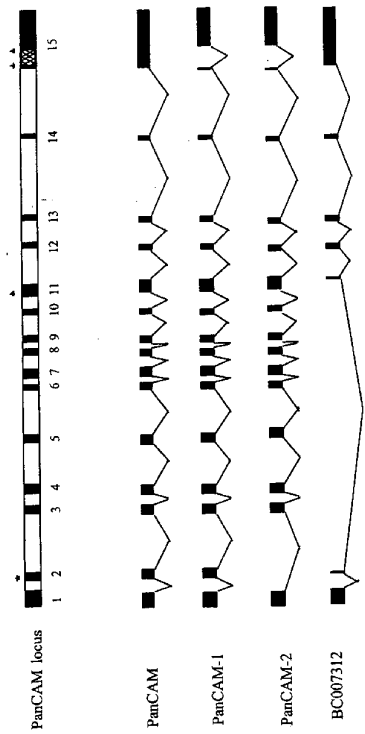




FIGURE 10A

```
1  ACTGGGGGCTAGTTTCAGTGGACTAAGGATGGCTGGCTCTAGGGGGCGAAAGACCTTC 60
   W G L V Q W T K D G L A L G G E R D L P
61  CAGGGTGGTCCCGGTACTGGATATCGGGGAATTCAGCCAGTGGCCAGCATGACCTCCACA 120
   G W S R Y W I S G N S A S G Q H D L H I
121  TTAAGCCTGTGGAATTGGAAGATGAGGCATCGTATGAGTCCAGGCCTTCGCAAGCAGGTC 180
   K P V E L E D E A S Y E C Q A S Q A G L
181  TCCGATCACGACAGCCCACTGCACGTGATGTTCTCGGAAATCTGACCTGTGGAGTC 240
   R S R P A Q L H V M V P G H L T C R S R
241  GTGGGATTCCTGACCTGCCCTGAACTACTGTGGTTCGAGATGGGATCCGGCTGGATG 300
   G D S R P A P E L L W F R D G I R L D A
301  CGAGCAGCTTCACCAGACCAGCTGAAGGACAAGGCCACTGGAAACAGTGGAAACACCT 360
   S S F H Q T T L K D K A T G T V E N T L
361  TATTCCTGACCCCTTCAGTCATGATGATGGCCACCTTGATCTGCGAGAGCGGAGGCC 420
   F L T P S S H D D G A T L I C R A R S Q
421  AGGCCCTGCCACAGGGAGGGACACAGCTGTTACACTGAGCCTTCAGTATCCCCCAATGG 480
   A L P T G R D T A V T L S L Q Y P P M V
481  TGACTCTGTCTGCTGAGCCCGAGACTGTGAGGAGGAGAGAGGTGACTTTCCTGTGTGTC 540
   T L S A E P Q T V Q E G E K V T F L C Q
541  AAGCCACTGCCAGCCTCCTGTCACTGGCTACAGTGGGCGAAGGGGGGATCCCCGTGC 600
   A T A Q P P V T G Y R W A K G G S P V L
601  TCGGGGCACGTGGCCCAAGTTGGAGTCTGTGCAGATGCCACTTTCCTGACTGAGCCGG 660
   G A R G P R L E V V A D A T F L T E P V
661  TGTCTGCGAGGTGAGCAACGGTGGAGAGGCGCAACCGCAGCACGGCGCTGGAGTGT 720
   S C E V S N A V G S A N R S T A L E V L
721  TGTATGAGCCATTCTGCAGGCAAAACCTAAGTCCGTGTCCGTGGACGTGGGGAAGATG 780
   Y G P I L Q A K P K S V S V D V G F D A
781  CCTCTTCAGCTGTGTGCGCGGGGAACCTTCCACGGATAACCTGGACCGCGATGG 840
   S F S C V W R G N P L P R I T W T R M G
841  GTGGCTCTCAGGTGCTGAGCTCCGGGCCACGCTGCGGCTTCGGTCCGTGGCACTGGAGG 900
   G S Q V L S S G P T L R L P S V A L E D
901  ATGCGGGCGACTATGTGTGAGAGCTGAGGCTGA 934
   A G D Y V C R A E A
```

