

(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2004-535174

(P2004-535174A)

(43) 公表日 平成16年11月25日(2004.11.25)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00 Z N A A	2 G O 4 5
A 6 1 K 31/7088	A 6 1 K 31/7088	4 B O 2 4
A 6 1 K 38/00	A 6 1 K 39/395 D	4 B O 2 9
A 6 1 K 39/395	A 6 1 K 39/395 N	4 B O 6 3
A 6 1 K 45/00	A 6 1 K 45/00	4 B O 6 4
審査請求 未請求 予備審査請求 未請求 (全 312 頁) 最終頁に続く		

(21) 出願番号	特願2002-583584 (P2002-583584)	(71) 出願人	301005050 インサイト・ゲノミックス・インコーポレイテッド アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) (22) 出願日	平成14年4月19日 (2002.4.19)	(74) 代理人	100089266 弁理士 大島 陽一
(85) 翻訳文提出日	平成15年10月17日 (2003.10.17)	(72) 発明者	クラマー、アーロン・エイ アメリカ合衆国カリフォルニア州94040・マウンテンビュー・#1・オルテガアベニュー 565
(86) 国際出願番号	PCT/US2002/012464		
(87) 国際公開番号	W02002/086069		
(87) 国際公開日	平成14年10月31日 (2002.10.31)		
(31) 優先権主張番号	60/285, 207		
(32) 優先日	平成13年4月20日 (2001.4.20)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/287, 114		
(32) 優先日	平成13年4月27日 (2001.4.27)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/288, 640		
(32) 優先日	平成13年5月3日 (2001.5.3)		
(33) 優先権主張国	米国 (US)		
最終頁に続く			

(54) 【発明の名称】 分泌タンパク質

(57) 【要約】

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

【特許請求の範囲】

【請求項 1】

以下の(a)乃至(i)からなる群から選択した単離されたポリペプチド。

(a)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むポリペプチド

(b)SEQ ID NO:2、SEQ ID NO:4-5、SEQ ID NO:7-12、SEQ ID NO:14-22、SEQ ID NO:24-30からなる群から選択した或るアミノ酸配列に対して少なくとも90%が同一であるような天然アミノ酸配列を含むポリペプチド

(c)SEQ ID NO:1、SEQ ID NO:3からなる群から選択した或るアミノ酸配列に対して少なくとも90%が同一であるような天然アミノ酸配列を含むポリペプチド

(d)SEQ ID NO:6のアミノ酸配列に対して少なくとも98%が同一であるような天然アミノ酸配列を含むポリペプチド 10

(e)SEQ ID NO:13のアミノ酸配列に対して少なくとも92%が同一であるような天然アミノ酸配列を含むポリペプチド

(g)SEQ ID NO:23のアミノ酸配列に対して少なくとも93%が同一であるような天然アミノ酸配列を含むポリペプチド

(h)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片

(i)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片。

【請求項 2】

SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含む請求項 1 に記載の単離されたポリペプチド。 20

【請求項 3】

請求項1に記載のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項2に記載のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列を有する、請求項 4 に記載の単離されたポリヌクレオチド。

【請求項 6】

請求項 3 に記載のポリヌクレオチドに機能的に連結したプロモーター配列を含む組換えポリヌクレオチド。 30

【請求項 7】

請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを含む遺伝形質転換体。

【請求項 9】

請求項 1 に記載のポリペプチドを生産する方法であって、

(a)前記ポリペプチドの発現に好適な条件下で、請求項 1 に記載されたポリペプチドをコードするポリヌクレオチドに機能的に連結されたプロモーター配列を含む組換えポリヌクレオチドで形質転換される細胞を培養する過程と、 40

(b)そのように発現した前記ポリペプチドを回収する過程とからなる方法。

【請求項 10】

前記ポリペプチドが、SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を含むことを特徴とする、請求項 9 に記載の方法。

【請求項 11】

請求項 1 に記載のポリペプチドと特異的に結合する単離された抗体。

【請求項 12】

以下の(a)乃至(i)からなる群から選択した単離されたポリヌクレオチド。

(a)SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチ 50

ド

(b)SEQ ID NO:32、SEQ ID NO:34-35、SEQ ID NO:37-60からなる群から選択した或るポリヌクレオチド配列に対して少なくとも90%が同一であるような天然ポリヌクレオチド配列を含むポリヌクレオチド

(c)SEQ ID NO:31とSEQ ID NO:33からなる群から選択したポリヌクレオチド配列に対して少なくとも94%が同一であるような天然ポリヌクレオチド配列を含むポリヌクレオチド

(d)SEQ ID NO:36のポリヌクレオチド配列に対して少なくとも98%が同一であるような天然ポリヌクレオチド配列を含むポリヌクレオチド

(e)(a)のポリヌクレオチドに相補的なポリヌクレオチド

(f)(b)のポリヌクレオチドに相補的なポリヌクレオチド

(g)(c)のポリヌクレオチドに相補的なポリヌクレオチド

(h)(d)のポリヌクレオチドに相補的なポリヌクレオチド

(i)(a)～(h)のRNA等価物

10

【請求項13】

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

【請求項14】

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

(a)前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を持つ少なくとも20の連続したヌクレオチドを含むプローブを用いて前記サンプルをハイブリダイズする過程であって、前記プローブと前記標的ポリヌクレオチドあるいはその断片との間でハイブリダイゼーション複合体が形成される条件下で、プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズすることを特徴とする過程と、

20

(b)前記ハイブリダイゼーション複合体の有無を検出し、該複合体が存在する場合にはオプションでその量を検出する過程を含む方法。

【請求項15】

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

【請求項16】

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

30

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

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

【請求項17】

請求項1に記載のポリペプチドと、薬剤として許容できる賦形剤とを含む組成物。

【請求項18】

前記ポリペプチドが、SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むことを特徴とする請求項17に記載の組成物。

40

【請求項19】

機能的なSECPの発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項17の組成物を投与することを含むことを特徴とする治療方法。

【請求項20】

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

(a)請求項1に記載のポリペプチドを含むサンプルを化合物に曝すステップと、

(b)前記サンプルにおいてアゴニスト活性を検出する過程を含むことを特徴とする方法。

50

【請求項 2 1】

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

【請求項 2 2】

機能的なSECPの発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項21の組成物を投与することを含むことを特徴とする治療方法。

【請求項 2 3】

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

- (a) 請求項 1 に記載のポリペプチドを含むサンプルを化合物に曝すステップと、
- (b) 前記サンプルにおいてアンタゴニスト活性を検出する過程を含むことを特徴とする方法。

