(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2004-500872
(P2004-500872A)
(43)公表日　平成16年1月15日(2004.1.15)

(51)Int.Cl.[7]	F I		テーマコード（参考）
C12N 15/09	C12N 15/00	ZNAA	2G045
A61K 39/395	A61K 39/395	D	4B024
A61K 48/00	A61K 39/395	N	4B063
A61P 1/18	A61K 48/00		4B064
A61P 25/00	A61P 1/18		4B065

審査請求 未請求　予備審査請求 有　（全 261 頁）　最終頁に続く

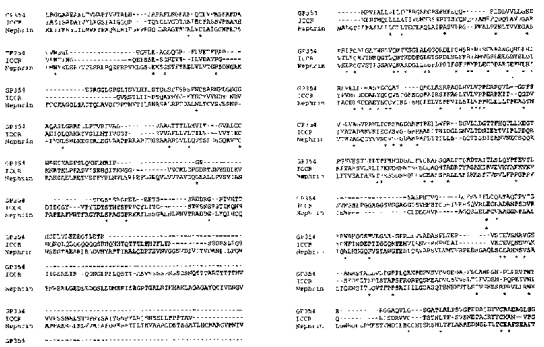
(21)出願番号　特願2002-504315(P2002-504315)
(86)(22)出願日　平成13年6月22日(2001.6.22)
(85)翻訳文提出日　平成14年12月19日(2002.12.19)
(86)国際出願番号　PCT/US2001/019904
(87)国際公開番号　WO2001/098360
(87)国際公開日　平成13年12月27日(2001.12.27)
(31)優先権主張番号　60/213,611
(32)優先日　平成12年6月22日(2000.6.22)
(33)優先権主張国　米国(US)

(71)出願人　592221528
バイオジェン　インコーポレイテッド
ＢＩＯＧＥＮ　ＩＮＣＯＲＰＯＲＡＴＥＤ
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
(74)代理人　100078282
弁理士　山本　秀策
(74)代理人　100062409
弁理士　安村　高明
(74)代理人　100113413
弁理士　森下　夏樹

最終頁に続く

(54)【発明の名称】ＧＰ３５４核酸およびポリペプチド

(57)【要約】
ＧＰ３５４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。ＧＰ３５４は、推定一回膜貫通ドメイン、および５つの免疫グロブリン（Ｉｇ）ドメインを、タンパク質の細胞外部分に有する。ＧＰ３５４のタンパク質構造および組織分布は、ＧＰ３５４が、膵臓および中枢神経系（ＣＮＳ）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ＧＰ３５４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、ＧＰ３５４を指向する抗体、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】
単離されたポリヌクレオチドであって、配列番号１２をコードするｃＤＮＡ配列または配列番号１２の対立遺伝子改変体を含む、単離されたポリヌクレオチド。
【請求項２】
前記ｃＤＮＡ配列が、配列番号１１またはその対立遺伝子改変体である、請求項１に記載のポリヌクレオチド。
【請求項３】
配列番号８をコードするｃＤＮＡ配列および配列番号８の対立遺伝子改変体を含む、請求項１に記載のポリヌクレオチド。
【請求項４】
前記ｃＤＮＡ配列が、配列番号７またはその対立遺伝子改変体である、請求項３に記載のポリヌクレオチド。
【請求項５】
前記ｃＤＮＡ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、請求項１に記載のポリヌクレオチド。
【請求項６】
異種ポリペプチドをコードする核酸配列をさらに含む、請求項１に記載のポリヌクレオチド。
【請求項７】
請求項１に記載の単離されたポリヌクレオチドを含む、ベクター。
【請求項８】
プラスミドベクターである、請求項７に記載のベクター。
【請求項９】
ウイルスベクターである、請求項７に記載のベクター。
【請求項１０】
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、請求項９に記載のベクター。
【請求項１１】
請求項１に記載のポリヌクレオチドを含む、宿主細胞。
【請求項１２】
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１１に記載の宿主細胞。
【請求項１３】
前記宿主細胞がヒト細胞である、請求項１１に記載の宿主細胞。
【請求項１４】
請求項１に記載のポリヌクレオチドによってコードされる、単離されたポリペプチド。
【請求項１５】
異種配列をさらに含む、請求項１４に記載のポリペプチド。
【請求項１６】
請求項１に記載のポリヌクレオチドおよび薬学的に受容可能なキャリアを含む、組成物。
【請求項１７】
請求項１４に記載のポリヌクレオチドおよび薬学的に受容可能なキャリアを含む、組成物。
【請求項１８】
請求項１４に記載のポリペプチドに結合する、抗体。
【請求項１９】
前記抗体が、モノクローナル抗体である、請求項１８に記載の抗体。
【請求項２０】

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１９に記載の抗体。
【請求項２１】
請求項１８～２０に記載のいずれか１つの抗体および薬学的に受容可能なキャリアを含む、組成物。
【請求項２２】
ポリペプチドを産生する方法であって、該方法は、
請求項１１に記載の宿主細胞を、前記ポリ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。
【請求項２３】
サンプル中のｇｐ３５４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：
該サンプルを、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、請求項１に記載の単離されたポリヌクレオチドと接触させる工程、および
該単離されたポリヌクレオチドの、該サンプルにおける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ｇｐ３５４コード配列の存在を示す、工程、
を包含する、方法。
【請求項２４】
サンプル中のＧＰ３５４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１８、１９または２０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ＧＰ３５４タンパク質の存在を示す、工程、
を包含する、方法。
【請求項２５】
ＧＰ３５４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：
ＧＰ３５４タンパク質を、試験化合物と接触させる工程；および
該ＧＰ３５４タンパク質および該試験化合物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ＧＰ３５４タンパク質に結合することを示す、工程、
を包含する、方法。
【請求項２６】
ＧＰ３５４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ＧＰ３５４タンパク質を、試験化合物と接触させる工程；および
該ＧＰ３５４タンパク質の活性に対する試験化合物の効果を検出する工程であって、ここで、接触工程の後の該活性の変化が、該試験化合物が該ＧＰ３５４タンパク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２７】
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、該被験体からの組織サンプル中のＧＰ３５４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ＧＰ３５４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、コントロールと比較した該組織サンプル中の該ＧＰ３５４タンパク質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２８】
前記疾患状態が、膵臓に関連する、請求項２７に記載の方法。

【請求項２９】
前記疾患状態が、中枢神経系に関連する、請求項２７に記載の方法。
【請求項３０】
被験体における疾患状態を診断する方法であって、該方法が、該被験体からの組織サンプル中のｇｐ３５４　ｍＲＮＡの量を、コントロールサンプル中のｇｐ３５４　ｍＲＮＡの量と比較する工程を包含し、ここで、コントロールと比較した該組織サンプル中の該ｍＲＮＡ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。
【請求項３１】
前記疾患状態が、膵臓に関連する、請求項３０に記載の方法。
【請求項３２】
前記疾患状態が、中枢神経系に関連する、請求項３０に記載の方法。
【請求項３３】
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ための診断アッセイであって、該遺伝的な病変または変異は、以下：（ｉ）ＧＰ３５４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）ＧＰ３５４タンパク質をコードする遺伝子の調節不良；および（ｉｉｉ）ＧＰ３５４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、該アッセイは、以下の工程：
該試験細胞、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１、３、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これらの一部を含む核酸プローブと、高いストリンジェンシーの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
該参照細胞の核酸と比較して、該核酸プローブが、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、診断アッセイ。
【請求項３４】
膵臓損傷の処置のための、請求項１６、１７または２１に記載の組成物の使用。
【請求項３５】
膵臓に関連した異常または疾患状態の処置のための、請求項１６、１７または２１に記載の組成物の使用。
【請求項３６】
前記状態が、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、脾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３５に記載の使用。
【請求項３７】
中枢神経系に対する損傷の処置のための、請求項１６、１７または２１に記載の組成物の使用。
【請求項３８】
中枢神経系に関連する異常または疾患状態の処置のための、請求項１６、１７または２１に記載の組成物の使用。
【請求項３９】
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３８に記載の使用。
【発明の詳細な説明】
【０００１】
（関連出願）
本出願は、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し、この仮特許出願の開示は、本明細書中で参考として援用される

。
（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。
（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キ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：リガンド依存性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。
【０００２】
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る三次構造的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、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、各レイヤーは、５～１０アミノ酸残基の３～５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。
【０００３】
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、広く分岐する細胞質末端（ｃｙｔｏｐｌａｓｍｉｃ　ｔａｉｌ）において１つ以上の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。これらの総括に例外は存在する。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。
【０００４】
Ｉｇスーパーファミリーの多数のメンバーは、脊椎動物の免疫軽における細胞認識、接着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。Ｉｇスーパーファミリーを産生する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Ｉｇ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（ＦｃｇＲＩＩ）、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）（Ａ

ｂｂａｓら、前出）。
【０００５】
いくつかのＩｇ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Ｉｇ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Ｉｇスーパーファミリーメンバーは、診断の目的のために細胞マーカーおよび組織マーカーに有用である。組織特異的Ｉｇスーパーファミリーメンバーはまた、特に、例えば、組織損傷後または組織傷害後の組織発生の間または組織再生の間の、組織における不適切な細胞－細胞接着およびシグナル伝達に関連した異常な状態、障害および／または疾患を処置するための治療用標的に適している。
【０００６】
（発明の要旨）
本発明は、これまで未知であったＩｇスーパーファミリーメンバー（ＧＰ３５４と称される）をコードする遺伝子の発見に（少なくとも一部）基づく。（他に示さない限り、名前が小文字の場合（ｇｐ３５４）は、本発明の新規核酸をいい、これに対して名前が大文字の場合（ＧＰ３５４）は、本発明の新規ポリペプチドをいう）。このヒトｇｐ３５４　ｃＤＮＡによってコードされるタンパク質（ＧＰ３５４）は、膵臓に豊富な内在性膜タンパク質である。ＧＰ３５４は、推定一回膜貫通型ドメインおよびこのタンパク質の細胞外部分に５つの免疫グロブリン（Ｉｇ）ドメインを有する。ＧＰ３５４タンパク質は、先に記載されたタンパク質のいずれかと、全体で３０％以下のアミノ酸の同一性を共有する。ＧＰ３５４構造および組織分布は、これが膵臓および中枢神経系（ＣＮＳ）中の細胞－細胞相互作用において役割を果たすことを示す。
【０００７】
本発明は、ＧＰ３５４またはその生物学的に活性な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。本発明はまた、ＧＰ３５４をコードするポリヌクレオチドの検出のための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の使用に適切なポリヌクレオチドフラグメントを提供する。他に特別に示さない限り、「ＧＰ３５４」、「ＧＰ３５４」タンパク質および「ＧＰ３５４」ポリペプチドは、ヒト遺伝子産物または他の非ヒト哺乳動物もしくは他の脊椎動物種におけるこのタンパク質の相同体をいう。
【０００８】
本発明は、（推定ｇｐ３５４　ｃＤＮＡによってコードされる）配列番号２のアミノ酸配列；（部分的ｇｐ３５４膵臓ｃＤＮＡによってコードされる）配列番号４のアミノ酸配列；（誘導ｇｐ３５４　ｃＤＮＡによってコードされた）配列番号８のアミノ酸配列；（部分的誘導ｇｐ３５４　ｃＤＮＡによってコードされる）配列番号１０のアミノ酸配列；もしくは（ｇｐ３５４膵臓ｃＤＮＡによってコードされる）配列番号１２のアミノ酸配列；または配列番号２、４、８、１０および１２のいずれか１つのうち少なくとも１つのＩｇドメインに対して、少なくとも８０％（８５％、９５％または９８％）同一である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。
【０００９】
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（ｇｐ３５４　ｃＤＮＡ）の配列、または少なくとも１７個の核酸単位（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（ｇｐ３５４を含むゲノムＤＮＡ）の配列、または少なくとも１７個の核酸単位を有するそのフラグメントを含む。例示的なフラグメントは、配列番号６（ｇｐ３５４上流のゲノム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（誘導ｇｐ３５４　ｃＤＮＡ）の配列、または少なくとも１７個の核酸単位を有するそのフラグメントを含む。例示的なフラグメントは、配列番号９のフラグメント（誘導ｇｐ３５４　ｃＤＮＡのＣ末端のフラグメン

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１１（膵臓のｇｐ３５４　ｃＤＮＡ）の配列、または少なくとも１７個の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、少なくとも１つの細胞外Ｉｇドメインおよび／またはＧＰ３５４の細胞内ドメインの一部または全てをコードする。

【００１０】

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の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し、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の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し、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。

【００１１】

配列番号２、４、８、１０もしくはｌ２のアミノ酸配列；またはこれらの配列のいずれか１つによってコードされるＩｇドメインに対して、少なくとも８０％（８５％、９５％または９８％）同一であるアミノ酸配列を含む単離されたＧＰ３５４タンパク質もまた本発明によって提供される。

【００１２】

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これらの配列のいずれか１つの一部に対して、少なくとも約６５％、好ましくは７５％、８５％または９５％同一である配列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または配列番号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。

【００１３】

本発明は、Ｉｇスーパーファミリーの他のメンバーをコードする核酸と比較してｇｐ３５４核酸を特異的に検出するｇｐ３５４ポリヌクレオチドを提供する。本発明はまた、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、ターゲティングベクターまたは発現ベクター）を提供し、このベクターは、本発明のｇｐ３５４ポリヌクレオチドを含む。

【００１４】

このような核酸構築物を含む宿主細胞はまた、提供され、ＧＰ３５４ポリペプチドを産生するように組換え発現構築物を含む本発明の宿主細胞を適切な培地中で培養することによってＧＰ３５４ポリペプチドを提供するための方法も同様に提供される。

【００１５】

単離されたＧＰ３５４タンパク質または組換えＧＰ３５４タンパク質および単離されたＧＰ３５４ペプチドまたは組換えＧＰ３５４ペプチドは、本発明によって提供される。好ましいＧＰ３５４タンパク質およびＧＰ３５４ポリペプチドは、天然のヒトＧＰ３５４によって保有される以下の（重複する）生物学的活性の少なくとも１つを保有する：（１）ＧＰ３５４タンパク質に自然に結合するリガンド（例えば、タンパク質レセプター、多糖類など）と相互作用する（例えば、結合する）能力；（２）天然のヒトＧＰ３５４に対する自己抗体または天然のヒトＧＰ３５４に対して生じる抗体に結合する能力；（３）膵臓の機能（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工程）に関連する能力；（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工程）に関連する能力；ならびに（５）細胞－細胞相互作用（例えば、認識、結合および／または接着）を媒介する能力。

【００１６】

ＧＰ３５４タンパク質またはその生物学的に活性な部分は、非ＧＰ３５４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での送達を促進する配列）に作動可能に連結して、ＧＰ３５４融合タンパク

質を形成し得る。

【００１７】

本発明はさらに、以下を含む抗体（例えば、ポリクローナル抗体またはモノクローナル抗体）を特徴とする：キメラ抗体およびヒト化抗体（これは、ＧＰ３５４タンパク質またはその一部に特異的に結合する）。

【００１８】

本発明は、上記のｇｐ３５４関連の単離されたポリヌクレオチド、ＧＰ３５４タンパク質もしくはその生物学的に活性な部分、抗体または融合タンパク質のうち少なくとも１つを含む薬学的組成物を提供し；この薬学的組成物は、必要に応じて薬学的に受容可能なキャリアを含む。このような組成物は、異常なＧＰ３５４の細胞局在化、発現および／または活性に関連した被験体における状態を改善するための治療方法に有用である。

【００１９】

このように、本発明はまた、本発明のｇｐ３５４関連化合物またはｇｐ３５４関連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、異常な状態、発生途中もしくは成人の膵臓系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機能に関連する異常な状態、障害または疾患を処置するために有用である。

【００２０】

膵臓の豊富なタンパク質（ＧＰ３５４）は、特に、例えば、膵臓の損傷、傷害または変性状態の後の、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の、発生途中および成人の膵臓における、不適切な細胞－細胞の結合、接着ならびにシグナル伝達に関連する異常な状態、障害ならびに／または疾患を処置するための適切な治療的標的である。これはまた、ＧＰ３５４が免疫状態、自己免疫状態、および変性状態に関連した膵臓の細胞死を阻害するための適切な治療的標的であると想定される。ＧＰ３５４の神経の形態は、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、ならびに中枢神経系における組織変性および細胞死を阻害するための類似の適切な治療的標的である。

【００２１】

本発明は、ＧＰ３５４の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、標的細胞は、ＧＰ３５４の活性またはＧＰ３５４の発現を調節（例えば、阻害または刺激）する因子と接触することによって、ＧＰ３５４の活性またはＧＰ３５４の発現を変更する。いくつかの実施形態において、この因子は、ＧＰ３５４に特異的に結合する抗体である。いくつかの実施形態において、この因子は、ｇｐ３５４遺伝子の転写、ｇｐ３５４　ＲＮＡのスプライシング、またはｇｐ３５４　ｍＲＮＡの翻訳を調節することによってＧＰ３５４の活性またはＧＰ３５４の発現を調節する。なお他の実施形態において、この因子は、ｇｐ３５４　ｍＲＮＡまたはｇｐ３５４遺伝子のコード鎖に対する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、この因子は、ＧＰ３５４タンパク質、ＧＰ３５４タンパク質をコードする核酸、あるいはＧＰ３５４タンパク質のアンタゴニストまたはアゴニスト（例えば、ペプチド、擬ペプチド（ｐｅｐｔｉｄｏｍｉｍｅｔｉｃ）、もしくは他の低分子）であり得る。

【００２２】

本発明はまた、ＧＰ３５４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、ＧＰ３５４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更するこれらの化合物を同定する工程を包含する。コンビナトリアルライブラリーは、これらの方法において候補化合物の供給源として用いられ得る。

【００２３】

本発明は、生物学的サンプル（例えば、患者由来の流体サンプルもしくは組織サンプル）中のｇｐ３５４ポリヌクレオチド、ＧＰ３５４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、ｇｐ３５４ポリヌクレオチド配列

、ＧＰ３５４タンパク質またはその活性の存在の指標を検出し得る因子と接触させることによる。

【００２４】

診断アッセイは、以下：（ｉ）ＧＰ３５４タンパク質をコードする遺伝子の異常な修飾または変異；（ｉｉ）ＧＰ３５４タンパク質をコードする遺伝子の調節不良（ｍｉｓ－ｒｅｇｕｌａｔｉｏｎ）（転写、スプライシングまたは翻訳）；および（ｉｉｉ）ＧＰ３５４タンパク質の異常な翻訳後修飾もしくはＧＰ３５４タンパク質の局在化、の少なくとも１つによって特徴付けられるｇｐ３５４関連遺伝子の損傷または変異の存在または非存在を同定するために提供され、ここで、この遺伝子の野生型形態は、ＧＰ３５４の生物学的活性を有するタンパク質をコードする。

【００２５】

本発明は、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）細胞の少なくともいくつかを提供する。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つかのみがこのポリヌクレオチドを保有するキメラ体であり得る。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、代替的にトランスジェニックであり得る

本発明はまた、ｇｐ３５４遺伝子の内因性オルソログが遺伝子ターゲティングによって破壊される（すなわち「ノックアウト」）非ヒト動物を提供する。ｇｐ３５４ポリヌクレオチド、生物学的サンプル（例えば、組織および流体）ならびにこれら由来のＧＰ３５４関連産物を含む細胞ならびに上述の動物はまた、本発明の範囲内である。

【００２６】

本発明は、本発明の核酸配列およびアミノ酸配列をソーティングするコンピュータで読み取り可能な手段を提供する。コンピュータで読み取り可能な手段の記録は、配列の読み取りおよび表示について、ならびに本発明の配列の他の配列との比較、アライメントおよび順序付けについて評価し得る。

【００２７】

本発明の他の特徴および利点は、以下の詳細な説明、図面および上記の特許請求の範囲から明白である。

【００２８】

（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（ＧＰ３５４）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（ｇｐ３５４）を、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染色体上の残基である。この遺伝子から転写されたｍＲＮＡは、１７７９個の塩基対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主に発現される。

【００２９】

（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、Ｂａｎｅｒら，Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．１２：１１－１５（２００１）；Ｅｓｃｕｄｅ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．Ｃｈｅｍｉ．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。

【００３０】

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、この単離された核酸は、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。

【００３１】

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た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。

【００３２】

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。

【００３３】

本発明のポリヌクレオチドは、全長ＧＰ３５４タンパク質をコードし得る場合に、「全長」とみなされる。

【００３４】

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。

【００３５】

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。

【００３６】
そう規定される場合に、用語「マイクロアレイ」は、Ｓｃｈｅｎａ（編）、ＤＮＡ Ｍｉｃｒｏａｒｒａｙｓ：Ａ Ｐｒａｃｔｉｃａｌ Ａｐｐｒｏａｃｈ（Ｐｒａｃｔｉｃａｌ Ａｐｐｒｏａｃｈ Ｓｅｒｉｅｓ）、Ｏｘｆｏｒｄ Ｕｎｉｖｅｒｓｉｔｙ Ｐｒｅｓｓ（１９９９）（ＩＳＢＮ：０１９９６３７７６８）；Ｎａｔｕｒｅ Ｇｅｎｅｔ．２１（１）（補遺）：１－６０（１９９９）；およびＳｃｈｅｎａ（編）、Ｍｉｃｒｏａｒｒａｙ Ｂｉｏｃｈｉｐ：Ｔｏｏｌｓ ａｎｄ Ｔｅｃｈｎｏｌｏｇｙ，Ｅａｔｏｎ Ｐｕｂｌｉｓｈｉｎｇ Ｃｏｍｐａｎｙ／ＢｉｏＴｅｃｈｎｉｑｕｅｓ Ｂｏｏｋｓ Ｄｉｖｉｓｉｏｎ（２０００）（ＩＳＢＮ：１８８１２９９３７６）；Ｂｒｅｎｎｅｒら、Ｐｒｏｃ．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。

【００３７】
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、基材に結合される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。

【００３８】
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、決定される。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決定される。

【００３９】
本明細書中での目的のために、「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℃の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℃の０．２×ＳＳＣ、０．１％ ＳＤＳで２０分間洗浄することとして規定される。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。

【００４０】
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℃の加湿オーブン中において、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中で一晩ハイブリダイゼーションし、その後５５℃の１×ＳＳＣ、０．２％ ＳＤＳ中で５分間、次いで５５℃の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にクロスハイブリダイゼーションするのに適しており、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。

【００４１】
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、アミノ酸残基２５以下を有する、タンパク質、ポリペプチドまたはペプチドを示す。

【００４２】

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／また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。

【００４３】

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、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、「実質的に精製されたとみなされる。

【００４４】

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。用語「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。

【００４５】

「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。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）。

【００４６】

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たフラグメント、化学的切断および／または化学的解離によって生じたフラグメント、ならびに組換え的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’）$_2$フラグメントおよび単鎖Ｆｖ（ｓｃＦｖ）フラグメントがある。

【００４７】

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（ｄｉａｂｉｄｙ）（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援用される）を参照のこと）。

【００４８】

「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。

【００４９】

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。

【００５０】

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。

【００５１】

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。

【００５２】

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。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。

【００５３】

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）；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：Ｈａｒｒｉｓｏｎ’ｓ ＰＲＩＮＣＩＰＬＥＳ ＯＦ ＩＮＴＥＲＮＡＬ ＭＥＤＩＣＩＮＥ，第１４版，（Ａｎｔｈｏｎｙ Ｓ．Ｆａｕｃｉら編），ＭｃＧｒａｗ－Ｈｉｌｌ Ｃｏｍｐａｎｉｅｓ，Ｉｎｃ．，１９９８。

【００５４】

（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日に寄託された）で同定した。この寄託物は、ｇｐ３５４のヒトゲノム配列（図６および配列番号５）を有し、これは、５’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。

【００５５】

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。

【００５６】

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１、５、７または１１のヌクレオチド配列、（ｉｉ）配列番号１、５、７または１１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。配列番号１は、推定されるｇｐ３５４　ｃＤＮＡ配列を示し、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７は、配列番号１のスプライス変異体であり得る誘導ｇｐ３５４ｃＤＮＡ配列を示し、そして配列番号１１は、膵臓ｇｐ３５４ｃＤＮＡ配列を示す。

【００５７】

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８または配列番号１２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（ｉｉ）配列番号２、配列番号８または配列番号１２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。配列番号２は、配列番号１のｃＤＮＡによってコードされるＧＰ３５４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ＧＰ３５４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ｃＤＮＡによってコードされるＧＰ３５４のアミノ酸配列を示す（図８）。

【００５８】

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、（ｉｉ）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（ｉｉｉ）（ｉ）または（ｉｉ）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ｃＤＮＡによってコードされるＧＰ３５４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ＧＰ３５４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ｃＤＮＡによってコードされるＧＰ３５４のアミノ酸配列を示す。

【００５９】

（ＧＰ３５４の一部をコードする核酸）

本発明はまた、ＧＰ３５４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エレメントとして、ＧＰ３５４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。

【００６０】

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、（ｉｉ）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（ｉｉｉ）（ｉ）または（ｉｉ）の相補体。配列番号３は、ｇｐ３５４膵臓ＲＮＡ由来の７８５塩基対ＲＴ－ＰＣＲフラグメントを示す。配列番号６は、ｇｐ３５４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、ｇｐ３５４膵臓ＲＮＡ由来の１７８２塩基対ＲＴ－ＰＣＲフラグメントを示す。

【００６１】

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、ＧＰ３５４のＮ末端細胞外部分に特徴的なＩｇ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Ｉｇドメインは、それぞれ、配列番号１のｇｐ３５４　ｃＤＮＡ配列のヌクレオチド１０３～３０６、４０６～６０９、７１５～８７０、９６７～１１２２および１２２８～１４４５（図１参照）、ならびに、それぞれ、配列番号８のｇｐ３５４　ｃＤＮＡ配列のヌクレオチド３０７～５１０、６１０～８１３、９１９～１０７４、１１７１～１３２６および１４３２～１６５９（図７参照）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。

【００６２】

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７参照）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７参照）によってコードされるＮ末端プロリン残基を有する。

【００６３】

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７参照）によってコードされる。

【００６４】

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７参照）によってコードされる。

【００６５】

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５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８）およびＧＸ１－２１９（配列番号９）を用いて、膵臓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。

【００６６】

（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、またはオルトログについても同様）に対するプローブとして、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。

【００６７】

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；または少なくとも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。

【００６８】

（好ましい核酸）

膵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。

【００６９】

（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。

【００７０】

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、１８ヌクレオチド、２０ヌクレオチド、２４ヌクレオチド、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。

【００７１】

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。

【００７２】

（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。

【００７３】

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、１８個連続するヌクレオチド、２０個連続するヌクレオチド、２４個連続するヌクレオチド、２５個連続するヌクレオチド、または５０個連続するヌクレオチドを含む。

【００７４】

【表１】

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

	cDNA-1	cDNA-2	genomic
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

【００７５】

【表２】

表2： gp354 cDNA4(配列番号11)およびゲノム(配列番号5)配列

	cDNA	ゲノム
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

（転写制御核酸）

別の局面において、本発明は、ｇｐ３５４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ｇｐ３５４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、ｇｐ３５４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、ｇｐ３５４遺伝子自体の発現パターンを変化させる。

【００７６】

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１～６４８３；配列番号５のヌクレオチド１４８３～６４８２；配列番号５のヌクレオチド２４８３～６４８２；配列番号５のヌクレオチド３４８３～６４８２；配列番号５のヌクレオチド４４８３～６４８２；配列番号５のヌクレオチド５４８３～６４８２；もしくは配列番号５のヌクレオチド５９８３～６４８２；またはこのような配列の相補鎖を含む、単離されたポリヌクレオチドを提供する。

【００７７】

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１～６４８３；配列番号５のヌクレオチド１４８３～６４８２；配列番号５のヌクレオチド２４８３～６４８２；配列番号５のヌクレオチド３４８３～６４８２；配列番号５のヌクレオチド４４８３～６４８２；配列番号５のヌクレオチド５４８３～６４８２；もしくは配列番号５のヌクレオチド５９８３～６４８２；またはこのような配列の相補鎖の、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。

【００７８】

配列番号５のヌクレオチド１～６４８３；配列番号５のヌクレオチド１４８３～６４８２；配列番号５のヌクレオチド２４８３～６４８２；配列番号５のヌクレオチド３４８３～６４８２；配列番号５のヌクレオチド４４８３～６４８２；配列番号５のヌクレオチド５４８３～６４８２；もしくは配列番号５のヌクレオチド５９８３～６４８２；またはこの

ような配列の相補鎖を含む単離されたポリヌクレオチドの各々は、ｇｐ３５４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。

【００７９】

（ｇｐ３５４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物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。

【００８０】

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く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。

【００８１】

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、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トログ）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。

【００８２】

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。

【００８３】

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：ｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖ／Ｂｌａｓｔ／ｂｌ２ｓｅｑ／ｂｌ２．ｈｔｍｌ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。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 ｆｏ

ｒ ａ ｍｉｓｍａｔｃｈ）：－２；（ｉｉｉ）オープンギャップ５およびエクステンションギャップ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１０ワードサイズ１１フィルター。

【００８４】

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。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変化しか生じない。

【００８５】

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。

【００８６】

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、進化的に関連する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）。

【００８７】

【表３】

	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

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Ｐ

ＡＭ２５０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。

【００８８】

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ＧＰ３５４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、ＧＰ３５４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。

【００８９】

ポリヌクレオチドの関係はまた、機能的試験を用いて、規定のハイブリダイゼーションストリンジェンシーにおいて２つのポリヌクレオチドが互いに塩基対を形成する能力が特徴付けられ得る。従って、本発明は、本明細書中で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であるだけでなく、本発明の種々の単離されたｇｐ３５４ポリヌクレオチド（「参照核酸」）の全てまたは一部分に高ストリンジェンシー条件（本明細書で規定されるような）下でハイブリダイズする単離されたポリヌクレオチド（「クロスハイブリダイゼーションする核酸」）を提供するための、単離されたポリヌクレオチドを提供する。

【００９０】

このようなクロスハイブリダイゼーションする核酸は、特に、本発明のタンパク質に関連するタンパク質（例えば、選択的スプライス改変体およびホモログ（例えば、オルトログおよびパラログ））のためのプローブ、およびこのようなタンパク質の発現を駆動するために有用である。特に有用なオルトログは、他の霊長類種由来（例えば、チンパンジー、アカゲザル、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）；ウサギ目由来（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）のオルトログである。

【００９１】

参照核酸のハイブリダイズする部分は、代表的に、少なくとも１５ヌクレオチド長であり、しばしば少なくとも１７、２０、２５、３０、３５、４０、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、１００ｎｔ、１５０ｎｔ、２００ｎｔ、２５０ｎｔ、３００ｎｔ、３５０ｎｔ、４００ｎｔ、４５０ｎｔ、５００ｎｔ以上であり、参照核酸の全長以下の部分）にハイブリダイズするクロスハイブリダイゼーションする核酸もまた、有用である。

【００９２】

クロスハイブリダイゼーションする核酸のハイブリダイズする部分は、参照核酸の少なくとも一部分と配列において少なくとも７５％同一である。代表的に、クロスハイブリダイゼーションする核酸のハイブリダイズする部分は、参照核酸の少なくとも一部分と配列において、少なくとも８０％、しばしば少なくとも８５％、８６％、８７％、８８％、８９％、または少なくとも９０％さえも同一である。しばしば、クロスハイブリダイゼーションする核酸のハイブリダイズする部分は、参照核酸配列の少なくとも一部分と配列において、少なくとも９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、または９９％同一である。時々、クロスハイブリダイゼーションする核酸のハイブリダイズする部分は、参照核酸の少なくとも一部分と配列において、少なくとも９９．５％同一である。

【００９３】

本発明はまた、本発明の種々の単離されたポリヌクレオチドまたは核酸のフラグメントを提供する。参照核酸の「フラグメント」により、本明細書ではこの参照核酸配列の一部分に対するヌクレオチド配列同一性を有する単離されたポリヌクレオチドまたは核酸（どのくらい獲得されていても）が意図される。この一部分は、少なくとも１７ヌクレオチドであり、そして参照核酸全体より短い。

【００９４】

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に１つ未満の頻度で無作為に生じるのに十分な長さのオリゴヌクレオチドであり、従って、哺乳動物ゲノムの複雑度を有する核酸混合物中の参照配列を特異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のオリゴヌクレオチドである。さらなる特異性は、部分ゲノムの複雑度の核酸サンプルをプローブすることにより、および／または１７ヌクレオチドと同じくらい短い長さの複数のフラグメントを集合的に用いて、核酸の増幅をプライムすることにより（例えば、ポリメラーゼ連鎖反応（ＰＣＲ）により）獲得され得る。

【００９５】

本発明の核酸プローブは、ホモログまたは同一な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するために使用され得る。このプローブは、それら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タンパク質を誤発現する（例えば、異常なスプライシング、異常なｍＲＮＡレベル）か、または（ｉｉ）ｇｐ３５４遺伝子中に変異（例えば、欠失、挿入、または点変異）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、標識試薬およびそれらの使用のための説明書を含む。

【００９６】

本発明の単離されたポリヌクレオチドはまた、ＰＣＲ、プライマー伸長などにおいてプライマーとして用いられ得る。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、８、９、または１０ヌクレオチド長）であり得る。このプライマーは、例えば、ｇｐ３５４　ｍＲＮＡまたはｃＤＮＡの増幅のために、ｇｐ３５４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、このプライマーは、ｇｐ３５４遺伝子のゲノム構造の非転写部分（例えば、調節部分）を利用するために、ｇｐ３５４遺伝子のイントロン配列または上流もしくは下流調節配列にハイブリダイズし得る。

【００９７】

本発明の核酸プライマーはまた、例えば、ＳＮＰ検出のための一塩基伸長（ＳＢＥ）をプライムするために使用され得る（例えば、本明細書中で、開示がその全体において参考として援用される、米国特許第６，００４，７４４号を参照のこと）。等温増幅アプローチ（例えば、ローリングサークル増幅）もまた、現在十分に記載されている。例えば、Ｓｃｈｗｅｉｔｚｅｒら、Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．１２（１）：２１－７（２００１）；米国特許第５，８５４，０３３号および同第５，７１４，３２０号、ならびに国際特許公開ＷＯ　９７／１９１９３およびＷＯ　００／１５７７９を参照のこと（これらの開示は、本明細書中でその全体において参考として援用される）。ローリングサークル増幅は、ＳＮＰ検出を容易にする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。

【００９８】

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）は、参照核酸により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示する上で有用である。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号および同第５，５９５，９１５号を参照のこと。

【００９９】

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フラグメント（すなわち、２４ヌクレオチド以上のフラグメント）は、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示する上で有用である。例えば、Ｌｅｒｎｅｒ、「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。