|     |                                                     |     |
|-----|-----------------------------------------------------|-----|
| 51  | WGLVQMTKSGLAGQGRDLPGWSYVTSIGNAANGQDLHTRFVPELEDAAS   | 100 |
| 1   | WGLVQMTKSGLAGQGRDLPGWSYVTSIGNAANGQDLHTRFVPELEDAAS   | 50  |
| 101 | YECQATQAGLSRRPQALHVLVPPAEQVQLGGPSSVLNAGVPAHLTRCSRSG | 150 |
| 51  | YECQASQAGLSRRPQALHVLVPPAEQVQLGGPSSVLNAGVPAHLTRCSRSG | 81  |
| 151 | DARPTPELLWFRDGVLLDGASTFQTLIKBGTGSGVESTLTLPSPSHDDGA  | 200 |
| 62  | DSRPQELLWFRDGVLLDGASTFQTLIKBGTGSGVESTLTLPSPSHDDGA   | 131 |
| 201 | TLICRARSQALPGTRDATITLSLQYPPVETLSASPHTPQBGKVLFLCQA   | 250 |
| 132 | TLICRARSQALPGTRDATITLSLQYPMVLTASAPTPQBGKVLFLCQA     | 181 |
| 251 | TAQPPVTGYRWAKGSGFVLGARGPRLVEVDATFLTEPVSCEVNAVSGA    | 300 |
| 182 | TAQPPVTGYRWAKGSGFVLGARGPRLVEVDATFLTEPVSCEVNAVSGA    | 231 |
| 301 | NRSTALDVLPGFLIQAQKPEPVSVDVGEDASFSCAWRGNPLRVITWRKGG  | 350 |
| 232 | NRSTALEVLKGPLIQAQKPPSVSDVGEDASFCSCWVGWNPRLRTWTRKGG  | 281 |
| 379 | AGVLGSGATLRLPSVGPEDASDVTVCRAEA                      | 379 |
| 310 | SQVLSGGFTLRLPSVVALEADGTVTCRAEA                      | 310 |

PCT/US02/2075

|                                          |           |                                                                   |
|------------------------------------------|-----------|-------------------------------------------------------------------|
| Applicant's or agent's<br>file reference | A097 PCT2 | International Application No. JP 2005-508147 A<br>New Application |
|------------------------------------------|-----------|-------------------------------------------------------------------|

INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 135is)

|                                                                                                                                                                                                                                                                              |                                                                                                                                             |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| A. The indications made below relate to the deposited microorganism or other biological material referred to in the description<br>on page <u>page 6</u> , line <u>lines 1-10</u>                                                                                            |                                                                                                                                             |
| B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>                                                                                                                                                      |                                                                                                                                             |
| Name of depositary institution<br>American Type Culture Collection (ATCC)                                                                                                                                                                                                    |                                                                                                                                             |
| Address of depositary institution (including postal code and country)<br>10801 University Boulevard<br>Manassas, Virginia 20110-2209<br>United States of America                                                                                                             |                                                                                                                                             |
| Identification Reference by Depositor: FL86-3                                                                                                                                                                                                                                |                                                                                                                                             |
| Date of deposit<br>11 June 2002                                                                                                                                                                                                                                              | Accession Number                                                                                                                            |
| C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>                                                                                                                                      |                                                                                                                                             |
| D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)                                                                                                                                                                   |                                                                                                                                             |
| E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)<br>The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")<br>Accession Number of Deposit |                                                                                                                                             |
| For receiving Office use only<br><input checked="" type="checkbox"/> This sheet was received with the international application<br>Authorized officer<br>Sandra D. Barnes<br>PCT/Interna'l Appl Processing Div<br>(703) 305-3665                                             | For International Bureau use only<br><input type="checkbox"/> This sheet was received by the International Bureau on:<br>Authorized officer |

Form PCT/RQ/134 (July 1998)

PCT/US02/20275

Applicants : BIOGEN, INC., et al.  
Filed : Concurrently Herewith  
Title : PANCAM NUCLEIC ACIDS AND  
POLYPEPTIDES  
Agent's Ref. : A097 PCT2

Indications Relating to the Expert Solution in  
Respect of Deposited Biological Material  
Referred to in the Description

The indications relating to deposited biological material are all contained in the description. The following additional indications are not required to be part of the description and should be treated as "separate indications." They relate only to the expert solution.

The additional indications made below relate to the deposited biological material referred to as FL86-3 in the description on

page 6, lines 1-10.

The deposit was made in:

American Type Culture Collection (ATCC)  
10801 University Boulevard  
Manassas, Virginia 20110-2209  
United States of America

on June 11, 2002, under Accession Number \_\_\_\_\_.

The additional indications are:

For CA (Canada) designation:

In respect of the designation of Canada, samples of the deposited biological material will be made available until the grant of the Canadian patent, or until the date on which the application is refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, as provided in Rules 107 and 108 of the Patent Rules under the Canadian Patent Act, only by

PCT/US02/20275

- 2 -

the issue of a sample to an independent expert nominated by the Commissioner (Rule 104(4)).

For EP (European Patent) designation:

In respect of the designation of the EPO, samples of the deposited biological material will be made available until the publication of the mention of the grant of the European patent, or until 20 years from the date of filing if the application is refused or withdrawn or is deemed to be withdrawn, as provided in Rule 28(3) of the Implementing Regulations under the EPC, only by the issue of a sample to an expert nominated by requester (Rule 28(4) EPC).

For FI (Finland) designation:

In respect of the designation of Finland, until the publication of the mention of grant by the National Board of Patents and Registration, or for 20 years from the date of filing if the application has been finally decided upon without resulting in the grant of a patent by the National Board of Patents and Registration, the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For GB (United Kingdom) designation:

In respect of the designation of the United Kingdom, the applicant(s) hereby give(s) notice of my/our intention that the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For IS (Iceland) designation:

In respect of the designation of Iceland, until a patent has been granted by the Icelandic Patent Office, or a final decision taken by the Icelandic Patent Office if the application has been finally decided upon without resulting in the grant of a patent, the furnishing of samples of the deposited biological material will only be effected to an expert in the art.