10

【請求項 2 4】

請求項23に記載の方法によって同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを含む組成物。

【請求項 2 5】

機能的なSECPの過剰な発現に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項24の組成物を投与することを含むことを特徴とする治療方法。

【請求項 2 6】

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

20

- (a) 請求項 1 に記載のポリペプチドを適切な条件下で少なくとも1つの試験化合物と混合する過程と、
- (b) 請求項 1 に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 に記載のポリペプチドに特異結合する化合物を同定する過程を含む方法。

【請求項 2 7】

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

- (a) 請求項 1 に記載のポリペプチドの活性が許容される条件下で、請求項 1 に記載のポリペプチドを少なくとも1つの試験化合物と混合する過程と、
- (b) 請求項 1 に記載のポリペプチドの活性を試験化合物の存在下で算定する過程と、
- (c) 試験化合物の存在下での請求項 1 に記載のポリペプチドの活性を、試験化合物の不存在下での請求項 1 に記載のポリペプチドの活性と比較する過程を含み、試験化合物の存在下での請求項 1 に記載のポリペプチドの活性の変化が、請求項 1 に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。

30

【請求項 2 8】

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

- (a) 前記標的ポリヌクレオチドの発現に好適な条件下で、該標的ポリヌクレオチドを含むサンプルを化合物に曝露する過程と、
- (b) 前記標的ポリヌクレオチドの発現改変を検出する過程と、
- (c) 可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程とを含むことを特徴とする方法。

40

【請求項 2 9】

試験化合物の毒性を算定する方法であって、

- (a) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、
- (b) 処理した前記生物学的サンプルの核酸と、請求項12に記載のポリヌクレオチドの少なくとも20の連続するヌクレオチドを持つプローブをハイブリダイズさせる過程であって、このハイブリダイゼーションセッションが、前記プローブと前記生物学的サンプル中の標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下

50

で行われ、前記標的ポリヌクレオチドが、請求項12に記載のポリヌクレオチドまたはその断片のポリヌクレオチド配列を持つポリヌクレオチドである、前記過程と、

(c)ハイブリダイゼーション複合体の収量を定量する過程と、

(d)前記処理された生物学的サンプル中のハイブリタイゼーション複合体の量を、処理されていない生物学的サンプル中のハイブリタイゼーション複合体の量と比較する過程とを含み、前記処理された生物学的サンプル中のハイブリタイゼーション複合体の量の差が、前記試験化合物の毒性を標示するような方法。

【請求項30】

生物学的サンプル中のSECPの発現に関連する症状または疾患に対する診断試験法であって、

(a)前記生物学的サンプルと請求項11に記載の抗体との混合を、前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体を形成するのに適した条件下で行う過程と

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

【請求項31】

請求項11に記載の抗体であって、

(a)キメラ抗体

(b)単鎖抗体

(c)Fab断片

(d) $F(ab')_2$ 断片、あるいは

(e)ヒト化抗体のいずれかである抗体。

【請求項32】

請求項11に記載の抗体と、許容できる賦形剤とを含む組成物。

【請求項33】

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

【請求項34】

前記抗体が標識されることを特徴とする請求項32に記載の組成物。

【請求項35】

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

【請求項36】

請求項11に記載の抗体の特異性を有するポリクローナル抗体を調製する方法であって、

(a)抗体反応を誘発する条件下で、SEQ ID NO:1-30からなる群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて動物を免疫化する過程と、

(b)前記動物から抗体を単離する過程と、

(c)前記単離された抗体を前記ポリペプチドでスクリーニングし、それによって、SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むポリペプチドに特異的に結合するポリクローナル抗体を同定する過程とを含むような方法。

【請求項37】

請求項36に記載の方法で産出したポリクローナル抗体。

【請求項38】

請求項37に記載のポリクローナル抗体及び適切なキャリアを含む組成物。

【請求項39】

請求項11に記載の抗体の特異性を有するモノクローナル抗体を作製する方法であって、

(a)抗体反応を誘発する条件下で、SEQ ID NO:1-30からなる群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて動物を免疫化する過程と、

(b)前記動物から抗体産出細胞を単離する過程と、

(c)前記抗体産出細胞と不死化した細胞とを融合して、モノクローナル抗体を産出するハ

10

20

30

40

50

イブリドーマ細胞を形成する過程と、

(d)前記ハイブリドーマ細胞を培養する過程と、

(e)SEQ ID NO: 1-30からなる群から選択したアミノ酸配列を含むポリペプチドに特異的に結合するようなモノクローナル抗体を前記培養物から単離する過程とを含むことを特徴とする方法。

【請求項40】

請求項39に記載の方法で産出したモノクローナル抗体。

【請求項41】

請求項40に記載のモノクローナル抗体及び適切なキャリアを含む組成物。

【請求項42】

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

【請求項43】

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

【請求項44】

SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むポリペプチドをサンプル中に検出する方法であって、

(a)請求項11に記載の抗体と前記ポリペプチドとの特異結合を許容する条件下で、前記抗体と1サンプルとをインキュベートする過程と、

(b)特異結合を検出する過程とを含み、該特異結合が、SEQ ID NO: 1-30からなる群から選択したアミノ酸配列を含むポリペプチドがサンプル中に存在することを標示することを特徴とする方法。

【請求項45】

SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むポリペプチドをサンプルから精製する方法であって、

(a)請求項11に記載の抗体と前記ポリペプチドとの特異結合を許容する条件下で、前記抗体と1サンプルとをインキュベートする過程と、

(b)前記サンプルから前記抗体を分離し、SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含む精製したポリペプチドを得る過程とを含むことを特徴とする方法。

【請求項46】

マイクロアレイの少なくとも1つのエレメントが請求項13に記載のポリヌクレオチドであるマイクロアレイ。

【請求項47】

ポリヌクレオチドを含むサンプルの発現プロファイルを作製する方法であって、

(a)サンプル中のポリヌクレオチドを標識化する過程

(b)ハイブリダイゼーション複合体が形成されるのに適した条件下で請求項46に記載のマイクロアレイのエレメントとサンプル中の標識化ポリヌクレオチドを接触させる過程

(c)サンプル中のポリヌクレオチドの発現を定量する過程を含む方法

【請求項48】

固体基板上の別個の物理的位置に固定した異なるヌクレオチド分子群を含むアレイであり、該ヌクレオチド分子の少なくとも1つが標的ポリヌクレオチドの少なくとも30の連続するヌクレオチドと特異的にハイブリダイゼーション可能な最初のオリゴヌクレオチドまたはポリヌクレオチド配列を含み、前記標的ポリヌクレオチドが請求項12に記載のポリヌクレオチドであるアレイ。