【０１００】
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、代表的には少なくとも１８ヌクレオチド長、そしてしばしば少なくとも２４、２５、３０、３５、４０、または４５ヌクレオチド（ｎｔ）長である。当然、少なくとも５０ｎｔ、１００ｎｔ、１５０ｎｔ、２００ｎｔ、２５０ｎｔ、３００ｎｔ、３５０ｎｔ、４００ｎｔ、４５０ｎｔ、５００ｎｔまたはそれ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、そして当業者に認識されるように、時には好ましい。
【０１０１】
発現されるメッセージの調査に基づくというよりもゲノム配列の採掘（ｍｉｎｉｎｇ）に基づき、本発明はさらに、ｇｐ３５４遺伝子の部分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ンプローブを提供する。このプローブは、エキソン（「参照エキソン」）の少なくとも一部分を含み、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、この単一エキソンプローブは、高ストリンジェンシー条件下で、参照エキソンを欠くがゲノムにおいて参照エキソンに隣接して見いだされる１つ以上のエキソンを含む核酸に検出可能にハイブリダイズしない。
【０１０２】
本発明はまた、ｇｐ３５４遺伝子の転写を制御する核酸配列エレメント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。転写制御配列として、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、サイレンサーなどが挙げられる。
【０１０３】
転写または翻訳のアンチセンス阻害における使用が所望される場合、あるいは酵素的核酸分子（例えば、リボザイム）のアンチセンス媒介標的化が所望される場合、本発明の単離されたポリヌクレオチドおよび核酸は、有用に、１つ以上の改変塩基（以下を参照のこと）および／または１つ以上の改変もしくは変更されたヌクレオシド間結合を含み得、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（Ｐｅｒｓｐｅｃｔｉｖｅｓ ｉｎ Ａｎｔｉｓｅｎｓｅ Ｓｃｉｅｎｃｅ），Ｋｌｕｗｅｒ Ｌａｗ Ｉｎｔｅｒｎａｔｉｏｎａｌ（１９９９）（ＩＳＢＮ：０７９２３８５３９Ｘ）；Ｓｔｅｉｎら（編），Ａｐｐｌｉｅｄ Ａｎｔｉｓｅｎｓｅ Ｏｌｉｇｏｎｕｃｌｅｏｔｉｄｅ Ｔｅｃｈｎｏｌｏｇｙ，Ｗｉｌｅｙ－Ｌｉｓｓ（ｃｏｖｅｒ（１９９８）（ＩＳＢＮ：０４７１１７２７９０）；Ｃｈａｄｗｉｃｋら（編），Ｏｌｉｇｏｎｕｃｌｅｏｔｉｄｅｓ ａｓ Ｔｈｅｒａｐｅｕｔｉｃ Ａｇｅｎｔｓ－Ｓｙｍｐｏｓｉｕｍ Ｎｏ．２０９，Ｊｏｈｎ Ｗｉｌｅｙ ＆ Ｓｏｎ Ｌｔｄ（１９９７）（ＩＳＢＮ：０４７１９７２７９７）を参照のこと。このような改変された塩基およびヌクレオシド間結合は、Ｇａｍｐｅｒら，Ｎｕｃｌ．Ａｃｉｄｓ Ｒｅｓ．２８（２１）：４３３２－９（２０００）（この開示は、本明細書中でその全体において参考として援用される）に記載されるように、本発明の単離された核酸が標的遺伝子修正のために使用される場合にもまた、しばしば所望される。
【０１０４】
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た酵素的核酸）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、その活性が、このベクターが導入される細胞型によって決定され得る。アンチセンス遺伝子を使用する遺伝子発現の調節に議論については、Ｗｅｉｎｔｒａｕｂら，Ａｎｔｉｓｅｎｓｅ ＲＮＡ ａｓ ａ ｍｏｌｅｃｕｌａｒ ｔｏｏｌ ｆｏｒ ｇｅｎｅｔｉｃ ａｎａｌｙｓｉｓ，ＲＥＶＩＥＷＳ－－ＴＲＥＮＤＳ ＩＮ ＧＥＮＥＴＩＣＳ，１巻（１）（１９８６）を参照のこと。
【０１０５】
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、相補性領域を有す

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、リボザイムは、ｇｐ３５４　ｍＲＮＡ転写物を触媒的に切断するために使用され、それにより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が、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（すなわち、配列番号１または３）に基づいて、設計され得る。

【０１０６】

ｇｐ３５４のオリゴヌクレオチド模倣物（例えば、ペプチド核酸（ＰＮＡ））が、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（１９９６）Ｂｉｏｏｒｇ．Ｍｅｄ．Ｃｈｅｍ．Ｌｅｔｔ．４：５－２３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、アミド含有骨格で（特に、アミド結合によって連結された反復Ｎ－（２－アミノエチル）グリシンユニットによって）置換される。例えば、ＰＮＡは、例えば、転写または翻訳の停止を誘導することによるか、または複製を阻害することによる、遺伝子発現の配列特異的調節のためのアンチセンス薬剤または抗遺伝子薬剤として使用され得る。ｇｐ３５４のＰＮＡはまた、例えば、ＰＮＡ指向性ＰＣＲクランピング（ＰＮＡ　ｄｉｒｅｃｔｅｄ　ＰＣＲ　ｃｌａｍｐｉｎｇ）による遺伝子中の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）と組み合わせて用いられる場合の人工制限酵素として；またはＤＮＡ配列決定およびハイブリダイゼーションのためのプローブまたはプライマーとして、用いられ得る（Ｈｙｒｕｐら、前出；およびＰｅｒｙ－Ｏ’Ｋｅｅｆｅ，前出）。ｇｐ３５４のＰＮＡは、親油性基または他のヘルパー基をＰＮＡに結合することによって、ＰＮＡ－ＤＮＡキメラの形成によってか、またはリポソームの使用もしくは当該分野で公知の薬物送達の他の技術（下記を参照のこと）によって、例えば、それらの安定性もしくは細胞取り込みを増強するために改変され得る。

【０１０７】

本発明のオリゴヌクレオチドは、ペプチド（例えば、インビボで宿主細胞レセプターを標的化するため）、または細胞膜もしくは血液脳関門を横切る輸送を容易にする薬剤のような他の付加基（ａｐｐｅｎｄｅｄ　ｇｒｏｕｐ）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因性切断剤（ｈｙｂｒｉｄｉｚａｔｉｏｎ　ｔｒｉｇｇｅｒｅｄ　ｃｌｅａｖａｇｅ　ａｇｅｎｔ）またはインターカレート剤で改変され得る。この目的のために、この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因性架橋剤、輸送剤、ハイブリダイゼーション誘因性切断剤など（下記を参照のこと））に結合体化され得る。

【０１０８】

天然に見出される核酸組成物との差異（例えば、非ネイティブ塩基、改変されたヌクレオシド間結合、合成後修飾）は、ｇｐ３５４ポリヌクレオチドの長さ全体を通して存在し得るか、または有用に、それらの異なる部分に配置され得る。後者の例として、別々のＤＮＡドメインおよびＲＮＡドメインを有するキメラ核酸が合成され得、米国特許第５，７６０，０１２号および同第５，７３１，１８１号にさらに記載されるように（これらの開示は、本明細書中でその全体において参考として援用される）、標的化された遺伝子の修復のための利用性を実証され得る。ＤＮＡおよびＰＮＡの両方を含むキメラ核酸は、改変ＰＣＲ反応において有用性を有することが実証されている。Ｍｉｓｒａら、Ｂｉｏｃｈｅｍ．３７：１９１７－１９２５（１９９８）を参照のこと；Ｆｉｎｎら、Ｎｕｃｌ．Ａｃｉｄｓ　Ｒｅｓ．２４：３３５７－３３６３（１９９６）（本明細書中に参考として援用される）もまた参照のこと。

【０１０９】

本発明のポリヌクレオチドおよび核酸はまた、有用に、基材に結合され得る。この基材は、多孔性または固体性、平板性または非平板性、一体的または分割された基板であり得；結合は、共有結合または非共有結合であり得る。基材に結合させた場合、本発明の核酸は、それらが標識されていない状態でプローブとして用いられ得る。例えば、本発明の核酸は、有用に、多孔性基材（通常は、膜（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を含む））に結合され得；このように結合された本発明の核酸は、標識された核酸サンプル内（ゲノム核酸のサンプルまたは転写物に由来する核酸のサンプルのいずれか）に存在するｇｐ３５４核酸を、例えば、逆ドットブロットにより検出するために用いられ得る。

【０１１０】

本発明の核酸はまた、有用に、固体基材（例えば、ガラス）に結合され得るが、他の固体材料（例えば、非晶質ケイ素、結晶性ケイ素、またはプラスチック）もまた用いられ得る。本発明の核酸は、支持基材の表面に共有結合され得るか、または推定される非共有結合的相互作用によって変性および接着を促進するカオトロピック剤において誘導体化した表面に適用され得るか、またはそれらのいくつかの組み合わせに適用され得る。

【０１１１】

本発明の核酸は、基材に結合され得、この基材に対して、複数の他の核酸が同時に結合され、この複数の結合した核酸の各々に対するハイブリダイゼーションは、別個に検出可能である。低密度では（例えば、多孔性膜上で）、これらの基材に結合した収集物は、代表的には、マクロアレイとよばれる；高密度では（代表的には、ガラスのような固体支持体上）、これらの基材に結合した複数の核酸の収集物は、通称、マイクロアレイと呼ばれる。本明細書中で用いられる場合、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。

【０１１２】

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用いて、ゲノム核酸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ｇｐ３５４核酸を検出し、特徴付け、そして定量し、そしてこれらのサンプルからｇｐ３５４核酸を単離し得る。溶液中で遊離している場合、このようなプローブは、代表的には（しかし、常にではない）、検出可能に標識されるか、マイクロアレイにおけるのと同様に基材に結合され、このようなプローブは、代表的には（しかし、常にではない）、標識されない。

【０１１３】

例えば、本発明の単離された核酸をプローブとして用いて、染色体スプレッド（ｃｈｒｏｍｏｓｏｍｅ　ｓｐｒｅａｄ）に対する蛍光インサイチュハイブリダイゼーション（ＦＩＳＨ）を通じて、ｇｐ３５４ゲノム遺伝子座における全改変（例えば、ｇｐ３５４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し得る。例えば、Ａｎｄｒｅｅｆｆら（編），Ｉｎｔｒｏｄｕｃｔｉｏｎ　ｔｏ　Ｆｌｕｏｒｅｓｃｅｎｃｅ　Ｉｎ　Ｓｉｔｕ　Ｈｙｂｒｉｄｉｚａｔｉｏｎ：Ｐｒｉｎｃｉｐｌｅｓ　ａｎｄ　Ｃｌｉｎｉｃａｌ　Ａｐｐｌｉｃａｔｉｏｎｓ，Ｊｏｈｎ　Ｗｉｌｅｙ＆Ｓｏｎｓ（１９９９）（ＩＳＢＮ：０４７１０１３４５５）を参照のこと（この開示は、その全体が本明細書中に参考として援用される）。本発明の単離された核酸は、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用いて、より小さなゲノム改変を評価するためにプローブとして用いられ得る。本発明の単離された核酸は、本発明の核酸を含むゲノムクローンを単離するためにプローブとして用いられ得る。このゲノムクローンは、その後、制限マッピングされ得、そして配列レベルで欠失、挿入、転座、および置換（一ヌクレオチド多型、すなわちＳＮＰ）を同定するために配列決定され得る。

【０１１４】

本発明の単離された核酸はまた、転写物由来の核酸サンプル中のｇｐ３５４核酸を検出し、特徴付け、そして定量し、そしてこのサンプルからｇｐ３５４核酸を単離するためにプローブとして用いられ得る。例えば、本発明の単離された核酸は、総ＲＮＡサンプルもしくはポリＡ＋で選択されたＲＮＡサンプルのノーザンブロットによりｇｐ３５４　ｍＲＮＡを検出し、長さにより特徴付けし、そして定量するために、ハイブリダイゼーションプローブとして用いられ得る。本発明の単離された核酸はまた、組織切片に対するインサイチュハイブリダイゼーションによりｇｐ３５４メッセージを検出し、位置により特徴付け、そして定量するために、ハイブリダイゼーションプローブとして用いられ得る（例えば、Ｓｃｈｗａｒｃｈｚａｃｈｅｒら，Ｉｎ　Ｓｉｔｕ　Ｈｙｂｒｉｄｉｚａｔｉｏｎ，Ｓ

ｐｒｉｎｇｅｒ－Ｖｅｒｌａｇ　Ｎｅｗ　Ｙｏｒｋ（２０００）（ＩＳＢＮ：　０３８７９１５９６６）（この開示は、その全体が本明細書中に参考として援用される）を参照のこと）。

【０１１５】

さらに、本発明の単離された核酸は、ｃＤＮＡライブラリーにおいてｇｐ３５４クローンの提示を測定するために、ハイブリダイゼーションプローブとして用いられ得る。例えば、本発明の単離された核酸は、ｃＤＮＡライブラリーからｇｐ３５４核酸を単離するために、ハイブリダイゼーションプローブとして用いられ得る。このことにより、ｇｐ３５４ＲＮＡメッセージの配列レベルでの特徴付け（欠失、挿入、短縮化の同定（選択的スプライシング形態におけるエクソンの欠失、挿入、および短縮化を含む）および一ヌクレオチド多型を含む）が可能になる。

本明細書中以下の実施例において記載されるように、本発明の核酸を使用して、ｇｐ３５４遺伝子の発現を測定するためにも、転写物由来のサンプル中のｇｐ３５４核酸を検出および定量し得る。ｇｐ３５４発現の測定は、異常なｇｐ３５４発現と関連した状態、障害および疾患（ｇｐ３５４が正常に発現される膵臓組織および神経組織においてｇｐ３５４が正常に発現される様式で、ならびに本明細書中以下の実施例においてさらに記載されるように、誤って発現され得る組織において）についての診断アッセイにおいて特定の有用性を有する。

【０１１６】

従って、当業者に容易に明らかであるように、ｇｐ３５４核酸プローブ各々は、標識されていようと、基材に結合されていようと、その両方であろうと、膵臓組織および神経組織（これらの組織では、発現が既に確認されている）におけるｇｐ３５４発現のレベルを測定するためのツールとして使用するために現在利用可能である。

【０１１７】

ｇｐ３５４を発現することがなおも実証されていない組織に関して、本発明のｇｐ３５４核酸プローブは、ｇｐ３５４核酸の存在を検出するために、例えば、膵臓もしくは神経系における異常なｇｐ３５４細胞発現または他の組織における異常な組織分布に関係する状態、障害または疾患を有して存在する患者の組織でのｇｐ３５４ＲＮＡ発現を検出するために、このような組織を調査するためのツールとして現在利用可能である。

【０１１８】

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ことにおいて有用であり；続いてこのマイクロアレイは、例えば、薬物発見および標的確認のプログラムにおいて遺伝子発現を測定し、そして調査するために有用な製品である。マイクロアレイ上に含まれる場合、ｇｐ３５４核酸プローブ各々により、ｇｐ３５４遺伝子の一部分がこのプローブ内に含まれることを検出し、従って、シグナルを検出する能力をマイクロアレイデバイスに与えるために、このマイクロアレイが特に有用となる。ここで、このようなプローブがない場合、このマイクロアレイデバイスは、シグナルがないことを報告する。

【０１１９】

ｇｐ３５４発現のレベルの変化は、有用性を有するための発現の測定に観察される必要はない。遺伝子発現分析が、細胞に対する化学的薬剤の毒性を評価するために用いられる場合、例えば、薬剤が遺伝子発現レベルを変化させることができないことは、この薬物が、その遺伝子の発現されたタンパク質が一部である経路に影響を与えないようであるという証拠である。同様に、遺伝子発現分析が、薬理学的薬剤の副作用を評価するために用いられる場合－主役（ｌｅａｄ）化合物の発見におけるか、または主役化合物の誘導体の引き続くスクリーニングにおけるかのいずれかにもかかわらず－薬剤が、遺伝子の発現レベルを改変できないことは、この薬物が遺伝子の発現されたタンパク質が一部分である経路に影響を与えないという証拠である。その全体が本明細書に参考として援用されるＷＯ９９／５８７２０は、その発現が計算で用いられる遺伝子の同一性または機能を問わずに、第１および第２の遺伝子発現プロフィールの関連性を定量する方法、および複数の遺伝子発現プロフィールの関連性を順序付ける方法を提供する。

【０１２０】
本発明のゲノム由来単一エキソンプローブおよびゲノム由来単一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可能にするさらなる有用性を有している。
【０１２１】
プロモーターをともな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ＲＮＡのインビトロ発現を駆動するために用いられ得る。このＲＮＡは、ｃＤＮＡ－ｍＲＮＡ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。ＧＰ３５４タンパク質またはその一部をコードするような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、ＧＰ３５４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。
【０１２２】
タンパク質発現がゲノムＤＮＡから行われる場合、代表的には、発現は、最初のＲＮＡ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ＤＮＡ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。
【０１２３】
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形態に特異的なイムノ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の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。
【０１２４】
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。
【０１２５】
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。
【０１２６】
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；６，２０４，０６１号；５，６３１，１５３号；５，６２７，０５９号；５，４８７，９９２号；５，４６４，７６４号；５，６１４，３９６号；５，５２７，６９５号および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 ＦａｃｔｓＢｏｏｋ、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。

【０１２７】

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、６，０４８，５２４号；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、５，８８８，９８３号、５，８７１，９８４号、５，７９５，９７２号、５，７８０，２９６号、５，７６０，０１２号、５，７５６，３２５号、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）；Ｇａｍｐｅｒ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。

【０１２８】

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。

【０１２９】

（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。

【０１３０】

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。

【０１３１】

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 Ｃｏｎｉｎｇ：Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｎｕａｌ（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。

【０１３２】

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、選択マーカーを欠いているが、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つの選択マーカーを含む。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ーは、所望の宿主細胞（単数または複数）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。

【０１３３】

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、（λ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、選択マーカーは、同様に、グラム陰性細菌中の選択性のために選択される：例えば、代表的なマーカーは、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；栄養要求性マーカーもまた用いられ得る。

【０１３４】

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ーを含む。

【０１３５】

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ＹＥｐ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。

【０１３６】

酵母ベクターにおける選択マーカーとしては、特異的な自己栄養性変異（例えば、ｕｒａ３－５２、ｈｉｓ３－Ｄ１、ｌｅｕ２－Ｄ１、ｔｒｐｌ－Ｄｌおよびｌｙｓ２－２０１）を補完する種々の自己栄養性マーカー（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α－アミノアジピン酸（αＡＡ）に、それぞれ基づくネガティブ選択をさらに可能にする。他の選択マーカーは、例えば、ゼオシン（ｚｅｏｃｉｎ）に対する耐性を与える。

【０１３７】

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細胞株、ならびにｅｘｐｒｅｓＳＦ^{TM}細胞（Ｐｒｏｔｅｉｎ Ｓｃｉｅｎｃｅｓ Ｃｏｒｐ．，Ｍｅｒｉｄｅｎ，ＣＴ，ＵＳＡ））に由来する場合、ベクター複製ストラテジーは、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。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、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。

【０１３８】

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、乳頭腫ウイルス起源、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、ウイルス複製ストラテジーに従って複製する。

【０１３９】

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）、ブラスチシジン、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。

【０１４０】
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。

【０１４１】
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、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。

【０１４２】
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、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）；Ｈｅｎｎｉｎｇ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。

【０１４３】
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。

【０１４４】
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。

【０１４５】
膵臓、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）、膵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。

【０１４６】
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、およびツー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。

【０１４７】
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。

【０１４８】
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。

【０１４９】
ＦｃＲｎレセプター（これはまた、ＦｃＲｐ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。

【０１５０】
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。

【０１５１】
（Ｂ．宿主細胞）
本発明は、細胞内でエピソーム的に存在するか、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。

【０１５２】
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。

【０１５３】
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　ｍｅｔｈａｎｏｌｉｃａ）；昆虫細胞株（例えば、Ｓｐｏｄｏｐｔｅｒａ　ｆｉｕｇｉｐｅｒｄａ由来

の細胞（例えば、Ｓｆ９細胞株およびＳｆ２１細胞株、ならびにｅｘｐｒｅｓＳＦ^{TM}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。

【０１５４】

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。

【０１５５】

（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のために適切なＧＰ３５４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ＧＰ３５４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、発色団対する）結合体がまた、提供される。

【０１５６】

いくつかの実施形態において、本発明は、全長ｇｐ３５４ ｃＤＮＡ（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ＧＰ３５４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ｇｐ３５４ ｃＤＮＡ（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ＧＰ３５４ポリペプチドを提供する。

【０１５７】

本発明また、ｇｐ３５４ ｃＤＮＡ（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ＧＰ３５４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ＧＰ３５４ポリペプチドは、少なくとも１つのＧＰ３５４の生物学的活性を有する。

【０１５８】

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ＧＰ３５４アミノ酸配列を含む、単離されたＧＰ３５４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のいずれかの全長ＧＰ３５４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ＧＰ３５４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個のアミノ酸、８個のアミノ酸、１５個のアミノ酸のから配列全

体までを有するフラグメントを提供する。

【０１５９】

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ＧＰ３５４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。

【０１６０】

従って、本発明は、ＧＰ３５４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ＧＰ３５４ポリペプチドを提供する（ＧＰ３５４ドメインおよびＧＰ３５４配列について、図１および図７ならびに配列番号２および配列番号８を参照のこと）。ＧＰ３５４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。

【０１６１】

本発明はまた、ＧＰ３５４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ＧＰ３５４ポリペプチドを提供する（ＧＰ３５４ドメインおよびＧＰ３５４配列について、図１および図７ならびに配列番号２および配列番号８を参照のこと）。ＧＰ３５４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、含まれる。

【０１６２】

好ましい実施形態において、単離されたＧＰ３５４ポリペプチドは、ＧＰ３５４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ＧＰ３５４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。ＧＰ３５４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。

【０１６３】

他の好ましい実施形態において、単離されたＧＰ３５４ポリペプチドは、異種タンパク質ドメインに融合されるＧＰ３５４Ｎ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ＧＰ３５４ポリペプチドは、少なくとも１つの細胞外Ｉｇドメイン、より好ましくは、２つのＧＰ３５４細胞外Ｉｇドメイン、そして最も好ましくは、３つ、４つ、または５つのＧＰ３５４細胞外Ｉｇドメインを含む。

【０１６４】

ＧＰ３５４の膜貫通ドメインおよびＧＰ３５４のＣ末端細胞質領域からなる群より選択されるＧＰ３５４フラグメントを含む、単離されたＧＰ３５４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ＧＰ３５４ポリペプチドは、異種タンパク質ドメインに融合される、ＧＰ３５４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。

【０１６５】

本発明のＧＰ３５４フラグメントは、ネイティブＧＰ３５４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ＧＰ３５４遺伝子およびタンパク質配列の知見が、ネイティブＧＰ３５４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。

【０１６６】

本発明はまた、ＧＰ３５４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；Ｉｇ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ＧＰ３５４抗体を惹起するために使用され得る。

【０１６７】

このような実施形態の第１連として、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ＧＰ３５４のＮ末端細胞外部分に特徴的なＩｇドメインを含む。具体的には、５つの細胞外Ｉｇドメインが、それぞれ、配列番号２のＧＰ３５４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ＧＰ３５４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２参照のこと）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、少なくともの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つ全てのドメイをコードする。

【０１６８】

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ＧＰ３５４シグナル分泌配列は、配列番号２のＧＰ３５４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ＧＰ３５４アミノ酸配列のアミノ酸１～２１（図７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、ＧＰ３５４由来である。

【０１６９】

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、ＧＰ３５４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。

【０１７０】

非常に所望される場合、上記の好ましいタンパク質フラグメントは、特定のシグナル伝達反応が、ＧＰ３５４結合相互作用に応答して所望される場合に、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、ＧＰ３５４に由来し得る（以下を参照のこと）か、またこのドメインは、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。

【０１７１】

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、ＧＰ３５４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ＧＰ３５４膜貫通ドメインは、配列番号２のＧＰ３５４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。

【０１７２】

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、ＧＰ３５４のＣ末端細胞内ドメインを有する。具体的には、ＧＰ３５４の１つの細胞内ドメインは、配列番号２のＧＰ３５４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。ＧＰ３５４の細胞内ドメインの別の形態は、配列番号８のＧＰ３５４アミノ酸配列のアミノ酸５３４～７０８（図７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ると考えられている。

【０１７３】

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ｇｐ３５４　ｃＤＮＡ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、ＧＸ１－２１８（配列番号１６）およびＧＸ１－２１９（配列番号１７；実施例２を参照のこと）を使用して膵臓ＲＮＡから増幅したＲＴ－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、ＧＰ３

５４の第１のＮ末端Ｉｇ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Ｉｇドメインのほとんどをコードしている。

【０１７４】

上述のように、本発明は、好ましいＧＰ３５４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）（ＳＥＱ　ＩＤ　ＮＯｓ：）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をさらに提供する。

【０１７５】

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、ＧＰ３５４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。

【０１７６】

（ＧＰ３５４タンパク質の特徴の他の定義付け）

図１は、ｇｐ３５４　ｃＤＮＡ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（ＳＮＰ）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。

【０１７７】

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に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に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に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って人の介入により生じ得る。

【０１７８】

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　ｓｅｑｕｅｎｃｅ－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、以下：ｈｔｔｐ　：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖ／Ｂｌａｓｔ／ｂｌ２ｓｅｑ／ｂｌ２．ｈｔｍｌ，
からオンラインで利用可能である。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ＢｌａｓｔＰ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。

【０１７９】

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に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。中程度に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。

【０１８０】

対立遺伝子改変は、例えば、位置１９５、１９６、５３９、および５４０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。

【０１８１】

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において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に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の改変体（「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」）を提供することは、本発明の別の局面である。

【０１８２】

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（ｌａｇｏｇｍｏｒｐｈ）由来；およびウシ、ブタ、ウマ、ヤギのような家畜由来のオルソログである。

【０１８３】

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。

【０１８４】

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）。

【０１８５】

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™ Ｉｎ－Ｆｒａｍｅ Ｌｉｎｋｅｒ Ｉｎｓｅｒｔｉｏｎ Ｋｉｔ，カタログ番号ＥＺＩ０４ＫＮ，Ｅｐｉｃｅｎｔｒｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ Ｃｏｒｐｏｒａｔｉｏｎ，Ｍａｄｉｓｏｎ，ＷＩ，ＵＳＡ）。

【０１８６】

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、ＧＰ３５４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ＧＰ３５４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使用され得る。

【０１８７】

本発明の単離されたタンパク質はまた、本発明のＧＰ３５４タンパク質の濃度および／または量を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ＥＬＩＳＡ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された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γ ＯｐｔＥＩＡ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γインターフェロンを含む）。

【０１８８】

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（ＳＥＬＤＩ）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（ＷＯ９８／５９３６２；ＷＯ９８／５９３６０；ＷＯ９８／５９３６１；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、ＢＩＡＣＯＲＥ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。

【０１８９】

本発明の単離されたタンパク質はまた、ＧＰ３５４の産生およびＧＰ３５４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。

【０１９０】

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、ＧＰ３５４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分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。

【０１９１】

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の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。

【０１９２】

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。

【０１９３】

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））。

【０１９４】

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。

【０１９５】

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、天然には、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ように、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。

【０１９６】

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、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示する細胞を操作する際に有用である。これは、順に、ＧＰ３５４または結合パートナーを発現する膵臓細胞またはＣＮＳ由来細胞；またはＧＰ３５４または結合パートナーを発現する膵臓細胞またはＣＮＳ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。

【０１９７】

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、ＩｇＧ　Ｆｃ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ＧＦＰ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。

【０１９８】

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、ＨＰＬＣ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、ＧＰ３５４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。

【０１９９】

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。

【０２００】

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。

【０２０１】

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表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、ＩｇＧ Ｆｃ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。

【０２０２】

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、炭素菌毒素致死因子、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。

【０２０３】

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。

【０２０４】

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。

【０２０５】

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（１

９９３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。

【０２０６】

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。

【０２０７】

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－－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）－－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。

【０２０８】

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。

【０２０９】

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．ＮａｔｌＡｃａｄ．Ｓｃｉ．ＵＳＡ　９６（９）：４７８０－５（１９９９）；Ｗａｎｇら、Ｓｃｉｅｎｃｅ　２９２（５５１６）；４９８－５００（２００１）。

【０２１０】

本発明の単離されたＧＰ３５３４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。

【０２１１】

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。別の例としては、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。

【０２１２】

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ε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。

【０２１３】

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）。

【０２１４】

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、ＰＭＰＩ、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。

【０２１５】

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。

【０２１６】

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。

【０２１７】

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。

【０２１８】

本発明の単離されたタンパク質の産生の後は、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、Ｔｈｏｒｎｅｒら（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。

【０２１９】

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。

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）。本発明のタンパク質はなお、ＨＰＬＣ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。

【０２２０】

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。

【０２２１】

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。

【０２２２】

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。

【０２２３】

本発明のＧＰ３５４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（ＰＶＤＦ）または陽イオン誘導体化された親水性ＰＶＤＦ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、本発明の固体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。

【０２２４】
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。
【０２２５】
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。
【０２２６】
（抗体および抗体産生細胞）
本発明は、本発明の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か、または本発明の単離されたＧＰ３５４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（フラグメントおよびその誘導体を含む）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、ＳＤＳ中での可溶化によって変性されるタンパク質に存在するかの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。
【０２２７】
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。
【０２２８】
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。
【０２２９】
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。
【０２３０】
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。
【０２３１】
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。
【０２３２】
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特異的な結合が可能なままそのフラグメントまたは誘導体をいう。
【０２３３】

「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。

【０２３４】

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、ヒト膵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。

【０２３５】

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1×10^{-6}モル濃度（Ｍ）、典型的には少なくとも約5×10^{-7}Ｍ、通常少なくとも約1×10^{-7}Ｍであり、少なくとも1×10^{-8}Ｍ、5×10^{-9}Ｍおよび1×10^{-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かっている。

【０２３６】

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。

【０２３７】

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。

【０２３８】

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。

【０２３９】

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

【０２４０】

本発明のＩｇＧ、ＩｇＭ、ＩｇＤ、ＩｇＥおよびＩｇＡ抗体はまた、通常、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、ウシおよびウマも）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。

【０２４１】

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。

【０２４２】

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体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）。

【０２４３】

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））。

【０２４４】

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。

【０２４５】

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 Ｂａｃｋｇｒｏｕｎｄ 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：０

８７９６９３１４２）；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。

【０２４６】

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）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。

【０２４７】

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ＩＩＩタンパク質（ｐＩＩＩ）または遺伝子ＶＩＩＩタンパク質（ｐＶＩＩＩ）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，Ｃｕｒｒｅｎｔ Ｏｐｉｎｉｏｎ ｉｎ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ ８：５０３－５０８（１９９７）；Ａｕｊａｍｅら，Ｈｕｍａｎ Ａｎｔｉｂｏｄｉｅｓ ８：１５５－１６８（１９９７）；Ｈｏｏｇｅｎｂｏｏｍ，Ｔｒｅｎｄｓ ｉｎ Ｂｉｏｔｅｃｈｎｏｌ．１５：６２－７０（１９９７）；ｄｅ Ｋｒｕｉｆら，１７：４５３－４５５（１９９６）；Ｂａｒｂａｓら，Ｔｒｅｎｄｓ ｉｎ Ｂｉｏｔｅｃｈｎｏｌ．１４：２３０－２３４（１９９６）；Ｗｉｎｔｅｒら，Ａｎｎ．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）（ＩＳＢＮ０－８７９６９－５４６－３）；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（１９９６，５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、ｓｃＦｖ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。

【０２４８】

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，Ｔａｋａｈａｓｈｉら，Ｂｉｏｓｃｉ．Ｂｉｏｔｅｃｈｎｏｌ．Ｂｉｏｃｈｅｍ．６４（１０）：２１３８－４４（２０００）；Ｆｒｅｙｒｅら，Ｊ．Ｂｉｏｔｅｃｈｎｏｌ．７６（２－３）：１５７－６３（２０００）；Ｆｉｓｃｈｅｒら，Ｂｉｏｔｅｃｈｎｏｌ．Ａｐｐｌ．Ｂｉｏｃｈｅｍ．３０（Ｐｔ２）：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）；Ｓｈｕｓｔａ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。

【０２４９】

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Ｎｅｓｂｉｔ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）。

【０２５０】

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）；Ｇａｖｉｌｏｎｄｏら，Ｂｉｏｔｅｃｈｎｉｑｕｅｓ ２９（１）：１２８－３８（２０００）；Ｆｉｓｃｈｅｒら，Ｊ Ｂｉｏｌ．Ｒｅｇｕｌ．Ｈｏｍｅｏｓｔ．Ａｇｅｎｔｓ １４（２）：８３－９２（２０００）；Ｆｉｓｃｈｅｒら，Ｂｉｏｔｅｃｈｎｏｌ．Ａｐｐｌ．Ｂｉｏｃｈｅｍ．３０（Ｐｔ２）：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）。

【０２５１】

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細胞および骨髄腫細胞が挙げられる。Ｖｅｒｍａら，Ｊ．Ｉｍｍｕｎｏｌ．Ｍｅｔｈｏｄｓ ２１６（１－２）：１６５－８１（１９９８）は、抗体の発現のための細菌、酵母、昆虫および哺乳動物発現系を概説および比較する。

【０２５２】

本発明の抗体はまた、Ｍｅｒｋら，Ｊ．Ｂｉｏｃｈｅｍ．（Ｔｏｋｙｏ）．１２５（２）：３２８－３３（１９９９）およびＲｙａｂｏｖａら，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Ｐｏｌｌｏｃｋ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。

【０２５３】

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。

【０２５４】

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）’$_2$、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。

【０２５５】

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、人間において免疫原性は低い。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インビボ投与にさらに適切である。

【０２５６】

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；Ｍｏｒｒｉｓｏｎ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。

【０２５７】

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））。

【０２５８】

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。

【０２５９】

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を提供することである。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、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。

【０２６０】

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、有用なためには、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、アルカリホスファターゼ、β－ガラクトシダーゼ、グルコースオキシダーゼ、セイヨウ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。

【０２６１】

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。

【０２６２】

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、有用なために、フルオロフォア（ｆｌｕｏｒｏｐｈｏｒｅ）で標識される。本発明の抗体に有用であるように結合し得る広範な種々のフルオロフォア標識が存在する。多くは、例えば、Ｍｏｌｅｃｕｌａｒ Ｐｒｏｂｅｓ，Ｉｎｃ．、Ｅｕｇｅｎｅ、ＯＲ、ＵＳＡから入手可能である。

【０２６３】

標識されたアビジン、ストレプトアビジン、カプタビジン（ｃａｐａｔａｖｉｄｉｎ）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、有用なためには、ビオチンで標識され得る。

【０２６４】

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、有用なためには、放射性同位体（例えば、^{33}Ｐ、^{32}Ｐ、^{35}Ｓ、^{3}Ｈおよび^{125}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}Ｔｈ、^{227}Ａｃ、^{225}Ａｃ、^{223}Ｒａ、^{213}Ｂｉ、^{212}Ｐｂ、^{212}Ｂｉ、^{211}Ａｔ、^{203}Ｐｂ、^{194}Ｏｓ、^{188}Ｒｅ、^{186}Ｒｅ、^{153}Ｓｍ、^{149}Ｔｂ、^{131}Ｉ、^{125}Ｉ、^{111}Ｉｎ、^{105}Ｒｈ、^{99m}Ｔｃ、^{97}Ｒｕ、^{90}Ｙ、^{90}Ｓｒ、^{88}Ｙ、^{72}Ｓｅ、^{67}Ｃｕまたは^{47}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。

【０２６５】

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、Ｐｓｅｄｏｍｏｎａｓ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。

【０２６６】

本発明の抗体は、有用なために、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、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。

【０２６７】

例えば、本発明の抗体は、有用なことには、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、ＮＨＳ活性化ＳｅｐｈａｒｏｓｅまたはＣＮＢｒ－活性化Ｓｅｐｈａｒｏｓｅ）に結合体化され得る。

【０２６８】

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－ストレプトアビジン相互作用によって、常磁性マイクロスフェアに結合体化され得、次いで、このマイクロスフェアは、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。

【０２６９】

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。

【０２７０】

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。

【０２７１】

（薬学的組成物および治療方法）

ＧＰ３５４は、主に膵臓で、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、膵臓、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、結合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。

【０２７２】

ＧＰ３５４に最も密接に関連する２つのタンパク質（Ｄｒｏｓｏｐｈｉｌａ不規則交差タンパク質（ＩＣＣＲ）およびヒトネフリン（ｎｅｐｈｒｉｎ）タンパク質（図２を参照のこと））は、両方共、発達のパターニング（ｐａｔｔｅｒｎｉｎｇ）および細胞－細胞連結に関連する。ＤｒｏｓｏｐｈｉｌａにおけるＩＣＣＲ遺伝子座での変異は、異常なアポトーシス活性におそらく少なくとも一部起因して、ハエにおいて感覚器官の発達に影響する（Ｒａｍｏｓ，Ｒ．Ｇら（１９９３）Ｇｅｎｅｓ Ｄｅｖ．７：２５３３～４７）。ネフリン遺伝子における変異は、ヒトにおいて先天性の腎炎を引き起こす（Ｋｅｓｔｉｌａ，Ｍ．ら（１９９８）Ｍｏｌ．Ｃｅｌｌ １：５７５～５８２）。ネフリンは、糸球体の細隙の横隔膜に位置し、そして、細胞接着において役割を果たすと考えられる（Ｒｕｏｔｓａｌａｉｎｅｎ，Ｖ．ら、（１９９９）Ｐｒｏｃ Ｎａｔｌ Ａｃａｄ Ｓｃｉ ９６：７９６２～７９６７）。ＧＰ３５４とこれら２つのタンパク質との間の類似性は、ＧＰ３５４もまた、同様の発達経路、および特に、膵臓および神経系における器官および組織の、発達および／または維持に関係する、シグナル伝達経路を誘発する、細胞－細胞相互作用において役割を果たすことを示唆する。

【０２７３】

膵臓に豊富なタンパク質として、ＧＰ３５４は、膵臓において、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、膵臓の損傷、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膵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。