PCT/US02/20275

- 3 -

For SE (Sweden) designation:

In respect of the designation of Sweden, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For SG (Singapore) designation:

Applicant(s) hereby give notice of my/our intention that samples of the above-identified culture shall be available only to experts in accordance with paragraph 3 of the Fourth Schedule to the Patents Rules 1995.

PCT/US02/2075

|                                          |           |                               |                 |
|------------------------------------------|-----------|-------------------------------|-----------------|
| Applicant's or agent's<br>file reference | A097 PCT2 | International Application No. | New Application |
|------------------------------------------|-----------|-------------------------------|-----------------|

INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

|                                                                                                                                                                                                                                                                              |                                                                                                        |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| A. The indications made below relate to the deposited microorganism or other biological material referred to in the description<br>on page _____, lines _____, line _____, lines 25-27 and 10-11 respectively                                                                |                                                                                                        |
| B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>                                                                                                                                                      |                                                                                                        |
| Name of depository institution<br>American Type Culture Collection (ATCC)                                                                                                                                                                                                    |                                                                                                        |
| Address of depository institution (including postal code and country)<br>10801 University Boulevard<br>Manassas, Virginia 20110-2209<br>United States of America                                                                                                             |                                                                                                        |
| Identification Reference by Depositor: CS0026                                                                                                                                                                                                                                |                                                                                                        |
| Date of deposit<br>11 June 2002                                                                                                                                                                                                                                              | Accession Number                                                                                       |
| C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>                                                                                                                                      |                                                                                                        |
| D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)                                                                                                                                                                   |                                                                                                        |
| E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)<br>The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")<br>Accession Number of Deposit |                                                                                                        |
| <input checked="" type="checkbox"/> This sheet was received with the international application<br>Authorized officer<br>Sonya D Barnes<br>PCT/Internat'l Appl Processing Div<br>(703) 305-3666                                                                               | <input type="checkbox"/> This sheet was received by the International Bureau on:<br>Authorized officer |

Form PCT/RO/134 (July 1998)

PCT/US02/20275

Applicants : BIOGEN, INC., et al.  
Filed : Concurrently Herewith  
Title : PANCAM NUCLEIC ACIDS AND  
POLYPEPTIDES  
Agent's Ref. : A097 PCT2

Indications Relating to the Expert Solution in  
Respect of Deposited Biological Material  
Referred to in the Description

The indications relating to deposited biological material are all contained in the description. The following additional indications are not required to be part of the description and should be treated as "separate indications." They relate only to the expert solution.

The additional indications made below relate to the deposited biological material referred to as CS0026 in the description on

page 10, lines 25-27; and  
page 101, lines 10-11.

The deposit was made in:

American Type Culture Collection (ATCC)  
10801 University Boulevard  
Manassas, Virginia 20110-2209  
United States of America

on June 11, 2002, under Accession Number \_\_\_\_\_.

The additional indications are:

For CA (Canada) designation:

In respect of the designation of Canada, samples of the deposited biological material will be made available until the grant of the Canadian patent, or until the date on which the application is refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, as provided in Rules 107 and 108 of



PCT/US02/20275

- 2 -

the Patent Rules under the Canadian Patent Act, only by the issue of a sample to an independent expert nominated by the Commissioner (Rule 104(4)).

For EP (European Patent) designation:

In respect of the designation of the EPO, samples of the deposited biological material will be made available until the publication of the mention of the grant of the European patent, or until 20 years from the date of filing if the application is refused or withdrawn or is deemed to be withdrawn, as provided in Rule 28(3) of the Implementing Regulations under the EPC, only by the issue of a sample to an expert nominated by requester (Rule 28(4) EPC).

For FI (Finland) designation:

In respect of the designation of Finland, until the publication of the mention of grant by the National Board of Patents and Registration, or for 20 years from the date of filing if the application has been finally decided upon without resulting in the grant of a patent by the National Board of Patents and Registration, the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For GB (United Kingdom) designation:

In respect of the designation of the United Kingdom, the applicant(s) hereby give(s) notice of my/our intention that the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For IS (Iceland) designation:

In respect of the designation of Iceland, until a patent has been granted by the Icelandic Patent Office, or a final decision taken by the Icelandic Patent Office if the application has been finally decided upon without resulting in the grant of a patent, the furnishing of samples of the deposited biological material will only be effected to an expert

PCT/US02/20275

- 3 -

in the art.

For SE (Sweden) designation:

In respect of the designation of Sweden, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For SG (Singapore) designation:

Applicant(s) hereby give notice of my/our intention that samples of the above-identified culture shall be available only to experts in accordance with paragraph 3 of the Fourth Schedule to the Patents Rules 1995.