【請求項49】

請求項48に記載のアレイであり、前記最初のオリゴヌクレオチドまたはポリヌクレオチド配列は前記標的ポリヌクレオチドの少なくとも30の連続するヌクレオチドに完全に相補的であるアレイ。

【請求項50】

10

20

30

40

50

請求項48に記載のアレイであり、前記最初のオリゴヌクレオチドまたはポリヌクレオチド配列は前記標的ポリヌクレオチドの少なくとも60の連続するヌクレオチドに完全に相補的であるアレイ。

【請求項51】

請求項48に記載のアレイであり、前記最初のオリゴヌクレオチドまたはポリヌクレオチド配列が前記標的ポリヌクレオチドに完全に相補的であるアレイ。

【請求項52】

請求項48に記載のアレイであり、マイクロアレイであるアレイ。

【請求項53】

請求項48に記載のアレイであり、前記最初のオリゴヌクレオチドまたはポリヌクレオチド配列を含むヌクレオチド分子にハイブリダイズされた前記標的ポリヌクレオチドをも含むアレイ。

10

【請求項54】

請求項48に記載のアレイであり、リンカーが少なくとも1つの前記ヌクレオチド分子を前記固体基板に接合するアレイ。

【請求項55】

請求項48に記載のアレイであり、基板上のそれぞれ固有の物理的位置には複数のヌクレオチド分子を含み、またいかなる単一の固有の物理位置の該複数のヌクレオチド分子も同じ配列を有し、基板上の各固有物理的位置には基板上の別の固有の物理的位置でのヌクレオチド分子の配列と異なる配列を持つヌクレオチド分子が含まれるアレイ。

20

【請求項56】

SEQ ID NO:1のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項57】

SEQ ID NO:2のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項58】

SEQ ID NO:3のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項59】

SEQ ID NO:4のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項60】

SEQ ID NO:5のアミノ酸配列を含む請求項1に記載のポリペプチド。

30

【請求項61】

SEQ ID NO:6のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項62】

SEQ ID NO:7のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項63】

SEQ ID NO:8のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項64】

SEQ ID NO:9のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項65】

SEQ ID NO:10のアミノ酸配列を含む請求項1に記載のポリペプチド。

40

【請求項66】

SEQ ID NO:11のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項67】

SEQ ID NO:12のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項68】

SEQ ID NO:13のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項69】

SEQ ID NO:14のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項70】

SEQ ID NO:15のアミノ酸配列を含む請求項1に記載のポリペプチド。

50

- 【請求項 7 1】
SEQ ID NO:16のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 7 2】
SEQ ID NO:17のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 7 3】
SEQ ID NO:18のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 7 4】
SEQ ID NO:19のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 7 5】
SEQ ID NO:20のアミノ酸配列を含む請求項1に記載のポリペプチド。 10
- 【請求項 7 6】
SEQ ID NO:21のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 7 7】
SEQ ID NO:22のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 7 8】
SEQ ID NO:23のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 7 9】
SEQ ID NO:24のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 8 0】
SEQ ID NO:25のアミノ酸配列を含む請求項1に記載のポリペプチド。 20
- 【請求項 8 1】
SEQ ID NO:26のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 8 2】
SEQ ID NO:27のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 8 3】
SEQ ID NO:28のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 8 4】
SEQ ID NO:29のアミノ酸配列を含む請求項1に記載のポリペプチド。
- 【請求項 8 5】
SEQ ID NO:30のアミノ酸配列を含む請求項1に記載のポリペプチド。 30
- 【請求項 8 6】
SEQ ID NO:31のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。
- 【請求項 8 7】
SEQ ID NO:32のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。
- 【請求項 8 8】
SEQ ID NO:33のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。
- 【請求項 8 9】
SEQ ID NO:34のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。
- 【請求項 9 0】
SEQ ID NO:35のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。 40
- 【請求項 9 1】
SEQ ID NO:36のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。
- 【請求項 9 2】
SEQ ID NO:37のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。
- 【請求項 9 3】
SEQ ID NO:38のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。
- 【請求項 9 4】
SEQ ID NO:39のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。
- 【請求項 9 5】
SEQ ID NO:40のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。 50

【請求項 9 6】

SEQ ID NO:41のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 9 7】

SEQ ID NO:42のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 9 8】

SEQ ID NO:43のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 9 9】

SEQ ID NO:44のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 0 0】

SEQ ID NO:45のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

10

【請求項 1 0 1】

SEQ ID NO:46のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 0 2】

SEQ ID NO:47のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 0 3】

SEQ ID NO:48のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 0 4】

SEQ ID NO:49のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 0 5】

SEQ ID NO:50のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

20

【請求項 1 0 6】

SEQ ID NO:51のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 0 7】

SEQ ID NO:52のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 0 8】

SEQ ID NO:53のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 0 9】

SEQ ID NO:54のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 1 0】

SEQ ID NO:55のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

30

【請求項 1 1 1】

SEQ ID NO:56のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 1 2】

SEQ ID NO:57のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 1 3】

SEQ ID NO:58のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 1 4】

SEQ ID NO:59のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

【請求項 1 1 5】

SEQ ID NO:60のポリヌクレオチド配列を含む請求項12に記載のポリヌクレオチド。

40

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、分泌タンパク質の核酸配列およびアミノ酸配列に関する。本発明はまた、これらの配列を利用した、肝臓の疾患、細胞増殖異常、自己免疫/炎症疾患、心血管障害、神経障害、および発生障害の、診断・治療・予防に関する。本発明は更に、分泌タンパク質の核酸配列およびアミノ酸配列の発現における外来性化合物の効果についての評価に関する。