【０２７４】

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、ＧＰ３５４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ＧＰ３５４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。ＧＰ３５４発現は、膵臓において比較的集中しているので、ＧＰ３５４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。

【０２７５】

本発明の組成物の投与が有効であ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、その全体において本明細書中で援用される）を参照のこと。

【０２７６】

ＧＰ３５４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ＣＮＳ組織においても検出される。従って、ＧＰ３５４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ＣＮＳ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ＧＰ３５４媒介神経移動、神経変性（例えば、ＧＰ３５４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。

【０２７７】

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、ＧＰ３５４調節因子（ＧＰ３５４アンチセンス試薬、ＧＰ３５４リガンドおよび抗ＧＰ３５４抗体を含む）の使用は、本発明の局面である。

【０２７８】

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、メチルセルロール、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[TM]）、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、蝋、油およびコロイド状シリカが挙げられる。

【０２７９】

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、ケルギン（ｋｅｌｇｉｎ）、カラゲナン、アカシア、ポリビニルピロリン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・矯臭剤を一緒に含む、溶液、乳濁液、シロップ、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。

【０２８０】

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。

【０２８１】

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。

【０２８２】

吸入および経皮的処方物もまた、容易に調製され得る。

【０２８３】

薬学的処方は、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：Ｇｅｎｎａｒｏ（編）、Ｒｅｍｉｎｇｔｏｎ：Ｔｈｅ Ｓｃｉｅｃｎｃｅ ａｎｄ Ｐｒａｃｔｉｃｅ ｏｆ Ｐｈａｒｍａｃｙ、第２０版、Ｌｉｐｐｉｎｃｏｔｔ、Ｗｉｌｌｉａｍｓ＆Ｗｌｉｋｉｎｓ（２０００）（ＩＳＢＮ：０６８３３０６４７２）；Ａｎｓｅｌら、Ｐｈａｒｍａｃｅｕｔｉｃａｌ Ｄｏｓａｇｅ Ｆｏｒｍｓ ａｎｄ Ｄｒｕｇ Ｄｅｌｉｖｅｒｙ Ｓｙｓｔｅｍｓ、第７版、Ｌｉｐｐｉｎｃｏｔｔ、Ｗｉｌｌｉａｍｓ＆Ｗｌｉｋｉｎｓ Ｐｕｂｌｉｓｈｅｒｓ（１９９９）（ＩＳＢＮ：０６８３３０５７２７）；およびＫｉｂｂｅ（編）、Ｈａｎｄｂｏｏｋ ｏｆ Ｐｈａｒｍａｃｅｕｔｉｃａｌ Ｅｘｃｉｐｉｅｎｔｓ、Ａｍｅｒｉｃａｎ Ｐｈａｒｍａｃｅｕｔｉｃｌ Ａｓｓｏｃｉａｔｉｏｎ、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。

【０２８４】

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、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投与され得る。

【０２８５】

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１ｋｇ当たり、０．０１ｍｇ～３０ｍｇの毎日の投薬量（例えば、１ｍｇ／ｋｇ～５ｍｇ／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。

【０２８６】

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ラベルおよび手引書を有するキットの一部として、含まれ得る。

【０２８７】

（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。

【０２８８】

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。

【０２８９】

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。

【０２９０】

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。

【０２９１】

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。

【０２９２】

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。

【０２９３】

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、または１１のヒトｇｐ３５４ＤＮＡ配列を含む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）。

【０２９４】

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、他のトランシ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。

【０２９５】

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。

【０２９６】

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遺伝子が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、数キロベースの隣接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（１９８７）Ｃｅｌｌ　５１：５０３を参照のこと。

【０２９７】

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（１９９２）Ｃｅｌｌ　６９：９１５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、ＴＥＲＡＴＯＣＡＲＣＩＮＯＭＡＳ　ＡＮＤ　ＥＭＢＲＹＯＮＩＣ　ＳＴＥＭ　ＣＥＬＬＳ：Ａ　ＰＲＡＣＴＯＣＡＬ　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。

【０２９８】

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（１９９１）Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．２：８２３－８２９；ＰＣＴ国際公開番号ＷＯ９０／１１３５４；ＷＯ９１／０１１４０；ＷＯ９２／０９６８；およびＷＯ９３／０４１６９。

【０２９９】

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（１９９７）Ｎａｔｕｒｅ　３８５：８１０－８１３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。

【０３００】

インビボの導入遺伝子の調節された発現を、制御可能な組換え系（例えば、ｃｒｅ／ｌｏｘＰ組換え系（例えば、Ｌａｋｓｏら（１９９２）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ　８９：６２３２－６２３６を参照のこと）およびＦＬＰ組換え系（Ｏ’Ｇｏｒｍｅｎら（１９９１）Ｓｃｉｅｎｃｅ　２５１：１３１５－１３５５））を使用して達成し得る。ｃｒｅ／ｌｏｘＰ組換え系を使用して導入遺伝子の発現を調節する場合、Ｃｒｅリコンビナーゼおよび選択されたタンパク質の両方をコードする導入遺伝子を含む動物が必要である。この系の両方のエレメントを含むトランスジェニック動物を、例えば、2匹のトランスジェニックマウス（各々が、選択されたタンパク質をコードする導入遺伝子またはリ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）を交尾させることによって、獲得し得る。

【０３０１】

（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。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、および／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。

【０３０２】

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ー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ｇｐ３５４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ＧＰ３５４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ＧＰ３５４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。

【０３０３】

本発明のアンチセンス核酸分子は、ＧＰ３５４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ＧＰ３５４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）。

【０３０４】

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、ＴＡＴＡ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、ｇｐ３５４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、ｇｐ３５４　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、ｇｐ３５４　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。

【０３０５】

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、天然の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。

【０３０６】

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、目的の標的核酸へのアンチセンス方向である）。

【０３０７】

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、ＧＰ３５４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ｇｐ３５４発現の抑制は、異常なｇｐ３５４発現によって特徴付けられる疾患／状態に対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。

【０３０８】

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、５’末端に、ポリ－Ｌ－リジン、トランスフェリン、ポリリジン、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。

【０３０９】

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、組織部位へ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極力な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。

【０３１０】

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－アノマー核酸分子である。ａ－アノマー核酸分子は、通常のｂ－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（１９８７）Ｎｕｃｌｅｉｃ　Ａｃｉｄｓ　Ｒｅｓ　１５：６６２５－６６４１）。アンチセンス核酸分子はまた、２’－Ｏ－メチルリボヌクレオチド（Ｉｎｏｕｅら（１９８７）Ｎｕｃｌｅｉｃ　Ａｃｉｄｓ　Ｒｅｓ　１５：６１３１－６１４８）またはキメラＲＮＡ－ＤＮＡアナログ（Ｉｎｏｕｅら、（１９８７）ＦＥＢＳ　Ｌｅｔｔ　２１５：３２７－３３０）を含み得る。

【０３１１】

（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。

【０３１２】

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　ｍＲＡＮ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、（１９９３）Ｓｃｉｅｎｃｅ　２６１：１４１１－１４１８を参照のこと。

【０３１３】

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、標的細胞におけるｇｐ３５４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．（１９９１）Ａｎｔｉｃａｎｃｅｒ　Ｄｒｕｇ　Ｄｅｓ．６：５６９－８４；Ｈｅｌｅｎｅ．ら、（１９９２）Ａｎｎ．Ｎ．Ｙ．Ａｃａｄ．Ｓｃｉ．６６０：２７－３６；およびＭａｈｅｒ（１９９２）Ｂｉｏａｓｓａｙｓ　１４：８０７－１５を参照のこと。

【０３１４】

（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、（１９９６）Ｂｉｏｏｒｇ．Ｍｅｄ．Ｃｈｅｍ．Ｌｅｔｔ．４：５－２３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。

【０３１５】

ｇｐ３５４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原剤として使用され得る。ｇｐ３５４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）。

【０３１６】

他の実施形態において、ｇｐ３５４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、ｇｐ３５４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、Ｎｕｃ．Ａｃｉｄｓ　Ｒｅｓ．２４：３３５７－６３（１９９６）に記載されるように実施され得る。

【０３１７】

例えば、ＤＮＡ鎖は、標準的なホスホラ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－チミジンホスホラミダイト）は、ＰＮＡとＤＮＡの５’末端との間に使用され得る（Ｍａｇら、Ｎｕｃ．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。

【０３１８】

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、Ｌｅｔｓｉｎｇｅｒら、Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ　８６：６５５３－６５５６（１９８９）；Ｌｅｍａｉｔｒｅら、Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ　８４：６４８－６５２（１９８７）；ＰＣＴ公開番号ＷＯ８８／０９８１０を参照のこと）または血液－脳関門（例えば、ＰＣＴ公開番号ＷＯ８９／１０１３４を参照のこと）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（ｔｒｉｇｇｅｒｅｄ）切断剤（例えば、Ｋｒｏｌら、ＢｉｏＴｅｃｈｎｉｑｕｅｓ　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、Ｚｏｎ、Ｐｈａｒｍ．Ｒｅｓ．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。

【０３１９】

ＰＮＡ化学および適用は、特に、Ｒａｙら、ＦＡＳＥＢ　Ｊ．１４（９）：１０４１－６０（２０００）；Ｎｉｅｌｓｅｎら、Ｐｈａｒｍａｃｏｌ　Ｔｏｘｉｃｏｌ．８６（１）：３－７（２０００）；Ｌａｒｓｅｎら、Ｂｉｏｃｈｉｍ　Ｂｉｏｐｈｙｓ　Ａｃｔａ　１４８９（１）：１５９－６６（１９９９）；Ｎｉｅｌｓｅｎ，Ｃｕｒｒ．Ｏｐｉｎ．Ｓｔｒｕｃｔ．Ｂｉｏｌ．９（３）：３５３－７（１９９９）、およびＮｉｅｌｓｅｎ，Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．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）において総説がなされる。

【０３２０】

（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、正常に発現される組織（例えば、膵臓およびＣＮＳ）、およびこのような異常組織ミス発現が疑われる場合、正常に発現されない組織におけるｇｐ３５４　ｍＲＮＡのレベルを評価するための核酸プローブとして使用され得る。

【０３２１】

従って、本発明は、サンプルを、ｇｐ３５４ポリヌクレオチド配列をハイブリダイゼーションによって特異的に検出し得る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、生物学的サンプル（例えば、細胞抽出物、患者由来の流体または組織サンプル）におけるｇｐ３５４ポリヌクレオチドの存在を検出するための方法を提供する。

【０３２２】

好ましくは、この方法は、高度に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、サンプルを単離された核酸と接触させる工程、および単離されたポリヌクレオチドのサンプル中の核酸へ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存在は、サンプル中のｇｐ３５４コード配列の存在を示す。

【０３２３】

本発明の単離されたポリヌクレオチドは、これらの組織においてｇｐ３５４の特異的発現に基づく、膵臓および中枢神経系において特定の細胞型に特異的な核酸プローブとして使

用され得る。従って、本発明は、サンプルを、ｇｐ３５４ポリヌクレオチド配列をハイブリダイゼーションによって特異的に検出し得る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、生物学的サンプル（例えば、細胞抽出物、患者由来の流体または組織サンプル）におけるｇｐ３５４ポリヌクレオチドの存在を検出することによって、膵臓細胞または神経細胞として細胞を同定するための方法を提供する。

【０３２４】

本発明はまた、以下の少なくとも１つによって特徴付けられる遺伝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する：（ｉ）ＧＰ３５４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）ＧＰ３５４タンパク質をコードする遺伝子のミス調節；および（ｉｉｉ）ＧＰ３５４タンパク質の異常な翻訳後修飾。ここで、野生型形態の遺伝子は、ＧＰ３５４生物学的活性を有するタンパク質をコードする。

【０３２５】

本発明は、ヒトｇｐ３５４遺伝子のホモログを同定する方法をさらに提供し、この方法は、核酸ライブラリーを、配列番号１、３、５、６、７、９または１１を含む核酸プローブ、または少なくとも１７ヌクレオチドを有するその部分と、中程度または高度に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ハイブリダイズする工程；およびこの核酸プローブがライブラリー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する。ライブラリー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ｇｐ３５４遺伝子のホモログである。

【０３２６】

（Ｂ．抗体診断）

本発明の抗体は、正常に発現される組織（例えば、膵臓およびＣＳＮ）、およびこのような異常組織ミス発現が疑われる場合、正常に発現されない組織において、ＧＰ３５４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。

【０３２７】

従って、本発明は、サンプルを、ＧＰ３５４タンパク質の存在またはその活性のインジケーターを検出し得る薬剤と接触させることによって、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）においてＧＰ３５４タンパク質の存在またはその活性を検出するための方法を提供する。好ましくは、この薬剤は、ＧＰ３５４タンパク質の少なくとも１つのエピトープに特異的な抗体である。

【０３２８】

従って、本発明は、ＧＰ３５４タンパク質がサンプル中に存在するか否かを決定するための方法を提供し、この方法は、サンプルを、少なくとも１つのＧＰ３５４エピトープを有する抗体と接触させる工程、および抗体の抗原への特異的結合を検出する工程を包含し、これは、サンプル中のＧＰ３５４タンパク質の存在を示す。

【０３２９】

上記方法はまた、被験体試験細胞由来の生物学的サンプル（例えば、細胞抽出物、被験体由来の流体または組織サンプル）に存在するＧＰ３５４ポリペプチドの量を、非膵臓組織かつ非神経組織由来の平行な生物学的サンプルに存在するＧＰ３５４ポリペプチドの量を比較することによって、被験体の試験細胞を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。

【０３３０】

（Ｃ．疾患を診断するための方法）

本発明のｇｐ３５４単離ポリヌクレオチド、タンパク質およびＧＰ３５４特異的抗体は、異常なｇｐ３５４発現と関連する種々の障害状態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。

【０３３１】

従って、本発明は、被験体の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ＧＰ３５４タンパク質の量または活性を、コントロールサンプル（例えば、健康な被験体由来の等価なもの）におけるＧＰ３５４ポリペプチドの量また

は活性と比較する工程を包含し、ここで、コントロールサンプルにおけるＧＰ３５４ポリペプチドの量または活性と比較した組織サンプルにおけるＧＰ３５４ポリペプチドの量または活性における有意な違いは、この被験体が疾患状態を有することを示す。

【０３３２】

好ましい実施形態において、組織サンプルにおけるＧＰ３５４タンパク質の量または活性は、本発明のＧＰ３５４ポリペプチドまたはフラグメントを使用する競合結合アッセイによって、あるいは本発明のＧＰ３５４特異的抗体を使用するイムノアッセイによって評価される。好ましくは、この方法は、膵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。

【０３３３】

異なる組織における相対的なｇｐ３５４　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提供される。好ましくは、この方法は、この被験体由来の試験組織サンプルにおけるｇｐ３５４　ｍＲＮＡの量を、コントロールサンプルにおけるｇｐ３５４　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における量と比較した試験サンプルにおけるｍＲＮＡの量における有意な差は、この被験体が疾患状態を有することを示す。

【０３３４】

好ましい実施形態において、組織サンプルにおけるｇｐ３５４　ｍＲＮＡの量は、本発明の単離されたｇｐ３５４ポリヌクレオチドまたは核酸フラグメントを使用して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、この方法は、膵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。

【０３３５】

（コンピューター読み取り可能な手段）

本発明のさらなる局面は、本発明のｇｐ３５４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な手段である。好ましい実施形態において、本発明は、本明細書中に記載される配列番号を、配列の完全なセットとしてまたは任意の組み合わせで保存するためのコンピューター読み取り可能な手段を提供する。コンピューター読み取り可能な手段の記録は、読み取りおよびディスプレイのため、特定の基準に合うデータについての問い合わせ（ｑｕｅｒｙ）におけるデータの位置付け、配列の比較、セットの基準に合う配列のアライメントまたは順序付けを可能にするプログラムの適用のためのコンピュータシステムとのインターフェースのためなどにアクセスされ得る。

【０３３６】

本発明の核酸配列およびアミノ酸配列は、検索分析および配列分析アルゴリズムにおいて有用なデータベースの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。

【０３３７】

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。

【０３３８】

従って、本発明は、ヒトｇｐ３５４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ｇｐ３５４遺伝子のホモログである。

【０３３９】

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（ｃｌａｄｉｓｔｉｃ）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（ｃａｌｌｉｎｇ）、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。

【０３４０】

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。

【０３４１】

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。

【０３４２】

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。

【０３４３】

（実施例）

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、本発明の精神および範囲内であることは、当業者に明らかである。

【０３４４】

以下に記載される実験について、全てのＲＴ－ＰＣＲ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取り、そして製造者の説明書（Ｑｉａｇｅｎ、Ｖａｌｅｎｃｉａ、ＣＡ）に従い、ＱＩＡＥＸ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、コンピテントなＴＯＰ１０　Ｅ．ｃｏｌｉ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。

【０３４５】

（実施例１．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およびＫａｒｌｉｎ、Ｊ．Ｍｏｌ．Ｂｉｏｌ．２６８：７８－９４（１９９７））ならびにＧＥＮＥＭＡＲＫＨＭＭ（ＬｕｋａｓｈｉｎおよびＢｏｒｏｄｏｖｓｋｙ、Ｎｕｃ．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タウンロードし、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）。

【０３４６】
ＧＥＮＳＣＡＮおよびＧＥＮＥＭＡＲＫＨＭＭから同様の結果が得られた遺伝子を、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およびＳｗｉｓｓｐｒｏｔの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。

【０３４７】
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（１９９９））。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（ｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖ／）を介して公共利用可能である。このアルゴリズムは、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、最大の達成値から問い合わせ配列Ｘに出くわす場合；（２）この累積的スコアが、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））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。

【０３４８】
ＢＬＡＳＴ（Ｋａｒｌｉｎら、Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ、９０、５８７３－５７８７（１９９３））およびＧＡＰＰＥＤ ＢＬＡＳＴは、２つの配列間の類似性の統計的な分析を行う。ＢＬＡＳＴ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ｇｐ３５４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ｇｐ３５４の遺伝子またはｃＤＮＡと類似しているとみなされる。

【０３４９】
このｇｐ３４５遺伝子（ＯＲＦ）を、ＧｅｎＢａｎｋ登録番号ＡＣ０２２３１５（これは

、２０００年２月１０日に寄託された）のＢＡＣ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ＧＥＮＳＣＡＮ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。

【０３５０】

【表４】

表３：GENSCANの結果

エキソン	始め	終り	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

公共利用可能なＥＳＴのデータベースに対するｇｐ３５４遺伝子のＢＬＡＳＴ分析は、ＥＳＴが予測された遺伝子と一致しないことを示した。

【０３５１】

（実施例２：ｇｐ３５４の増幅）

ｇｐ３５４ｃＤＮＡの配列を、ＭＡＲＡＴＨＯＮ－ＲＥＡＤＹ　ＲＡＣＥキット（Ｃｌｏｎｔｅｃｈ、Ｐａｌｏ　Ａｌｔｏ、ＣＡ）を使用して、ｃＤＮＡ末端の迅速な増幅（ＲＡＣＥ）を実施することによって得る。ＭＡＲＡＴＨＯＮ－ＲＥＡＤＹ　ｃＤＮＡは、ヒトの組織のｍＲＮＡから合成した二重鎖ｃＤＮＡであり、これを、標準的なアダプターのセット（Ｃｌｏｎｔｅｃｈ）に連結した。全てのＲＡＣＥ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ＡＰ－１、５’－ＣＣＡＴＣＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣ－３’（配列番号１４）を使用する。ｇｐ３５４についての３’ＲＡＣＥは、順方向プライマーＧＸ１－２１８、５’－ＴＡＣＴＧＧＧＧＧＣＴＡＧＴＴＣＡＧＴＧＧＡＣＴＡＡ－３’（配列番号１６）、または逆方向プライマーＧＸ１－２１９の相補体、５’－ＣＣＡＡＡＣＡＧＣＡＣＡＴＣＣＡＧＣＧＣＡＧＴＡＣ－３’（配列番号１７）と共にＡＰ－１を使用し得る。ｇｐ３５４についての５’ＲＡＣＥは、逆方向プライマーＧＸ１－２１９、または逆方向プライマーＧＸ１－２１８の相補体と共にＡＰ－１を使用し得る。ＡＤＶＡＮＴＡＧＥ　２　ＤＮＡポリメラーゼ（Ｃｌｏｎｔｅｃｈ）を、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、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；そして６８℃で１０分間。

【０３５２】

（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。このＰＣＲを、複数の組織のｃＤＮＡパネル（逆転写ＰＣＲ（「ＲＴ－ＰＣＲ」）によって、これらの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。複数の組織のｃＤＮＡパネルより、ＰＣＲのためのテンプレートとして二重鎖のヒトｃＤＮＡを提供した。ＧＸ１－２１８およびＧＸ１－２１９（前出）を、ＰＣＲのためのプライマーとして使用した。ＰＣＲのための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）。

【０３５３】

この複数のヒト組織のｃＤＮＡパネルは、以下の組織由来のｃＤＮＡを含んだ：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。これらの結果を、表３に示す。約７８５ｂｐのバンドは、膵臓において観測され、そして他の組織中において観測されなかった。

【０３５４】

膵臓由来のＰＣＲフラグメントを、ＰＣＲ２．１プラスミドベクター（Ｉｎｖｉｔｒｏｇｅｎ、Ｃａｒｌｓｂａｄ、ＣＡ）中にクローニングした。得られたプラスミド構築物ＣＳ００２６（ＡＴＣＣ登録番号ＰＣＴ－４４５０；２００２年６月１１日に寄託された）を増殖し、そして挿入物を、上記のように配列決定した。この配列を、配列番号３として示す。

【０３５５】

（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配列の比率を得るために、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ｇｐ３５４ｃＤＮＡ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。

【０３５６】

（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、２μｇのポリＡ ＲＮＡを含むブロット（Ｃｌｏｎｔｅｃｈカタログ番号７７６０－１）の各レーンで処理した。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。５０ｎｇ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℃、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×ＳＳＣ／０．０５％ＳＤＳで洗浄し、続いて、５０℃の０．１×ＳＳＣ／０．１％ＳＤＳで２回洗浄した。膵臓ｃＤＮＡのＰＣＲにおける場合、約７８５ｂｐのバンドを、ノーザンブロット中で観測した。他の組織は、ＧＰ３５４の発現を示さなかった（図４）。

【０３５７】

（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コード領域のサブスクリーニング）

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、２μｌのヒトゲノムライブラリー（約10^8 ＰＦＵ／ｍｌ）（Ｃｌｏｎｔｅｃｈ）を、６ｍｌのＫ８０２細胞（一晩培養）（Ｃｌｏｎｔｅｃｈ）に添加し、次いで、２５０ｍｌ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℃でインキュベートする。７ｍｌ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＋１０ｍｌの$MgSO_4$プレート上に注ぎ、そして３７℃で一晩インキュベートした。５ｍｌのＳＭファージ緩衝液（０．１Ｍ ＮａＣｌ、８．１ｍＭ $MgSO_4$・$7H_2O$、５０ｍＭ Ｔｒｉｓ・Ｃｌ（ｐＨ７．５）、０．０１％ゼラチン）を各プレート（５ｍｌ）に添加し、そしてアガロースの上部を、顕微鏡用のスライドを用いて取り除き、そして５０ｍｌの遠心管に配置する。一滴のクロロホルムを添加し、そしてこのチューブを、３７℃の振盪器に１５分間配置し、次いで、２０分間、４０００ｒｐｍ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℃で保存した。

【０３５８】

次いで、ＰＣＲを、以下を含む２０ｍｌ中で実施した：８．８ｍｌのH_2O、４ｍｌの５×ＲＡＰＩＤ－ＬＯＡＤ ＢＵＦＦＥＲ（Ｏｒｉｇｅｎｅ）、２ｍｌの１０×ＰＣＲ ＢＵＦＦＥＲ ＩＩ（Ｐｅｒｋｉｎ Ｅｌｍｅｒ）、２ｍｌの２５ｍＭの$MgCl_2$、０．８ｍｌの１０ｍＭ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、０．１２ｍｌのプライマー（１ｍｇ／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、０．１２ｍｌのプライマー（１ｍｇ／ｍｌ）、０．２ｍｌ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２ｍｌのファージ溶液。このＰＣＲ反応は、以下を含む：８０℃で２０分間、９５℃で１０分間の１サイクル；次いで、９５℃で３０秒間、７２℃で４分間（サイクル毎に１℃低下させる）、６８℃で２分間の２２サイクル、続いて、９５℃で３０秒間、５５℃で３０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。

【０３５９】

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＋１０ｍｌの$MgSO_4$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。フィルターを取り除き、ペトリ皿の上側にファージを配置し、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細胞および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。８ｍ

ｌのＳＭを、各プレートに添加し、そしてアガロースの上部を、顕微鏡用のスライドで擦り落し、そして遠心管中に収集する。

【０３６０】

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。

【０３６１】

ｇｐ３５４遺伝子のコード配列をサブクローニングするために、ＰＣＲを、以下を含む５０μｌの反応物中で実施する：３３μｌのH_2O、５μｌの１０×ＴＴ緩衝液（１４０ｍＭの硫化アンモニウム、０．１％のゼラチン、０．６ＭのＴｒｉｓ－トリシン　ｐＨ８．４）、５μｌの１５ｍＭｎＭｇSO_4、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。

【０３６２】

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、ゲルから切り取り、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、エタノールで沈殿させ、そしてライゲーションのために１２μｌのH_2O中で再懸濁させた。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。

【０３６３】

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。

【０３６４】

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。

【０３６５】

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た。サイクル－配列決定反応物は、１４ｍｌのH_2O、１６ｍｌのＢＥＧＤＹＥ　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熱サイクラー中で、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秒間の９９サイクル、そして６０℃で４分間。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。

【０３６６】

ｐＣＲＩＩにサブクローニングした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－配列決定反応物は、６ｍｌのH_2O、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熱サイクラー中で、９６℃で１０秒間、５０℃で１０秒間の２５サイクル、そして６０℃で４分間により実施する。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。

【０３６７】

（実施例７：脳におけるＧＰ３６４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、インサイチュの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。膵臓における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、矢状および冠状のラットの膵臓凍結切片（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）に凍結固定し、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空気乾燥させる。他の組織を、同様の様式でアッセイし得る。

【０３６８】

ｇｐ３５４特異的プローブを、ＰＣＲ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。ＰＣＲ増幅後、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ｐＢｌｕｅｓｃｒｉｐｔ　ＩＩにクローン化する。ｇｐ３５４のセンス鎖に特異的なプローブの生成のために、ｐＢｌｕｅｓｃｒｉｐｔ　ＩＩにクローン化されたクローン化ｇｐ３５４フラグメントを、適切な制限酵素で直線化し得、これはベクター性Ｔ７プロモーターおよび市販のＴ７　ＲＮＡポリメラーゼを使用して、標識されたランオフ（ｒｕｎ－ｏｆｆ）転写物（すなわち、ｃＲＮＡリボプローブ）に対する基質を提供する。ｇｐ３５４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、Ｔ３プロモーターおよび同族ポリメラーゼを使用して、標識されたランオフｃＲＮＡ転写物の生成のための直線化基質を生成することにより、ｐＢｌｕｅｓｃｒｉｐｔ中のｇｐ３５４を使用して容易に調製し得る。

【０３６９】

リボプローブを、［^{35}Ｓ］－ＵＴＰで標識して、アンチセンスリボプローブについて、約０．４０×10^6ｃｐｍ／ｐｍｏｌの比活性、およびセンス鎖リボプローブについて、

約０．６５×１０^{6} ｃｐｍ／ｐｍｏｌの比活性を生じ得る。続いて、各リボプローブを変性し、そして５０％ ホルムアミド、１０％ デキストラン、０．３Ｍ ＮａＣｌ、１０ｍＭ Ｔｒｉｓ（ｐＨ８．０）、１ｍＭ ＥＤＴＡ、１×デンハルト溶液、および１０ｍＭ 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する（２ｐｍｏｌ／ｍｌ）。

【０３７０】

得られた膵臓凍結切片を含む顕微鏡用スライドを、スライド１つあたり４５μｌのハイブリダイゼーション溶液に単独で暴露し、そしてハイブリダイゼーション溶液に暴露されている部分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℃で一晩（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、等量の凍結切片を、センスまたはアンチセンスのｇｐ３５４特異的ｃＲＮＡリボプローブに暴露する。

【０３７１】

ハイブリダイゼーション期間の後、このスライドのカバーガラスを１×ＳＳＣ中で洗浄し、続いてこのスライドを、３７℃で４５分間、１０ｍＭ Ｔｒｉｓ・ＨＣｌ（ｐＨ７．４）、０．５Ｍ ＥＤＴＡおよび０．５Ｍ ＮａＣｌを含む緩衝液中２０μｇ／ｍｌのＲＮａｓｅ Ａに暴露することによって、ＲＡａｓｅ 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×ＳＳＣ中、５２℃の３回の高い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暴露し、続いて空気乾燥し、そしてＫＯＤＡＫ ＢＩＯＭＡＸ ＭＲ－１フィルムに暴露することによって、この凍結切片を脱水する。暴露の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。

【０３７２】

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に、ＫＯＤＡＫ ＮＴＢ－２核トラックエマルジョンをコーティングし、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部分を、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グレイン（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グレインは、プローブの非特異的結合を示す。

【０３７３】

膵臓および脳におけるＧＰ３５４の発現（下記）は、ＧＰ３５４活性の調節は、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効能を提供する。

【０３７４】

（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行って、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、^{32}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は、以下のように行う：

（１）以下の試薬を混合する：

１μｌ ｇｐ３５４含有プラスミド

２μｌ 順方向プライマー

２μｌ 逆方向プライマー

１０μｌ 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

１μｌ　１０ｍＭ　ｄＮＴＰ（例えば、Ｂｏｅｈｒｉｎｇｅｒ　Ｍａｎｎｈｅｉｍカタログ番号１９６９０６４）
０．５μｌ　Ｔａｑポリメラーゼ（例えば、Ａｍｅｒｓｈａｍ　Ｐｈａｒｍａｃｉａ　Ｂｉｏｔｅｃｈカタログ番号２７－０７９９－６２）
８３．５μｌ　水。
【０３７５】
（２）以下のプログラムを使用して、サーモサイクラーでＰＣＲを行う：
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；および７２℃で１０分。
【０３７６】
ＰＣＲ産物を、ＱＩＡＱＵＩＣＫ　ＰＣＲ　Ｐｕｒｉｆｉｃａｔｉｏｎ　Ｋｉｔ（Ｑｉａｇｅｎ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２－９２４０－０１）を使用するランダムプライミングによって、^{32}Ｐ－ｄＣＴＰ（Ａｍｅｒｓｈａｍ　Ｐｈａｒｍａｃｉａ　Ｂｉｏｔｅｃｈカタログ番号ＡＡ００５／２５０）で標識する。メーカーのプロトコルに従って、Ｃｌｏｎｔｅｃｈ由来のヒト多組織ノーザンブロットについてハイブリダイゼーションを行う。Ｍｏｌｅｃｕｌａｒ　Ｄｙｎａｍｉｃｓ　ＰＨＯＳＰＨＯＲ　ＩＭＡＧＥＲスクリーン（カタログ番号ＭＤ１４６－８１４）で一晩露光した後、約１．３５ｋｂのバンドを可視化する。
【０３７７】
（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する。例えば、実施例１に記載の配列を、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。
【０３７８】
ＧＰ３５４を安定に発現する細胞を、１００μｇ／ｍｌのゼオシン（ｚｅｏｃｉｎ）（Ｓｔｒａｔａｇｅｎｅ，ＬａＪｏｌｌａ，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、細胞から精製し得る。精製を容易にするために、ＧＰ３５４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ＧＰ３５４を親和性精製する。このＧＰ３５４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ＦＬＡＧ）とインフレームで発現して、精製を促進する。さらに、ＧＰ３５４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ＧＰ３５４の精製を必要としないことが理解される。
【０３７９】
（Ｂ．２９３細胞におけるＧＰ３５４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ＧＰ３５４の発現のために、関連のｇｐ３５４コード配列を有するプラスミドを、ベクターｐＳｅｃＴａｇ２Ａ（Ｉｎｖｉｔｒｏｇｅｎ）を使用して調製する。ベクターｐＳｅｃＴａｇ２Ａは、分泌のためのマウスＩｇＫ鎖リーダー配列、抗ｍｙｃ抗体を用いて組み替えタンパク質を検出するためのｃ－ｍｙｃ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ｇｐ３５４　ｃＤＮＡの増幅のための順方向プライマーを、標準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、

ＨｉｎｄＩＩＩ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ｇｐ３５４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ＸｈｏＩ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ｇｐ３５４配列の逆相補体に対応するヌクレオチドを含む。ＰＣＲ条件は、アニーリング温度として５５℃である。ＰＣＲ産物をゲル精製し、そしてベクターのＨｉｎｄＩＩＩ－ＸｈｏＩにクローニングする。ＱＩＡＧＥＮクロマトグラフィーカラムを使用して、ＤＮＡを精製し、そしてＤＯＴＡＰトランスフェクション培地（Ｂｏｅｈｒｉｎｇｅｒ　Ｍａｎｎｈｅｉｍ）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Ｈｉｓおよび抗ＧＰ３５４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。

【０３８０】

永久的にトランスフェクトされた細胞を、Ｚｅｏｃｉｎを用いて選択肢、そして増殖する。組換えタンパク質の産生を、抗Ｈｉｓ、抗Ｍｙｃまたは抗ＧＰ３５４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。

【０３８１】

（Ｃ．ＣＯＳ細胞におけるＧＰ３５４の発現）

ＣＯＳ７細胞におけるＧＰ３５４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ｐ３－ＣＩにクローニングし得る。このベクターは、ｂＧＨ（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ＨＣＭＶ（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、ｐＵＣ１８由来プラスミドである。さらに、プラスミドは、安定な形質転換体の選択のためのメトトレキセート（ＭＴＸ）の存在下における選択を提供するｄｈｒｆ（ジヒドロホレートレダクターゼ）を含む。

【０３８２】

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ＸｂａＩ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、ＳａｌＩ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。

【０３８３】

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分の伸長を包含する。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。

【０３８４】

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。この精製は、例えば、クロマトグラフィーによって達成し得る。

【０３８５】

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。

【０３８６】

（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、続いて入れて配列番号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。

【０３８７】

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（Ｐｈａｒｍｉｎｇｅｎ，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カリフォルニア核多角体病ウイルス（ＡｃＭＮＰＶ）の強力な多角プロモーター、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。

【０３８８】

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。

【０３８９】

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、ＡＭＡＮＵＡＬ ＯＦ ＭＥＴＨＯＤＳ ＦＯＲ ＢＡＣＵＬＯＶＩＲＵＳ ＶＥＣＴＯＲＳ 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ｐＡｃＨＬＴ－Ａを、ＢＡＣＵＬＯＧＯＬＤトランスフェクションキット（Ｐｈａｒｍｉｎｇｅｎ）を使用して、バキュロウイルスに導入する。感染の２４時間後に、感染細胞を^{35}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、そしてスケールアップした発現のために使用する。

【０３９０】

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ＡｃＨＬＴ－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の予備的ウエ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。

【０３９１】

（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、Ｆｉｅ

ｌｄｓら、Ｎａｔｕｒｅ　３４０：２４５（１９８９）において記載された。プロトコルは、Ａｕｓｕｂｅｌら、ＳＨＯＲＴ　ＰＲＯＴＯＣＯＬＳ　ＩＮ　ＭＯＬＥＣＵＬＡＲ　ＢＩＯＬＯＧＹ，Ｊｏｈｎ　Ｗｉｌｅｙ　＆　Ｓｏｎｓ，ＮＹ（１９９９）およびＡｕｓｕｂｅｌ．Ｆ．Ｍ．ら、ＣＵＲＲＥＮＴ　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。

【０３９２】

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）融合ライブラフィーを、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。

【０３９３】

ＤＮＡ－ＢＤ／ＧＰ３５４融合構築物を、配列決定によって確認し、そして自立性レセプター遺伝子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可能にする）について試験する。類似のコントロールを、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、酵母細胞に、ＧＰ３５４およびライブラフィー融合プラスミドの両方を形質転換する（約１０５形質転換体／ｍｇＤＮＡ）。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。酵母細胞を栄養不足培地に入れて、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増殖した後、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－ＢＤ融合タンパク質を含む他の株に再導入して、特定のＧＰ３５４／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、ＧＡＬ４　ＡＤ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ＧＰ３５４結合タンパク質の同一性を決定する。