PCT/US02/2075

|                                          |           |                               |             |                 |
|------------------------------------------|-----------|-------------------------------|-------------|-----------------|
| Applicant's or agent's<br>file reference | A097 PCT2 | International application No. | 5112/2002/5 | New Application |
|------------------------------------------|-----------|-------------------------------|-------------|-----------------|

INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

|                                                                                                                                                                                                                                                                              |                                                                                                                                                 |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| A. The indications made below relate to the deposited microorganism or other biological material referred to in the description<br>on page <u>page 11</u> , line <u>lines 1-7 and 10-11</u>                                                                                  |                                                                                                                                                 |
| B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>                                                                                                                                                      |                                                                                                                                                 |
| Name of depositary institution<br>American Type Culture Collection (ATCC)                                                                                                                                                                                                    |                                                                                                                                                 |
| Address of depositary institution (including postal code and country)<br>10801 University Boulevard<br>Manassas, Virginia 20110-2209<br>United States of America                                                                                                             |                                                                                                                                                 |
| Identification Reference by Depositor: CS0067-1                                                                                                                                                                                                                              |                                                                                                                                                 |
| Date of deposit<br>11 June 2002                                                                                                                                                                                                                                              | Accession Number                                                                                                                                |
| C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>                                                                                                                                      |                                                                                                                                                 |
| D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)                                                                                                                                                                   |                                                                                                                                                 |
| E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)<br>The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")<br>Accession Number of Deposit |                                                                                                                                                 |
| <input checked="" type="checkbox"/> For receiving Office use only<br>This sheet was received with the international application<br><br>Sonya D. Barnes<br>PCT/Internat'l Appl Processing Div<br>(703) 305-3666                                                               | <input type="checkbox"/> For International Bureau use only<br>This sheet was received by the International Bureau on:<br><br>Authorized officer |

Form PCT/RO/134 (July 1998)

PCT/US02/2075

Applicants : BIOGEN, INC., et al.  
Filed : Concurrently Herewith  
Title : PANCAM NUCLEIC ACIDS AND  
POLYPEPTIDES  
Agent's Ref. : A097 PCT2

Indications Relating to the Expert Solution in  
Respect of Deposited Biological Material  
Referred to in the Description

The indications relating to deposited biological material are all contained in the description. The following additional indications are not required to be part of the description and should be treated as "separate indications." They relate only to the expert solution.

The additional indications made below relate to the deposited biological material referred to as CS0067-1 in the description on

page 11, lines 1-7; and  
page 11, lines 10-11.

The deposit was made in:

American Type Culture Collection (ATCC)  
10801 University Boulevard  
Manassas, Virginia 20110-2209  
United States of America

on June 11, 2002, under Accession Number \_\_\_\_\_.

The additional indications are:

For CA (Canada) designation:

In respect of the designation of Canada, samples of the deposited biological material will be made available until the grant of the Canadian patent, or until the date on which the application is refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, as provided in Rules 107 and 108 of

PCT/US02/20275

- 2 -

the Patent Rules under the Canadian Patent Act, only by the issue of a sample to an independent expert nominated by the Commissioner (Rule 104(4)).

For EP (European Patent) designation:

In respect of the designation of the EPO, samples of the deposited biological material will be made available until the publication of the mention of the grant of the European patent, or until 20 years from the date of filing if the application is refused or withdrawn or is deemed to be withdrawn, as provided in Rule 28(3) of the Implementing Regulations under the EPC, only by the issue of a sample to an expert nominated by requester (Rule 28(4) EPC).

For FI (Finland) designation:

In respect of the designation of Finland, until the publication of the mention of grant by the National Board of Patents and Registration, or for 20 years from the date of filing if the application has been finally decided upon without resulting in the grant of a patent by the National Board of Patents and Registration, the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For GB (United Kingdom) designation:

In respect of the designation of the United Kingdom, the applicant(s) hereby give(s) notice of my/our intention that the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For IS (Iceland) designation:

In respect of the designation of Iceland, until a patent has been granted by the Icelandic Patent Office, or a final decision taken by the Icelandic Patent Office if the application has been finally decided upon without resulting in the grant of a patent, the furnishing of samples of the deposited biological material will only be effected to an expert

PCT/US02/20275

- 3 -

in the art.

For SE (Sweden) designation:

In respect of the designation of Sweden, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For SG (Singapore) designation:

Applicant(s) hereby give notice of my/our intention that samples of the above-identified culture shall be available only to experts in accordance with paragraph 3 of the Fourth Schedule to the Patents Rules 1995.

PCT/US02/2075

|                                          |           |                               |                 |
|------------------------------------------|-----------|-------------------------------|-----------------|
| Applicant's or agent's<br>file reference | A087 PCT2 | International Application No. | New Application |
|------------------------------------------|-----------|-------------------------------|-----------------|