【背景技術】

【0 0 0 2】

50

タンパク質の輸送および分泌は、細胞機能のために必須である。タンパク質の輸送は、タンパク質のアミノ末端のシグナルペプチドによって媒介されて、輸送若しくは分泌される。シグナルペプチドは、約10から20の疎水性アミノ酸群からなり、新生タンパク質をリボソームから小胞体（ER）のような特定の膜で囲まれた区画まで導く。ERを標的とするタンパク質には、分泌経路を進むものや、ER、ゴルジ体若しくはリソソームなど、分泌細胞器官に残るものがある。分泌経路を進むタンパク質は、細胞外空間へ分泌されるか、または原形質膜内に残る。原形質膜内に保持されるタンパク質は1つ以上の膜貫通ドメインを有し、その各々が約20の疎水性アミノ酸残基からなる。分泌タンパク質は、通常、不活性の前駆体として合成され、分泌経路を移動する際に翻訳後プロセッシングイベントによって活性化される。そのようなイベントの例として、グリコシル化、タンパク質分解、およびシグナルペプチダーゼによるシグナルペプチドの除去が挙げられる。タンパク質の運搬の間に起こり得るその他のイベントの例として、新生タンパク質のシャペロン依存アンフォールディングおよびフォールディング、および受容体若しくは細孔複合体とタンパク質の相互作用が挙げられる。アミノ末端シグナルペプチドを有する分泌タンパク質の例は以下に述べられるが、これらには細胞間シグナル伝達において重要な役割を有するタンパク質を含む。そのようなタンパク質の例として、膜貫通受容体および細胞表面マーカー、細胞外基質分子、サイトカイン、ホルモン、成長因子および分化因子、酵素、ニューロペプチド、血管介在物質（vasomediators）がある（Alberts, B.他（1994）Molecular Biology of The Cell, Garland Publishing, New York, NY, 557-560、582-592ページの概説を参照）。

10

20

【0003】

細胞表面マーカーの例として、免疫系の白血球細胞上で同定される細胞表面抗原がある。これらの抗原を同定するには、系統的な、モノクローナル抗体（mAb）ベースの「ショットガン（shot gun）」技術が利用される。これらの技術によって、数百種のmAbが、未知の細胞表面白血球抗原類に対して産生されている。それらの抗原は「分化のクラスター群」へとグループ分けされている。分類は、多様な分化白血球細胞型と未分化白血球細胞型とで共通の、免疫細胞化学的な局在パターンに基づく。或るクラスター内の抗原群は或る単一の細胞表面タンパク質を同定すると仮定され、「分化クラスター」すなわち「CD（cluster of differentiation）」名が指定されている。CD抗原類が同定するタンパク質類をコードするいくつかの遺伝子が、標準的な分子生物学技術によってクローン化され、確認されている。CD抗原類の特徴は、膜貫通タンパク質であることと、細胞表面タンパク質であって脂肪酸含有糖脂質（例えばグリコシルホスファチジルイノシトール（GPI））への共有結合性の付着を介して原形質膜に固着されていることである（Barclay, A. N. 他（1995）The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, 17-20ページに概説されている）。

30

40

【0004】

基質タンパク質（MP）は、膜貫通タンパク質および細胞外タンパク質であり、組織の形成、成長、再形成、および維持において機能し、炎症反応の重要な介在物質および制御因子として働く。MPの発現およびバランスは、先天的疾患、後成的疾患、若しくは感染性疾患の結果として生じる生化学的変化によって乱されることがある。また、MPは、免疫応答における白血球の遊走、増殖、分化、および活性化に影響を与える。MPは、1つ以上のドメインの存在によってしばしば特徴づけられ、それらのドメインの内にはコラーゲン様ドメイン、EGF様ドメイン、免疫グロブリン様ドメイン、およびフィブロネクチン様ドメインが含まれ得る。加えて、MPは重度にグリコシル化されることがある。また、接着相互作用の役割を果たすことのあるアルギニン - グリシン - アスパラギン酸（RGD）トリペプチドモチーフを含むことがある。MPの例としては、細胞外タンパク質として、フィブロネクチン、コラーゲン、ガレクチン（galectin）、ピトロネクチンおよびそのタンパク質分解誘導体ソマトメジンBがあり、また細胞接着受容体として、細胞接着分子（CAM）、カドヘリン、およびインテグリンがある（Ayad, S. 他（1994）The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, 2-16ページ、Ruoslahti, E.（1997）Kidney Int

50

. 51: 1413-1417、Sjaastad, M. D.および Nelson, W. J. (1997) *BioEssays* 19:47-55に概説されている)。

【0005】

ムチン類は高度にグリコシル化された糖タンパク質であり、粘液ゲルの主要な構成成分である。ムチン類の生理機能は、細胞保護、機械的保護、分泌物の粘性の保持、および細胞認識である。MUC6はヒトの胃ムチンであり、胆嚢、膵臓、精液小胞、および女性の生殖管にも見つけられる(Toribara, N.W. 他 (1997) *J. Biol. Chem.* 272:16398-16403)。MUC6の遺伝子はヒトの11番染色体にマップされている(Toribara, N.W., 他 (1993) *J. Biol. Chem.* 268:5879-5885)。ヘモムチンは新規のショウジョウバエの表面ムチンであり、抗菌エフェクター分子の誘導に関与している可能性がある(Theopold, U., 他 (1996) *J. Biol. Chem.* 271:12708-12715)。

【0006】

タフテリン類(tuftelin)は、これまでに同定された、異なる4つのエナメル基質タンパク質の1つである。他の3つの既知のエナメル基質タンパク質は、アメロゲニン、エナメルリンおよびアメロプラスチンである。これらの成分タンパク質群からのエナメル細胞外基質のアセンブリは、無機質置換を受けるのに適格な基質の産生において重要であると考えられている(Paine, C.T. 他 (1998) *Connect Tissue Res.* 38:257-267)。タフテリンmRNAは、非ミネラル化歯原性腫瘍であるヒトエナメル上皮腫において発現されることが発見されている(Deutsch D. 他 (1998) *Connect. Tissue Res.* 39:177-184)。

【0007】

オルファクトメジン(olfactomedin)関連タンパク質群は、保存されたC末端モチーフを有する分泌性糖タンパク質たる細胞外基質である。これらは、多様な組織に、それも線虫からヒトにいたる広範囲の種にわたって発現される。オルファクトメジン関連タンパク質群には、ヒトにおいて少なくとも5つのファミリーメンバーを持つ遺伝子ファミリーを有する。この5つのうちの1つであるTIGR/ミオシリン(myocilin)タンパク質は目に発現され、緑内障の病原と関連している(Kulkarni, N.H. 他 (2000) *Genet. Res.* 76:41-50)。Yokoyama 他 (1996)による研究では、AMYと呼ぶ135アミノ酸のタンパク質は、神経芽細胞腫細胞系cDNAライブラリにおけるラットの神経細胞のオルファクトメジン関連ER局在性タンパク質に96%の配列同一性を有することが発見され、これはAMYの神経組織における役割が重要であることを示している(Yokoyama, M. 他 (1996) *DNA Res.* 3:311-320)。ラット脳