【０３９４】

（実施例１１：ＧＰ３５４ポリペプチドに対する抗体）

ＧＰ３５４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」改変体および全ヒト抗体を含む）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、Ｓａｍｂｒｏｏｋら（前出）、およびＨａｒｌｏｗら編、ＡＮＴＩＢＯＤＩＥＳ，Ａ　ＬＡＢＯＲＡＴＯＲＹ　ＭＡＮＵＡＬ、Ｃｏｌｄ　Ｓｐｒｉｎｇ　Ｈａｒｂｏｒ　Ｌａｂｏｒａｔｏｒｙ，Ｃｏｌｄ　Ｓｐｒｉｎｇ　Ｈａｒｂｏｒ，ＮＹ（１９８８）に見出され得る。いくつかの実施形態において、組換えＧＰ３５４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。いくつかの実施形態において得、ＧＰ３５４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。ＧＰ３５４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、またはハプテンに結合して、抗体産生を増加し得る

。

【０３９５】

（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ＧＰ３５４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）を免役する。抗原性を増大するために、ペプチドを、例えば、Ｐｉｅｒｃｅ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、ＧＰ３５４抗原のさらなるアリコートを、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、ＧＰ３５４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ＧＰ３５４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。

【０３９６】

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ＲＰＭＩ　１６４０に入れ、そして２ｍＭ　Ｍ－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（ＲＰＭＩ）（Ｂｉｂｃｏ，Ｃａｎａｄａ）を補充した血清を含まないＲＰＭＩ　１６４０中でこの脾臓を粉砕することによって、単一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ＲＰＭＩに再懸濁する。ネイティブのＢａｌｂ／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（ＦＢＳ）（Ｈｙｃｌｏｎｅ　Ｌａｂｏｒａｔｏｒｉｅｓ，Ｉｎｃ．，Ｌｏｇａｎ，Ｕｔａｈ）を含むＲＰＭＩ中で、融合前の３日間対数期に保たれたＮＳ－１骨髄腫細胞を、同様に遠心分離し、そして洗浄する。

【０３９７】

ハイブリドーマ融合物を生成するために、免役したマウス由来の脾臓細胞をＮＳ－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ＰＥＧ　１５００（７５ｍＭ　ＨＥＰＥＳ（ｐＨ８．０）中５０％）をこのペレットに加えて撹拌し、続いて血清を含まないＲＰＭＩを添加する。その後、この細胞を遠心分離し、１５％　ＦＢＳ、１００μＭ　ナトリウムハイポキサンチン、０．４μＭ　アミノプテリン、１６μＭ　チミジン（ＨＡＴ）（Ｇｉｂｃｏ）、２５単位／ｍｌ　ＩＬ－６（Ｂｏｅｈｒｉｎｇｅｒ－Ｍａｎｎｈｅｉｍ）および１．５×10^6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。

【０３９８】

融合の２、４および６日後に、１００μｌ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。

【０３９９】

（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、および治療的介在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_4サブタイプのものであることが好ましい。

【０４００】

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、Ｍｏｒｒｉｓｏｎ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。

【０４０１】

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）；Ｖｅｒｈｏｅｙｅｎ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β－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。Ｋｅｔｔｌｅｂｏｒｏｕｇｈ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。

【０４０２】

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。

【０４０３】

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。

【０４０４】

（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，Ｈｕｍａｎ Ａｎｔｉｂｏｄｉｅｓ ８（４）：１５５－１６８（１９９７）；Ｈｏｏｇｅｎｂｏｏｍ，ＴＩＢＴＥＣＨ １５：６２－７０（１９９７）；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。このような構築物を含むファージライブラリーは、細菌において発現され、そしてこのライブラリーは、抗原プローブとして標識ＧＰ３５４または固定化ＧＰ３５４を使用して、ＧＰ３５４特異的ファージ抗体に関してスクリーニングされる。

【０４０５】

（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Ｂｒｕｉｇｇｅｍａｎｎら，Ｉｍｍｕｎｏｌ．Ｔｏｄａｙ １７（８）：３９１－９７（１９９６）およびＢｒｕｇｇｅｍａｎｎ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）。

【０４０６】

（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うちでも、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である。

【０４０７】

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。

【０４０８】

（Ａ．ｃＡＭＰアッセイ）

１つの型のアッセイにおいて、環状アデノシン一リン酸（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 Ｓｃｉｅｎｃｅｓ １８：５２９－５４１（１９７６）；Ｄｏｏｌｅｙら，Ｊ．ｏｆ Ｐｈａｒｍａｃｏｌ．Ｅｘｐ．Ｔｈｅｒａｐ．２８３（２）：７３５－４１（１９９７）；およびＧｅｏｒｇｅら，Ｊ．ｏｆ Ｂｉｏｍｏｌ．Ｓｃｒｅｅｎｉｎｇ ２（４）：２３５－４０（１９９７）を参照のこと。ＮＥＮ Ｌｉｆｅ Ｓｃｉｅｎｃｅ Ｐｒｏｄｕｃｔｓのアデニルシクラーゼ活性化ＦＬＡＳＨＰＬＡＴＥアッセイを使用する、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。

【０４０９】

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アッセイキットの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、化合物の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}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}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の調節活性の指標である。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レベルの顕著な（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レベルの顕著な減少（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。

【０４１０】

（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。補因子セレントラジンの存在下で、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。

【０４１１】

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。

【０４１２】

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，Ｇａｉｔｈｅｒｓｂｕｒｇ，ＭＤ）中で、３７℃で２４時間培養する。次いで、この培養培地を、５μＭセレントラジン（Ｍｏｌｅｃｕｌａｒ Ｐｒｏｂｅｓ）を含む無血清ＭＥＭと交換する。培養を、３７℃でさらに2時間続ける。引き続いて、ＶＥＲＳＥＮ（Ｇｉｂｃｏ／ＢＲＬ）を使用してこれらの細胞をプレートから剥し、洗浄し、そして無血清ＭＥＭ中に2×10^5細胞／ｍｌで再懸濁させる。

【０４１３】

候補ＧＰ３５４調節因子化合物の希釈物を、無血清ＭＥＭ中で調製し、そして不透明な９６ウェルアッセイプレートのウェルに、５０μＬ／ウェルで分配する。次いで、このプレートをＭＬＸマイクロタイタープレートルミノメーター（Ｄｙｎｅｘ Ｔｅｃｈｎｏｌｏｇｉｅｓ，Ｉｎｃ．，Ｃｈａｎｔｉｌｌｙ，ＶＡ）に設置する。この装置は、５０μｌの細胞懸濁物を、１回に１つのウェルで各ウェルに分配し、そして即座に、発光を１５秒間読み取るようプログラムされている。候補調節因子についての用量－応答曲線を、各光シグナルピークについての曲線下面積を使用して構築する。データを、１部位リガンドについての式を使用して、ＳＬＩＤＥＷＲＩＴＥを用いて分析し、そしてＥＣ５０値を得る。これらの化合物によって引き起こされる発光の変化を、調節活性の指標とみなす。Ｇタン

パク質のＧｑサブタイプに結合するレセプターにおいて、アゴニストとして働く調節因子は、１００倍までの発光の増加を与える。インバースアゴニストとして働く調節因子は、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のレセプターにおいて、この効果を逆にする。

【０４１４】

（Ｃ．ルシフェラーゼレポーター遺伝子アッセイ）

発光タンパク質ルシフェラーゼは、ＧＰ３５４調節因子を同定するための別の有用なツールを提供する。細胞（例えば、ＣＨＯ細胞またはＣＯＳ７細胞）を、ｇｐ３５４発現構築物および転写因子結合部位（例えば、ｃＡＭＰ応答エレメント（ＣＲＥ）、ＡＰ－１、またはＮＦ－κＢ）の下流のルシフェラーゼタンパク質に対する遺伝子を含むレポーター構築物で、一時的に同時トランスフェクトする。ルシフェラーゼの発現レベルは、シグナル伝達事象の活性化状態を反映する。一般的には、Ｇｅｏｒｇｅら，Ｊ．Ｂｉｏｍｏｌ．Ｓｃｒｅｅｎｉｎｇ　２（４）：２３５－２４０（１９９７）；およびＳｔｒａｔｏｗａら，Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．６：５７４－５８１（１９９５）を参照のこと。ルシフェラーゼ活性を、例えば、Ｐｒｏｍｅｇａ（Ｍａｄｉｓｏｎ，ＷＩ）から入手可能なルシフェラーゼアッセイ試薬を使用して、定量的に測定し得る。

【０４１５】

１つの例示的なアッセイにおいて、ＣＨＯ細胞を、２４ウェル培養プレートに、トランスフェクションの１日前に10^5細胞／ウェルの密度でプレートし、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１０μｇ／ｍｌストレプトマイシンを補充したＭＥＭ（Ｇｉｂｃｏ／ＢＲＬ）中３７℃で培養する。細胞を、ｇｐ３５４発現構築物およびルシフェラーゼ遺伝子を含むレポーター構築物で、一時的に同時トランスフェクトする。レポータープラスミド構築物のＣＲＥ－ルシフェラーゼ、ＡＰ－１－ルシフェラーゼおよびＮＦ－κＢ－ルシフェラーゼを、Ｓｔｒａｔａｇｅｎｅ（ＬａＪｏｌｌａ，ＣＡ）から購入し得る。トランスフェクションを、ＦＵＧＥＮＥ６トランスフェクション試薬（Ｂｏｅｈｒｉｎｇｅｒ－Ｍａｎｎｈｅｉｍ）を使用して、供給元の指示書に従って実施する。レポーター構築物単独でトランスフェクトした細胞を、コントロールとして使用する。

【０４１６】

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分間のインキュベーションの後に、１５μｌの溶解物を、不透明白色の９６ウェルプレートで５０μｌ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。

【０４１７】

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。

【０４１８】

（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た指標であり、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

×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、３６ｍｇ／Ｌピルベートおよび１ｇ／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４μＭの濃度で添加した改変ダルベッコＰＢＳ（Ｄ－ＰＢＳ）中、３７℃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℃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。

【０４１９】

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（１０μＭ；陽性コントロール）、またはＡＴＰ（４μＭ；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（４８８ｎｍ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、候補アゴニストの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。

【０４２０】

（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．ａｎｄ　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℃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、８０μｌ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２０μｌ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。

【０４２１】

１６～１８時間の培養の後に、１μＣｉの［^{3}Ｈ］－チミジン（２Ｃｉ／ｍｍｏｌ）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Ｂｅｔａｐｌａｔｅカウンターで計数する。無血清試験ウェルへの［^{3}Ｈ］－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×［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、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、単体の最適化（Ｓｉｍｐｌｅｘ　ｏｐｔｉｍｉｚａｔｉｏｎ）によって決定される。

【０４２２】
レセプターに結合するアゴニストは、細胞への［^{3}Ｈ］－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。
【０４２３】
（Ｆ．［^{35}Ｓ］ＧＴＰｇＳ結合アッセイ）
ＧＰ３５４がＧタンパク質媒介経路を通っ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（これの活性は、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］ＧＴＰｇＳ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。
【０４２４】
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×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×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。
【０４２５】
ホモジネートの上清を２０，０００×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　ＭｇＣｌ２、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。
【０４２６】
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　ＭｇＣｌ２、１ｍＭ　ＥＤＴＡ、１２０ｍＭ　ＮａＣｌ、１０μＭ　ＧＤＰ、および０．２ｍＭアスコルビン酸を含む緩衝液中に、１０～５０μｇ／ｍｌの濃度で希釈する。９０μｌ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１００μＭ　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、１０μｌグアノシン５’－Ｏ－（３［^{35}Ｓ］チオ）三リン酸（ＮＥＮ、１２００Ｃｉ／ｍｍｏｌ；［^{35}Ｓ］－ＧＴＰｇＳ）を、１００～２００ｐＭ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、１ｍｌの１０ｍＭ　ＨＥＰＥＳ（ｐＨ７．４）、１０ｍＭ　ＭｇＣｌ２（４℃）を添加し、そして濾過によって反応を停止させる。
【０４２７】
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２０ｍｌの氷冷１０ｍＭ　ＨＥＰＥＳ（ｐＨ７．４）、１０ｍＭ　$MgCl_2$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。［^{35}Ｓ］－ＧＴＰｇＳの非特異的結合を、１００μＭ　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［^{35}Ｓ］－ＧＴＰｇＳ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、［^{35}Ｓ］－ＧＴＰｇＳ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。
【０４２８】

（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，Ｊ．Ｐｈａｒｍａｃｏｌ．Ｅｘｐ．Ｔｈｅｒａｐ．２６７（３）：１５７３－１５８１（１９９３）およびＢｏｕｌｔｏｎら，Ｃｅｌｌ　６５：６６３－６７５（１９９１）を参照のこと。いくつか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7×10^{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、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。

【０４２９】
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２００ｎＭ　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、１ｍＭ　ＥＤＴＡを含む１ｍｌの氷冷ＰＢＳでリンスする。その後、２００μｌの細胞溶解緩衝液（１２．５ｍＭ　ＭＯＰＳ（ｐＨ７．３）、１２．５ｍＭグリセロホスフェート、７．５ｍＭ　$MgCl_2$、０．５ｍＭ　ＥＧＴＡ、０．５ｍＭバナジン酸ナトリウム、１ｍＭベンズアミジン、１ｍＭジチオトレイトール、１０μｇ／ｍｌロイペプチン、１０μｇ／ｍｌアプロチニン、２μｇ／ｍｌペプスタチンＡ、および１μＭ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　３／４Ｇ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×ｇで１５分間の遠心分離によって、調製する。

【０４３０】
細胞質ゾルのアリコート（１～５μｇのタンパク質を含む５～１０μｌ）を、１ｍＭ　ＭＡＰＫ基質ペプチド（ＡＰＲＴＰＧＧＲＲ（配列番号９）、Ｕｐｓｔａｔｅ　Ｂｉｏｔｅｃｈｎｏｌｏｇｙ，Ｉｎｃ．，ＮＹ）および５０μＭ［$g-^{32}P$］ＡＴＰ（ＮＥＮ、３０００Ｃｉ／ｍｍｏｌ）と混合し、約２０００ｃｐｍ／ｐｍｏｌの最終比活性に、総容量２５μｌ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、$2cm^2$の正方形のＷｈａｔｍａｎ　Ｐ８１ホスホセルロース紙上に２０μｌ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_3PO_4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。

【０４３１】
（Ｈ．　［^3H］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、Ｍｏｌｅｃｕｌａｒ　Ｐｈａｒｍａｃｏｌｏｇｙ　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、1.5×10^{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１０μｇ／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０．５μＣｉ／ｍｌの［

^{3}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１０μＭ　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}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。

【０４３２】

（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、４×１０^{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１０μｇ／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_2中３７℃で、２４時間インキュベートする。

【０４３３】

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、１０μｇ／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、１００μｌ／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、レセプターのアンタ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。

【図面の簡単な説明】

【図１】

図１は、ＧＰ３５４のヌクレオチド配列およびアミノ酸配列である。配列番号１または配列番号２を参照のこと。細胞外部位における免疫グロブリン（Ｉｇ）ドメインに下線を引き、そして膜貫通ドメインは、四角で囲んでいる。

【図２】

図２は、ＧＰ３５４アミノ酸配列（上）（配列番号２）と、Ｄｒｏｓｏｐｈｉｌａの不規則な交差（ｉｒｒｅｇｕｌａｒ　ｃｈｉａｓｍ）（ＩＣＣＲ）の配列（配列番号１３）およびヒトネフリン（配列番号１４）タンパク質とのアライメントである。ダッシュは、任意の配列中のギャップを示す。アスタリスクは、３本の配列中において同定されるアミノ酸を示す。

【図３】

図３は、逆転写ポリメラーゼ鎖反応（ＲＴ－ＰＣＲ）によって決定される場合のヒト組織

におけるＧＰ３５４の発現である。ＲＴ－ＰＣＲは文字において記載されるように実施した。ＧＰ３５４の発現は、膵臓においてのみ検出される。Ｂ＝脳のレーン、Ｈ＝心臓のレーン、Ｋ＝腎臓のレーン、Ｌｖ＝肝臓のレーン、Ｌｇ＝肺のレーン、Ｐｎ＝膵臓のレーン、Ｐｔ＝胎盤のレーン、Ｍｓ＝骨格筋のレーン、Ｃ＝結腸のレーン、Ｏｖ＝卵巣のレーン、Ｌｅ＝末梢血リンパ球のレーン、Ｐｒ＝前立腺のレーン、Ｓｉ＝小腸のレーン、Ｓｐ＝脾臓のレーン、Ｔｅ＝精巣のレーン、Ｔｙ＝胸腺のレーン、－＝テンプレートを含まないコントロールのレーン、Ｇ＝ゲノムＤＮＡコントロールのレーン

【図４】

図４は、ノーザンブロットによって決定される場合のヒト組織におけるＧＰ３５４のＲＮＡの発現である。ノーザンブロットは、ｇｐ３５４配列から調製したプローブとハイブリダイズさせた。およそ３．２ｋｂのハイブリダイズするＲＮＡを、膵臓において観察したが、試験した他の組織のいずれにおいても観察しなかった。Ｈ＝心臓、Ｂ＝脳、Ｐ＝胎盤、Ｌｎ＝肺、Ｌ＝肝臓、Ｍ＝骨格筋、Ｋ＝腎臓、Ｐｃ＝胎盤。

【図５】

図５は、プライマーＧＸ１－２１８およびＧＸ１－２１９を用いて得たＲＴ－フラグメントの配列である。（配列番号３を参照のこと）。

【図６】

図６は、ヒトゲノムｇｐ３５４のヌクレオチド配列である、エキソン（Ｅｘｏｎ）には、下線を引いている。配列番号５を参照のこと。

【図７】

図７は、発現されたＧＰ３５４のヌクレオチドおよび誘導アミノ酸配列である。配列番号７および８を参照のこと。

【図８】

図８は、膵臓のｇｐ３５４　ｃＤＮＡのヌクレオチドおよび推定アミノ酸配列である。配列番号１１および１２を参照のこと。

【図２】

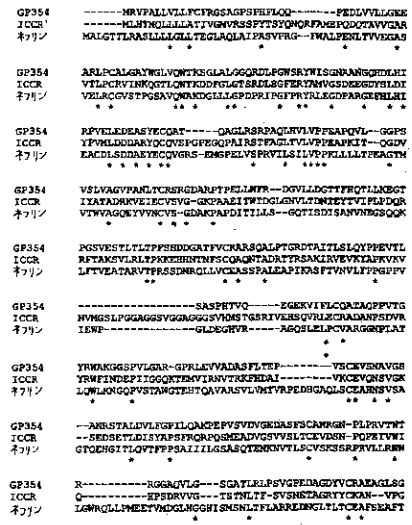


FIG. 2

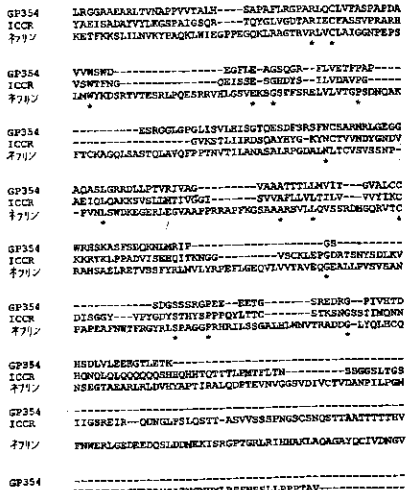


FIG. 2（続き）

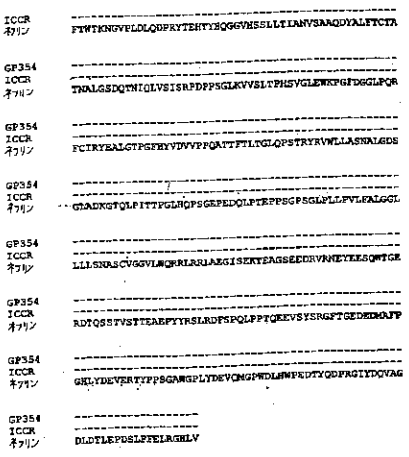


FIG. 2 (図2のつづき)③

【図６】

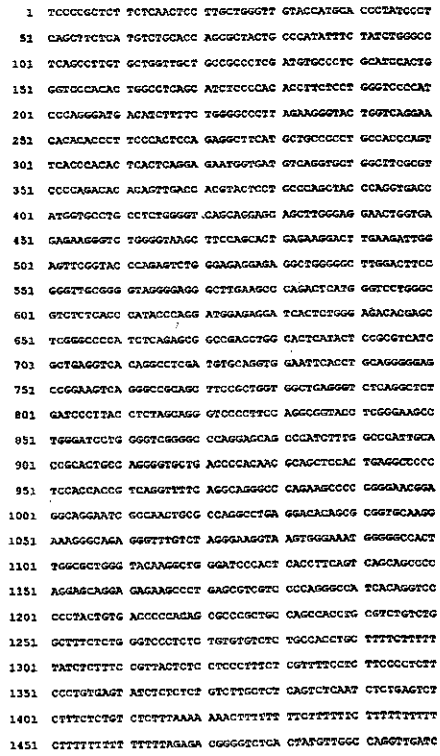


FIG. 6

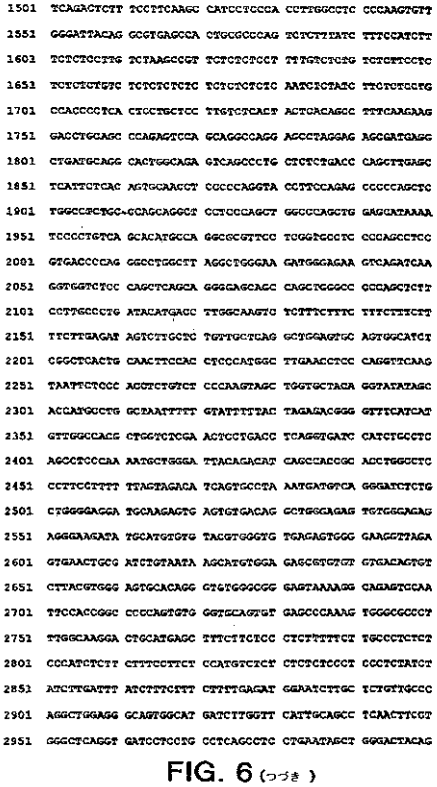


FIG. 6 (つづき)

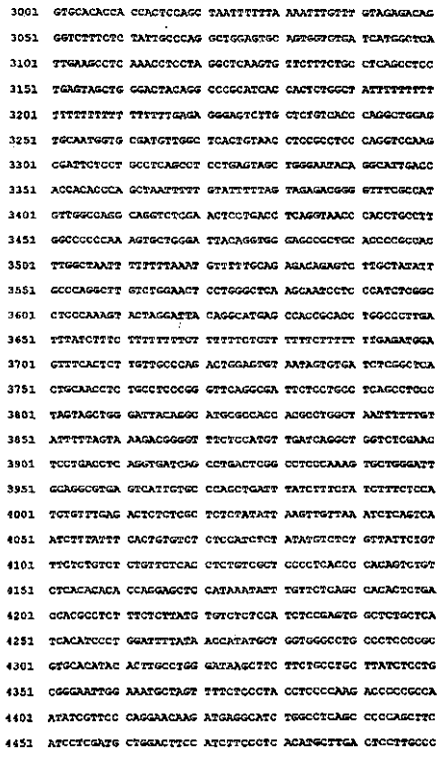


FIG. 6 (つづき)

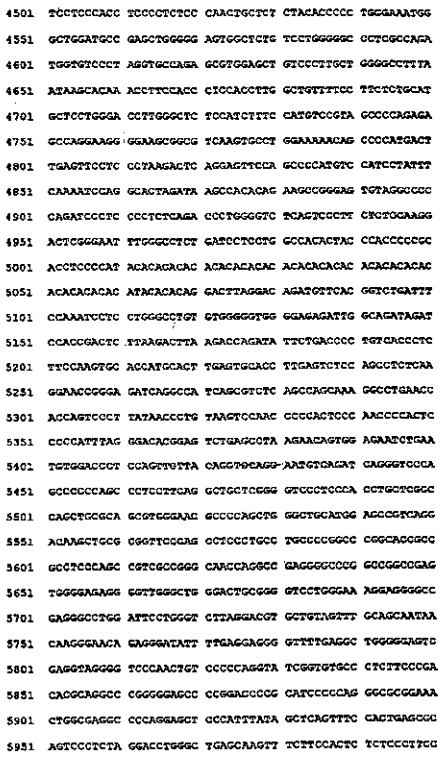


FIG. 6(つづき)

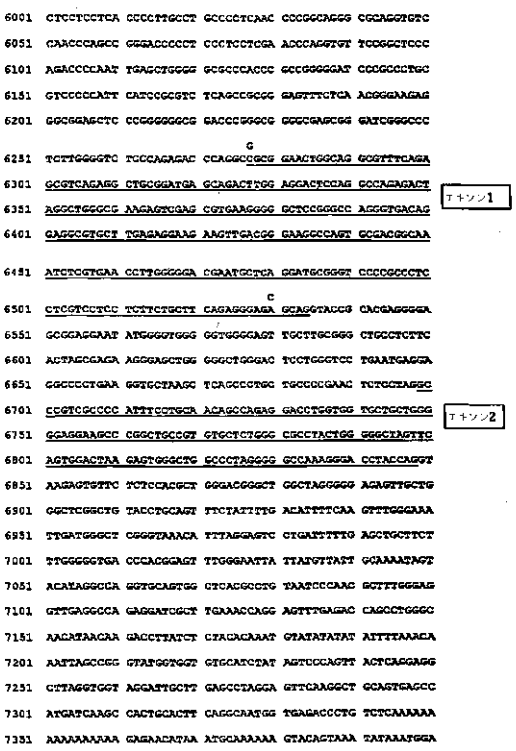


FIG. 6(つづき)

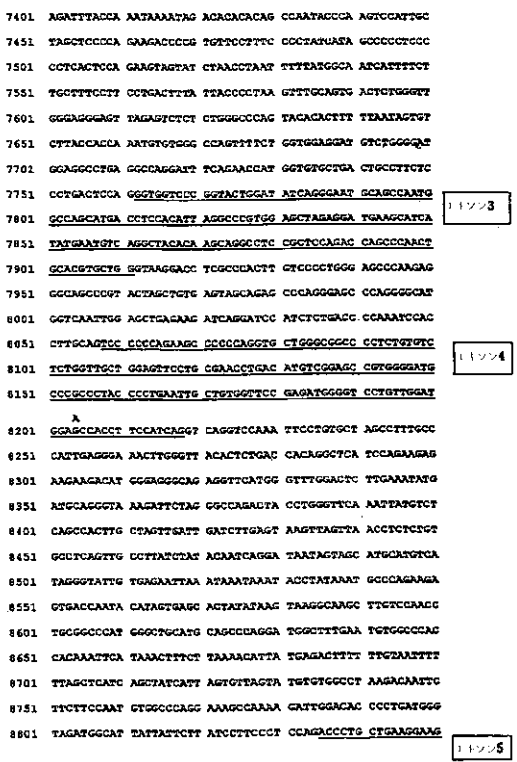


FIG. 6(つづき)

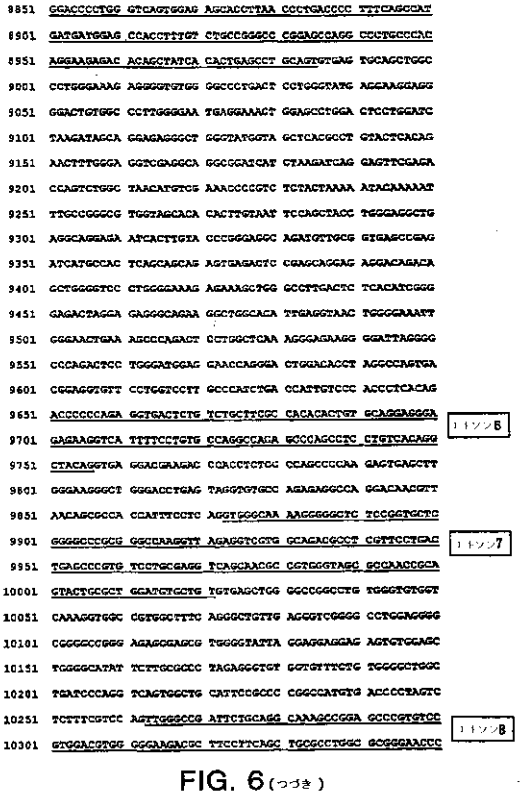
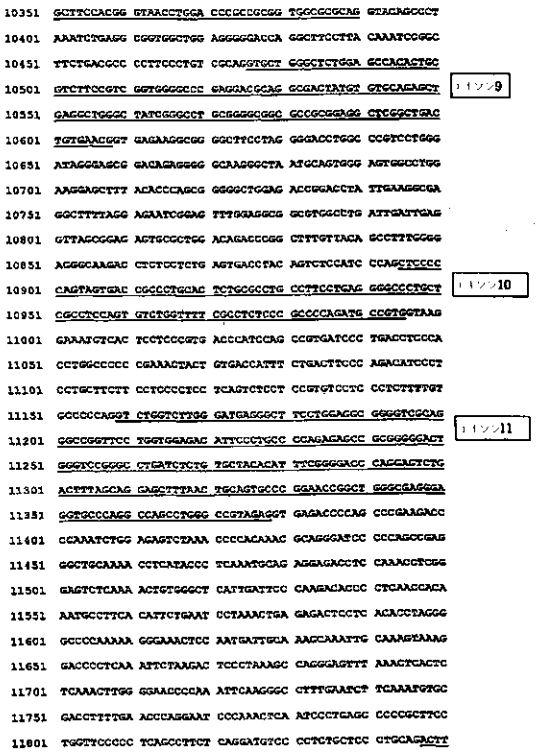


FIG. 6(つづき)

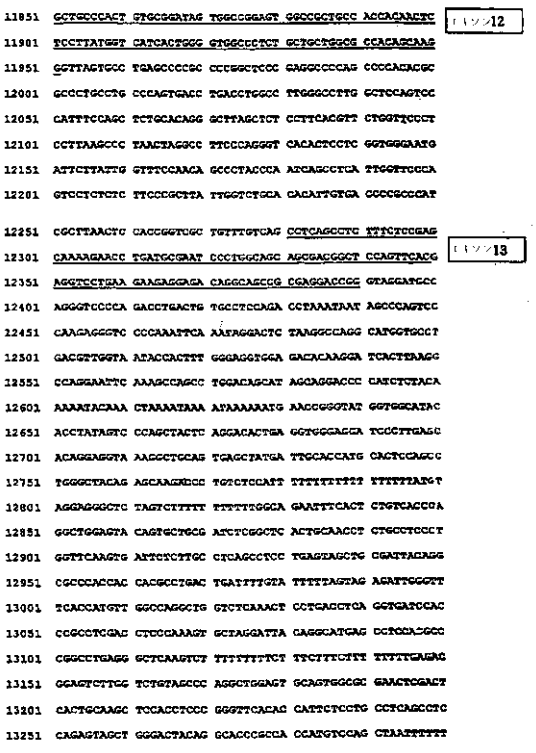


イントロン9

イントロン10

イントロン11

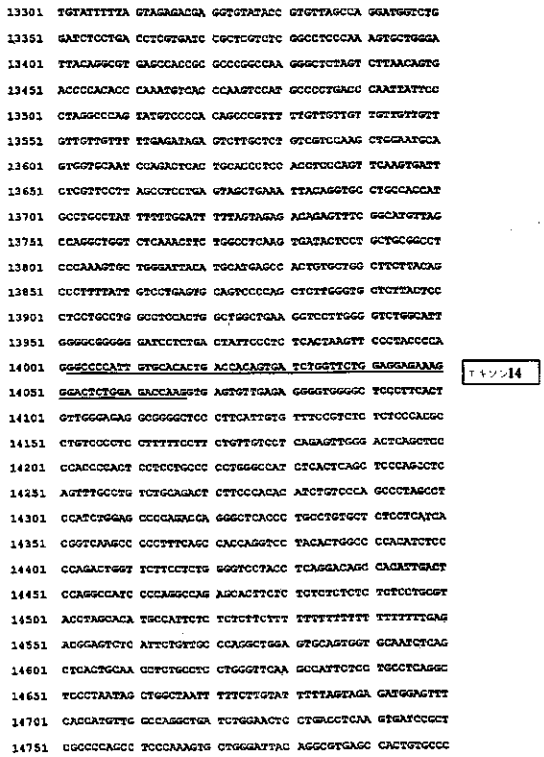
FIG. 6 (つづき)



イントロン12

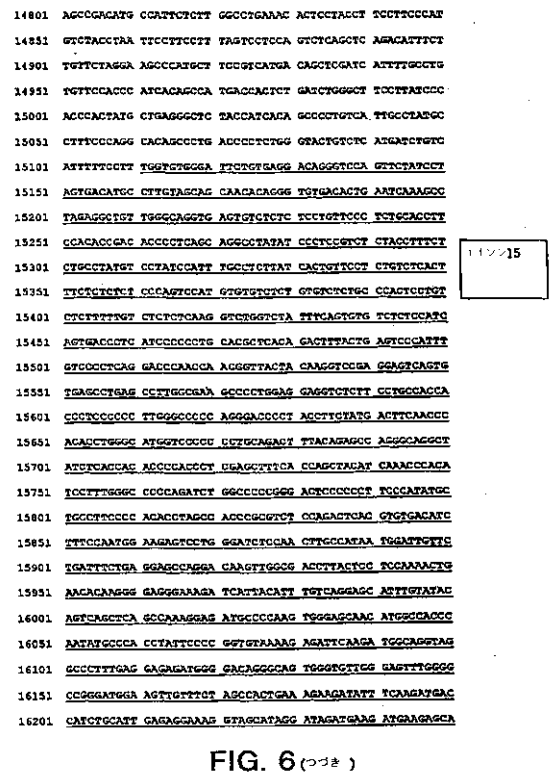
イントロン13

FIG. 6 (つづき)



イントロン14

FIG. 6 (つづき)



イントロン15

FIG. 6 (つづき)

16251 TACCAGGCCC CACCCTGGCT CTCCCTGAGG GGAACTTTGC TCGGCCAATG
16301 GAAATGCAGC CAAGATGGCC ATATACTCCC TAGGAACCCA AGATGGCCAC
16351 CATCTTGATT TTACTTTCCT TAAAGACTCA GAAAGACTTG GACCCAAGGA
16401 GTGGGGATAC AGTGAGAATT ACCACTGTTG GGGCAAAATA TTGGGATAAA
16451 AATATTTATG TTTAATAATA AAAAAAAGTC AAAGAGGCAA GTGTGTCTTA
16501 GGGAGTCTAC TGGCATTATC ACTCTCCACC AAGGAAGGGG TCCCTTAGAC
16551 CTGTCCCAAG GTCCCTCCTC TACCCTAGCC TATGAGGTGG CTGTAGGAGT
16601 AAAACTGTGA GCCACCTCTC AGCCTCTTGC TACCTGCAAA GCACTCTAGG
16651 CTCTTTTTTT TTTTTTCTTG AGACAAGATC TGGCTCTATG GCCCACATTG
16701 GAGTGCAGTG GCATGATCTC AGCCCACTGC TACCTCTGCA TCCTGGGGTC
16751 AAGCCATCCT TCCACCTCAG CCTCCCAAGT AGCTGGGACT ACAGGTGCAT
16801 GCCACCACAC CCAGCTAATT TTTGTATTTG TTTGTAGACA GGGTTTCACC
16851 ATGTTGGCCA GGCTGGTCTC AAACTCCTGA CCTCAAGTGA TCCGCCCACC
16901 TAGGCCTCCC AATGTGCTGG GATTACAGGC ATGAGCCACT GTGCCCAGCC
16951 ATGGGCTCTT TTAATATACA TCTTCACACA CACACACACA CACACACACA
17001 GGCACACACA CACATGAGTT GCAAACAGAA AAGACACACA CATAGGCATG
17051 TATGCACAGA CACACGCATA GATGTCCACA CAGTTGCACA CAAGTGACAG
17101 GGCTGCCCCA GGGGTCCTGG GGAAGACTGA ATTCTAACTC TCATTAGAGG
17151 AGACAAACAA GTGAGCCCTG AAGTGGAGCA GGGAAGGGGA GACTATGGGT
17201 AGGAAAATGG CAATCCCCTG GTCCTTACAG CAAGCGTGGA GATCCAGACC
17251 CTAATCCTGA GGTGCTGCAT CCACAGTGGG CATGGTGCTG GTGCCTGCTT
17301 GGATGATCCT TAAAGAAAGG TCCTGGGGGC TTTGGTTCAT CGATCCTTGA
17351 GCTAGGAGTT AAAGGTCCAG GCCCCTGGGA CCCTTGGGAA GCAGAGCAAG
17401 AAGAGTGAAC TCCTGGGTCT GAAGGAGAAT GGGCTGGGGG CTTGGTCTCT
17451 GGTCCTGAGA GAGAAGGTGC CCAGACTTCT GGATCTGAAA GAGGAAGGGA
17501 CTAGGTCTCA ACTGCTGCCT TCTTGACTGG GGACATTTTG GAGGCCTGTA
17551 TTCCTGAGCC CTCAACAGAG GAATGTACTA GGGGATGGGG GTCTCTGATG
17601 CTTGCATCCT TGGAAAAGGA CAAAACTGTG AGTGTCTGGG TCTAAAGAGG
17651 GTGAGAGTCC TGCGGGAGGA CTCAAAATCC ACAACGGGGG GAGCCCATAG
17701 CCGGACTCCT GGCTGGGCCC TTCATGGGGC GGGACGCCTG GAATCTCGAG