INDICATIONS RELATING TO DEPOSITED MICROORGANISM  
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

|                                                                                                                                                                                   |                                                                                  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| A. The indications made below relate to the deposited microorganism or other biological material referred to in the description<br>on page <u>page 11</u> , line <u>lines 1-7</u> |                                                                                  |
| B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>                                                                      |                                                                                  |
| Name of depositary institution<br>American Type Culture Collection (ATCC)                                                                                                         |                                                                                  |
| Address of depositary institution (including postal code and country)<br>10801 University Boulevard<br>Manassas, Virginia 20110-2209<br>United States of America                  |                                                                                  |
| Identification Reference by Depositor: CS0067-3                                                                                                                                   | Accession Number                                                                 |
| Date of deposit<br>11 June 2002                                                                                                                                                   |                                                                                  |
| C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>                                           |                                                                                  |
|                                                                                                                                                                                   |                                                                                  |
| D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)                                                                        |                                                                                  |
|                                                                                                                                                                                   |                                                                                  |
| E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)                                                                                                             |                                                                                  |
| The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g. "Accession Number of Deposit")               |                                                                                  |
| Accession Number of Deposit                                                                                                                                                       |                                                                                  |
|                                                                                                                                                                                   |                                                                                  |
| For receiving Office use only                                                                                                                                                     | For International Bureau use only                                                |
| <input checked="" type="checkbox"/> This sheet was received with the international application                                                                                    | <input type="checkbox"/> This sheet was received by the International Bureau on: |
| Authorized officer:<br>Sonya D. Barnes<br>PCT/Internat'l Appl Processing Div<br>703) 305-3665                                                                                     | Authorized officer:                                                              |

Form PCT/RO/134 (July 1998)

PCT/US02/20275

Applicants : BIOGEN, INC., et al.  
Filed : Concurrently Herewith  
Title : PANCAM NUCLEIC ACIDS AND  
POLYPEPTIDES  
Agent's Ref. : A097 PCT2

Indications Relating to the Expert Solution in  
Respect of Deposited Biological Material  
Referred to in the Description

The indications relating to deposited biological material are all contained in the description. The following additional indications are not required to be part of the description and should be treated as "separate indications." They relate only to the expert solution.

The additional indications made below relate to the deposited biological material referred to as CS0067-3 in the description on

page 11, lines 1-7.

The deposit was made in:

American Type Culture Collection (ATCC)  
10801 University Boulevard  
Manassas, Virginia 20110-2209  
United States of America

on June 11, 2002, under Accession Number \_\_\_\_\_.

The additional indications are:

For CA (Canada) designation:

In respect of the designation of Canada, samples of the deposited biological material will be made available until the grant of the Canadian patent, or until the date on which the application is refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, as provided in Rules 107 and 108 of the Patent Rules under the Canadian Patent Act, only by



PCT/US02/20275

- 2 -

the issue of a sample to an independent expert nominated by the Commissioner (Rule 104(4)).

For EP (European Patent) designation:

In respect of the designation of the EPO, samples of the deposited biological material will be made available until the publication of the mention of the grant of the European patent, or until 20 years from the date of filing if the application is refused or withdrawn or is deemed to be withdrawn, as provided in Rule 28(3) of the Implementing Regulations under the EPC, only by the issue of a sample to an expert nominated by requester (Rule 28(4) EPC).

For FI (Finland) designation:

In respect of the designation of Finland, until the publication of the mention of grant by the National Board of Patents and Registration, or for 20 years from the date of filing if the application has been finally decided upon without resulting in the grant of a patent by the National Board of Patents and Registration, the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For GB (United Kingdom) designation:

In respect of the designation of the United Kingdom, the applicant(s) hereby give(s) notice of my/our intention that the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For IS (Iceland) designation:

In respect of the designation of Iceland, until a patent has been granted by the Icelandic Patent Office, or a final decision taken by the Icelandic Patent Office if the application has been finally decided upon without resulting in the grant of a patent, the furnishing of samples of the deposited biological material will only be effected to an expert in the art.

PCT/US02/20275

- 3 -

For SE (Sweden) designation:

In respect of the designation of Sweden, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of samples of the deposited biological material will be made available only to an expert in the art.

For SG (Singapore) designation:

Applicant(s) hereby give notice of my/our intention that samples of the above-identified culture shall be available only to experts in accordance with paragraph 3 of the Fourth Schedule to the Patents Rules 1995.

## 【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  
3 January 2003 (03.01.2003)

PCT

(10) International Publication Number  
WO 03/00021 A3(51) International Patent Classification: C07K 14/00, 16/46, 16/28, C07H 21/04, C12P 21/06, C12N 5/02, 15/74, C12Q 1/68, G01N 33/53, A61K 48/00  
(74) Agents: HALEY, James, F. et al.; Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

(21) International Application Number: PCT/US02/20275

(22) International Filing Date: 24 June 2002 (24.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: PCT/US01/19904 22 June 2001 (22.06.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application: PCT/US01/19904 (CIP) Filed on 22 June 2001 (22.06.2001)

(71) Applicant (for all designated States except US): BIO-GEN, INC. (US/US); 14 Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): SUN, Chao [CN/US]; 11 Cimino Road, Needham, MA 02494 (US). CARULLI, John, P. [US/US]; 9 Harris Drive, Southborough, MA 01772 (US). LUKASHIN, Alexander, V. [US/US]; 42 8th Street, #4403, Charlestown, MA 02129 (US). KILBURN, Daniel, R. [US/US]; 6 Hancock Avenue, Lexington, MA 02420 (US). GARDNER, Humphrey, Athelstan, Roy [CA/US]; 14 Phelps Street, Salem, MA 01970 (US).