【0008】

Mac-2結合タンパク質は90キロダルトンの血清タンパク質(90K)であり、ヒト乳癌細胞系SK-BR-3およびヒト母乳の両方から単離された別の分泌性糖タンパク質である。これは、ヒトマクロファージ関連レクチンであるMac-2に特異的に結合する。構造的には、成熟したタンパク質は長さが567アミノ酸で、18アミノ酸リーダーが、その前に付いている。16のシステイン群と、7つの潜在的N連結グリコシル化部位とがある。最初の106アミノ酸はマクロファージスカベンジャー受容体高システインドメインによって定義される古いタンパク質スーパーファミリーに極めて類似するドメインを示す(Kohts, K. 他 (1993) *J. Biol. Chem.* 268:14245-14249)90Kはエイズ患者の亜集団群の血清中で上昇し、原発腫瘍サンプルと腫瘍細胞株とにおいて多様なレベルで発現される。Ullrich 他 (1994)は90Kが宿主防御系を刺激し、インターロイキン2の分泌を誘発し得ることを実証した。この免疫刺激は、発癌性形質転換、ウイルス感染または病原性侵襲の結果であると提起されている(Ullrich, A., 他 (1994) *J. Biol. Chem.* 269:18401-18407)。

【0009】

セマフォリン(semaphorin)類は軸索誘導分子の1大グループで少なくとも30種のメンバーがあり、脊椎動物、無脊椎動物、更には数種のウイルスでも発見されている。全てのセ

マフォリンは、長さが約500アミノ酸のセマ(sema)ドメインを有する。セマフォリン受容体のニューロピリン(neuropilin)は*in vitro*で神経突起の進展を促進することが示されている。ニューロピリンの細胞外領域は、3種のドメインからなる。CUB、ジスコイジン(discoidin)、およびMAMドメインである。ニューロピリンのCUBおよびMAMモチーフはタンパク質間相互作用に役割を果たすことが示唆されており、さらにセマドメインおよびC末端ドメインを通じてセマフォリンの結合に関与していることが示唆されている(Raper, J.A. (2000) *Curr. Opin. Neurobiol.* 10:8894の概説を参照)。プレキシン(plexin)類は神経細胞表面分子であり、カルシウムイオンの存在下で同種親和性結合メカニズムにより細胞接着を媒介する。プレキシン類は、特定の感覚系の受容体や神経細胞において発現することがわかっている(Ohta, K. 他(1995) *Cell* 14:1189-1199)。いくつかのプレキシンが、発生段階の神経系において運動神経軸索や中枢神経軸索の誘導を制御するよう機能することを示す証拠もある。プレキシンは、それ自体完全なセマフォリンドメインを有しており、古典的セマフォリンとセマフォリンの結合パートナーの両方の祖先であり得る(Winberg, M.L. 他(1998) *Cell* 95:903-916)。

10

【0010】

ヒト妊娠特異的 1糖タンパク質(PSG)は、分子量が72KDa、64KDa、62KDaおよび54KDaの、密接に関連する糖タンパク質類の1ファミリーである。癌胎児抗原と共に、PSGは免疫グロブリンスーパーファミリー内にサブファミリーを構成する(Plouzek C.A及びChou J.Y. *Endocrinology* 129:950958)。PSGの異なる亜集団が、ヒト胎盤の栄養胚葉および羊膜や漿膜によって産生されることが発見されている(Plouzek C.A. 他(1993) *Placenta* 14:277-285)。

20

【0011】

自己分泌モチーフ因子(AMF)は運動サイトカイン調節性腫瘍細胞遊走の一つであり、したがって、それと結合するシグナル伝達経路の識別は決定的に重要である。自己分泌型運動促進因子受容体(AMFR)の発現は、胸腺腫における腫瘍の進行に関連することがわかっている(Ohta Y. 他(2000) *Int. J. Oncol.* 17:259264)。AMFRは、分子量78KDaの細胞表面糖タンパク質である。

【0012】

ホルモンは、血液循環により運ばれ、標的細胞の表面のまたは内部の特異的受容体に結合する分泌分子である。ホルモンは多様な生化学的組成や作用機構を持っているが、2つのカテゴリーに分類し得る。第1のカテゴリーは小さい親油性ホルモンで、標的細胞の原形質膜を通過して拡散し、サイトゾル内受容体または核内受容体に結合し、複合体を形成して遺伝子発現を改変する。これらの分子としては、レチノイン酸やチロキシンが、更には、プロゲステロン、エストロゲン、テストステロン、コルチゾール、アルドステロンなどの、コレステロール由来ステロイドホルモンがある。第2のカテゴリーに含まれる親水性ホルモンは、原形質膜の内側へ信号を伝達する細胞表面受容体に結合することで機能する。そのようなホルモンの例としては、アミノ酸の誘導體としてカテコールアミン類(エピネフリン、ノルエピネフリン)およびヒスタミンが、ペプチドホルモンとしてグルカゴン、インスリン、ガストリン、セクレチン、コレシストキニン、副腎皮質刺激ホルモン、卵胞刺激ホルモン、黄体形成ホルモン、甲状腺刺激ホルモン、およびバソプレッシンがある(Lodish 他(1995) *Molecular Cell Biology*, Scientific American Books Inc., New York, NY, 856-864ページ等を参照)。

30

40

【0013】

プロオピオメラノコルチン(POMC)は、脳下垂体前葉によって合成されるホルモンであるコルチコトロピン(ACTH)の前駆体ペプチドで、副腎皮質の刺激において作用する。POMCはまた、リポトロピンホルモン(-LPH)の前駆体ペプチドでもある。各ホルモンは、固有の生物活性を持つより小さいペプチド群を含む。メラニン細胞刺激ホルモン(-MSH)と副腎皮質刺激ホルモン様中葉ペプチド(CLIP)はACTHから形成され、-リポトロピン(-LPH)と-エンドルフィン(-LPH)のペプチド成分であるが、-MSHは-LPH中に含まれる。POMCのエキソン2および3における遺伝子突然変異に起因する、ACTH欠乏による副腎

50

不全は、早期発症肥満症、副腎不全および赤毛色素沈着を特徴とする内分泌障害を定義する (Chretien, M.他 (1979) *Canad. J. Biochem.* 57:1111-1121、Krude, H. 他 (1998) *Nature Genet.* 19:155-157、Online Mendelian Inheritance in Man (OMIM) 176830)。Johns Hopkins University, Baltimore, MD. MIM Number:176830: 2000年8月1日. World Wide Web URL: www.ncbi.nlm.nih.gov/omim/。