FIG. 6 (つづき)

17751 GGGCGGGGGC CTGGGCCAGG CTCCCGCCCG GGGTTCCCGA GCTGCTCCAC
17801 TCTGCGCGAA GCCGCCAGGC TATTGTCCTG ACCAGGAAGG CGGGGCCGGC
17851 GCGGGGCGGG GCTGGCGGCG CCGGCGCAGC CCGGGGGCGG CGGGAGGAGG
17901 AGGTGGCGGC GGTGGCGCTG GGAGCTCCTG TCACCGCTGG GGCCGGGCCG
17951 GGCGGGAGTG CAGGGGACGT GAGGGCGCAA GGGCCGGGAC ATGGGGCCCG
18001 CCAGCCCCGC TGCTCGCGGT CTAAGTCGCC GCCCGGGCCA GCCGCCGCTG
18051 CCGCTGCTGC TGCCACTATT GCTGCTGCTT CTGCGCGCGC AGCCCGCCAT
18101 CGGGAGCCTG GCCGGTGGGA GCCCCGGCGC GGCCGAGGTG AGGCCGGCCC
18151 GGGTCCTGGG GGATGGGGGA AGGGGCGGGA CCGGGTCTCT GGACGCCGGC
18201 GCGGACATGT CCAGGGCAGA AAGCGCGGTC TTTCCAGCCA GGTGGTCAGC
18251 GCCCAGGCGC CCCCAATCAC ATTTATGAAC CCAGGGTTCC AGGCCCCAGC
18301 TCCCCCATCA TGCGACGTCC CAGCCCCCTC CCATCTCGAG CATAGGAACT
18351 GGTCTATTCA GAGCCCCTGG TCCCAGAAGT CCAGCCCCCT CTCCAGACCC
18401 AGGTGACTCG GCCCCAACCC CCTCCCGCCT GGACATAGGA CCCACCAAGC
18451 AGCGAGGCAT TTAGATCCAA TAATCCAGAC CCCTTGTATT CTCTGGACCC
18501 ATATGGAGGC CCTTGCAGCC TCCCAGGACC CAGGAGTCCA GTCCTTCAGT
18551 CACCAGCCAC GCCAACCAGA TGTAGCTCTC CAGTCCTCAA GGACCTGGTG
18601 TCCAGGACTG TAGGCCCCTG AAGCCAGGCC TTGTCAGCTT TGCATCCTGC
18651 AACGGGAGCC TGAGCAAGGG ATGGAGGGAG GAGGGGCCAG AACTCCTGGG
18701 TTCTGGCCTC CTCCTCCGCG ATTCAGGTTT AACCCCTTCG GGCTCCAGAG
18751 CGGCTGCGCT GGGGTGGGGG CGGAGTCTGT CTCCGCGGCA ACAAGGCAGA
18801 AAGAATCCCG GGGGACCCAG GTGGCCATAG CAACGGGAGC GCTGGGGCGC
18851 CCCCGCCCTA CGGGAGCTGT TCCCAGGGA ACGGTGCCTC CATGGAGGCG
18901 GTGTGCGGTG CTTGGGGGAG GGGGCTGGTG CTGGGGGTCT CGGTCCTAGG
18951 GAGCAAAGAA CCAGGGGACC CTCATGCCAA CGCCCCCCGA GCCCTCACTG
19001 TCCTTTCCAC TTCCATCCAG GCCCCGGGGT CGGCCCAGGT GGCTGGACTA
19051 TCCGGGCGCC TAACCCTTCA CCGGGACCTG CGCACCGGCC GCTGGGAACC
19101 AGACCCACAG CGCTCTCGAC GCTGTCTCCG GGACCCGCAG CGCGTGCTGG
19151 AGTACTGCAG ACAGGTGGGC GGGGCCGAAC GGGAGAGGCG GGGCCGCCCA
19201 TAGAAAGCTA GACTTGAAAA AGGCGTGGTC CAGGGTGCTG CGCGATCTAA

FIG. 6 (つづき)

19251 GGCGTGGAGG CTGGGGGGCG TGGCCAATAA AGAGGCGCAA CTATGCTAGG
19301 GGCAGGGGAC CTGTTTTGAG ATACTAAGTC AGGAAAAGGG GAGAGCCGCG
19351 AGATAGCCAG AGAGGAAGTG GAATTTAGGA ATCTGGTCGT CTTTGTAAAG
19401 AGTAGAGGTG TAGGGGCGAG TGGCGAAAGG ATAGGCGGGG CTAAGACAGA
19451 AAGAGACCTT AAGGACCAGC AAGATGGGGA AAGGGGTGGA GCCCAATGAG
19501 AGCGCGGAGA GCTGGGGGGC CGTGGCCATG AAAAGACAAA TTTATAACGG
19551 GAAGGGAGAG TTTTGGAGAG GCGGAATAGA GGAAAAGGCG GGGCCTAAAG
19601 GAGGGTGAGA CCTTGGGGGA GACGAATCTG ACTGCGGGGA GGGGTGACCA
19651 GAGAGGTGGG CTTAGAGGGA CGTTCAGAAA GAAACAGCAC AGGAAAAGAG
19701 ATAGGGCTTA AAGATGACGG GACTTTTAAG GGAAAACTGC TAGTGGGCGT
19751 GGCCAATGAG CACAAGGAGC TTGGATATCT AAGGCTGGTG CTAGGGAGAA
19801 GCAGGGCCTA GGGAAGCGAT GTCCTCATGA ATACTAGAGC CTTGAAAACG
19851 GACCTGGCCG GGCGCGGTGG CTCACGCCTG TAATCGCAGC ACTTGGGGAG
19901 GCCGAGGCAG GCGGATCAGC TGAGGTCAGA AGTTCGAGAC CAGCCTGGCC
19951 AACACGGCGA AACTCCGTCT CTACTAAAAA TACAAAAATT AGCCTGGCAT
20001 GGTGGTGCGT GCCTGTAATC CCAGCTACTC AGGAGGCTGA GACAGGAGAA

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/98360 A2

(51) International Patent Classification⁷: C07K 16/00

(21) International Application Number: PCT/US01/19904

(22) International Filing Date: 22 June 2001 (22.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/213,611 22 June 2000 (22.06.2000) US

(71) Applicant (*for all designated States except US*): **BIOGEN, INC.** [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **CARULLI, John, P.** [US/US]; 9 Harris Drive, Southborough, MA 01772 (US). **LUKASHIN, Alexander, V.** [US/US]; 42 8th Street, #4403, Charlestown, MA 02139 (US). **KILBURN, Daniel, R.** [US/US]; 6 Hancock Avenue, Lexington, MA 02420 (US). **SUN, Chan** [CN/US]; 11 Chimino Road, Needham, MA 02494 (US).

(74) Agents: **HALEY, James, F., Jr.** et al.; Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

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

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

WO 01/98360 A2

Declarations under Rule 4.17:

as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)

as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)

as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)

as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY,

[Continued on next page]

(54) Title: GP354 NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: An isolated polynucleotide encoding a novel immunoglobulin superfamily member named GP354 is provided. GP354 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 GP354 indicate that it plays a role in cell-cell recognition, binding, signaling and adhesion events in the peripheral and central nervous system (CNS). Provided by the invention are isolated GP354 related polynucleotides and polypeptides, vectors, and host cells comprising any of the above, antibodies directed to GP354, cells which produce such antibodies, and related diagnostic and therapeutic methods.

WO 01/98360 A2

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, KH, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)

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

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

GP354 NUCLEIC ACIDS AND POLYPEPTIDES

RELATED APPLICATIONS

The present application claims priority from United States Provisional Application No. 60/213,611, filed June 22, 2000, the disclosure of which is incorporated herein by reference.

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 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 β-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 (FcgRII), 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, 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.

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). The protein encoded by this human gp354 cDNA (GP354) is a pancreas-enriched integral membrane protein. It is also detected in low levels in central nervous system (CNS) tissue. GP354 has a predicted single membrane spanning domain and five immunoglobulin (Ig) domains in the extracellular portion of the protein. The GP354 protein shares no more than 30% amino acid identity overall with any previously described proteins. The protein structure and tissue distribution of GP354 indicate that it plays a role in cell-cell interactions in the pancreas and central nervous system (CNS).

The invention provides isolated polynucleotides encoding GP354 or biologically active portions thereof. This invention also provides polynucleotide fragments suitable for use as primers or hybridization probes for the detection of GP354-encoding polynucleotides. Unless otherwise specified, "GP354," "GP354" protein and "GP354" polypeptide refer to a human gene product or a homolog of this protein in other non-human mammalian or other vertebrate species.

The invention 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); 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 examplary 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 examplary 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 NOS:4, 8, 10 or 12, wherein the nucleic acid hybridizes to SEQ ID NO:1 or 11 under stringent conditions.

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 NOS:2, 4, 8, 10 or12; or to an Ig domain encoded by any one of those sequences.

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 hybridizes under stringent conditions to a nucleic acid having the sequence of SEQ ID NOS: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 recombinant vector (e.g., a cloning, targeting or expression vector), comprising a gp354 polynucleotide of the invention.

Host cells containing such nucleic acid constructs are also provided, as is a method for producing a GP354 polypeptide by culturing, in a suitable medium, a host cell of the invention containing a recombinant expression construct such that a GP354 polypeptide is produced.

Isolated or recombinant GP354 proteins and polypeptides are provided by the invention. Preferred GP354 proteins and polypeptides possess at least one of the following (overlapping) biological activities possessed by naturally occurring human GP354: (1) the ability to interact with (e.g., bind to) a ligand (e.g., a protein receptor, a polysaccharide, etc.) that naturally binds to GP354 protein; (2) the ability to bind to an auto-antibody to naturally occurring human GP354 or an antibody raised against naturally occurring human GP354; (3) the ability to participate in a pancreatic function (e.g., a signal transduction function in the pancreas or a step in the organ development of the pancreas); (4) the ability to participate 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) the ability to mediate cell-cell interactions such as recognition, binding and/or adhesion.

The GP354 proteins or biologically active portions thereof can be operably linked to a non-GP354 polypeptide (e.g., heterologous amino acid sequences, such as sequences that facilitate protein stability, detection, purification, or *in vivo* delivery to target cells) to form GP354 fusion proteins.

The invention further features antibodies (e.g., polyclonal or monoclonal antibodies), including chimeric and humanized antibodies, that specifically bind to GP354 proteins or portions thereof.

The invention provides pharmaceutical compositions comprising at least one of the above-described gp354-related isolated polynucleotides, GP354 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 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 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 inhibiting pancreatic cell death associated with immune, auto-immune, and degenerative conditions. 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 degeneration and cell death in the central nervous system.

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

stimulates) GP354 activity or expression such that the GP354 activity or expression is altered. In some embodiments, the agent is an antibody that 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 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"). 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.

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.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 Nucleotide and deduced amino acid sequences of GP354. See SEQ ID NOS:1 and 2. The immunoglobulin (Ig) domains in the extracellular portion are underlined and the transmembrane domain is boxed.

FIG. 2 The alignment of GP354 amino acid sequences (top) (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.

FIG. 3 Expression of GP354 in human tissues as determined by reverse transcription polymerase chain reaction (RT-PCR). RT-PCR was performed as described in the text. GP354 expression is detected only in the pancreas. B = brain, H = heart, K = kidney, Lv = liver, Lg = lung, Pn = pancreas, Pt = placenta, Ms = skeletal muscle, C = colon, Ov = ovary, Le = peripheral blood leukocytes, Pr = prostate, Si = small intestine, Sp = spleen, Te = testis, Ty = thymus, - = no template control, G = genomic DNA control lane.

FIG. 4 Expression of GP354 RNA in human tissues as determined by Northern blot analysis. A Northern blot was hybridized with a probe prepared from gp354 sequences. A hybridizing RNA of approximately 3.2 kilobases is observed in the pancreas but not in any of the other tissues tested. H = heart, B= brain, P = placenta, Ln = lung, L = liver, M = skeletal muscle, K = kidney, Pc = placenta.

FIG. 5 Sequence of the RT-PCR fragment obtained using primers GX1-218 and GX1-219. (See SEQ ID NO:3).

FIG. 6 The nucleotide sequence of human genomic gp354. Exons are underlined. See SEQ ID NO:5.

FIG. 7 A nucleotide and derived amino acid sequence of an expressed GP354. See SEQ ID NOS:7 and 8.

FIG. 8 Nucleotide and deduced amino acid sequences of a pancreatic gp354 cDNA. See SEQ ID NOS:11 and 12.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based, at least in part, on the discovery of a novel human gene encoding a heretofore unknown protein, GP354. 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 transcribed from this gene has an open reading frame of 1779 base pairs, 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 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, 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 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 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.

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 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 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 "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 free of other cellular material, or substantially free of culture medium when 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.

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.

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.

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* 97(4):1665-1670 (2000). The disclosures of all of these references are incorporated herein by reference in their entireties.

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.

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 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.

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 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 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 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-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 denotes a protein, polypeptide, or peptide having 25 or fewer amino acid residues.

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 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', $F(ab')_2$, 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 (bispecific 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^{-7} M (e.g., at least about 10^{-8} M or 10^{-9} 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 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, 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 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 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 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 those of skill in the art include: Harrison's PRINCIPLES OF INTERNAL MEDICINE, 14th 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 GenBank accession number AC022315, which was deposited on February 10, 2000. That deposit has the human genomic sequence of gp354 (Fig. 6 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 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.

In preferred embodiments, the invention provides an isolated polynucleotide comprising (i) the nucleotide sequence of SEQ ID NOS:1, 5, 7 or 11; (ii) a degenerate variant of the nucleotide sequence of SEQ ID NOS:1, 5, 7 or 11; or (iii) the complement of (i) or (ii). SEQ ID NO:1 presents a predicted gp354 cDNA sequence, SEQ ID NO:5 presents the genomic DNA sequence comprising the gp354 coding sequences, including 5' and 3' non-transcribed regions, SEQ ID NO:7 presents a derived gp354 cDNA sequence which may be a splice variant of SEQ ID NO:1, and SEQ ID NO:11 presents a pancreatic gp354 cDNA sequence.

In other embodiments, the invention provides an isolated polynucleotide comprising (i) a nucleotide sequence that encodes a polypeptide with the amino acid sequence of SEQ ID NOS:2, 8 or 12; or (ii) the complement of a nucleotide sequence that encodes a polypeptide with the amino acid sequence of SEQ ID NOS:2, 8 or 12. SEQ ID NO:2 presents the amino acid sequence of GP354 encoded by the cDNA of SEQ ID NO:1. SEQ ID NO:8 present 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 (Fig.8).

In other embodiments, the invention provides an isolated polynucleotide having a nucleotide sequence that (i) encodes a polypeptide having the sequence of SEQ ID NOS:2, 8 or 12, (ii) encodes a polypeptide having the sequence of SEQ ID NOS:2 , 8 or 12 with conservative amino acid substitutions, or (iii) that is the complement of (i) or (ii), where SEQ ID NO:2 present the amino acid sequence of GP354 encoded by the cDNA of SEQ ID NO:1; SEQ ID NO:8 present 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.

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 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 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 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 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. 7). 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 is 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. 7). 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. 7).

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. 7).

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 (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 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. 7).

One preferred isolated polynucleotide of the invention is shown in Fig. 5 (see SEQ ID NO:3) and comprises nucleotides 139-923 of the gp354 cDNA 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:8) and GX1-219 (SEQ ID NO:9). 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 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 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 polynucleotide comprising a sequence that hybridizes under high stringency 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.

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 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 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 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 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 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

	cDNA-1	cDNA-2	genomic
exon 1	1-52	1-52	6483-6534
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
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
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

	cDNA	genomic
Exon 1	1-256	6278-6534
Exon 2	257-406	6699-6848
Exon 3	407-556	7762-7911
Exon 4	557-717	8058-8218
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
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
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 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.

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 NO:5; or the complements of such sequences.

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.

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 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% 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.

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 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:

http://www.ncbi.nlm.nih.gov/Blast/bl2seq/bl2.html.

To assess percent identity of nucleic acid sequences, the BLASTN 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 entireties.

The isolated polynucleotides of the present invention being useful for expression of GP354 proteins and protein fragments, the present invention thus provide isolated polynucleotides that encode GP354 proteins and portions thereof 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
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

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 paralogs). 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 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 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 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 (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.

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 regulation of gene expression using antisense genes, see Weintraub et al., Antisense RNA as a molecular tool for genetic analysis, REVIEWS--TRENDS IN GENETICS, Vol. 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 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 NOs: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. (1996) Bioorg. Med. Chem. Lett. 4:5-23. 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. PNAs 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*). 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 (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 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 — 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 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 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 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 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 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 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 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.

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 typically, but not invariably unlabeled.

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^+- 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.*, Schwarchzacher *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, 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 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.

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 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 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 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 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. 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 (1998) (ISBN: 071677013X); Tymms *et al.* (eds.), Gene Knockout Protocols, Humana Press (2000) (ISBN: 0896035727); Mak *et al.* (eds.), The Gene Knockout 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

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); Gacesa *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 (3rd 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 (4th ed.), John Wiley & Sons, 1999 (ISBN: 047132938X), the disclosures of which are incorporated herein by reference in their entireties. 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 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 YEp vectors containing centromere sequences, CEN, and autonomously replicating sequences, ARS. YACs are based on yeast linear plasmids, denoted YLp, 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.

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 α-aminoadipic acid (αAA), 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 expreSF™ 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 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 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 larger than can readily be accomodated 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 accomodated in vectors derived from plasmids or 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 (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, 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 variety

of other genetic elements operatively linked to the protein-encoding heterologous nucleic acid insert, typically genetic elements that drive and regulate 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 (1989) Proc. Natl. Acad. Sci. USA 86:5473-5477), a pancreas-specific promoter (Edlund et al. (1985) Science 230:912-916), 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 (1990) Science 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) Genes Dev. 3:537-546). 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 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 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^{TM} 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 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 7; SEQ ID NO: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 7; SEQ ID NOS:2 and 8 for GP354 domains and 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.

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 511-533 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 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 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 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-protein interaction; to inhibit protein-protein interactions between Ig domain 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. 7). 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. 7). 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. 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.7). 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 sequences presented in SEQ ID NO: 4, 8, 10 and 12 are deduced from the nucleotide sequences presented in SEQ ID NO: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. 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 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 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%, 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 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:

http://www.ncbi.nlm.nih.gov/Blast/bl2seq/bl2.html,

To assess percent identity of amino acid sequences, the BlastP 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 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 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; FZ::TN™ In-Frame Linker Insertion Kit, catalogue 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 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 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 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-γ OptEIA kit, catalog no. 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, 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.*, *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.

The invention also provides fragments of various of the proteins of the present invention. The protein fragments are useful as antigenic and 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 entireties. 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 entireties. 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 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.

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 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 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 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, 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 intracellular domain have particular utility. Fusions comprising GP354 Ig domains will be useful in engineering fusion proteins that will recognize other Ig domain-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 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.

Other useful protein fusions include those that permit display of the encoded protein on the surface of a phage or cell, fusions to intrinsically delectable 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.

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.

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.

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 nonnatural analogues, nonnative inter-residue bonds, or post-synthesis modifications will be limited to those that do not interfere with the immunogenicity 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 biotinoyl--(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 *t*BOC 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) 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 (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).

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).

The isolated GP3534 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.

For example, when produced by recombinant expression in eukaryotic cells, the isolated proteins, fragments, and fusion proteins of the present 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 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, BSOCOES, DFDNB, DMA, DMP, DMS, DPDPB, DSG, DSP (Lomant's Reagent), DSS, DST, DTBP, DTME, DTSSP, EGS, HBVS, Sulfo-BSOCOES, 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, KMUH, GMBS, LC-SMCC, LC-SPDP, MBS, M2C2H, MPBH, MSA, NHS-ASA, PDPH, PMPI, 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 (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 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.

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); 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 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 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 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 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 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 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.

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. 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 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.

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.

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 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 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 covalent or noncovalent. Porous substrates, commonly membranes, typically comprise nitrocellulose, polyvinylidene fluoride (PVDF), or cationically 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 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 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 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.

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 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 derivatives thereof, the binding of which can be competitively inhibited by one or 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.

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 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, 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.

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 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 1×10^{-6} molar (M), typically at least about 5×10^{-7} M, usefully at least about 1×10^{-7} M, with affinities and avidities of at least 1×10^{-8} M, 5×10^{-9} M, and 1×10^{-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.

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 immunohistochemical 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.*, *Current Opinion in Biotechnology* 8:503-508 (1997); Aujame *et al.*, *Human Antibodies* 8:155-168 (1997); Hoogenboom, *Trends in Biotechnol.* 15:62-70 (1997), de Kruif *et al.*, 17:453-455 (1996); Barbas *et al.*, *Trends in Biotechnol.* 14:230-234 (1996); Winter *et al.*, *Ann. 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 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-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, 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 (1998); Eldin *et al.*, *J. Immunol. Methods.* 201(1):67-75 (1997); and in

Saccharomyces cerevisiae, Franken *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 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.

Antibodies and fragments and derivatives thereof of the present invention may also be produced in plant cells, Giddings *et al.*, *Nature Biotechnol.* 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 (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* 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), 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

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, $F(ab)'_2$, 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 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 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 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 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, β-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}P, ^{32}P, ^{35}S, ^{3}H, and ^{125}I. As another example, when the antibodies of the present invention are used for radioimmunotherapy, the label can usefully be ^{228}Th, ^{227}Ac, ^{225}Ac, ^{223}Ra, ^{213}Bi, ^{212}Pb, ^{212}Bi, ^{211}At, ^{203}Pb, ^{194}Os, ^{188}Re, ^{186}Re, ^{153}Sm, ^{149}Tb, ^{131}I, ^{125}I, ^{111}In, ^{105}Rh, ^{99m}Tc, ^{97}Ru, ^{90}Y, ^{90}Sr, ^{88}Y, ^{72}Se, ^{67}Cu, or ^{47}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* 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 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 (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 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 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-activated Sepharose for purposes of immunoaffinity chromatography.

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, and 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.

The two proteins that are the most closely related to GP354 – *Drosophila* irregular chiasm protein (ICCR) and human nephrin protein (see FIG. 2) – are both involved in developmental patterning and cell-cell communication. Mutations at the ICCR locus in *Drosophila* affect sensory organ development in the fly, apparently due at least in part to abnormal apoptotic activity (Ramos, R.G.et al. (1993) Genes Dev. 7:2533-47). Mutations in the nephrin gene cause congenital nephritis in humans (Kestila, M. et al. (1998) Mol. Cell 1:575-582). Nephrin is localized to the glomerula slit diaphragm and is thought to play a role in cell adhesion (Ruotsalainen, V. et al. (1999) Proc Natl Acad Sci. 96:7962-7967). The similarity between GP354 and these two 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 nervous system.

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, 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 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 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 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 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 which affect the pancreas have also been described. For a more detailed description of diagnoses and treaments of pancreatic disorders and diseases, see Harrison's PRINCIPLES OF INTERNAL MEDICINE, 14th 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.

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 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, carragcenan, acacia, polyvinylpyrrolidone, and polyvinyl alcohol. The liquid formulations can also 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.

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, 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 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 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 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 described in Gennaro (ed.), Remington: The Science and Practice of Pharmacy, 20th ed., Lippincott, Williams & Wilkins (2000) (ISBN: 0683306472), Ansel *et al.*, Pharmaceutical Dosage Forms and Drug Delivery Systems, 7th ed., Lippincott Williams & Wilkins Publishers (1999) (ISBN: 0683305727); and Kibbe (ed.), Handbook of Pharmaceutical Excipients American Pharmaceutical Association, 3rd 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

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 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 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 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

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 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. (1987) Cell 51:503 for an exemplary description of homologous recombination vectors.

The vector is introduced into an embryonic stem cell line (e.g., by 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. (1992) Cell 69:915). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras. See e.g., Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene.

Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) Curr. Opin. Biotechnol. 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut et al. (1997) Nature 385:810-813. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G0 phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to

morula or 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.

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. (1992) Proc. Natl. Acad. Sci. USA 89:6232-6236) and the FLP recombinase system(O'Gorman et al. (1991) Science 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. 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

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 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 or11).

Antisense oligonucleotides are preferably directed to a regulatory region of a nucleotide sequence of SEQ ID NO:1, 7 or11, 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 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 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.

An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors

described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet 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 et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-O-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

B. Ribozymes and Catalytic Nucleic Acids

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 hammerhead, hairpin, Group I intron ribozymes, and the like) 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 (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., (1993) Science 261:1411-1418.

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. (1991) Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

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. (1996) Bioorg. Med. Chem. Lett. 4:5-23. 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 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, *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup., *supra* and Finn et al., Nuc. 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., Nuc. 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).

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 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 tissued. 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.

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. 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 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 polypeptides present in a parallel biological sample from non-pancreatic or non-neural tissue.

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.

The invention thus provides a method for diagnosing a disease condition in a subject, comprising the steps of comparing the amount or activity of 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 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.

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.

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: Gene Prediction and Sequence Analysis

The gene prediction software programs GENSCAN (Burge and Karlin, J. Mol. Biol. 268:78-94 (1997)) and GENEMARKHMM (Lukashin and Borodovsky, Nuc. 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*).

Genes for which GENSCAN and GENEMARKHMM yielded similar results were further analyzed. Specifically, the gene sequences were translated to protein sequences which were in turn used as queries in Blast analyses of the Genpept and Swissprot protein sequence databases.

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 at the website http://www.ncbi.nlm.nih.gov/. 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.

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.

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 adapters (Clontech). All RACE reactions use an adapter primer AP-1, 5'-CCATCCTAATACGACTCACTATAGGGC-3' (SEQ ID NO:14) provided with the kit. The 3' RACE for gp354 may use AP-1 together with the forward primer GX1-218, 5'-TACTGGGGGCTAGTTCAGTGGACTAA-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. The PCR was carried out using a multi-tissue cDNA panel (generated by reverse transcription PCR – "RT-PCR" – from mRNA isolated from these tissues) according to the manufacturer's specifications (Clontech). The multi-tissue cDNA panel provided double-stranded human cDNAs as templates for PCR. GX1-218 and GX1-219 (*supra*) were used as primers for the PCR. 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 at the last cycle.

The multi-tissue human cDNA panel contained cDNAs from the following tissues: brain, heart, kidney, liver, lung, pancreas, pituitary, skeletal muscle, colon, ovary, peripheral blood leukocyte, prostate, small intestine, spleen, testis, and thymus. The results are shown in Figure 3. A band of approximately 785 bp was observed in the pancreas and in no other tissues.

The PCR fragment from the pancreas was cloned into the PCR2.1 plasmid vector (Invitrogen, Carlsbad, CA). The resultant plasmid construct **[insert name]** was propagated and the insert was sequenced as described above. The sequence is shown as SEQ ID NO:3. Plasmid construct **[insert name]** was deposited on **[DATE]** according to the provisions of the Budapest Treaty, and was assigned the ATCC accession number designated: **[ATCC no.]**. All restrictions on the availability to the public of the above ATCC deposit will be irrevocably removed upon the granting of a patent on this application.

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

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. As in the PCR of pancreas cDNAs, a band of approximately 785 bp was observed in the Northern blot. No other tissues showed expression of GP354 (Figure 4).

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 (~10^6 PFU/ml) (Clontech) are added to 6 ml of an overnight culture of K802 cells (Clontech), and then distributed as 250 ml 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 $MgSO_4$ plates and incubated overnight at 37°C. To each plate 5 ml of SM phage buffer (0.1 M NaCl, 8.1 mM $MgSO_4 \cdot 7H_2O$, 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 ml containing 8.8 ml H_2O, 4 ml 5X RAPID-LOAD BUFFER (Origene), 2 ml 10X PCR BUFFER II (Perkin Elmer), 2 ml 25 mM MgCl2, 0.8 ml 10 mM dNTP, 0.12 ml of a primer comprising at least a portion of the sequence of the 5' end of the gp354 polynucleotide of SEQ 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 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 used to set up five 1:10 dilutions that are plated onto LB agar + 10 mM $MgSO_4$ 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 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 subdividions 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 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 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 in a 50 µl reaction containing 33 µl H_2O, 5 µl 10X TT buffer (140 mM ammonium sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 µl 15 mM $MgSO_4$, 2 µl 10 mM 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 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 ethanol-precipitated and resuspended in 12 µl H_2O 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 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 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 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 Cycle Sequencing Ready Reaction Kit. The cycle-sequencing reaction may contain 14 ml of H_2O, 16 ml of BIGDYE Terminator mix, 7 ml genomic phage DNA (0.1 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_2O, 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, 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 [^{35}S]-UTP to yield a specific activity of about 0.40×10^6 cpm/pmol for antisense riboprobes and about 0.65×10^6

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.

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 violate-stained cell bodies. Autoradio-graphic 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 ^{32}P-labeling. The PCR is performed as follows: (1) mix the following reagents:

1 µl	gp354-containing plasmid
2 µl	forward primer
2 µl	reverse primer
10 µl	10X PCR buffer provided by the manufacturer of the Taq polymerase (e.g., Amersham Pharmacia Biotech)
1 µl	10mM dNTP (e.g., Boehringer Mannheim catalogue no. 1 969 064)
0.5 µl	Taq polymerase (such as Amersham Pharmacia Biotech catalogue no. 27-0799-62)
83.5 µl	water

(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 labeled with ^{32}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 (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 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) 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 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 from the host cell.

B. Expression of GP354 in 293 cells

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 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' 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 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, 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.

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 hormone) polyadenylation sequence and a multiple cloning site. In addition, the plasmid contains the dhrf (dihydrofolate reductase) gene which provides selection 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-CI. 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 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 pAclM1, 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 ^{35}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 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 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 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 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) 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, 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. 105 transformants/mg of DNA) with both the GP354 and library fusion plasmids according to standard 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 (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.), ANTIBODIES, A LABORATORY MANUAL, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1988). 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 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 10^6 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

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 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 preferred to be of the IgG_4 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 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 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); Verhoeyen et al., Science 239:1534-36 (1988); and Tempest et al., Bio/Technology 9:266-71 (1991). If necessary, the β-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).

In an alternative approach, the surface of a non-human monoclonal antibody of interest is humanized by altering selected surface residues of the 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 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 techniques such as those described in Aujame et al., Human Antibodies 8(4):155-168 (1997); Hoogenboom, TIBTECH 15:62-70 (1997); 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 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.

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 Bruggemann et al., Curr. Opin. Biotechnol. 8:455-58 (1997). Transgenic mice carrying human V-gene segments in germline configuration and 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).

Example 12: Assays to Identify Modulators of GP354 Activity

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 Sciences 18:529-541 (1976); Dooley et al., J. of Pharmacol. Exp. Therap. 283(2): 735-41 (1997), and George et al., J. of 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 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 [^{125}I]-cAMP, and the plate is counted using a Packard TOPCOUNT 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells or from standards and fixed amounts of [^{125}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 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. Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

B. Acquorin Assays

In another assay, cells (e.g., CHO cells) are transiently co-transfected with a gp354 expression construct and a construct that encodes the photoprotein apoaquorin. In the presence of the cofactor coelenterazine, apoaquorin 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 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 expression construct and a construct that encodes the photoprotein apoaquorin (Molecular Probes) using the transfection reagent FUGENE 6 (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 $2X10^5$ 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 (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 modulators of GP354 activity. For example, CHO cells stably transfected with a gp354 expression vector are plated at a density of $4X10^4$ 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 μM. 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 μM; positive control), or ATP (4 μM; 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.

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. and 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 μl of fresh MEM, or MEM containing a known mitogen, is added along with 20 μl 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. Untransfected cells or cells transfected with vector alone also may serve as controls.

After culture for 16-18 hours, 1 μCi 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 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 x [C/ (D + C)] + G where A is the 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 serum. Antagonists that bind to the receptor will inhibit the stimulation seen with a known agonist by up to 100%.

F. [^{35}S]GTPgS Binding Assay

It is possible to evaluate whether GP354 signals through a G protein-mediated pathway. G protein-coupled receptors signal through 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]GTPgS 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).

In one exemplary assay, cells stably transfected with a gp354 expression vector are grown in 10 cm tissue culture dishes to subconfluence, rinsed once with 5 ml of ice-cold Ca^{2+}/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 TBE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM MgCl2, 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at -70°C until use.

Aliquots of cell membranes prepared as described above and stored at -70°C are thawed, homogenized, and diluted into buffer containing 20 mM HEPES, 10 mM MgCl2, 1 mM EDTA, 120 mM NaCl, 10 µM GDP, and 0.2 mM ascorbate, at a concentration of 10-50 µg/ml. In a final volume of 90 µl, homogenates are incubated with varying concentrations of candidate modulator compounds or 100 µM GTP for 30 minutes at 30°C and then placed on ice. To each sample, 10 µl guanosine 5'-O-(3[^{35}S]thio) triphosphate (NEN, 1200 Ci/mmol; [^{35}S]-GTPgS), was added to a final concentration of 100-200 pM. Samples are incubated at 30°C for an additional 30 minutes, 1 ml of 10 mM HEPES, pH 7.4, 10 mM MgCl2, at 4°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 $MgCl_2$. Filters are counted by liquid scintillation spectroscopy. Nonspecific binding of [^{35}S]-GTPgS is measured in the presence of 100 µM GTP and subtracted from the total. Compounds are selected that modulate the amount of [^{35}S]-GTPgS binding in the cells, compared to untransfected control cells. Activation of receptors by agonists gives up to a five-fold increase in [^{35}S]-GTPgS binding. This response is blocked by antagonists.

G. MAP Kinase Activity Assay

Evaluation of MAP kinase activity in cells expressing GP354 provides another assay to identify modulators of GP354 activity. See, e.g., *Lajiness* et al., J. Pharmacol. Exp. Therap. 267(3):1573-1581 (1993) 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 $7X10^4$ 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 $MgCl_2$, 0.5 mM EGTA, 0.5 mM sodium vanadate, 1 mM benzamidine, 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:9), Upstate Biotechnology, Inc., NY) and 50 µM [g-^{32}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^2 squares of Whatman P81 phosphocellulose paper. The filter squares are washed in 4 changes of 1% H_3PO_4, 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. 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.

H. [^{3}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., Molecular Pharmacology 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 $1.5X10^4$ 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 [^{3}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 [^{3}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 $4X10^5$ cells/cup in MEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 10 U/ml penicillin, and 10 μg/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO2 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.

CLAIMS

What is claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) SEQ ID NO:1, 3, 5, 7, 9 or 11 and allelic variants thereof;

(b) a fragment of (a), consisting of at least 17 nucleotides;

(c) a nucleotide sequence that encodes SEQ ID NO:2, 4, 8, 10, or 12, an allelic variant thereof, or fragments of either consisting of at least six amino acid residues;

(d) a nucleotide sequence that is at least 65% identical to SEQ ID NO:2, 8, 10, or 12;

(e) a nucleotide sequence that encodes a protein that is at least 80% homologous to SEQ ID NO:2, 4, 8, 10, or 12;

(f) a nucleotide sequence that hybridizes to SEQ ID NO:1, 3, 5, 7, 9 or 11 under high stringency conditions; and

(g) a complementary sequence of any of (a) through (f).