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

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

## Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- with (any) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description

(88) Date of publication of the international search report: 20 November 2003

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 03/00021 A3

(54) Title: PANCAM NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: An isolated polynucleotide encoding a novel immunoglobulin superfamily member named PanCAM (also named GP354) is provided. PanCAM has a predicted single membrane spanning domain and five immunoglobulin (Ig) domains in the extracellular portion of the protein. The protein structure and tissue distribution of PanCAM indicate that it plays a role in cell-cell recognition, binding, signaling and adhesion events in the pancreas (especially the beta islets), the pituitary gland, and the central nervous system (CNS). Provided by the invention are isolated PanCAM related polynucleotides and polypeptides, vectors, and host cells comprising any of the above, antibodies directed to PanCAM, cells which produce such antibodies, and related diagnostic and therapeutic methods.

## 【国際調査報告】

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                        |                                                                   |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|
| <b>INTERNATIONAL SEARCH REPORT</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                        | International application No.<br>PCT/US02/20275                   |
| <b>A. CLASSIFICATION OF SUBJECT MATTER</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                        |                                                                   |
| IPC(7) : C07K 14/00, 16/46, 16/28; C07H 21/04; C12P 21/06; C12N 5/02, 15/74; C12Q 1/68; G01N 33/53; A61K 48/0                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                                                        |                                                                   |
| US CL : 530/350, 387.3, 388.25, 389.1; 536/23.5; 435/69.1, 325, 320.1; 435/6, 7.1; 514/44; 424/141.1, 93.1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                        |                                                                   |
| <b>B. FIELDS SEARCHED</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                        |                                                                   |
| Minimum documentation searched (classification system followed by classification symbols)<br>U.S. : 530/350, 387.3, 388.25, 389.1; 536/23.5; 435/69.1, 325, 320.1; 435/6, 7.1; 514/44; 424/141.1, 93.1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                        |                                                                   |
| 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 practicable, search terms used)<br>Medline, STN, WEST                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                        |                                                                   |
| <b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                                                        |                                                                   |
| Category *                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Citation of document, with indication, where appropriate, of the relevant passages                                                     | Relevant to claim No.                                             |
| X.P                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | WO 01/68848 A2 (GENENTECH INC.) 20 September 2001 (20.09.2001), claims 11 and 16, Fig 584, pages 87, line 87 through page 92, line 25. | 1-56, 26-28                                                       |
| <input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                        |                                                                   |
| * Special categories of cited documents:<br>"A" document defining the general state of the art which is not considered to be of particular relevance<br>"E" earlier application or patent published on or after the international filing date<br>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)<br>"O" document referring to an oral disclosure, use, exhibition or other means<br>"P" document published prior to the international filing date but later than the priority date claimed<br>"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<br>"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<br>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other such documents, each contribution being obvious to a person skilled in the art<br>"R" document member of the same patent family |                                                                                                                                        |                                                                   |
| Date of the actual completion of the international search<br>16 July 2003 (16.07.2003)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                        | Date of mailing of the international search report<br>16 SEP 2003 |
| Name and mailing address of the ISA/US<br>Mail Stop PCT, Attn: ISA/US<br>Commissioner for Patents<br>P.O. Box 1450<br>Alexandria, Virginia 22313-1450<br>Facsimile No. 703 305-3230                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                                                                                                                                        | Authorized officer<br>M. M. Haddad<br>Telephone No. 703 308-0196  |

## 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 1/18             | A 6 1 K 39/395 | V 4 C 0 8 4 |
| A 6 1 P 3/10             | A 6 1 K 48/00  | 4 C 0 8 5   |
| A 6 1 P 5/00             | A 6 1 P 1/18   | 4 C 0 8 6   |
| A 6 1 P 5/06             | A 6 1 P 3/10   | 4 C 0 8 7   |
| A 6 1 P 5/10             | A 6 1 P 5/00   | 4 H 0 4 5   |
| A 6 1 P 25/00            | A 6 1 P 5/06   |             |
| A 6 1 P 25/06            | A 6 1 P 5/10   |             |
| A 6 1 P 25/08            | A 6 1 P 25/00  |             |
| A 6 1 P 25/16            | A 6 1 P 25/06  |             |
| A 6 1 P 25/18            | A 6 1 P 25/08  |             |
| A 6 1 P 25/28            | A 6 1 P 25/16  |             |
| A 6 1 P 35/00            | A 6 1 P 25/18  |             |
| A 6 1 P 37/00            | A 6 1 P 25/28  |             |
| A 6 1 P 43/00            | A 6 1 P 35/00  |             |
| C 0 7 K 14/47            | A 6 1 P 37/00  |             |
| C 0 7 K 16/18            | A 6 1 P 43/00  | 1 0 5       |
| C 0 7 K 19/00            | A 6 1 P 43/00  | 1 1 1       |
| C 1 2 N 1/15             | C 0 7 K 14/47  |             |
| C 1 2 N 1/19             | C 0 7 K 16/18  |             |
| C 1 2 N 1/21             | C 0 7 K 19/00  |             |
| C 1 2 N 5/06             | C 1 2 N 1/15   |             |
| C 1 2 N 5/10             | C 1 2 N 1/19   |             |
| C 1 2 P 21/02            | C 1 2 N 1/21   |             |
| C 1 2 Q 1/02             | C 1 2 P 21/02  | C           |
| C 1 2 Q 1/68             | C 1 2 Q 1/02   |             |
| G 0 1 N 33/15            | C 1 2 Q 1/68   | A           |
| 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           |
| // C 1 2 P 21/08         | G 0 1 N 33/53  | M           |
|                          | G 0 1 N 33/566 |             |
|                          | C 1 2 N 5/00   | A           |
|                          | C 1 2 N 5/00   | E           |
|                          | C 1 2 P 21/08  |             |