【0014】

成長因子および分化因子は、細胞間伝達の中で機能する分泌タンパク質である。いくつかの因子の活動には、オリゴマー形成が、または膜タンパク質との会合が必要である。それらの因子とそれらの受容体間の複雑な相互作用は、細胞分裂、細胞分化、細胞シグナル伝達および細胞運動性を刺激または抑制する、細胞内信号伝達経路の引き金となる。殆どの成長および分化因子は、その局所環境にある細胞に作用する (パラ分泌シグナル伝達)。成長因子および分化因子には3つの主要なクラスがある。第1のクラスには、上皮成長因子、線維芽細胞成長因子、トランスフォーミング成長因子、インシュリン様成長因子、および血小板由来の成長因子など、大きなポリペプチド成長因子を含む。第2のクラスには、コロニー刺激因子 (CSF) のような造血性成長因子を含む。造血性成長因子類は、Bリンパ球、Tリンパ球、赤血球、血小板、好酸球、好塩基球、好中球、マクロファージ、およびそれらの幹細胞前駆体など、血液細胞の増殖および分化を刺激する。第三のクラスは、小さなペプチド因子として、ポンペシン、バソプレッシン、オキシトシン、エンドセリン、トランスフェリン、アンジオテンシンII、血管作用性小腸ペプチド、及び増殖以外の細胞機能を制御するためにホルモンとして働くブラジキニンが含まれる。

10

20

【0015】

成長因子および分化因子は、*in vitro*での細胞の腫瘍性転換、および*in vivo*での腫瘍進行において、幾つかの重大な役割を果たす。腫瘍細胞による、成長因子の不適正な発現は、腫瘍の血管新生および転移に寄与し得る。造血時の成長因子の調節異常によっては、貧血、白血病、およびリンパ腫が生じ得る。インターフェロンなど幾つかの成長因子は、*in vitro*および*in vivo*の双方で、腫瘍細胞群に対し細胞毒性を持つ。更に、幾つかの成長因子および成長因子受容体は、構造的にも機能的にも腫瘍性タンパク質類と関連する。加えて、成長因子は癌原遺伝子群および腫瘍抑制遺伝子群の双方の転写調節に影響を与える (Pimentel, E. (1994) *Handbook of Growth Factors*, CRC Press, Ann Arbor, MI, 1-9ページの概説を参照)。

30

【0016】

最初にショウジョウバエにおいて同定されたSlitタンパク質は、中枢神経系正中線形成およびおそらく神経組織の組織発生と軸索の誘導において重要である。Itoh 他は、Slit遺伝子の哺乳類相同体 (ヒト Slit-1、Slit-2、Slit-3およびラットSlit-1) を同定した。コードされたタンパク質は、共に保存されたタンパク質間相互作用ドメインであるEFG様モチーフおよびロイシンリッチリピートを含む推定分泌性タンパク質である。Slit-1 mRNA、Slit-2 mRNAおよびSlit-3 mRNAは、それぞれ脳、脊髄、甲状腺に発現される (Itoh, A. 他 (1998) *Brain Res. Mol. Brain Res.* 62:175-186)。タンパク質のSlitファミリーは神経組織内のグリピカン (glypican) -1の機能性リガンドであることが示されており、このことは、それらの相互作用は中枢神経系の組織発生時の特定の段階において重要であることを意味する (Liang, Y. 他 (1999) *J. Biol. Chem.* 274:17885-17892)。

40

【0017】

神経ペプチドおよび血管介在物質 (NP/VM) は、内因性シグナル伝達分子の1大ファミリーを構成する。NP/VMファミリーに含まれる分子は、神経ペプチドおよび神経ペプチドホルモンとして、ポンペシン、神経ペプチドY、ニューロテンシン、ニューロメディンN、メラノコルチン類、オピオイド類、ガラニン、ソマトスタチン、タキキニン類、ウロテンシンIIおよび平滑筋刺激に関する関連ペプチド類、バソプレッシン、および血管作用性小腸ペプチドがある。また、循環系に運ばれるシグナル伝達分子として、アンジオテンシン、補体、カルシトニン、エンドセリン類、ホルミルメチオニルペプチド類、グルカゴン、コレシストキニン、およびガストリンがある。NP/VMは直接に信号を伝達し得る他、別の

50

神経伝達物質およびホルモンの活性若しくは放出をモジュレートしたり、またカスケードで触媒酵素として機能することができる。NP/VMの効果は、ごく短時間のものから長期間持続するものまで幅広い (Martin, C.R. 他 (1985) *Endocrine Physiology*, Oxford University Press, New York, NY, 57-62ページの概説を参照)。

【0018】

NP/VMは、多くの神経障害および心血管障害に関与する。例えば、神経ペプチドYは、高血圧症、鬱血心不全、情動障害、食欲調節に関与する。ソマトスタチンは、下垂体前葉での成長ホルモンおよびプロラクチンの分泌を阻害し、また、腸、膵臓腺房細胞、および膵臓細胞内の分泌も阻害する。ソマトスタチンレベルの低下は、アルツハイマー病およびパーキンソン病で報告された。バソプレッシンは、腎臓内で水分吸収およびナトリウム吸収を増加させ、また高濃度では、血管平滑筋の収縮、血小板活性化、および肝臓内でのグリコーゲン分解を刺激する。バソプレッシンおよびその類似体は、尿崩症の治療のため臨床的に用いられる。エンドセリンおよびアンジオテンシンは高血圧症に関与し、アンジオテンシンの血漿レベルを減少させるカプトプリルのような薬剤は、血圧を下げるのに用いられる (Watson, S. および S. Arkininstall, (1994) *The G-protein Linked Receptor Facts Book*, Academic Press, San Diego CA, 194, 252, 284, 55, および111の各ページ)。

10

【0019】

神経ペプチドはまた、侵害受容 (痛覚) にも役割を幾つか持つことが示されている。血管作用性小腸ペプチドは、慢性神経障害痛において重要な役割を果たすようである。オピオイド受容体様1受容体に対する内因性リガンドであるノシセプチンは、主に抗侵害受容作用を有すると考えられ、持続性の痛み若しくは慢性痛の、様々な動物モデルにおいて鎮痛特性を有している事が示された (Dickinson, T. および Fleetwood-Walker, S. M. (1998) *Trends Pharmacol. Sci.* 19:346-348)。