2. The polynucleotide of claim 1, comprising SEQ ID NO:1, 3, 5, 7, 9 or 11 and allelic variants thereof.

3. The polynucleotide of claim 1, comprising a sequence encoding SEQ ID NO:2, 4, 8, 10, or 12 or an allelic variant thereof.

4. The polynucleotide of claim 1, comprising a sequence encoding SEQ ID NO:8 or 10, or an allelic variant thereof.

5. The polynucleotide of claim 1, further comprising a transcription regulatory sequence operatively linked to any of (a) through (g).

6. The polynucleotide of claim 1, further comprising a nucleic acid sequence encoding a heterologous polypeptide.

7. The polynucleotide of claim 1, comprising a nucleotide sequence that is complementary to (i) SEQ ID NO:1, 3, 5, 7, 9 or 11, or (ii) a fragment thereof having at least 17 nucleotides.

8. A vector comprising a polynucleotide of claim 1.

9. A vector comprising a polynucleotide of any one of claims 2-7.

10. The vector of claim 8, which is a plasmid vector.

11. The vector of claim 8, which is a viral vector.

12. The vector of claim 11, selected from the group consisting of baculoviruses, adenoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, lentiviruses and retroviruses.

13. A host cell containing the polynucleotide of claim 1.

14. The host cell of claim 13, wherein the host cell is selected from the groups consisting of a bacterial cell, insect cell, yeast cell, plant cell and mammalian cell.

15. The host cell of claim 13, wherein the host cell is a human cell.

16. An isolated polypeptide encoded by the nucleic acid of claim 1.

17. The polypeptide of claim 16, comprising (i) SEQ ID NO:2, 4, 8, 10 or 12, or an allelic variant thereof, or (ii) a fragment of (i) consisting of at least six amino acid residues.

18. The polypeptide of claim 16 or 17, further comprising a heterologous sequence.

19. A composition comprising the nucleic acid of claim 1 and a pharmaceutically acceptable carrier.

20. A composition comprising the polypeptide of claim 16 or 17 and a pharmaceutically acceptable carrier.

21. An antibody that binds to an epitope on a polypeptide of claim 16.

22. The antibody of claim 21, wherein the antibody is a monoclonal antibody.

23. The antibody of claim 22, wherein the antibody is a humanized or a fully human antibody.

24. A composition comprising the antibody of any one of claims 21-23 and a pharmaceutically acceptable carrier.

25. A method of producing a polypeptide, comprising the steps of:

culturing the host cell of claim 13 in a medium under conditions in which said nucleic acid is expressed, and

recovering the polypeptide from the cell or from the culture medium.

26. A method of determining the presence of a gp354-encoding sequence in a sample, comprising the steps of:

contacting the sample with the isolated polynucleotide of claim 1 under high stringency hybridization conditions, and

detecting hybridization of said 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.

27. A method of determining the presence of a GP354 protein in a sample, comprising the steps of:

contacting the sample with the antibody of claims 21, 22 or 23; and

detecting specific binding of said antibody to an antigen, wherein the occurrence of said specific binding indicates the presence of a GP354 protein in the sample.

28. A method of identifying a compound that binds a GP354 protein, comprising the steps of:

contacting a GP354 protein with a test compound; and

detecting a complex formed by said GP354 protein and said test compound, wherein the presence of said complex indicates that said test compound binds to said GP354 protein.

29. A method of identifying a compound that modulates the activity of a GP354 protein, comprising the steps of:

contacting said GP354 protein with a test compound, and

determining the effect of the test compound on the activity of said GP354 protein, whereas a change of said activity after the contacting step indicates that said test compound modulates the activity of said GP354 protein.

30. A method of identifying a homolog of a human gp354 gene, comprising the steps of screening a nucleic acid database with a query sequence consisting of SEQ ID NO:1, 3, 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, 7, 9 or 11 or said portion thereof, if found, is a homolog of a human gp354 gene.

31. A method of identifying a homolog of a human gp354 gene, comprising the steps 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 consisting at least 17 nucleotides, under medium or high stringency hybridization conditions; and

determining whether said nucleic acid probe hybridizes to a nucleic acid sequence in the library,

wherein the nucleic acid sequence so hybridized is a homolog of a human gp354 gene.

32. A method of diagnosing a disease condition in a subject, comprising the step of comparing the amount or activity of a GP354 protein in a tissue sample from said subject to the amount or activity of the GP354 polypeptide in a control sample, wherein a significant difference in the amount or activity of said GP354 polypeptide in said tissue sample relative to the amount or activity of said GP354 polypeptide in said control sample indicates that the subject has a disease condition.

33. The method of claim 32, wherein the disease condition relates to the pancreas.

34. The method of claim 32, wherein the disease condition relates to the central nervous system.

35. A method of diagnosing a disease condition in a subject, comprising the step of comparing the amount of a gp354 mRNA in a tissue sample from the subject to the amount of said gp354 mRNA in a control sample, wherein a significant difference in the amount of said mRNA in said tissue sample relative to the amount of said mRNA in said control sample indicates that the subject has a disease condition.

36. The method of claim 35, wherein the disease condition relates to the pancreas.

37. The method of claim 35, wherein the disease condition relates to the central nervous system.

38. A diagnostic assay for 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 GP354 protein; (ii) mis-regulation of a gene encoding a GP354 protein; and (iii) aberrant post-translational modification of a GP354 protein; comprising the steps of

separately hybridizing nucleic acids from the test cell and from a reference cell that lacks said genetic lesion or mutation with a nucleic acid probe comprising SEQ ID NO:1, 3, 7, 9 or 11, or a portion thereof having at least 17 nucleotides, under high stringency hybridization conditions; and

separately washing said nucleic acid hybrids under high stringency wash conditions to allow dissociation of the hybrids; and

determining whether said nucleic acid probe dissociates more readily from the nucleic acids of the test cell compared to the nucleic acids of the reference cell.

39. The use of a composition of claims 18 or 19 for the treatment of a pancreatic injury.

40. The use of a composition of claims 18 or 19 for the treatment of an abnormal or disease condition that relates to the pancreas.

41. The use of claim 40, wherein the condition is selected from the 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.

42. The use of a composition of claims 18 or 19 for the treatment of an injury to the central nervous system.

43. The use of a composition of claims 18 or 19 for the treatment of an abnormal or disease condition that relates to the central nervous system.

44. The use of claim 43, 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.

```
    atgcgggtccccgccctcctcgtcctcctcttctgcttcagagggagcgcaggcccgtcg
1   ---------+---------+---------+---------+---------+---------+
    M  R  V  P  A  L  L  V  L  L  F  C  F  R  G  S  A  G  P  S

    ccccatttcctgcaacagccagaggacctggtggtgctgctgggggaggaagcccggctg
61  ---------+---------+---------+---------+---------+---------+
    P  H  F  L  Q  Q  P  E  D  L  V  V  L  L  G  E  E  A  R  L

    ccgtgtgctctgggcgcctactgggggctagttcagtggactaagagtgggctggccctа
121 ---------+---------+---------+---------+---------+---------+
    P  C  A  L  G  A  Y  W  G  L  V  Q  W  T  K  S  G  L  A  L

    ggggggccaaagggacctaccagggtggtcccgtactggatatcagggaatgcagccaat
181 ---------+---------+---------+---------+---------+---------+
    G  G  Q  R  D  L  P  G  W  S  R  Y  W  I  S  G  N  A  A  N

    ggccagcatgacctccacattaggcccgtggagctagaggatgaagcatcatatgaatgt
241 ---------+---------+---------+---------+---------+---------+
    G  Q  H  D  L  H  I  R  P  V  E  L  E  D  E  A  S  Y  E  C

    caggctacacaagcaggcctccgctccagaccagcccaactgcacgtgctggtcccccca
301 ---------+---------+---------+---------+---------+---------+
    Q  A  T  Q  A  G  L  R  S  R  P  A  Q  L  H  V  L  V  P  P

    gaagcccccaggtgctgggcggcccctctgtgtctctggttgctggagttcctgcgaac
361 ---------+---------+---------+---------+---------+---------+
    E  A  P  Q  V  L  G  G  P  S  V  S  L  V  A  G  V  P  A  N

    ctgacatgtcggagccgtggggatgcccgccctacccctgaattgctgtggttccgagat
421 ---------+---------+---------+---------+---------+---------+
    L  T  C  R  S  R  G  D  A  R  P  T  P  E  L  L  W  F  R  D

    ggggtcctgttggatggaaccaccttccatcagaccctgctgaaggaagggaccccctggg
481 ---------+---------+---------+---------+---------+---------+
    G  V  L  L  D  G  T  T  F  H  Q  T  L  L  K  E  G  T  P  G

    tcagtggagagcaccttaaccctgacccctttcagccatgatgatggagccacctttgtc
541 ---------+---------+---------+---------+---------+---------+
    S  V  E  S  T  L  T  L  T  P  F  S  H  D  D  G  A  T  F  V

    tgccgggcccggagccaggccctgcccacaggaagagacacagctatcacactgagcctg
601 ---------+---------+---------+---------+---------+---------+
    C  R  A  R  S  Q  A  L  P  T  G  R  D  T  A  I  T  L  S  L
```

FIG. 1

```
                cagtacccccagaggtgactctgtctgcttcgccacacactgtgcaggagggagagaag
661             ---------+---------+---------+---------+---------+---------+
                Q  Y  P  P  E  V  T  L  S  A  S  P  H  T  V  Q  E  G  E  K

                gtcattttcctgtgccaggccacagcccagcctcctgtcacaggctacaggtgggcaaaa
721             ---------+---------+---------+---------+---------+---------+
                V  I  F  L  C  Q  A  T  A  Q  P  P  V  T  G  Y  R  W  A  K

                gggggctctccggtgctcggggccccgcgggccaaggttagaggtcgtggcagacgcctcg
781             ---------+---------+---------+---------+---------+---------+
                G  G  S  P  V  L  G  A  R  G  P  R  L  E  V  V  A  D  A  S

                ttcctgactgagcccgtgtcctgcgaggtcagcaacgccgtgggtagcgccaaccgcagt
841             ---------+---------+---------+---------+---------+---------+
                F  L  T  E  P  V  S  C  E  V  S  N  A  V  G  S  A  N  R  S

                actgcgctggatgtgctgtttgggccgattctgcaggcaaagccggagcccgtgtccgtg
901             ---------+---------+---------+---------+---------+---------+
                T  A  L  D  V  L  F  G  P  I  L  Q  A  K  P  E  P  V  S  V

                gacgtgggggaagacgcttccttcagctgcgcctggcgcgggaacccgcttccacgggta
961             ---------+---------+---------+---------+---------+---------+
                D  V  G  E  D  A  S  F  S  C  A  W  R  G  N  P  L  P  R  V

                acctggacccgcctgcggtggcgcgcaggtgctgggctctggagccacactgcgtcttccg
1021            ---------+---------+---------+---------+---------+---------+
                T  W  T  R  R  G  G  A  Q  V  L  G  S  G  A  T  L  R  L  P

                tcggtggggcccgaggacgcaggcgactatgtgtgcagagctgaggctgggctatcgggc
1081            ---------+---------+---------+---------+---------+---------+
                S  V  G  P  E  D  A  G  D  Y  V  C  R  A  E  A  G  L  S  G

                ctgcggggcggcgccgcggaggctcggctgactgtgaacgctcccccagtagtgaccgcc
1141            ---------+---------+---------+---------+---------+---------+
                L  R  G  G  A  A  E  A  R  L  T  V  N  A  P  P  V  V  T  A

                ctgcactctgcgcctgccttcctgaggggccctgctcgcctccagtgtctggttttcgcc
1201            ---------+---------+---------+---------+---------+---------+
                L  H  S  A  P  A  F  L  R  G  P  A  R  L  Q  C  L  V  F  A

                tctcccgccccagatgccgtggtctggtcttgggatgagggcttcctggaggcggggtcg
1261            ---------+---------+---------+---------+---------+---------+
                S  P  A  P  D  A  V  V  W  S  W  D  E  G  F  L  E  A  G  S
```

FIG. 1(cont.)

```
                cagggccggttcctggtggagacattccctgccccagagagccgcgggggactgggtccg
1321            ---------+---------+---------+---------+---------+---------+
                Q  G  R  F  L  V  E  T  F  P  A  P  E  S  R  G  G  L  G  P

                ggcctgatctctgtgctacacatttcggggacccaggagtctgactttagcaggagcttt
1381            ---------+---------+---------+---------+---------+---------+
                G  L  I  S  V  L  H  I  S  G  T  Q  E  S  D  F  S  R  S  F

                aactgcagtgcccggaaccggctgggcgagggaggtgcccaggccagcctgggccgtaga
1441            ---------+---------+---------+---------+---------+---------+
                N  C  S  A  R  N  R  L  G  E  G  G  A  Q  A  S  L  G  R  R

                gacttgctgcccactgtgcggatagtggccggagtggccgctgccaccacaactctcctt
1501            ---------+---------+---------+---------+---------+---------+
                D  L  L  P  T  V  R  I  V  A  G  V  A  A  A  T  T  T  L  L

                atggtcatcactggggtggccctctgctgctggcgccacagcaaggcctctttctccgag
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

                gaagaggagacaggcagccgcgaggaccggggccccattgtgcacactgaccacagtgat
1681            ---------+---------+---------+---------+---------+---------+
                E  E  E  T  G  S  R  E  D  R  G  P  I  V  H  T  D  H  S  D

                ctggttctggaggaggaagggactctggagaccaag
1741            ---------+---------+---------+------  1776
                L  V  L  E  E  E  G  T  L  E  T  K
```

FIG. 1(cont.)

```
GP354     ------MRVPALLVLLFCFRGSAGPSPHFLQQ--------PEDLVVLLGEE
ICCR      ------MLHTMQLLLLATIVGMVRSSPYTSYQNQRFAMEPQDQTAVVGAR
Nephrin   MALGTTLRASLLLLGLLTEGLAQLAIPASVPRG---FWALPENLTVVEGAS
                     * *          *              *       *

GP354     ARLPCALGAYWGLVQWTKSGLALGGQRDLPGWSRYWISGNAANGQHDLHI
ICCR      VTLPCRVINKQGTLQWTKDDFGLGTSRDLSGFERYAMVGSDEEGDYSLDI
Nephrin   VELRCGVSTPGSAVQWAKDGLLLGPDPRIPGFPRYRLEGDPARGEFHLHI
            * *          ** *    **      *  **    *    *   * *

GP354     RPVELEDEASYECQAT-----QAGLRSRPAQLHVLVPPEAPQVL--GGPS
ICCR      YPVMLDDDARYQCQVSPGPEGQPAIRSTFAGLTVLVPPEAPKIT--QGDV
Nephrin   EACDLSDDAEYECQVGRS-EMGPELVSPRVILSILVPPKLLLLTPEAGTM
              * * * * **             *    *  ****          *

GP354     VSLVAGVPANLTCRSRGDARPTPELLWFR--DGVLLDGTTFHQTLLKEGT
ICCR      IYATADRKVEIECVSVG-GKPAAEITWIDGLGNVLTDNIEYTVIPLPDQR
Nephrin   VTWVAGQEYVVNCVS-GDARPAPDITILLS--GQTISDISANVNEGSQQK
            *       * * *   *

GP354     PGSVESTLTLTPFSHDDGATFVCRARSQALPTGRDTAITLSLQYPPEVTL
ICCR      RFTAKSVLRLTPKKEHHNTNFSCQAQNTADRTYRSAKIRVEVKYAPKVKV
Nephrin   LFTVEATARVTPRSSDNRQLLVCEASSPALEAPIKASFTVNVLFPPGPPV
                    **          * *   *              *

GP354     ------------------SASPHTVQ-----EGEKVIFLCQATAQPPVTG
ICCR      NVMGSLPGGAGGSVGGAGGGSVHMSTGSRIVEHSQVRLECRADANPSDVR
Nephrin   IEWF----------------GLDEGHVR-----AGQSLELPCVARGGNPLAT
                                                 *  *
                                                 *

GP354     YRWAKGGSPVLGAR-GPRLEVVADASFLTEP--------VSCEVSNAVGS
ICCR      YRWFINDEPIIGGQKTEMVIRNVTRKFHDAI--------VKCEVQNSVGK
Nephrin   LQWLKNGQPVSTAWGTEHTQAVARSVLVMTVRPEDHGAQLSCEAHNSVSA
            *     *                                ** * *

GP354     --ANRSTALDVLFGPILQAKPEPVSVDVGEDASFSCAWRGN-PLPRVTWT
ICCR      --SEDSETLDISYAPSFRQRPQSMEADVGSVVSLTCRVDSN-PQPEIVWI
Nephrin   GTQEHGITLQVTFPPSAIILLGSASQTENKNVTLSCVSKSSRPRVLLRWW
                  *     *                      *       *     *

GP354     R----------RGGAQVLG---SGATLRLPSVGPEDAGDYVCRAEAGLSG
ICCR      Q----------HPSDRVVG---TSTNLTF-SVSNETAGRYYCKAN--VPG
Nephrin   LGWRQLLPMEETVMDGLHGGHISMSNLTFLARREDNGLTLTCEAFSEAFT
                           *       *              * *
```

FIG. 2

```
GP354     LRGGAAEARLTVNAPFVVTALH----SAPAFLRGPARLQCLVFASFAPDA
ICCR      YAEISADAYVYLKGSPAIGSQR----TQYGLVGDTARIECFASSVPRARH
Nephrin   KETFKKSLILNVKYPAQKLWIEGPPEGQKLRAGTRVRLVCLAIGGNPEPS
                                          *  *

GP354     VVWSWD----------------EGFLE-AGSQGR--FLVETFPAP----
ICCR      VSWTFNG---------------QEISSE-SGHDYS--ILVDAVPG-----
Nephrin   LMWYKDSRTVTESRLPQESRRVHLGSVEKSGSTFSRELVLVTGPSDNQAK
            *                          *  *           *

GP354     ----------ESRGGLGPGLISVLHISGTQESDFSRSFNCSARNRLGEGG
ICCR      -------------------GVKSTLIIRDSQAYHYG-KYNCTVVNDYGNDV
Nephrin   FTCKAGQLSASTQLAVQFPPTNVTILANASALRPGDALNLTCVSVSSNP-
                                             *

GP354     AQASLGRRDLLPTVRIVAG---------VAAATTTLLMVIT---GVALCC
ICCR      AEIQLQAKKSVSLLNTIVGGI-------SVVAFLLVLPILV---VVYIKC
Nephrin   -PVNLSWDKEGERLEGVAAPPRRAPFKGSAAARSVLLQVSSRDHGQRVTC
            *                          *    *           *

GP354     WRHSKASFSEQKNLMRIP----------------------GS----------
ICCR      KKRTKLPPADVISEHQITKNGG----------VSCKLEPGDRTSNYSDLKV
Nephrin   RAHSAELRETVSSFYRLNVLYRPEFLGEQVLVVTAVEQGEALLPVSVSAN
                                                 *

GP354     -------------SDGSSSRGPEE--EETG-------SREDRG--PIVHTD
ICCR      DISGGY---VPYGDYSTHYSPPPQYLTTC-------STKSNGSSTIMQNN
Nephrin   PAPEAPNWTFRGYRLSPAGGPRHRILSSGALHLWNVTRADDG-LYQLHCQ
                        *    *                    *

GP354     HSDLVLEEEGTLETK----------------------------------
ICCR      HQNQLQLQQQQQQSHHQHHTQTTTLPMTFLTN----------SSGGSLTGS
Nephrin   NSEGTAEARLRLDVHYAPTIRALQDPTEVNVGGSVDIVCTVDANPILPGM

GP354     ----------------------------------------------
ICCR      IIGSREIR--QDNGLPSLQSTT-ASVVSSSPNGSCSNQSTTAATTTTTHV

Nephrin   FNWERLGEDEEDQSLDDMEKISRGPTGRLRIHHAKLAQAGAYQCIVDNGV

GP354     ----------------------------------------------
ICCR      VVPSSMALSVDPRYSAIYGNPYLRSSNSSLLPPPTAV-----------
Nephrin   APPARRLLRLVVRFAPQVEHPTPLTKVAAAGDSTSSATLHCRARGVPNIV

GP354     ----------------------------------------------
```

FIG. 2(cont.)

```
ICCR      ----------------------------------------------------
Nephrin   FTWTKNGVPLDLQDPRYTEHTYHQGGVHSSLLTIANVSAAQDYALFTCTA

GP354     ----------------------------------------------------
ICCR      ----------------------------------------------------
Nephrin   TNALGSDQTNIQLVSISRPDPPSGLKVVSLTPHSVGLEWKPGFDGGLPQR

GP354     ----------------------------------------------------
ICCR      ----------------------------------------------------
Nephrin   FCIRYEALGTPGFHYVDVVPPQATTFTLTGLQPSTRYRVWLLASNALGDS

GP354     ----------------------------------------------------
ICCR      ----------------------------------------------------
Nephrin   GLADKGTQLPITTPGLHQPSGEPEDQLPTEPPSGPSGLPLLPVLFALGGL

GP354     ----------------------------------------------------
ICCR      ----------------------------------------------------
Nephrin   LLLSNASCVGGVLWQRRLRRLAEGISEKTEAGSEEDRVRNEYEESQWTGE

GP354     ----------------------------------------------------
ICCR      ----------------------------------------------------
Nephrin   RDTQSSTVSTTEAEPYYRSLRDFSPQLPPTQEEVSYSRGFTGEDEIMAFP

GP354     ----------------------------------------------------
ICCR      ----------------------------------------------------
Nephrin   GHLYDEVERTYPPSGAWGPLYDEVQMGPWDLHWPEDTYQDPRGIYDQVAG

GP354     --------------------
ICCR      --------------------
Nephrin   DLDTLEPDSLPFELRGHLV
```

FIG. 2(cont.)

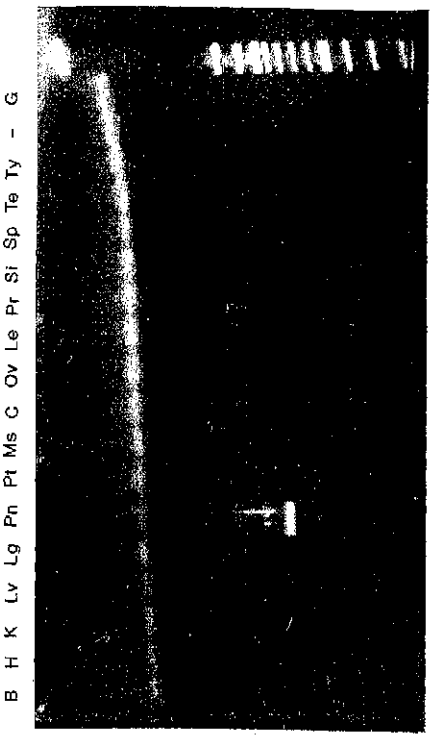


FIG. 3

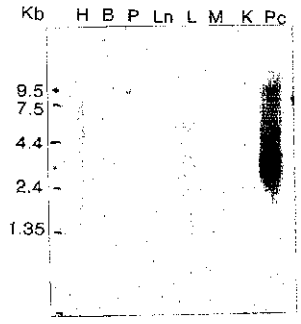


FIG. 4

```
  1  TACTGGGGGC TAGTTCAGTG GACTAAGAGT GGGCTGGCCC TAGGGGGCCA
 51  AAGGGACCTA CCAGGGTGGT CCCGGTACTG GATATCAGGG AATGCAGCCA
101  ATGGCCAGCA TGACCTCCAC ATTAGGCCCG TGGAGCTAGA GGATGAAGCA
151  TCATATGAAT GTCAGGCTAC ACAAGCAGGC CTCCGCTCCA GACCAGCCCA
201  ACTGCACGTG CTGGTCCCCC CAGAAGCCCC CCAGGTGCTG GGCGGCCCCT
251  CTGTGTCTCT GGTTGCTGGA GTTCCTGCGA ACCTGACATG TCGGAGCCGT
301  GGGGATGCCC GCCCTACCCC TGAATTGCTG TGGTTCCGAG ATGGGGTCCT
351  GTTGGATGGA ACCACCTTCC ATCAGACCCT GCTGAAGGAA GGGACCCCTG
401  GGTCAGTGGA GAGCACCTTA ACCCTGACCC CTTTCAGCCA TGATGATGGA
451  GCCACCTTTG TCTGCCGGGC CCGGAGCCAG GCCCTGCCCA CAGGAAGAGA
501  CACAGCTATC ACACTGAGCC TGCAGTACCC CCCAGAGGTG ACTCTGTCTG
551  CTTCGCCACA CACTGTGCAG GAGGGAGAGA AGGTCATTTT CCTGTGCCAG
601  GCCACAGCCC AGCCTCCTGT CACAGGCTAC AGGTGGGCAA AAGGGGCTC
651  TCCGGTGCTC GGGGCCCGCG GGCCAAGGTT AGAGGTCGTG GCAGACGCCT
701  CGTTCCTGAC TGAGCCCGTG TCCTGCGAGG TCAGCAACGC CGTGGGTAGC
751  GCCAACCGCA GTACTGCGCT GGATGTGCTG TTTGG
```

FIG. 5

```
   1  TCCCCGCTCT TCTCAACTCC TTGCTGGGTT GTACCATGCA CCCTATCCCT
  51  CAGCTTCTCA TGTCTGCACC AGCGCTACTG CCCATATTTC TATCTGGGCC
 101  TCAGCCTTGT GCTGGTTGCT GCCGCCCTCG ATGTGCCCTC GCATCCACTG
 151  GGTCCCACAC TGGCCTCAGC ATCTCCCCAC ACCTTCTCCT GGGTCCCCAT
 201  CCCAGGGATG ACATCTTTTC TGGGGCCCTT AGAAGGGTAC TGGTCAGGAA
 251  CACACACCCT TCCCACTCCA GAGGCTTCAT GCTGCCCCCT GCCACCCAGT
 301  TCACCCACAC TCACTCAGGA GAATGGTGAT GTCAGGTGCT GGCTTCGCGT
 351  CCCCAGACAC ACAGTTGACC ACGTACTCCT GCCCAGCTAC CCAGGTGACC
 401  ATGGTGCCTG CCTCTGGGGT CAGCAGGAGC AGCTTGGGAG GAACTGGTGA
 451  GAGAAGGGTC TGGGGTAAGC TTCCAGCACT GAGAAGGACT TGAAGATTGG
 501  AGTTCGGTAC CCAGAGTCTG GGAGAGGAGA GGCTGGGGGC TTGGACTTCC
 551  GGGTTGCGGG GTAGGGGAGG GCTTGAAGCC CAGACTCATG GGTCCTGGGC
 601  GTCTCTCACC CATACCCAGG ATGGAGAGGA TCACTCTGGG AGACACGAGC
 651  TCGGGCCCCA TCTCAGAGCG GCCGACCTGG CACTCATACT CCGCGTCATC
 701  GCTGAGGTCA CAGGCCTCGA TGTGCAGGTG GAATTCACCT GCAGGGGGAG
 751  CCGGAAGTCA GGGCCGCAGC TTCCGCTGGT GGCTGAGGGT CTCAGGCTCT
 801  GATCCCTTAC CTCTAGCAGG GTCCCCTTCC AGGCGGTACC TGGGGAAGCC
 851  TGGGATCCTG GGGTCGGGGC CCAGGAGCAG CCCATCTTTG GCCCATTGCA
 901  CCGCACTGCC AGGGGTGGTG ACCCCACAAC CCACCTCCAC TGAGGCCCCC
 951  TCCACCACCG TCAGGTTTTC AGGCAGGGCC CAGAAGCCCC GGGGAACGGA
1001  GGCAGGAATC GCCAACTGCG CCAGGCCTGA GGACACAGCG CGGTGCAAGG
1051  AAAGGGCAGA GGGTTTGTCT AGGGAAGGTA AGTGGGAAAT GGGGGCCACT
1101  TGGCGCTGGG TACAAGGCTG GGATCCCACT CACCTTCAGT CAGCAGCCCC
1151  AGGAGCAGGA GAGAAGCCCT GAGCGTCGTC CCCAGGGCCA TCACAGGTCC
1201  CCCTACTGTG ACCCCGACAG CGCCCGCTGC CAGCCACCTG CGTCTGTCTG
1251  GCTTTCTCTG GGTCCCTCTC TGTGTGTCTC TGCCACCTGC TTTTCTTTTT
1301  TATCTCTTTC CGTTACTCTC CTCCCTTTCT CGTTTTCCTC TTCCCCTCTT
1351  CCCTGTGAGT ATCTCTCTCT GTCTTGCTCT CAGTCTCAAT CTCTGAGTCT
1401  CTTTCTCTGT CTCTTTAAAA AAACTTTTTT TTCTTTTTTC TTTTTTTTTT
1451  CTTTTTTTTT TTTTTAGAGA CGGGGTCTCA CTATGTTGCC CAGGTTGATC
```

FIG. 6

```
1501  TCAGACTCTT TCCTTCAAGC CATCCTCCCA CCTTGGCCTC CCCAAGTGTT
1551  GGGATTACAG GCGTGAGCCA CTGCGCCCAG TCTCTTTATC TTTCCATCTT
1601  TCTCTCCTTG TCTAAGCCGT TCTCTCTCCT TTTGTCTCTG TCTCTTCCTC
1651  TCTCTCTGTC TCTCTCTCTC TCTCTCTCTC AATCTCTATC TTCTCTCCTG
1701  CCACCCCTCA CTCCTGCTCC TTGTCTCACT ACTCACAGCC TTTCAAGAAG
1751  GACCTGCAGC CCAGAGTCCA GCAGGCCAGG AGCCTAGGAG AGCGATGAGG
1801  CTGATGCAGG CACTGGCAGA GTCAGCCCTG CTCTCTGACC CAGCTTGAGC
1851  TCATTCTCAC AGTGCAACCT CCCCCAGGTA CCTTCCAGAG CCCCCAGCTC
1901  TGGGCCTCTG CCCAGCAGGCT CCTCCCAGCT GGCCCAGCTG GAGCATAAAA
1951  TCCCCTGTCA GCACATGCCA GGCGCGTTCC TCGGTGCCTC CCCAGCCTCC
2001  GTGACCCCAG GGCCTGGCTT AGGCTGGGAA GATGGGAGAA GTCAGATCAA
2051  GGTGGTCTCC CAGCTCAGCA GGGGAGCAGC CAGCTGGGCC CCCAGCTCTT
2101  CCTTGCCCTG ATACATGACC TTGGCAAGTC TCTTTCTTTC TTTCTTTCTT
2151  TTCTTGAGAT AGTCTTGCTC TGTTGCTCAG GCTGGAGTGC AGTGGCATCT
2201  CGGCTCACTG CAACTTCCAC CTCCCATGGC TTGAACCTCC CAGGTTCAAG
2251  TAATTCTCCC ACCTCTGTCT CCCAAGTAGC TGGTGCTACA GGTATATAGC
2301  ACCATGCCTG GCTAATTTTT GTATTTTTAC TAGAGACGGG GTTTCATCAT
2351  GTTGGCCACG CTGGTCTCGA ACTCCTGACC TCAGGTGATC CATCTGCCTC
2401  AGCCTCCCAA AATGCTGGGA TTACAGACAT CAGCCACCGC ACCTGGCCTC
2451  CCTTCCTTTT TTAGTAGAGA TCAGTGCCTA AATGATGTCA GGGATCTCTG
2501  CTGGCGAGGA TGCAAGAGTG AGTGTGACAG GCTGGGAGAG TGTGGGAGAG
2551  AGGGAAGATA TGCATGTGTG TACGTGGGTG TGAGAGTGGG GAAGGTTAGA
2601  GTGAACTGCG ATCTGTAATA AGCATGTGGA GAGCGTGTGT GTGACAGTGT
2651  CTTACGTGGG AGTGCACAGG GTGTGGGCGG GAGTAAAAGG CAGAGTCCAA
2701  TTCCACCGGC CCCCAGTGTG GGTGCAGTGT GAGCCCAAAG TGGGCGCCCT
2751  TTGGCAAGGA CTGCATGAGC TTTCTTCTCC CTCTTTTTCT TGCCCTCTCT
2801  CCCATCTCTT CTTTCCTTCT CCATGTCTCT CTCTCTCCCT CCCTCTATCT
2851  ATCTTGATTT ATCTTTCTTT CTTTTGAGAT GGAATCTTGC TCTGTTGCCC
2901  AGGCTGGAGG GCAGTGGCAT GATCTTGGTT CATTGCAGCC TCAACTTCCT
2951  GGGCTCAGGT GATCCTCCTG CCTCAGCCTC CTGAATAGCT GGGACTACAG
```

FIG. 6 (cont.)

```
3001  GTGCACACCA CCACTCCAGC TAATTTTTTA AAATTTGTTT GTAGAGACAG
3051  GGTCTTTCTC TATTGCCCAG GCTGGAGTGC AGTGGTGTGA TCATGGCTCA
3101  TTGAAGCCTC AAACCTCCTA GGCTCAAGTG TTCTTTCTGC CTCAGCCTCC
3151  TGAGTAGCTG GGACTACAGG CCCGCATCAC CACTCTGGCT ATTTTTTTTT
3201  TTTTTTTTTT TTTTTTGAGA GGGAGTCTTG CTCTGTCACC CAGGCTGGAG
3251  TGCAATGGTG CCATCTTGGC TCACTGTAAC CTCCGCCTCC CAGGTCCAAG
3301  CGATTCTCCT GCCTCAGCCT CCTGAGTAGC TGGGAATACA GGCATTGACC
3351  ACCACACCCA GCTAATTTTT GTATTTTTAG TAGAGACGGG GTTTCGCCAT
3401  GTTGGCCAGG CAGGTCTCGA ACTCCTGACC TCAGGTAACC CACCTGCCTT
3451  GGCCCCCCAA AGTGCTGGGA TTACAGGTGG GAGCCGCTGC ACCCCGGCAC
3501  TTGGCTAATT TTTTTTAAAT GTTTTTGCAG AGACAGAGTC TTGCTATATT
3551  GCCCAGGCTT GTCTGGAACT CCTGGGCTCA AGCAATCCTC CCATCTCGGC
3601  CTCCCAAAGT ACTAGGATTA CAGGCATGAG CCACCGCACC TGGCCCTTGA
3651  TTTATCTTTC TTTTTTTTCT TTTTCTCTTT TTTTCTTTTT TTGAGATGGA
3701  GTTTCACTCT TGTTGCCCAG ACTGGAGTGT AATAGTGTGA TCTCGGCTCA
3751  CTGCAACCTC TGCCTCCCGG GTTCAGGCGA TTCTCCTGCC TCAGCCTCCC
3801  TAGTAGCTGG GATTACAGGC ATGCGCCACC ACGCCTGGCT AATTTTTTGT
3851  ATTTTTAGTA AAGACGGGGT TTCTCCATGT TGATCAGGCT GGTCTCGAAC
3901  TCCTGACCTC AGGTGATCAG CCTGACTCGG CCTCCCAAAG TGCTGGGATT
3951  GCAGGCGTGA GTCATTGTGC CCAGCCGATT TATCTTTCTA TCTTTCTCCA
4001  TCTGTTTGAG ACTCTCTCGC TCTCTATATT AAGTTGTTAA ATCTCAGTCA
4051  ATCTTTATTT CACTGTGCCT CTCCATCTCT ATATGTCTCT GTTATTCTGT
4101  TTCTCTGTCT CTGTTCTCAC CTCTGTCGCT CCCCTCACCC CACAGTCTGT
4151  CTCACACACA CCAGGAGCTC CATAAATATT TGTTCTCAGC CACACTCTGA
4201  CCACGCCTCT TTCTCTTATG TGTCTCTCCA TCTCCGAGTG GCTCTGCTCA
4251  TCACATCCCT GCATTTTATA ACCATATGCT GGTGGGCCTG CCCTCCCCGC
4301  GTGCACATAC ACTTGCCTGG GATAAGCTTC TTCTGCCTGC TTATCTCCTG
4351  CGGGAATTGG AAATGCTAGT TTTCTCCCTA CCTCCCCAAG ACCCCCGCCA
4401  ATATCGTTCC CAGGAACAAG ATGAGGCATC TGGCCTCAGC CCCCAGCTTC
4451  ATCCTCGATG CTGGACTTCC ATCTTCCCTC ACATGCTTGA CTCCTTGCCC
```

FIG. 6 (cont.)

```
4501  TCCTCCCACC TCCCCTCTCC CAACTGCTCT CTACACCCCC TGGGAAATGG
4551  GCTGGATGCC GAGCTGGGGG AGTGGCTCTG TCCTGGGGGC CCTCGCCAGA
4601  TGGTGTCCCT AGGTGCCAGA GCGTGGAGCT GTCCCTTGCT GGGGCCTTTA
4651  ATAAGCACAA ACCTTCCACC CTCCACCTTG GCTGTTTTCC TTCTCTGCAT
4701  GCTCCTGGGA CCTTGGGCTC TCCATCTTTC CATGTCCGTA GCCCCAGAGA
4751  GCCAGGAAGG GGAAGCGGCG TCAAGTGCCT GGAAAAACAG CCCCATGACT
4801  TGAGTTCCTC CCTAAGACTC AGGAGTTCCA GCCCCATGTC CATCCTATTT
4851  CAAAATCCAG GCACTAGATA AGCCACACAG AAGCCGGGAG TGTAGGCCCC
4901  CAGATCCCTC CCCTCTCAGA CCCTGGGGTC TCAGTCGGTT CTCTCCAAGG
4951  ACTCGGGAAT TTGGGCCTCT GATCCTCCTG GCCACACTAC CCACCCCCGC
5001  ACCTCCCCAT ACACACACAC ACACACACAC ACACACACAC ACACACACAC
5051  ACACACACAC ATACACACAG GACTTAGGAC AGATGTTCAC GGTCTGATTT
5101  CCAAATCCTC CTGGGCCTGT GTGGGGGTGG GGAGAGATTG GCAGATAGAT
5151  CCACCGACTC TTAAGACTTA AGACCAGATA TTCTGACCCC TGTCACCCTC
5201  TTCCAAGTGC ACCATGCACT TGAGTGCACC TTGAGTCTCC AGCCTCTCAA
5251  GGAACCGGGA GATCAGGCCA TCAGCGTCTC AGCCAGCAAA GGCCTGAACC
5301  ACCAGTCCCT TATAACCCTG TAAGTCCAAC CCCCACTCCC AACCCCACTC
5351  CCCCATTTAG GGACACGGAG TCTGAGCCTA AGAACAGTGG AGAATCTGAA
5401  TGTGGACCCT CCAGTTGTTA CAGGTGCAGG AATGTCAGAT CAGGGTCCCA
5451  GCCCCCCAGC CCTCCTTCAG GCTGCTCGGG GTCGCTCCCA CCTGCTCGGC
5501  CAGCTGCGCA GCGTGGGAAC GCCCCAGCTG GGCTGCATGG AGCCGTCAGG
5551  ACAAGCTGCG CGGTTCCCAG CCTCCCTGCC TGCCCCGGCC CGGCACCGCC
5601  GCCTCCCAGC CGTCGCCGGG CAACCAGGCC GAGGGGCCCG GCCGGCCGAG
5651  TGGGGAGAGG GGTTGGGCTG GGACTGCGGG GTCCTGGGAA AGGAGGGGCC
5701  GAGGGCCTGG ATTCCTGGGT CTTAGGACGT GCTGTAGTTT GCAGCAATAA
5751  CAAGGGAACA GAGGGATATT TTGAGGAGGG GTTTTGAGGC TGGGGGAGTC
5801  GAGGTAGGGG TCCCAACTGT CCCCCAGGTA TCGGTGTGCC CTCTTCCCGA
5851  CACGCAGGCC CGGGGGAGCC CCGGACCCCG CATCCCCCAG GGCGCGGAAA
5901  CTGGCGAGGC CCCAGGAGCT CCCATTTATA GCTCAGTTTC CACTGAGCGC
5951  AGTCCCTCTA GGACCTGGGC TGAGCAAGTT TCTTCCACTC TCTCCCTTCC
```

FIG. 6 (cont.)