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

(72)発明者 サン, チャオ  
 アメリカ合衆国, マサチューセッツ 0 2 4 9 4, ニーダム, シミノ ロード 1 1  
 (72)発明者 カルーリ, ジョン ピー.  
 アメリカ合衆国, マサチューセッツ 0 1 7 7 2, サウスボロー, ハリス ドライブ 9  
 (72)発明者 ルカシン, アレクサンダー ヴィー.  
 アメリカ合衆国, マサチューセッツ 0 2 1 2 9, チャールズタウン, シャープ 4 4 0 3,

8 ティーエイチ ストリート 42

(72)発明者 キルバーン, ダニエル アール.

アメリカ合衆国, マサチューセッツ 02420, レキシントン, ハンコック アヴェニュー 6

(72)発明者 ガードナー, ハンフリー アセルスタン ロイ

アメリカ合衆国, マサチューセッツ 01970, セイラム, フェルプス ストリート 14

F ターム(参考) 2G045 AA34 AA35 BB05 BB10 BB46 BB51 DA13 DA36 FB02 FB03  
4B024 AA01 AA11 BA53 BA61 BA80 CA04 CA07 CA09 CA12 CA20  
DA02 DA03 DA12 EA02 EA04 GA13 HA03 HA11 HA13 HA14  
HA17  
4B063 QA01 QA13 QA17 QA19 QQ02 QQ08 QQ43 QQ53 QQ79 QR08  
QR32 QR35 QR40 QR42 QS16 QS25 QS34 QS36 QX01 QX02  
4B064 AG01 AG27 CA10 CA19 CA20 CC01 CC24 DA01 DA12  
4B065 AA90X AA93X AA93Y AB01 AC14 BA30 CA24 CA43 CA44 CA46  
4C084 AA13 NA14 ZA021 ZA061 ZA081 ZA151 ZA161 ZA181 ZA661 ZB021  
ZB071 ZB111 ZB211 ZB212 ZB261 ZC011 ZC012 ZC031 ZC041 ZC351  
4C085 AA13 AA14 AA16 AA34 BB31 CC03 CC22 CC29 DD62 EE01  
EE03  
4C086 AA01 EA16 MA01 MA04 NA14 ZA02 ZA06 ZA08 ZA15 ZA16  
ZA18 ZA66 ZB02 ZB07 ZB11 ZB21 ZB26 ZC01 ZC03 ZC04  
ZC35  
4C087 AA01 BB44 BB51 BB65 CA12 NA14 ZA02 ZA06 ZA08 ZA15  
ZA16 ZA18 ZA66 ZB02 ZB07 ZB11 ZB21 ZB26 ZC01 ZC03  
ZC04 ZC35  
4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA75 DA76 EA20  
EA50 FA71 FA72 FA74