20

【0020】

シグナルペプチドを有する他のタンパク質として、酵素活性を持つ分泌タンパク質類がある。酵素活性としては、例えば、オキシドレダクターゼ/デヒドロゲナーゼ活性、トランスフェラーゼ活性、加水分解酵素活性、リアーゼ活性、異性化酵素活性、若しくはリガーゼ活性がある。例えば、基質メタロプロテイナーゼは細胞外基質を分解する分泌性加水分解酵素であり、したがって、腫瘍転移、組織形態形成、および関節炎において重要な役割を果たす (Reponen, P. 他 (1995) *Dev. Dyn.* 202 :388-396、Firestein, G. S. (1992) *Curr. Opin. Rheumatol.* 4 : 348-354、Ray, J. M. 及び Stetler-Stevenson, W. G. (1994) *Eur. Respir. J.* 7 : 2062-2072、Mignatti, P. 及び Rifkin, D. B. (1993) *Physiol. Rev.* 73: 161-195)。別の例は、脂質合成にまたはエネルギー生成に用いる酢酸を活性化させるアセチルCoA合成酵素である (Luong, A. 他 (2000) *J. Biol. Chem.* 275:26458-26466)。アセチルCoA合成酵素が活性化されると、酢酸とCoAからアセチルCoAが形成される。アセチルCoA合成酵素は、AMP結合ドメインシグネチャとして同定される或る配列類似性領域を共有する。アセチルCoA合成酵素は高血圧と関連性があることが示されている (H. Toh (1991) *Protein Seq. Data Anal.* 4:111-117、Iwai, N. 他 (1994) *Hypertension* 23:375-380)。

30

40

【0021】

多くの異性化酵素は、タンパク質の折りたたみ、光伝達、および種々のタンパク質同化経路や異化経路における諸段階を触媒する。或るクラスの異性化酵素は、ペプチジルプロリルシス-トランス異性化酵素 (PPIアーゼ) として知られる。PPIアーゼ類は、タンパク質内の特定のプロリンイミド結合の、シスからトランスへの異性化を触媒する。PPIアーゼの2つのファミリーが、FK506結合タンパク質 (FKBP) とシクロフィリン (CyP) である。FKBP類が、共に強力な免疫抑制薬であるFK506とラパマイシンとに結合することにより、T細胞内の複数のシグナル伝達経路が抑制される。特に、FKBPのPPIアーゼ活性は、FK506あるいはラパマイシンの結合によって阻害される。算出した分子量にしたがって名付けたFKBPファミリーの5つのメンバー (FKBP12、FKBP13、FKBP25、FKBP52およびFKBP65) が

50

あり、それらは、様々なタンパク質複合体に会合する、細胞の様々な領域に局在している (Coss, M.他 (1995) J. Biol. Chem. 270:2933629341、Schreiber, S.L. (1991) Science 251:283-287)。

【0022】

CyPのペプチジルプロリル異性化酵素活性は、T細胞活性化につながるシグナル伝達経路の一部である可能性がある。CyP異性化酵素活性はタンパク質折りたたみおよびタンパク質輸送に関連しており、またタンパク質複合体のアセンブリまたは分解、およびタンパク質活性の調節にも関与している可能性がある。例えば、ショウジョウバエにおいてCyP NinaAはロドプシンの正確な局在化に必要であり、哺乳類のCyP (Cyp40) はステロイド受容体に結合するHsp90/Hsc70複合体の一部である。哺乳類CypAはヒト免疫不全ウイルス1 (HIV-1) からのgagタンパク質と結合するが、この相互作用はシクロスポリンによって抑制し得ることが示されている。シクロスポリンは強力な抗HIV-1活性を有するので、CypAはHIV-1の複製において重要な機能を果たし得る。最後に、Cyp40は転写因子c-Mybに結合し、これを不活化することが示されているが、この作用はシクロスポリンによって逆転される。この作用は、転写、形質転換および分化の調節でのCyp類の関与を意味する (Bergsma, D.J. 他 (1991) J. Biol. Chem. 266:23204 -23214、Hunter, T. (1998) Cell 92: 141-143、Levenson, J.D.およびNess, S.A. (1998) Mol. Cell. 1:203-211)。

【0023】

プロリンリッチな γ -カルボキシグルタミン酸 (Gla) タンパク質 (PRGP) 類は、ビタミンK依存性1回貫通型膜内在性タンパク質ファミリーのメンバーである。PRGPタンパク質はGlaリッチなほぼ45アミノ酸の細胞外アミノ末端ドメインを特徴とする。細胞内カルボキシル末端領域には配列PPXYの1つまたは2つのコピーがあり、PPXYは、シグナル伝達、細胞周期進行、およびタンパク質代謝回転のような、細胞の多様な機能に関与する種々のタンパク質に存在するモチーフである (Kulman, J.D. 他, (2001) Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375)。Glaを形成するグルタミン酸残基の翻訳後修飾過程は、ビタミンK依存性のカルボキシル化である。Glaを含むタンパク質としては、血液凝固に関与する血漿タンパク質類がある。これらのタンパク質とは、プロトロンビン、タンパク質C、SおよびZ、並びに血液凝固第VII因子、第IX因子および第X因子である。オステオカルシン (骨-Glaタンパク質、BGP) およびマトリックスGlaタンパク質 (MGP) もまた、Glaを含む (Friedman, P.A.およびC.T. Przywiecki (1987) Int. J. Biochem. 19:1-7、C. Vermeer (1990) Biochem. J. 266:625-636)。

【0024】

ショウジョウバエ遺伝子crossveinless 2の特徴は、まず、推定上のシグナル配列または膜貫通配列を有することである。また、部分的なVon Willebrand (フォンウィルブランド) 因子Dドメインをも持つが、このドメインは分子内結合と分子間結合の形成を調節することが知られるドメイン群に類似している。また、5ヶ所のシステインリッチドメインをも持つが、これらはBMP (bone morphogenetic proteins、骨形成因子) 様リガンド群を結合するとされている。これらの特徴は、crossveinless 2 が細胞外で作用するか、または分泌経路において、直接にリガンドのシグナル伝達を可能にするように作用し、したがって、翅脈特異化 (vein specification) において或る役割を果たすとされているBMP様シグナル伝達経路に関与することを示唆する (Conley, C.A. 他 (2000) Development 127:3947-3959)。脊椎動物およびショウジョウバエの胚における背腹軸形成 (dorsal-ventral patterning) には、位置情報勾配 (a positional informational gradient) を発生させるための細胞外タンパク質群の保存されたシステムを必要とする。