6001 CTCCTCCTCA CCCCTTGCCT GCCCCTCAAC CCCGGCAGGG CGCAGGTGTC
6051 CAACCCAGCC GGGACCCCCT CCCTCCTCGA ACCCAGGTGT TCCGGCTCCC
6101 AGACCCCAAT TGACCTGCCG GCGCCCACCC GCCCGGCGAT CCCCCCCTGC
6151 GTCCCCCATT CATCCGCGTC TCAGCCGCGG GAGTTTCTCA ACGGGAAGAG
6201 GGCGGAGCTC CCGGGGGGCG GACCCGGGCG GGGCGAGCGG GATCGGGCCC
6251 TCTTGGGGTC TCCCAGAGAC CCAGGCCGCG GAACTGGCAG GCGTTTCAGA
6301 GCGTCAGAGG CTGCGGATGA GCAGACTTGG AGGACTCCAG GCCAGAGACT
6351 AGGCTGGGCG AAGAGTCGAG CGTGAAGGGG GCTCCGGGCC AGGGTGACAG
6401 GAGGCGTGCT TGAGAGGAAG AAGTTGACGG GAAGGCCAGT GCGACGGCAA
6451 ATCTCGTGAA CCTTGGGGGA CGAATGCTCA GGATGCGGGT CCCCGCCCTC
6501 CTCGTCCTCC TCTTCTGCTT CAGAGGGAGA GCAGGTACCG CACGAGGGGA
6551 GCGGAGGAAT ATGGGGTGGG GGTGGGGAGT TGCTTGCGGG CTGCCTCTTC
6601 ACTAGCGAGA AGGGAGCTGG GGGCTGGGAC TCCTGGGTCC TGAATGAGGA
6651 GGCCCCTGAA GGTGCTAAGC TCAGCCCTGC TGCCCCGAAC TCTCCTAGGC
6701 CCGTCGCCCC ATTTCCTGCA ACAGCCAGAG GACCTGGTGG TGCTGCTGGG
6751 GGAGGAAGCC CGGCTGCCGT GTGCTCTGGG GGCCTAGTGG GGGCTAGTTC
6801 AGTGGACTAA GAGTGGGCTG GCCCTAGGGG GCCAAAGGGA CCTACCAGGT
6851 AAGAGTGTTC TCTCCACGCT GGGACGGGCT GGCTAGGGGG AGAGTTGCTG
6901 GGCTCGGCTG TACCTGCAGT TTCTATTTTG ACATTTTCAA GTTTGGGAAA
6951 TTGATGGGCT CGGGTAAAGA TTTAGGAGTC CTGATTTTTG AGCTGCTTCT
7001 TTGGGGGTGA CCCACGGAGT TTGGGAATTA TTATGTTATT GCAAAATAGT
7051 ACATAGGCCA GGTGCAGTGG CTCACGCCTG TAATCCCAAC GCTTTGGGAG
7101 GTTGAGGCCA GAGGATCGCT TGAAACCAGG AGTTTGAGAC CAGCCTGGGC
7151 AACATAACAA GACCTTATCT CTACACAAAT GTATATATAT ATTTTAAACA
7201 AATTAGCCGG GTATGGTGGT GTGCATCTAT AGTCCCAGTT ACTCAGGAGG
7251 CTTAGGTGGT AGGATTGCTT GAGCCTAGGA GTTCAAGGCT GCAGTGAGCC
7301 ATGATCAAGC CACTGCACTT CAGGCAATGG TGAGACCCTG TCTCAAAAAA
7351 AAAAAAAAAA GAGAACATAA ATGCAAAAAA GTACAGTAAA TATAAATGGA

Exon 1

Exon 2

FIG. 6(cont.)

7401 AGATTTACCA AATAAAATAG ACACACACAG CCAATACCCA AGTCCATTGC
7451 TAGCTCCCCA GAAGACCCCG TGTTCCTTTC CCCTATCATA GCCCCCTCCC
7501 CCTCACTCCA GAAGTAGTAT CTAACCTAAT TTTTATGGCA ATCATTTTCT
7551 TGCTTTCGTT CCTGACTTTA TTACCCCTAA GTTTGCAGTG ACTCTGGGTT
7601 GGGAGGGAGT TAGAGTCTCT CTGGGCCCAG TACACACTTT TTAATAGTGT
7651 CTTACCACCA AATGTGTGGG CCAGTTTTCT GGTGGAGGAT GTCTGGGGAT
7701 GGAGGCCTGA GGCCAGGATT TCAGAACCAT GGTGTGCTGA CTGCCTTCTC
7751 CCTGACTCCA GGGTGGTCCC GGTACTGGAT ATCAGGGAAT GCAGCCAATG
7801 GCCAGCATGA CCTCCACATT AGGCCCGTGG AGCTAGAGGA TGAAGCATCA
7851 TATGAATGTC AGGCTACACA AGCAGGCCTC CGCTCCAGAC CAGCCCAACT
7901 GCACGTGCTG GGTAAGGACC TCGCCCACTT GTCCCCTGGG AGCCCAAGAG
7951 GGCAGCCCGT ACTAGCTGTG AGTAGCAGAG CCCAGGGAGC CCAGGGGCAT
8001 GGTCAATTGG AGCTGAGAAG ATCAGGATCC ATCTCTGACC CCAAATCCAC
8051 CTTGCAGTCC CCCCAGAAGC CCCCCAGGTG CTGGGCGGCC GCTCTGTGTC
8101 TCTGGTTGCT GGAGTTCCTG CGAACCTGAC ATGTCGGAGC CGTGGGGATG
8151 CCCGCCCTAC CCCTGAATTG CTGTGGTTCC GAGATGGGGT CCTGTTGGAT
A
8201 GGAGCCACCT TCCATCAGGT CAGGTCCAAA TTCCTGTGCT AGCCTTTGCC
8251 CATTGAGGGA AACTTGGGTT ACACTCTGAC CACAGGCTCA TCCAGAAGAG
8301 AAGAAGACAT GGGAGGGCAG AGGTTCATGG GTTTGGACTC TTGAAATATG
8351 ATGCAGGGTA AAGATTCTAG GGCCAGACTA CCTGGGTTCA AATTATGTCT
8401 CAGCCACTTG CTAGTTGATT GATCTTGAGT AAGTTAGTTA ACCTCTCTGT
8451 GCCTCAGTTG CCTTATCTAT ACAATCAGGA TAATAGTAGC ATGCATGTCA
8501 TAGGGTATTG TGAGAATTAA ATAAATAAAT ACCTATAAAT GCCCAGAAGA
8551 GTGACCAATA CATAGTGAGC ACTATATAAG TAAGGCAAGC TTGTCCAACC
8601 TGCGGCCCAT GGGCTGCATG CAGCCCAGGA TGGCTTTGAA TGTGGCCCAC
8651 CACAAATTCA TAAACTTTCT TAAAACATTA TGAGACTTTT TTGTAATTTT
8701 TTAGCTCATC AGCTATCATT AGTGTTAGTA TGTGTGGCCT AAGACAATTC
8751 TTCTTCCAAT GTGGCCCAGG AAAGCCAAAA GATTGGACAC CCCTGATGGG
8801 TAGATGGCAT TATTATTCTT ATCGTTCCCT CCAGACCCTG CTGAAGGAAG

Exon 3

Exon 4

Exon 5

FIG. 6 (cont.)

8851 GGACCCCTGG GTCAGTGGAG AGCACCTTAA CCCTGACCCC TTTCAGCCAT
8901 GATGNNTGGAG CCACCTTTGT CTGCCGGGCC CGGAGCCAGG CCCTGCCCAC
8951 AGGAAGAGAC ACAGCTATCA CACTGAGCCT GCAGTGTGAG TGCAGCTGGC
9001 CCTGGGAAAG AGGGGTGTGG GGCCCTGACT CCTGGCTATG AGGAAGGAGG
9051 GGACTGTGGC CCTTGGGGAA TGAGGAAACT GGAGCCTGGA CTCGTGGATC
9101 TAAGATAGCA GGAGAGGGCT GGGTATGGTA GCTCACGCCT GTAGTCACAG
9151 AACTTTGGGA GGTCGAGGCA GGCGGATCAT CTAAGATCAG GAGTTCGAGA
9201 CCAGTCTGGC TAACATGTCG AAACCCCGTC TCTACTAAAA ATACAAAAAT
9251 TTGCCGGGCG TGGTAGCACA CACTTGTAAT TCCAGCTACC TGGGAGGCTG
9301 AGGCAGGAGA ATCACTTGTA CCCGGGAGGC AGATGTTGCG GTGAGCCGAG
9351 ATCATGCCAC TCAGCAGCAG AGTGAGACTC CGAGCAGGAG AGGACAGACA
9401 GCTGGGGTCC CTGGGGAAAG AGAAAGCTGG GCCTTGACTC TCACATGGGG
9451 GAGACTAGGA GAGGGCAGAA GGCTGGCACA TTGAGGTAAC TGGGGAAATT
9501 GGGAACTGAA AGCCCAGACT CCTGGCTCAA AGGGAGAAGG GGATTAGGGG
9551 CCCAGACTCC TGGGATGGAG GAACCAGGGA CTGGACACCT AGGCCAGTGA
9601 CGGAGGTGTT CCTGGTCGTT GCCCATCTGA CCATTGTCCC ACCCTCACAG
9651 ACCCCCCAGA CGTGACTCTG TCTGCTTCGC CACACACTGT GCAGGAGGGA Exon 6
9701 GAGAAGGTCA TTTTCCTGTG CCAGGCCACA GCCCAGCCTC CTGTCACAGG
9751 CTACAGGTGA GGACGAAGAC CCACCTCTCC CCAGCCCCAA GAGTGAGCTT
9801 GGGAAGGGCT GGGACCTGAG TAGCTGTGCC AGAGAGGCCA GGACAACGTT
9851 AACAGCGCCA CCATTTCCTC AGGTGGGCAA AAGGGGGCTC TCCGGTGCTC
9901 GGGGCCCGCG GGCCAAGGTT AGAGGTCGTG GCAGACGCCT CGTTCCTGAC Exon 7
9951 TGAGCCCGTG TCCTGCGAGG TCAGCAACGC CGTGGGTAGC GCCAACCGCA
10001 GTACTGCGCT GGATGTGCTG TGTGAGCTGG GGCCGGCCTG TGGGTGTGGT
10051 CAAAGGTGGC CGTGGCTTTC AGGGCTGTTG AGGGTCGGGG CCTGGAGGGG
10101 CGGGGCCGGG AGAGCGAGCG TGGGGTATTA GGAGGAGGAG AGTGTGGAGC
10151 TGGGGCATAT TCTTGCGCCC TAGAGGGTGT GGTGTTTCTG TGGGGCTGGC
10201 TGATCCCAGG TCAGTGGCTG CATTCCGCCC CGGCCATGTG ACCCGTAGTC
10251 TCTTTCGTCC AGTTGGGCCG ATTCTGCAGG CAAAGCCGGA GCCCGTGTCC Exon 8
10301 GTGGACGTGG GGGAAGACGC TTCCTTCAGC TGCGCCTGGC GCGGGAACCC

FIG. 6 (cont.)

10351 GCTTCCACGG GTAACCTGGA CCCGCCGCGG TGGCGCGCAG GTACAGCCCT
10401 AAATCTGAGG CGGTGGCTGG AGGGGGACCA GGCTTCCTTA CAAATCCGGC
10451 TTCTGACGCC CCTTCCCTGT CGCAGGTGCT GGGCTCTGGA GCCACACTGC
10501 GTCTTCCGTC GGTGGGGCCC GAGGACGCAG GCGACTATGT GTGCAGAGCT Exon 9
10551 GAGGCTGGGC TATCGGGCCT GCGGGGCGGC GCCGCGGAGG CTCGGCTGAC
10601 TGTGAACGGT GAGAAGGCGG GGCTTCCTAG GGGACCTGGC CCGTCCTGGG
10651 ATAGGGAGCG GACAGAGGGG GCAAGGGCTA ATGCAGTGGG AGTGGCCTGG
10701 AAGGAGCTTT ACACCCAGCG GGGGCTGGAG ACCGGACCTA TTGAAGGCGA
10751 GGCTTTTAGG AGAATCGGAG TTTGGAGGCG GCGTGGCCTG ATTGATTGAG
10801 GTTAGCGGAG AGTGCGGTGG ACAGACCCGG CTTTGTTACA GCCTTTGGGG
10851 AGGGCAAGAC CTCTCCTCTG AGTGACCTAC AGTCTCCATC CCAGCTCCCC
10901 CAGTAGTGAC CGCCCTGCAC TCTGCGCGTG CCTTCCTGAG GGGCCCTGCT Exon 10
10951 CGCCTCCAGT GTCTGGTTTT CGCCTCTCCC GCCCCAGATG CGGTGGTAAG
11001 GAAATGTCAC TCCTCCCGTC ACCCATCCAG CCGTGATCCC TGACCTCCCA
11051 CCTGGCCCCC CGAAACTACT GTGACGATTT CTGACTTCCC AGACATCCCT
11101 CCTGCTTCTT CCTCCCCTCC TCAGTCTCCT CCGTGTCCTC CCTCTTTTGT
11151 GCCCCCAGGT CTGGTCTTGG GATGAGGGCT TCCTGGAGGC GGGGTCGCAG
11201 GGCCGGTTCC TGGTGGAGAC ATTCCCTGCC CCAGAGAGCC GCGGGGGACT Exon 11
11251 GGGTCCGGGC CTGATCTCTG TGCTACACAT TTCGGGGACC CAGGAGTCTG
11301 ACTTTAGCAG GAGCTTTAAC TGCAGTGCCC GGAACCGGCT GGGCGAGGGA
11351 GGTGCCCAGG CCAGCCTGGG CCGTAGAGGT CAGACCCCAG CCCGAAGACC
11401 GCAAATCTGG AGAGTCTAAA CCCCACAAAC GCAGGGATCC CCCAGCCGAG
11451 GGCTGCAAAA CCTCATACCC TCAAATGCAG AGGAGACCTC CAAACCTCGG
11501 GAGTCTCAAA AGTGTGGGCT CATTGATTCC CAAGACACCC CTCAACCACA
11551 AATGGCTTCA CATTCTGAAT CGTAAACTGA GAGACTCCTC ACACCTAGGG
11601 GCCCCAAAAA GGGAAACTCC AATGATTGCA AAGCAAATTG CAAAGTAAAG
11651 GACCCCTCAA ATTCTAAGAC TCCCTAAAGC CAGGGAGTTT AAACTCACTC
11701 TCAAACTTGG GGAACCCCAA ATTCAAGGGC CTTTGAATCT TCAAATGTGC
11751 GACCTTTTGA ACCCAGGAAT CCCAAACTCA ATCCCTGAGC CGCCGGTTCC
11801 TGGTTCCCCC TCAGCGTTCT CAGGATGTCC CCTCTGCTCC CTGCAGACTT

FIG. 6 (cont.)

11851 GCTGCCCACT GTGCGGATAG TGGCCGGAGT GGCCGCTGCC ACCACAACTC Exon 12
11901 TCCTTATGGT CATCACTGGG GTGGCCCTCT GCTGCTGGCG CGACAGCAAG
11951 GGTTAGTGCC TGAGCCCCGC CCCGGCTCCC GAGGCCCCAG CCCCACAGGC
12001 GCCCTGCCTG CCCAGTGACC TGACCTGGCC TTGGGCCTTG GCTCCAGTCC
12051 CATTTCCAGC TCTGCAGAGG GCTTAGCTCT CCTTCACGTT CTGGTTCGCT
12101 CCTTAAGCCC TAACTAGGCC TTCCCAGGGT CACAGTCCTC GGTGGGAATG
12151 ATTCTTATTG GTTTCCAACA GCCCTAGCCA ATGAGCCTCA TTGGTTCGCA
12201 GTCCTCTCTC TTCCCGCTTA TTGGTCTGCA CACATTGTGA CCCCGCCCAT

12251 CGCTTRACTC CACCGGTCGC TGTTTGTCAG CCTCAGCCTC TTTCTCCGAG
12301 CAAAAGAACC TGATGCGAAT CCCTGGCAGC AGCGACGGCT CCAGTTCACG Exon 13
12351 AGGTCCTGAA GAAGAGGAGA CAGGCAGCCG CGAGGACCGG GTAGGATGCC
12401 AGGGTCCCCA GACCTGACTG TGCCTCCAGA CCTAAATAAT AGCCCAGTCC
12451 CAAGAGGGTC CCCAAATTCA AATAGGACTC TAAGGCCAGG CATGGTGCCT
12501 GACGTTGGTA ATACCACTTT GGGAGGTGGA GACACAAGGA TCACTTAAGG
12551 CCAGGAATTC AAAGCCAGCC TGGACAGCAT AGCAGGACCC CATCTCTACA
12601 AAAATACAAA CTAAAATAAA ATAAAAAATG AACCGGGTAT GGTGGCATAC
12651 ACCTATAGTC CCAGCTACTC AGGACACTGA GGTGGGAGGA TCCCTTGAGC
12701 ACAGGAGGTA AAGGCTGCAG TGAGCTATGA TTGCACCATG CACTCCAGCC
12751 TGGGCTACAG AGCAAGACCC TGTCTCGATT TTTTTTTTTT TTTTTTATGT
12801 AGGAGGGGTC TAGTCTTTTT TTTTTTGGCA GAATTTCACT CTGTCACCCA
12851 GGCTGGAGTA CAGTGGTGCG ATCTCGGGTC ACTGCAACCT CTGCCTCCCT
12901 GGTTCAAGTG ATTCTCTTGC CTCAGCCTCC TGAGTAGCTG GGATTACAGG
12951 CGCCCACCAC CACGCCTGAC TGATTTTGTA TTTTTAGTAG AGATTGGGTT
13001 TCACCATGTT GGCCAGGCTG GTCTCAAACT CCTGACCTCA GGTGATCCAC
13051 CGGCCTCGAC CTCCCAAAGT GCTAGGATTA CAGGCATGAG CCTCCACGCC
13101 CGGCCTGAGG GCTCAAGTCT TTTTTTTTCT TTCTTTCTTT TTTTTGAGAC
13151 GGAGTCTTGC TCTGTAGCCC AGGCTGGAGT GCAGTGGCGC GAACTCGACT
13201 CACTGCAAGC TCCACCTCCC GGGTTCACAC CATTCTCCTG CCTCAGCCTC
13251 CAGAGTAGCT GGGACTACAG GCACCCGCCA CCATGTCCAG CTAATTTTTT

FIG. 6(cont.)

13301 TGTATTTTTA GTAGAGACGA GGTGTATAGC GTGTTAGCCA GGATGGTCTG
13351 GATCTCCTGA CCTCGTGATC CGCTCGTCTC GGCCTCCCAA AGTGCTGGGA
13401 TTACAGGCGT GAGCCACCGC GCCCGGCCAA GGGCTCTAGT CTTAACAGTG
13451 ACCCCACAGC GAAATGTCAC CCAAGTCCAT GCCCCTGACC CAATTATTCC
13501 CTAGGCCCAG TATGTCCCCA CAGCCCGTTT TTGTTGTTGT TGTTGTTGTT
13551 GTTGTTGTTT TTGAGATAGA GTCTTGCTCT GTCGTCCAAG CTGGAATGCA
13601 GTGGTGCAAT CCAGACTCAC TGCACCCTCC ACCTCCCAGT TCAAGTGATT
13651 CTCGTTCCTT AGCCTCCTGA GTAGCTGAAA TTACAGGTGC CTGCCACCAT
13701 GCCTGCCTAT TTTTTGGATT TTTAGTAGAG ACAGAGTTTC GGCATGTTAG
13751 CCAGGCTGGT CTCAAACTTC TGGCCTCAAG TGATACTCCT GCTGCGGCCT
13801 CCCAAAGTGC TGGGATTACA TGCATGAGCC ACTGTGCTGG CTTCTTACAG
13851 CCCTTTTATT GTCCTGAGTG CAGTCCCCAG CTCTTGGGTG CTCTTACTCC
13901 CTCCTGCCTG GCCTCCACTG GCTGGCTGAA GGTCCTTGGG GTCTGGCATT
13951 GGGGCGGGGG GATCCTCTGA CTATTCCCTC TCACTAAGTT CCCTACCCCA
14001 GGGCCCCATT GTGCACACTG ACCACAGTGA TCTGGTTCTG GAGGAGAAAG
14051 GGACTCTGGA GACCAAGGTG AGTGTTGAGA GGGGTGGGGC TCCCTTCACT
14101 GTTGGGAGAG GCGGGGCTCC CTTCATTGTG TTTCCGTCTC TCTCCCACGC
14151 CTGTCCCCTC CTTTTTCCTT CTGTTGTCCT CAGAGTTGGG ACTCAGCTCC
14201 CCACCCGACT GCTCGTGCGC CGTGGGCCAT CTCACTCAGC TCCCAGJCTC
14251 AGTTTGCCTG TCTGCAGACT CTTCCCACAC ATCTGTCCCA GCCCTAGCCT
14301 CCATCTGGAG CCCCAGACCA GGGCTCACCG TGCCTGTGCT CTCCTCATCA
14351 CGGTCAAGCC CCCTTTCAGC CACCAGGTCC TACAGTGGGC CCACATCTGC
14401 CCAGACTGGT TCTTCCTCTG GGGTCCTACC TCAGGACAGC CACATTGACT
14451 CCAGGCCATC CCCAGGCCAG AGCACTTCTC TCTCTCTCTC TCTCCTGCGT
14501 ACCTAGCACA TGCCATTCTC TCTCTTCTTT TTTTTTTTTT TTTTTTGAG
14551 ACGGAGTCTC ATTCTGTTGC CCAGGCTGGA GTGCAGTGGT GCAATCTCAG
14601 CTCACTGCAA CCTCTGCCTC CTGGGTTCAA GCCATTCTCC TGCCTCAGGC
14651 TCCCTAATAG CTGGCTAATT TTTCTTGTAT TTTTAGTAGA GATGGAGTTT
14701 CACCATGTTG GCCAGGCTGA TCTGGAACTC CTGACCTCAA GTGATCCGCT
14751 CGCCCCAGCC TCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGTGCCC

Exon 14

FIG. 6(cont.)

```
14901  AGCCGACATG CCATTCTCTT GGCCTGAAAC ACTCCTACCT TCCTTCCCAT
14951  GTCTACCTAA TTCCTTCCTT TAGTCCTCCA GTCTCAGCTC AGACATTTCT
15001  TGTTCTAGGA AGCCCATGCT TCCGTCATGA CAGCTCGATC ATTTTGCCTG
15051  TGTTCCACCC ATCACAGCCA TGACCACTCT GATCTGGGCT TCCTTATCGC
15101  ACCCACTATC CTGAGGGCTC TACCATCACA GCCCCTGTCA TTGCCTATGC
15151  CTTTCCCAGG CACAGCCCTG ACCCCTCTGG GTACTGTCTC ATGATCTGTC
15201  ATTTTTCCTT TGGTGTGGGA TTCTGTGAGG ACAGGGTCCA GTTCTATCCT
15251  AGTGACATGC CTTGTAGCAG CAACACAGGG TGTGACACTG AATCAAAGCC
15301  TAGAGGCTGT TGGGCAGGTG AGTGTCTCTC TCCTGTTCCC TCTGCACCTT
15351  CCACACCGAC ACCCCTCAGC AGGCCTATAT CCCTCCGTCT CTACCTTTCT
15401  CTGCCTATGT CCTATCCATT TGCCTCTTAT CACTGTTCCT CTGTCTCACT
15451  TTCTCTCTCT CCCAGTCCAT GTGTGTCTGT GTGTCTCTGG CCACTGCTGT
15501  CTCTTTTTGT CTCTCTCAAG GTCTGGTCTA TTTCAGTGTG TCTCTCCATC
15551  AGTGACCGTC ATCDCCCCTG CACGGTCACA GACTTTACTG AGTCCCATTT
15601  GTCCCCTCAG GACCCAACCA ACGGTTACTA CAAGGTCCGA GGAGTCAGTG
15651  TGAGCCTGAG CCTTGGCGAA GCCCCTGGAG GAGGTCTCTT CCTGCCACCA
15701  CCCTCCCCCC TTGGGCCCCC AGGGACCCCT ACCTTCTATG ACTTCAACGC
15751  ACACCTGGGC ATGGTCCCCC CCTGCAGACT TTACAGAGCC AGGGCAGGGT
15801  ATCTCACCAC ACCCCACCCT CGAGCTTTCA CCAGCTACAT CAAACCCAGA
15851  TCCTTTGGGC CCCCAGATCT GGCCCCCGGG AGTCCCCCCT TCCCATATGC
15901  TGCCTTCCCC ACACCTAGCC ACCCGCGTCT CCAGACTCAC GTGTGACATC
15951  TTTCCAATGG AAGAGTCCTG GGATCTCCAA CTTGCCATAA TGGATTGTTC
16001  TGATTTCTGA GCAGCCAGGA CAAGTTGGCG ACCTTACTCC TCCAAAACTG
16051  AACACAAGGG GAGGGAAAGA TCATTACATT TGTCAGGAGC ATTTGTATAC
16101  AGTCAGCTCA GCCAAAGGAG ATGCCCCAAG TGGGAGCAAC ATGGCCACCC
16151  AATATGCCCA CCTATTCCCC GGTGTAAAAG AGATTCAAGA TGGCAGGTAG
16201  GCCCTTTGAG GAGAGATGGG GACAGGGCAG TGGGTGTTGG GAGTTTGGGG
16251  CGGGGATGGA AGTTGTTTCT AGCCACTGAA AGAAGATATT TCAAGATGAC
16301  CATCTGCATT GAGAGGAAAG GTAGCATAGG ATAGATGAAG ATGAAGAGCA
```

Exon 15

FIG. 6 (cont.)

```
16251  TACCAGGCCC CACCCTGGCT CTCCCTGAGG GGAACTTTGC TCGGCCAATG
16301  GAAATGCAGC CAAGATGGCC ATATACTCCC TAGCAJCCCA AGATGGCCAC
16351  CATCTTGATT TTACTTTCCT TAAAGACTCA GAAAGACTTG GACCCAAGGA
16401  GTGGGGATAC AGTGAGAATT ACCACTGTTG GGGCAAAATA TTGGGATAAA
16451  AATATTTATG TTTAATAATA AAAAAAAGTC AAAGAGGCAA GTGTGTCTTA
16501  GGGAGTCTAC TGGCATTATC ACTCTCCACC AAGGAAGGGG TCCCTTAGAC
16551  CTGTCCCAAG GTCCGTCCTC TACCCTAGCC TATGAGGTGG CTGTAGGAGT
16601  AAAACTGTGA GCCACCTCTC AGCCTCTTGC TACCTGCAAA GCACTCTAGG
16651  CTCTTTTTTT TTTTTCTTG AGACAAGATC TGGCTCTATG GCCCACATTG
16701  GAGTGCAGTG GCATGATCTC AGCCCACTGC TACCTCTGCA TCCTGGGCTC
16751  AAGCCATCCT TCCACCTCAG CCTCCCAAGT AGCTGGGACT ACAGGTGCAT
16801  GCCACCACAC CCAGCTAATT TTTGTATTTG TTTGTAGACA GGGTTTCACC
16851  ATGTTGGCCA GGCTGGTCTC AAACTCCTGA CCTCAAGTGA TCCGCCCACC
16901  TAGGCCTCCC AATGTGCTGG GATTACAGGC ATGAGCCACT GTGCCCAGCC
16951  ATGGGCTCTT TTAATATACA TCTTCACACA CACACACACA CACACACACA
17001  CGCACACACA CACATGAGTT GCAAACAGAA AAGACACACA CATAGGCATG
17051  TATGCACAGA CACACGCATA GATGTCCACA CAGTTGCACA CAAGTGACAG
17101  GGCTGCCCCA GGGGTCCTGG GGAAGACTGA ATTCTAACTC TCATTAGAGG
17151  AGACAAACAA GTGAGCCCTG AAGTGGAGCA GGGAAGGGGA GACTATGGGT
17201  AGGAAAATGG CAATCCCCTG GTCCTTACAG CAAGCGTGGA GATCCAGACC
17251  CTAATCCTGA GGTGCTGCAT CCACAGTGGG CATGGTGCTG GTGCCTGCTT
17301  GGATGATCCT TAAAGAAAGG TCCTGGGGGC TTTGGTTCAT GGATCCTTGA
17351  GCTAGGAGTT AAAGGTCCAG GCCCCTGGGA CCCTTGGGAA GCAGAGCAAG
17401  AAGAGTGAAC TCCTGGGTCT GAAGGAGAAT GGGCTGGGGG CTTGGTCTCT
17451  GGTCCTGAGA GAGAAGGTGC CCAGACTTCT GGATCTGAAA GAGGAAGGGA
17501  CTAGGTCTCA ACTGCTGCCT TCTTGACTGG GGACATTTTG GAGGCCTGTA
17551  TTCCTGAGCC CTCAACAGAG GAATGTACTA GGGGATGGGG GTCTCTGATG
17601  CTTGCATCCT TGGAAAAGGA CAAAACTGTG AGTGTCTGGG TCTAAAGAGG
17651  GTGAGAGTCC TGCGGGAGGA CTCAAAATCC ACAACGGGCG GAGCCCATAG
17701  CCGGACTCCT GGCTGGGCCC TTCATGGGGC GGGACGCCTG GAATCTCGAG
```

FIG. 6 (cont.)

```
17751  GGGCGGGGGC CTGGCGCAGG CTCCCGCCCG GGGTTCCCGA GCTGCTCCAC
17801  TCTGCGCGAA GCCGCCACGC TATTGTCCTG ACCAGGAAGG CGGGGCCGGC
17851  GCGGGGCGGG GCTGCCGGGG CCGCCGCAGC CCGGGGGCGG CGGGAGGAGG
17901  AGGTGGCGGC GGTGGCGCTG GGAGCTCCTG TCACCGCTGG GGCCGGGCCG
17951  GGCGGGAGTG CAGGGGACGT GAGGGCGCAA GGGCCGGGAC ATGGGGCCCG
18001  CCAGCCCCGC TGCTCGCGGT CTAAGTCGCC GCCCGGGCCA GCCGCCGCTG
18051  CCGCTGCTGC TGCCACTATT GCTGCTGCTT CTGCGCGCGC AGCCCGCCAT
18101  CGGGAGCCTG GCCGGTGGGA GCCCCGGCGC GGCCGAGGTG AGGCCGGGCC
18151  GGGTCCTGGG GGATGGGGGA AGGGGCGGGA CCGGGTCTCT GGACGCCGGC
18201  GCGGACATGT CCAGGGCAGA AAGCGCGGTC TTTCCAGCCA GGTGGTCAGC
18251  CCCCAGGCGC CCCCAATCAC ATTTATGAAC GCAGGGTTCC AGGCGCCAGC
18301  TCCCCCATCA TGCGACGTCC CAGCCCCCTC CCATCTCGAG CATAGGAACT
18351  GGTCTATTCA GAGCCCCTGG TCCCAGAAGT GCAGCCCGCT CTCCAGACCC
18401  AGGTGACTCG GCCCCAACCG CCTCCCGCCT GGACATAGGA CCCACCAAGC
18451  AGCGAGGCAT TTAGATCCAA TAATCCAGAC CCCTTGTATT CTCTGGACCC
18501  ATATGGAGGC CCTTGCAGCC TCCCAGGACC CAGGAGTCCA GTCCTTGAGT
18551  CACCAGCCAC CCCAACCAGA TGTAGCTCTC CAGTCCTCAA GGACCTGGTG
18601  TCCAGGACTG TAGGCCCCTG AAGCCAGGCC TTGTCAGCTT TGCATCCTGC
18651  AACGGGAGCC TGAGCAAGGG ATGGAGGGAG GAGGGGCCAG AACTCCTGGG
18701  TTCTGGCCTC CTCCTCCGCG ATTCAGGTTT AACCCCTTCG GGCTCCAGAG
18751  CGGCTGCGCT GGGGTGGGGG CGGAGTCTGT CTCCGCGGCA ACAAGGCAGA
18801  AAGAATCCCG GGGGACCCAG GTCGCCATAG CAACGGGAGC GCTGGGGCGC
18851  CCCCGCCCTA CGGGAGCTGT TTCCCAGGGA ACGGTGCCTC CATGGAGGCG
18901  GTGTGCGGTG CTTGGGGGAG GGGGCTGGTG CTGGGGGTCT CGGTCCTAGG
18951  GAGCAAAGAA CCAGGGGACC CTCATGCCAA CGCCCCCCGA GCCCTCACTG
19001  TCCTTTCCAC TTCCATCCAG GCCCCGGGGT CGGCCCAGGT GGCTGGACTA
19051  TGCGGGCGCC TAACCCTTCA CCGGGACCTG CGCACCGGCC GCTGGGAACC
19101  AGACCCACAG CGCTCTCGAC GCTGTCTCCG GGACCCGCAG CGCGTGCTGG
19151  AGTACTGCAG ACAGGTGGGC GGGGCCGAAC GGGAGAGGCG GGGCCGCCCA
19201  TAGAAAGCTA GACTTGAAAA AGCCGTGGTC CAGGGTGCTG CGCGATCTAA
```