|             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |         |            |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|------------|
| 专利名称(译)     | PanCAM核酸和多肽                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |         |            |
| 公开(公告)号     | <a href="#">JP2005508147A</a>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 公开(公告)日 | 2005-03-31 |
| 申请号         | JP2003506477                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 申请日     | 2002-06-24 |
| 申请(专利权)人(译) | Biogen Idec公司 , Emuei公司                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |         |            |
| [标]发明人      | サンチャオ<br>カルーリジョンピー<br>ルカシンアレクサンダーヴィー<br>キルバーンダニエルアール<br>ガードナーハンフリーアセルスタンロイ                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |         |            |
| 发明人         | サン, チャオ<br>カルーリ, ジョン ピー.<br>ルカシン, アレクサンダー ヴィー.<br>キルバーン, ダニエル アール.<br>ガードナー, ハンフリー アセルスタン ロイ                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |         |            |
| IPC分类号      | G01N33/50 A61K31/7088 A61K35/39 A61K38/00 A61K39/395 A61K48/00 A61P1/18 A61P3/10 A61P5/00 A61P5/06 A61P5/10 A61P25/00 A61P25/06 A61P25/08 A61P25/16 A61P25/18 A61P25/28 A61P35/00 A61P37/00 A61P43/00 C07K14/47 C07K14/705 C07K16/18 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/02 C12N5/06 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |         |            |
| CPC分类号      | A61K38/00 A61P1/18 A61P3/10 A61P5/00 A61P5/06 A61P5/10 A61P25/00 A61P25/06 A61P25/08 A61P25/16 A61P25/18 A61P25/28 A61P35/00 A61P37/00 A61P43/00 C07K14/70503                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |         |            |
| FI分类号       | C12N15/00.ZNA.A A61K31/7088 A61K35/39 A61K39/395.D A61K39/395.N A61K39/395.V A61K48/00 A61P1/18 A61P3/10 A61P5/00 A61P5/06 A61P5/10 A61P25/00 A61P25/06 A61P25/08 A61P25/16 A61P25/18 A61P25/28 A61P35/00 A61P37/00 A61P43/00.105 A61P43/00.111 C07K14/47 C07K16/18 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A C12N5/00.E C12P21/08                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |         |            |
| F-TERM分类号   | 2G045/AA34 2G045/AA35 2G045/BB05 2G045/BB10 2G045/BB46 2G045/BB51 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA53 4B024/BA61 4B024/BA80 4B024/CA04 4B024/CA07 4B024/CA09 4B024/CA12 4B024/CA20 4B024/DA02 4B024/DA03 4B024/DA12 4B024/EA02 4B024/EA04 4B024/GA13 4B024/HA03 4B024/HA11 4B024/HA13 4B024/HA14 4B024/HA17 4B063/QA01 4B063/QA13 4B063/QA17 4B063/QA19 4B063/QQ02 4B063/QQ08 4B063/QQ43 4B063/QQ53 4B063/QQ79 4B063/QR08 4B063/QR32 4B063/QR35 4B063/QR40 4B063/QR42 4B063/QS16 4B063/QS25 4B063/QS34 4B063/QS36 4B063/QX01 4B063/QX02 4B064/AG01 4B064/AG27 4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC01 4B064/CC24 4B064/DA01 4B064/DA12 4B065/AA90X 4B065/AA93X 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/BA30 4B065/CA24 4B065/CA43 4B065/CA44 4B065/CA46 4C084/AA13 4C084/NA14 4C084/ZA021 4C084/ZA061 4C084/ZA081 4C084/ZA151 4C084/ZA161 4C084/ZA181 4C084/ZA661 4C084/ZB021 4C084/ZB071 4C084/ZB111 4C084/ZB211 4C084/ZB212 4C084/ZB261 4C084/ZC011 4C084/ZC012 4C084/ZC031 4C084/ZC041 4C084/ZC351 4C085/AA13 4C085/AA14 4C085/AA16 4C085/AA34 4C085/BB31 4C085/CC03 4C085/CC22 4C085/CC29 4C085/DD62 4C085/EE01 4C085/EE03 4C086/AA01 4C086/EA16 4C086/MA01 4C086/MA04 4C086/NA14 4C086/ZA02 4C086/ZA06 4C086/ZA08 4C086/ZA15 4C086/ZA16 4C086/ZA18 4C086/ZA66 4C086/ZB02 4C086/ZB07 4C086/ZB11 4C086/ZB21 4C086/ZB26 4C086/ZC01 4C086/ZC03 4C086/ZC04 4C086/ZC35 4C087/AA01 4C087/BB44 4C087/BB51 4C087/BB65 4C087/CA12 4C087/NA14 4C087/ZA02 4C087/ZA06 4C087/ZA08 4C087/ZA15 4C087/ZA16 4C087/ZA18 4C087/ZA66 4C087/ZB02 4C087/ZB07 4C087/ZB11 4C087/ZB21 4C087/ZB26 4C087/ZC01 |         |            |

4C087/ZC03 4C087/ZC04 4C087/ZC35 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/BA41 4H045/CA40 4H045/DA75 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA71 4H045/FA72 4H045/FA74

|        |                                 |
|--------|---------------------------------|
| 代理人(译) | 夏木森下                            |
| 优先权    | PCT/US2001/019904 2001-06-22 WO |
| 其他公开文献 | JP4224395B2                     |
| 外部链接   | <a href="#">Espacenet</a>       |

摘要(译)

提供了分离的多核苷酸，其编码称为PanCAM（也称为GP354）的新型免疫球蛋白超家族成员。PanCAM在蛋白质的细胞外部分具有一个预期的跨膜结构域和五个免疫球蛋白（Ig）结构域。PanCAM的蛋白质结构和组织分布表明，它在胰腺（尤其是β胰岛），垂体和中枢神经系统（CNS）的细胞识别，结合，信号传导和粘附事件中起作用。本发明提供了分离的与PanCAM相关的多核苷酸和多肽，载体和包含上述任何一种的宿主细胞，针对PanCAM的抗体，产生这种抗体的细胞以及相关的诊断和治疗方法。

表 3: GENSCAN 結果

| エキソン | 始点   | 終点   | 長さ  |
|------|------|------|-----|
| 14   | 1844 | 1779 | 66  |
| 13   | 3567 | 3464 | 104 |
| 12   | 4007 | 3903 | 105 |
| 11   | 4695 | 4476 | 220 |
| 10   | 4959 | 4859 | 101 |
| 09   | 5378 | 5246 | 133 |
| 08   | 5591 | 5464 | 128 |
| 07   | 5981 | 5833 | 149 |
| 06   | 6203 | 6098 | 106 |
| 05   | 7019 | 6869 | 151 |
| 04   | 7796 | 7636 | 161 |
| 03   | 8092 | 7943 | 150 |
| 02   | 9157 | 9008 | 150 |
| 01   | 9373 | 9322 | 52  |