FIG. 6 (cont.)

```
19251  GGCGTGGNGG CTGGGGGGCG TGGCCAATAA AGAGGCGCAA CTATGCTAGG
19301  GGCAGGGGAC CTGTTTTGAG ATACTAAGTC AGGAAAAGGG GAGAGCCGCG
19351  AGATAGCCAG AGAGGAAGTG GAATTTAGGA ATCTGGTGGT CTTTGTAAAG
19401  AGTAGAGGTG TAGGGGGGAG TGGCGAAAGG ATAGGCGGGG CTAAGACAGA
19451  AAGAGACCTT AAGGACCAGC AAGATGGGGA AAGGGGTGGA GCCCAATGAG
19501  AGCGCGGAGA GCTGGGGGGG CGTGGCCATG AAAAGACAAA TTTATAACGG
19551  GAAGGGAGAG TTTTGGAGAG GCGGAATAGA GGAAAAGGCG GGGCCTAAAG
19601  GAGGGTGAGA CCTTTGGGGA GACGAATCTG ACTGCGGGGA GGGGTGACCA
19651  GAGAGGTGGG CTTAGAGGGA GGTTCAGAAA GAAACAGCAC AGGAAAAGAG
19701  ATAGGGCTTA AAGATGACGG GACTTTTAAG GGAAAACTGC TAGTGGGCGT
19751  GGCCAATGAG CACAAGGAGC TTGGATATCT AAGGCTGGTG CTAGGGAGAA
19801  GCAGGGCCTA GGGAAGCGAT GTCCTCATGA ATACTAGAGC CTTGAAAACG
19851  GACCTGGCCG GGCGCGGTGG CTCACGCCTG TAATCGCAGC ACTTGGGGAG
19901  GCCGAGGCAG GCGGATCACC TGAGGTCAGA AGTTCGAGAC CAGCCTGGCC
19951  AACACGGCGA AACTCCGTCT CTACTAAAAA TACAAAAATT AGCCTGGCAT
20001  GGTGGTGCGT GCCTGTAATC CCAGCTACTC AGGAGGCTGA GACAGGAGAA
```

FIG. 6 (cont.)

```
                GGGAACTGGCAGGCGTTTCAGAGCGTCAGAGGCTGCGGATGAGCAGACTTGGAGGACTCC
1               ---------+---------+---------+---------+---------+---------+   60

                AGGCCAGAGACTAGGCTGGGCGAAGAGTCGAGCGTGAAGGGGGCTCCGGGCCAGGGTGAC
61              ---------+---------+---------+---------+---------+---------+  120

                AGGAGGCGTGCTTGAGAGGAAGAAGTTGACGGGAAGGCCAGTGCGACGGCAAATCTCGTG
121             ---------+---------+---------+---------+---------+---------+  180

                AACCTTGGGGGACGAATGCTCAGGATGCGGGTCCCCGCCCTCCTCGTCCTCCTCTTCTGC
181             ---------+---------+---------+---------+---------+---------+  240
                                  M  L  R  M  R  V  P  A  L  L  V  L  L  F  C

                TTCAGAGGGAGAGCAGGCCCGTCGCCCCATTTCCTGCAACAGCCAGAGGACCTGGTGGTG
241             ---------+---------+---------+---------+---------+---------+  300
                F  R  G  R  A  G  P  S  P  H  F  L  Q  Q  P  E  D  L  V  V

                CTGCTGGGGGAGCAAGCCCGGCTGCCGTGTGCTCTGGGCGCCTACTGGGGGCTAGTTCAG
301             ---------+---------+---------+---------+---------+---------+  360
                L  L  G  E  E  A  R  L  P  C  A  L  G  A  Y  W  G  L  V  Q

                TGGACTAAGAGTGGGCTGGCCCTAGGGGGCCAAAGGGACCTACCAGGGTGGTCCCGGTAC
361             ---------+---------+---------+---------+---------+---------+  420
                W  T  K  S  G  L  A  L  G  G  Q  R  D  L  P  G  W  S  R  Y

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

                GAGGATGAAGCATCATATGAATGTCAGGCTACACAAGCAGGCCTCCGCTCCAGACCAGCC
481             ---------+---------+---------+---------+---------+---------+  540
                E  D  E  A  S  Y  E  C  Q  A  T  Q  A  G  L  R  S  R  P  A

                CAACTGCACGTGCTGGTCCCCCCAGAAGCCCCCCAGGTGCTGGGCGGCCCCTCTGTGTCT
541             ---------+---------+---------+---------+---------+---------+  600
                Q  L  H  V  L  V  P  P  E  A  P  Q  V  L  G  G  P  S  V  S

                CTGGTTGCTGGAGTTCCTGCGAACCTGACATGTCGGAGCCGTGGGGATGCCCGCCCTACC
601             ---------+---------+---------+---------+---------+---------+  660
                L  V  A  G  V  P  A  N  L  T  C  R  S  R  G  D  A  R  P  T

                CCTGAATTGCTGTGGTTCCGAGATGGGGTCCTGTTGGATGGAGCCACCTTCCATCAGAGC
661             ---------+---------+---------+---------+---------+---------+  720
                P  E  L  L  W  F  R  D  G  V  L  L  D  G  A  T  F  H  Q  T
```

FIG. 7 (cont.)

```
     CTGCTGAAGGAAGGGACCCCTGGGTCAGTGGAGAGCACCTTAACCCTGACCCCTTTCAGC
721  ---------+---------+---------+---------+---------+---------+  780
     L  L  K  E  G  T  P  G  S  V  E  S  T  L  T  L  T  P  F  S

     CATGATGATGGAGCCACCTTTGTCTGCCGGGCCCGGAGCCAGGCCCTGCCCACAGGAAGA
781  ---------+---------+---------+---------+---------+---------+  840
     H  D  D  G  A  T  F  V  C  R  A-R  S  Q  A  L  P  T  G  R

     GACACAGCTATCACACTGAGCCTGCAGTACCCCCCAGAGGTGACTCTGTCTGCTTCGCCA
841  ---------+---------+---------+---------+---------+---------+  900
     D  T  A  I  T  L  S  L  Q  Y  P  P  E  V  T  L  S  A  S  P

     CACACTGTGCAGGAGGGAGAGAAGGTCATTTTCCTGTGCCAGGCCACAGCCCAGCCTCCT
901  ---------+---------+---------+---------+---------+---------+  960
     H  T  V  Q  E  G  E  K  V  I  F  L  C  Q  A  T  A  Q  P  P

     GTCACAGGCTACAGGTGGCCAAAAGGGGGCTCTCCGGTGCTCGGGGCCCGCGGGCCAAGG
961  ---------+---------+---------+---------+---------+---------+ 1020
     V  T  G  Y  R  W  A  K  G  G  S  P  V  L  G  A  R  G  P  R

     TTAGAGGTCGTGGCAGACGCCTCGTTCCTGACTGAGCCCGTGTCCTGCGAGGTCAGCAAC
1021 ---------+---------+---------+---------+---------+---------+ 1080
     L  E  V  V  A  D  A  S  F  L  T  E  P  V  S  C  E  V  S  N

     GCCGTGGGTAGCGCCAACCGCAGTACTGCGCTGGATGTGCTGTTTGGGCCGATTCTGCAG
1081 ---------+---------+---------+---------+---------+---------+ 1140
     A  V  G  S  A  N  R  S  T  A  L  D  V  L  F  G  P  I  L  Q

     GCAAAGCCGGAGCCCGTGTCCGTGGACGTGGGGGAAGACGCTTCCTTCAGCTGCGCCTGG
1141 ---------+---------+---------+---------+---------+---------+ 1200
     A  K  P  E  P  V  S  V  D  V  G  E  D  A  S  F  S  C  A  W

     CGCGGGAACCCGCTTCCACGGGTAACCTGGACCCGCGCGGGTGGCGCTCAGGTGGTGGGG
1201 ---------+---------+---------+---------+---------+---------+ 1260
     R  G  N  P  L  P  R  V  T  W  T  R  R  G  G  A  Q  V  L  G

     TCTGGAGCCACACTGCGTCTTCCGTCGGTGGGGCCCGAGGACGCAGGCGACTATGTGTGC
1261 ---------+---------+---------+---------+---------+---------+ 1320
     S  G  A  T  L  R  L  P  S  V  G  P  E  D  A  G  D  Y  V  C

     AGAGCTGAGGCTGGGCTATCGGGCCTGCGGGGCCGCCGCCGGGAGGCTCGGCTGACTGTG
1321 ---------+---------+---------+---------+---------+---------+ 1380
     R  A  E  A  G  L  S  G  L  R  G  G  A  A  E  A  R  L  T  V
```

FIG. 7 (cont.)

```
                AACGCTCCCCCAGTAGTGACCGCCCTGCACTCTGCGCCTGCCTTCCTGAGGGGCCCTGCT
1381            ---------+---------+---------+---------+---------+---------+ 1440
                N  A  F  P  V  V  T  A  L  H  S  A  P  A  F  L  R  G  P  A

                CGCCTCCAGTGTCTGGTTTTCGCCTCTCCCGCCCCAGATGCCGTGGTCTGGTCTTGGGAT
1441            ---------+---------+---------+---------+---------+---------+ 1500
                R  L  Q  C  L  V  F  A  S  P  A  P  D  A  V  V  W  S  W  D

                GAGGGCTTCCTGGAGGCGGGGTCGCAGGGCCGGTTCCTGGTGGAGACATTCCCTGCCCCA
1501            ---------+---------+---------+---------+---------+---------+ 1560
                E  G  F  L  E  A  G  S  Q  G  R  F  L  V  E  T  F  P  A  P

                GAGAGCCGCGGGGGACTGGGTCCGGGCCTGATCTCTGTGCTACACATTTCGGGGACCCAG
1561            ---------+---------+---------+---------+---------+---------+ 1620
                E  S  R  G  G  L  G  P  G  L  I  S  V  L  H  I  S  G  T  Q

                GAGTCTGACTTTAGCAGGAGCTTTAACTGCAGTGCCCGGAACCGGCTGGGCGAGGGAGGT
1621            ---------+---------+---------+---------+---------+---------+ 1680
                E  S  D  F  S  R  S  F  N  C  S  A  R  N  R  L  G  E  G  G

                GCCCAGGCCAGCCTGGGCCGTAGAGACTTGCTGCCCACTGTGCGGATAGTGGCCGGAGTG
1681            ---------+---------+---------+---------+---------+---------+ 1740
                A  Q  A  S  L  G  R  R  D  L  L  P  T  V  R  I  V  A  G  V

                GCCGCTGCCACCACAACTCTCCTTATGGTCATCACTGGGGTGGCCCTCTGCTGCTGGCGC
1741            ---------+---------+---------+---------+---------+---------+ 1800
                A  A  A  T  T  T  L  L  M  V  I  T  G  V  A  L  C  C  W  R

                CACAGCAAGGCCTCAGCCTCTTTCTCCGAGCAAAAGAACCTGATGCGAATCCCTGGCAGC
1801            ---------+---------+---------+---------+---------+---------+ 1860
                H  S  K  A  S  A  S  F  S  E  Q  K  N  L  M  R  I  P  G  S

                AGCGACGGCTCGAGTTCACGAGGTCCTGAAGAAGAGGAGACAGCCAGCCGCGAGGACCGG
1861            ---------+---------+---------+---------+---------+---------+ 1920
                S  D  G  S  S  S  R  G  P  E  E  E  E  T  G  S  R  E  D  R

                GGCCCCATTGTGCACACTGACCACAGTGATCTGGTTCTGGAGGAGGAAGGGACTCTGGAG
1921            ---------+---------+---------+---------+---------+---------+ 1980
                G  P  I  V  H  T  D  H  S  D  L  V  L  E  E  E  G  T  L  E

                ACCAAGGACCCAACCAACGGTTACTACAAGGTCCGAGGAGTCAGTGTGAGCCTGAGCCTT
1981            ---------+---------+---------+---------+---------+---------+ 2040
                T  K  D  P  T  N  G  Y  Y  K  V  R  G  V  S  V  S  L  S  L
```

FIG. 7 (cont.)

```
                GGCGAAGCCCCTGGAGGAGGTCTCTTCCTGCCACCACCCTCCCCCCTTGGGCCCCCAGGG
2041            ---------+---------+---------+---------+---------+---------+ 2100
                G  E  A  P  G  G  G  L  F  L  P  P  P  S  P  L  G  P  P  G

                ACCCCTACCTTCTATGACTTCAACCCACACCTGGGCATGGTCCCCCCCTGGCAGACTTTAC
2101            ---------+---------+---------+---------+---------+---------+ 2160
                T  P  T  F  Y  D  F  N  P  H  L  G  M  V  P  P  C  R  L  Y

                AGAGCCAGGGCAGGCTATCTCACCACACCCCACCCTCGAGCTTTCACCAGCTACATCAAA
2161            ---------+---------+---------+---------+---------+---------+ 2220
                R  A  R  A  G  Y  L  T  T  P  H  P  R  A  F  T  S  Y  I  K

                CCCACATCCTTTGGGCCCCCAGATCTGGCCCCCGGGACTCCCCCCTTCCCATATGCTGCC
2221            ---------+---------+---------+---------+---------+---------+ 2280
                P  T  S  F  G  P  P  D  L  A  P  G  T  P  P  F  P  Y  A  A

                TTCCCCACACCTAGCCACCCGCGTCTCCAGACTCACGTGTGACATCTTTCCAATGGAAGA
2281            ---------+---------+---------+---------+---------+---------+ 2340
                F  P  T  P  S  H  P  R  L  Q  T  H  V  *

                GTCCTGGGATCTCCAACTTGCCATAATGGATTGTTCTGATTTCTGAGGAGCCAGGACAAG
2341            ---------+---------+---------+---------+---------+---------+ 2400

                TTGGCGACCTTACTCCTCCAAAACTGAACACAAGGGGAGGGAAAGATCATTACATTTGTC
2401            ---------+---------+---------+---------+---------+---------+ 2460

                AGGAGCATTTGTATACAGTCAGCTCAGCCAAACGAGATGCCCCAAGTGGGAGCAACATGG
2461            ---------+---------+---------+---------+---------+---------+ 2520

                CCACCCAATATGCCCACCTATTCCCCGGTGTAAAAGAGATTCAAGATGGCAGGTAGGCCC
2521            ---------+---------+---------+---------+---------+---------+ 2580

                TTTGAGGAGAGATGGGGACAGGGCAGTGGGTGTTGGGAGTTTGGGGCCGGGATGGAAGTT
2581            ---------+---------+---------+---------+---------+---------+ 2640

                GTTTCTAGCCACTGAAAGAAGATATTTCAAGATGACCATCTGCATTGAGAGGAAAGGTAG
2641            ---------+---------+---------+---------+---------+---------+ 2700

                CATAGGATAGATGAAGATGAAGAGCATACCAGGCCCCACCCTGGCTCTCCCTGAGGGGAA
2701            ---------+---------+---------+---------+---------+---------+ 2760
```

FIG. 7 (cont.)

```
      CTTTGCTCGGCCAATGGAAATGCAGCCAAGATGGCCATATACTCCCTAGGAACCCAAGAT
2761  ---------+---------+---------+---------+---------+---------+  2820

      GGCCACCATCTTGATTTTACTTTCCTTAAAGACACAGAAAGACTTGGACCCAAGGAGTGG
2821  ---------+---------+---------+---------+---------+---------+  2880

      GGATACAGTGAGAATTACCACTGTTGGGGCAAAATATTGGGATAAAAATATTTATGTTTA
2881  ---------+---------+---------+---------+---------+---------+  2940

      ATAATAAAAAAAAGTCAAA
2941  ---------+---------  2959
```

FIG. 7 (cont.)

```
   1 ATGCGGGTCCCCGCCCTCCTCGTCCTCCTCTTCTGCTTCAGAGGGAGAGCAGGCCCGTCG
     M  R  V  P  A  L  L  V  L  L  F  C  F  R  G  R  A  G  P  S

  61 CCCCATTTCCTGCAACAGCCAGAGGACCTGGTGGTGCTGCTGGGGGAGGAAGCCCGGCTG
     P  H  F  L  Q  Q  P  E  D  L  V  V  L  L  G  E  E  A  R  L

 121 CCGTGTGCTCTGGGCGCCTACTGGGGGCTAGTTCAGTGGACTAAGAGTGGGCTGGCCCTA
     P  C  A  L  G  A  Y  W  G  L  V  Q  W  T  K  S  G  L  A  L

 181 GGGGGCCAAAGGGACCTACGAGGGTGGTCCCGGTACTGGATATCAGGGAATGCAGCCAAT
     G  G  Q  R  D  L  P  G  W  S  R  Y  W  I  S  G  N  A  A  N

 241 GGCCAGCATGACCTCCACATTAGGCCCGTGGAGCTAGAGGATGAAGCATCATATGAATGT
     G  Q  H  D  L  H  I  R  P  V  E  L  E  D  E  A  S  Y  E  C

 301 CAGGCTACACAAGCAGGCCTCCGCTCCAGACCAGCCCAACTGCACGTGCTGGTCCCCCCA
     Q  A  T  Q  A  G  L  R  S  R  P  A  Q  L  H  V  L  V  P  P

 361 GAAGCCCCCCAGGTGCTGGCGGCCCCTCTGTGTCTCTGGTTGCTGGAGTTCCTGCGAAC
     E  A  P  Q  V  L  G  G  P  S  V  S  L  V  A  G  V  P  A  N

 421 CTGACATGTCGGAGCCGTGGGGATGCCCGCCCTACCCCTGAATTGCTGTGGTTCCGAGAT
     L  T  C  R  S  R  G  D  A  R  P  T  P  E  L  L  W  F  R  D

 481 GGGGTCCTGTTGGATGGAGCCACCTTCCATCAGACCCTGCTGAAGGAAGGGACCCCTGGG
     G  V  L  L  D  G  A  T  F  H  Q  T  L  L  K  E  G  T  P  G

 541 TCAGTGGAGAGCACCTTAACCCTGACCCCTTTCAGCCATGATGATGGAGCCACCTTTGTC
     S  V  E  S  T  L  T  L  T  P  F  S  H  D  D  G  A  T  F  V

 601 TGCCGGGCCCGGAGCCAGGCCCTGCCCACAGGAAGAGACACAGCTATCACACTGAGCCTG
     C  R  A  R  S  Q  A  L  P  T  G  R  D  T  A  I  T  L  S  L

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

FIG. 8 (cont.)

```
721  GTCATTTTCCTGTGCCAGCCCACAGCCCAGCCTCCTGTCACAGGCTACAGGTGGGCAAAA  780
     V  I  F  L  C  Q  A  T  A  Q  P  P  V  T  G  Y  R  W  A  K

781  GGGGGCTCTCCGGTGCTCGGGGCCCCCCGGGCCAAGGTTAGAGGTCGTGGCAGACGCCTCG  840
     G  G  S  P  V  L  G  A  R  G  P  R  L  E  V  V  A  D  A  S

841  TTCCTGACTGAGCCCGTGTCCTGCGAGGTCAGCAACGCCGTGGGTAGCGCCAACCCCAGT  900
     F  L  T  E  P  V  S  C  E  V  S  N  A  V  G  S  A  N  R  S

901  ACTGCGCTGGATGTGCTGTTTGGGCCGATTCTGCAGGCAAAGCCGGAGCCCGTGTCCGTG  960
     T  A  L  D  V  L  F  G  P  I  L  Q  A  K  P  E  P  V  S  V

961  GACGTGGGGGAAGACGCTTCCTTCAGCTGCGCCTGGCGCGGGAACCCGCTTCCACGGGTA  1020
     D  V  G  E  D  A  S  F  S  C  A  W  R  G  N  P  L  P  R  V

1021 ACCTGGACCCGCCGCGGTGGCGCTCAGGTGCTGGGCTCTGGAGCCACACTGCGTCTTCCG  1080
     T  W  T  R  R  G  G  A  Q  V  L  G  S  G  A  T  L  R  L  P

1081 TCGGTGGGGCCCGAGGACGGAGGCGACTATGTGTGCAGAGCTGAGGCTGGGCTATCGGGC  1140
     S  V  G  P  E  D  A  G  D  Y  V  C  R  A  E  A  G  L  S  G

1141 CTGCGGGGCGGCGCCGCGGAGGCTCGGCTGACTGTGAACGCTCCCCCAGTAGTGACCGCC  1200
     L  R  G  G  A  A  E  A  R  L  T  V  N  A  P  P  V  V  T  A

1201 CTGCACTCTGCGCCTGCCTTCCTGAGGGGCCGTGCTCGGCTCCAGTGTCTGGTTTTCGCC  1260
     L  H  S  A  P  A  F  L  R  G  P  A  R  L  Q  C  L  V  F  A

1261 TCTCCCGCCCCAGATGCCGTGGTCTGGTCTTGGGATGAGGCTTCCTGGAGGCGGGGTCG  1320
     S  P  A  P  D  A  V  V  W  S  W  D  E  G  F  L  E  A  G  S

1321 CAGGGCCGGTTCCTGGTGGAGACATTCCCTGCCCCAGAGAGCCGCGGGGGACTGGGTCCG  1380
     Q  G  R  F  L  V  E  T  F  P  A  P  E  S  R  G  G  L  G  P

1381 GGCCTGATCTCTGTGCTACACATTTCGGGGACCCAGGAGTCTGACTTTAGCAGGAGCTTT  1440
     G  L  I  S  V  L  H  I  S  G  T  Q  E  S  D  F  S  R  S  F
```

FIG. 8 (cont.)

```
     AACTGCAGTGCCCGGAACCGGCTGGGCGAGGGAGGTGCCCAGGCCAGCCTGGGCCGTAGA
1441 ---------+---------+---------+---------+---------+---------+ 1500
     N  C  S  A  R  N  R  L  G  E  G  G  A  Q  A  S  L  G  R  R

     GACTTGCTGCCCACTGTGCGGATAGTGGCCGGAGTGGCCGCTGCCACCACAACTCTCCTT
1501 ---------+---------+---------+---------+---------+---------+ 1560
     D  L  L  P  T  V  R  I  V  A  G  V  A  A  A  T  T  T  L  L

     ATGGTCATCACTGGGGTGGCCCTCTGCTGCTGGCGCCACAGCAAGGCCTCAGCCTCTTTC
1561 ---------+---------+---------+---------+---------+---------+ 1620
     M  V  I  T  G  V  A  L  C  C  W  R  H  S  K  A  S  A  S  F

     TCCGAGCAAAAGAACCTGATGCGAATCCCTGGCAGCAGCGACGGCTCCAGTTCACGAGGT
1621 ---------+---------+---------+---------+---------+---------+ 1680
     S  E  Q  K  N  L  M  R  I  P  G  S  S  D  G  S  S  S  R  G

     CCTGAAGAAGAGGAGACAGGCAGCCGCGAGGACCGGGGCCCCATTGTGCACACTGACCAC
1681 ---------+---------+---------+---------+---------+---------+ 1740
     P  E  E  E  E  T  G  S  R  E  D  R  G  P  I  V  H  T  D  H

     AGTGATCTGGTTCTGGAGGAGGAAGGGACTCTGGAGACCAAG
1741 ---------+---------+---------+---------+-- 1782
     S  D  L  V  L  E  E  E  G  T  L  E  T  K
```

FIG. 8 (cont.)

【国際公開パンフレット（コレクトバージョン）】

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/098360 A3

(51) International Patent Classification⁷: C12N 15/12, C07K 14/705, 16/28, C12Q 1/68, G01N 33/50, A61K 31/7088, 38/17, 39/395

(21) International Application Number: PCT/US01/19902

(22) International Filing Date: 22 June 2001 (22.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/213,611 22 June 2000 (22.06.2000) US

(71) Applicant (*for all designated States except US*): BIOGEN, INC. [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): CARULLI, John, P. [US/US]; 9 Harris Drive, Southborough, MA 01772 (US). LUKASHIN, Alexander, V. [US/US]; 42 8th Street, #4403, Charlestown, MA 02139 (US). KILBURN, Daniel, R. [US/US]; 6 Hancock Avenue, Lexington, MA 02420 (US). SUN, Chao [CN/US]; 11 Cimino Road, Needham, MA 02494 (US).

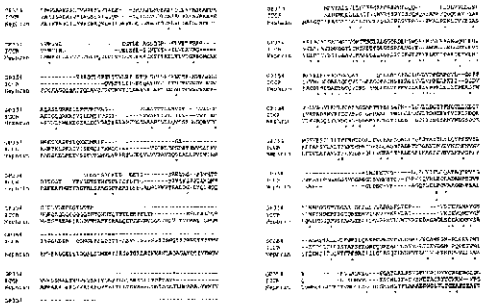
(74) Agents: HALEY, James, F., Jr. et al.; Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

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

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: GP354 NUCLEIC ACIDS AND POLYPEPTIDES



(57) Abstract: An isolated polynucleotide encoding a novel immunoglobulin superfamily member named GP354 is provided. GP354 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 GP354 indicate that it plays a role in cell-cell recognition, binding, signaling and adhesion events in the pancreas and central nervous system (CNS). Provided by the invention are isolated GP354 related polynucleotides and polypeptides, vectors, and host cells comprising any of the above, antibodies directed to GP354, cells which produce such antibodies, and related diagnostic and therapeutic methods.

WO 01/098360 A3

WO 01/098360 A3

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)*

as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(iv)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)

as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)

Published:

with international search report

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
28 November 2002

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

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/19904

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/705 C07K16/28 C12Q1/68 G01N33/50
A61K31/7088 A61K38/17 A61K39/395

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] EBI, Hinxton, UK; 31 January 2000 (2000-01-31) ABOLA A P ET AL: "Homo sapiens chromosome 19 clone RP11-38C1" Database accession no. AC022315 XP002212519 see nucleotides 81730 to 96300 abstract ---	1-10,13, 14
X	DATABASE EMBL [Online] EBI, Hinxton, UK; 11 June 1997 (1997-06-11) LAMERDIN J E ET AL: "Sequence analysis of a 1 Mb region in 19q13.1" Database accession no. AC002133 XP002212520 abstract --- -/--	1-10,13. 14

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 September 2002	20. 09. 2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL- 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax. (+31-70) 340-3016	Oderwald, H

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/19904

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RAMOS RICARDO G P ET AL: "The Irregular chiasm C-roughest locus of Drosophila. which affects axonal projections and programmed cell death, encodes a novel immunoglobulin-like protein." GENES & DEVELOPMENT, vol. 7, no. 12B, 1993, pages 2533-2547, XP008007783 ISSN: 0890-9369 the whole document ---	1-10,13, 14, 16-18, 25,26
P,X	DATABASE EMBL [Online] EBI, Hinxton, UK; 5 January 2001 (2001-01-05) DUESTERHOEFT A ET AL: "Homo sapiens mRNA" Database accession no. AL136654 XP002212521 abstract ---	1-10,13, 14, 16-18,25
E	WO 01 68848 A (BAKER KEVIN P ;CHEN JIAN (US); DESNOYERS LUC (US); GENENTECH INC ()) 20 September 2001 (2001-09-20) see SEQ ID NO: 583 and 584 page 1 -page 95; claims 1-24; figures 583,584; examples 1-16 -----	1-44

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/19904

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 39-44 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (1)) (July 1998)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/19004

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 0169848 A	20-09-2001	AU	2055401 A	12-06-2001
		AU	3864800 A	09-10-2000
		AU	5152700 A	02-01-2001
		AU	5459900 A	18-12-2000
		AU	6802801 A	24-09-2001
		AU	7079300 A	26-03-2001
		AU	7573000 A	26-03-2001
		EP	1169442 A2	09-01-2002
		EP	1208195 A2	29-05-2002
		EP	1238201 A2	29-05-2002
		EP	1238202 A2	29-05-2002
		WO	0077037 A2	21-12-2000
		WO	0073452 A2	07-12-2000
		WO	0116318 A2	08-03-2001
		WO	0116319 A2	08-03-2001
		WO	0140466 A2	07-06-2001
		WO	0168848 A2	20-09-2001
		US	2002090681 A1	11-07-2002
		AU	1916701 A	16-07-2001
		AU	2590901 A	03-07-2001
		AU	3514400 A	28-09-2000
		AU	3774300 A	18-12-2000
		AU	6391000 A	19-02-2001
		EP	1220905 A2	10-07-2002
		EP	1210418 A1	05-06-2002
		EP	1212417 A2	12-06-2002
		EP	1185648 A2	13-03-2002
		WO	0073454 A1	07-12-2000
		WO	0073445 A2	07-12-2000
		WO	0073348 A2	07-12-2000
		WO	0109327 A2	08-02-2001
		WO	0149715 A2	12-07-2001
		WO	0146420 A2	28-06-2001
		US	2002096506 A1	25-07-2002
		US	2002098506 A1	25-07-2002
		US	2002098507 A1	25-07-2002
		US	2002094416 A1	18-07-2002
		US	2002010137 A1	24-01-2002
		US	2002058309 A1	16-05-2002
		US	2002102647 A1	01-08-2002
		US	2002110859 A1	15-08-2002
		US	2002072496 A1	13-06-2002
		US	2002072067 A1	13-06-2002
		US	2002072092 A1	13-06-2002
		US	2002072497 A1	13-06-2002
		US	2002103125 A1	01-08-2002
		AU	3434601 A	12-06-2001
		AU	5516801 A	17-09-2001
		EP	1234036 A2	28-08-2002
		WO	0140465 A2	07-06-2001

Form PCT/ISA/210 (patent family annex) (July 1992)

フロントページの続き

(51)Int.Cl.[7]	FI		テーマコード（参考）
Ａ６１Ｐ 25/06	Ａ６１Ｐ 25/00		４Ｃ０８４
Ａ６１Ｐ 25/08	Ａ６１Ｐ 25/06		４Ｃ０８５
Ａ６１Ｐ 25/16	Ａ６１Ｐ 25/08		４Ｈ０４５
Ａ６１Ｐ 25/28	Ａ６１Ｐ 25/16		
Ａ６１Ｐ 35/00	Ａ６１Ｐ 25/28		
Ｃ０７Ｋ 16/18	Ａ６１Ｐ 35/00		
Ｃ０７Ｋ 16/42	Ｃ０７Ｋ 16/18		
Ｃ０７Ｋ 19/00	Ｃ０７Ｋ 16/42		
Ｃ１２Ｎ 5/10	Ｃ０７Ｋ 19/00		
Ｃ１２Ｐ 21/02	Ｃ１２Ｐ 21/02	Ｃ	
Ｃ１２Ｑ 1/68	Ｃ１２Ｑ 1/68	Ａ	
Ｇ０１Ｎ 33/15	Ｇ０１Ｎ 33/15	Ｚ	
Ｇ０１Ｎ 33/50	Ｇ０１Ｎ 33/50	Ｚ	
Ｇ０１Ｎ 33/53	Ｇ０１Ｎ 33/53	Ｄ	
Ｇ０１Ｎ 33/566	Ｇ０１Ｎ 33/53	Ｍ	
// Ｃ１２Ｐ 21/08	Ｇ０１Ｎ 33/566		
	Ｃ１２Ｎ 5/00	Ｃ	
	Ｃ１２Ｐ 21/08		

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

(72)発明者 カルーリ, ジョン ピー.
アメリカ合衆国 マサチューセッツ ０１７７２, サウスボロー, ハリス ドライブ ９
(72)発明者 ルカシン, アレキサンダー ブイ.
アメリカ合衆国 マサチューセッツ ０２１３９, チャールズタウン, ８ティーエイチ ストリート ４２, ナンバー４４０３
(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 アベニュー ６
(72)発明者 サン, チャオ
アメリカ合衆国 マサチューセッツ ０２４９４, ニードハム, シミノ ロード １１

Ｆターム(参考) 2G045 AA34 AA35 AA40 BA11 BB50 DA12 DA13 DA14 DA36 FB02 FB03
4B024 AA01 AA11 BA44 BA61 CA04 CA09 CA11 DA02 DA05 DA06 DA11 EA02 EA04 GA01 GA11 HA09 HA11 HA12
4B063 QA01 QA18 QQ02 QQ08 QQ21 QQ42 QQ53 QQ63 QQ79 QQ89 QR08 QR31 QR42 QR48 QR56 QR72 QR77 QS03 QS25 QS28 QS33 QS34 QS36 QX02
4B064 AG26 AG27 CA02 CA05 CA10 CA11 CA19 CA20 CC24 DA01
4B065 AA01X AA26X AA57X AA87X AA90Y AB01 AB02 BA02 BA08 CA24 CA25 CA44
4C084 AA06 AA07 AA13 BA22 CA53 CA56 DA39 MA23 MA35 MA52 MA66 NA14 ZA011 ZA021 ZA061 ZA081 ZA161 ZA661 ZB261

4C085 AA13 AA14 CC05 DD22 DD23 EE01 GG01
4H045 AA11 AA20 AA30 BA09 CA40 DA75 DA76 EA20 FA72 FA73
FA74

专利名称(译)	GP354核酸和多肽		
公开(公告)号	JP2004500872A	公开(公告)日	2004-01-15
申请号	JP2002504315	申请日	2001-06-22
[标]申请(专利权)人(译)	拜奥根有限公司		
申请(专利权)人(译)	生物遗传公司		
[标]发明人	カルーリジョンピー ルカシンアレキサンダーブイ キルバーンダニエルアール サンチャオ		
发明人	カルーリ, ジョン ピー. ルカシン, アレキサンダー ブイ. キルバーン, ダニエル アール. サン, チャオ		
IPC分类号	G01N33/50 A61K38/00 A61K39/395 A61K48/00 A61P1/18 A61P25/00 A61P25/06 A61P25/08 A61P25/16 A61P25/28 A61P35/00 C07K14/705 C07K16/18 C07K16/42 C07K19/00 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	A61K38/00 A61P1/18 A61P25/00 A61P25/06 A61P25/08 A61P25/16 A61P25/28 C07K14/70503		
FI分类号	C12N15/00.ZNA.A A61K39/395.D A61K39/395.N A61K48/00 A61P1/18 A61P25/00 A61P25/06 A61P25/08 A61P25/16 A61P25/28 A61P35/00 C07K16/18 C07K16/42 C07K19/00 C12P21/02.C C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.C C12P21/08		
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/AA40 2G045/BA11 2G045/BB50 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA44 4B024/BA61 4B024/CA04 4B024/CA09 4B024/CA11 4B024/DA02 4B024/DA05 4B024/DA06 4B024/DA11 4B024/EA02 4B024/EA04 4B024/GA01 4B024/GA11 4B024/HA09 4B024/HA11 4B024/HA12 4B063/QA01 4B063/QA18 4B063/QQ02 4B063/QQ08 4B063/QQ21 4B063/QQ42 4B063/QQ53 4B063/QQ63 4B063/QQ79 4B063/QQ89 4B063/QR08 4B063/QR31 4B063/QR42 4B063/QR48 4B063/QR56 4B063/QR72 4B063/QR77 4B063/QS03 4B063/QS25 4B063/QS28 4B063/QS33 4B063/QS34 4B063/QS36 4B063/QX02 4B064/AG26 4B064/AG27 4B064/CA02 4B064/CA05 4B064/CA10 4B064/CA11 4B064/CA19 4B064/CA20 4B064/CC24 4B064/DA01 4B065/AA01X 4B065/AA26X 4B065/AA57X 4B065/AA87X 4B065/AA90Y 4B065/AB01 4B065/AB02 4B065/BA02 4B065/BA08 4B065/CA24 4B065/CA25 4B065/CA44 4C084/AA06 4C084/AA07 4C084/AA13 4C084/BA22 4C084/CA53 4C084/CA56 4C084/DA39 4C084/MA23 4C084/MA35 4C084/MA52 4C084/MA66 4C084/NA14 4C084/ZA011 4C084/ZA021 4C084/ZA061 4C084/ZA081 4C084/ZA161 4C084/ZA661 4C084/ZB261 4C085/AA13 4C085/AA14 4C085/CC05 4C085/DD22 4C085/DD23 4C085/EE01 4C085/GG01 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA09 4H045/CA40 4H045/DA75 4H045/DA76 4H045/EA20 4H045/FA72 4H045/FA73 4H045/FA74		
代理人(译)	夏木森下		
优先权	60/213611 2000-06-22 US		
其他公开文献	JP2004500872A5		
外部链接	Espacenet		

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の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の結果

エキソン	始め	終り	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