【0025】

シグナルペプチドの存在はまた、前立腺癌と関連するタンパク質内でも見出されている。前立腺腫瘍におけるアンドロゲン非依存性細胞の発生は、疾患の進行を決定する方法として有用である。ほとんどの腫瘍と同様、前立腺癌は多段階の進行を通して発達して、最終的に侵略的な腫瘍の表現型をもたらす。腫瘍進行の初めの段階には、正常な内腔そして/あるいは基底上皮細胞の過剰増殖が関係する。アンドロゲン応答性細胞は過形成し、初期

10

20

30

40

50

段階の腫瘍に発展する。初期段階の腫瘍はしばしばアンドロゲンに対して感受性があり、アンドロゲン除去療法に反応するが、アンドロゲン非依存性細胞の集団が過形成の集団から発展する。これらの細胞は、浸襲性となり、また骨、脳、または肺に転移する可能性がある前立腺腫瘍のより進んだ形態を表す。

【0026】

発現プロファイル作成

アレイ技術は、単一の多型遺伝子の発現や、多数の関連遺伝子または無関係の遺伝子の発現プロファイルを探求する、簡単な方法を提供し得る。単一遺伝子の発現を試験するときには、アレイを用いて、或る特定遺伝子又はその変異体の発現を検出する。発現プロファイルを試験するときには、アレイは次のような遺伝子を同定するプラットフォームを提供する。即ちどの遺伝子が組織特異的か、毒性アッセイにおいてテストされる物質に影響されるか、シグナル伝達カスケードの一部であるか、ハウスキーピング機能を実行するか、又は、特定の遺伝的素因や、条件、疾患、又は障害に、特異的に関連する遺伝子であるかの同定である。

10

【0027】

泡沫細胞形成と分泌タンパク質

アテローム性動脈硬化症、それに関連する冠状動脈疾患、脳卒中は、先進国におけるもっとも一般的な死の原因を代表する。ある種の主要な危険因子は同定されてきたが、原因を解明して、この複雑な疾患を治療する全分子性質決定は達成されていない。アテローム性血管硬化病変部の成長と退化の分子の性質決定には、成長、安定、溶解、破裂、そして最も致命的な閉塞性血管血栓の誘発を含む病変部の特徴に寄与する遺伝子の同定を必要とする。

20

【0028】

アテローム性動脈硬化症発生の初期段階は、「脂肪線条」の形成である。コレステロールに富んだ低密度リポタンパク質(LDL)等のリポタンパク質は、血管内膜の細胞外空間に蓄積して、修飾を受ける。LDLの酸化は、循環抗酸化体防御の効果が少ない内皮下空間で最も激しい。LDL酸化の程度は、LDLと標的細胞との相互作用に影響する。「最小限に酸化した」LDL(MM-LDL)は、同定された酸化LDL(Ox-LDL)またはスカベンジャー受容体タイプAとB、CD36、CD68/macrosialin、LOX-1等の「スカベンジャー」受容体にはなく、LDL受容体に結合することができる(Navab 他. (1994) *Arterioscler Thromb Vasc Biol* 16:831-842、Kodama 他. (1990) *Nature* 343:531-535、Acton 他. (1994) *J Biol Chem* 269:21003-21009、Endemann 他. (1993) *J Biol Chem* 268:11811-11816、Ramprasad 他 (1996) *Proc Natl Acad Sci* 92:14833-14838、Kataoka 他 (1999) *Circulation* 99:3110-3117)。MM-LDLは単球の接着と浸透を増加し、内皮細胞による単球走化タンパク質1(MCP-1)の放出を刺激し、またマクロファージにおけるスカベンジャー受容体A(SRA)とCD36の発現を誘発することができる(Cushing 他. (1990) *Proc Natl Acad Sci* 87:5134-5138、Yoshida 他. (1998) *Arterioscler Thromb Vasc Biol* 18:794-802、Steinberg (1997) *J Biol Chem* 272:20963-20966)。SRA及び他のスカベンジャー受容体は、Ox-LDLに結合し、リポタンパク質粒子の取込みを強化することができる。

30

【0029】

単核食細胞は脈管内膜に入り、マクロファージに分化して、Ox-LDL等の修飾した脂質を取り込む。大抵のタイプの細胞では、コレステロール含有はLDL受容体と生合成酵素のフィードバック調節により厳密に制御されている(Brown 及び Goldstein (1986) *Science* 232:34-47)。しかしマクロファージでは、スカベンジャー受容体の追加によってコレステロールの制御不能な取込み(Brown及びGoldstein (1983) *Annu Rev Biochem* 52:223-261)、及び「泡沫細胞」表現型を作製する多数の細胞内脂質飛沫の蓄積へとつながる。コレステロール充満の死んだマクロファージは、初期「脂肪線条」プラークの塊および典型的な罹患した動脈の「進行した」損傷のはとんどの原因となる。多数の研究において、アテローム硬化性血管壁プラークの成長と破裂に寄与する多様な泡沫細胞応答について記載されている。これらの反応には、増殖と隣接細胞の接着を促す複数の成長因子とサイトカイン、

40

50

循環する単球を成長しているプラークにさらに引き寄せるケモカイン、細胞外マトリックスの再構築を引き起こすタンパク質、血栓症誘発の可能性のある組織因子の作製を含む(Ross (1993) Nature 362:801-809; Quin 他. (1987) Proc Natl Acad Sci 84:2995-2998)。よってアテローム斑形成のほとんどの段階において大量に発生するコレステロールを蓄積したマクロファージは、アテローム動脈硬化性(atherosclerotic)過程の開始、最終的なプラーク破裂と閉塞性血栓に寄与する。

【0030】

Ox-LDL取り込み中、マクロファージはサイトカインと成長因子を生成する。それらはさらに平滑筋細胞増殖と細胞外マトリックスの生成等の、アテローム発生を調節する細胞事象を誘発する。さらに、これらのマクロファージは、発現誘導型一酸化窒素シンターゼを含む炎症に關与する遺伝子を活性化する可能性がある。従って、泡沫細胞形成中に示差的に発現される遺伝子群がアテローム硬化性過程のマーカーであると期待することは道理的に可能である。

10

【0031】

ステロイド分子と分泌タンパク質

遺伝子発現プロファイル作成の潜在的応用は、特に治療の可能性のある化合物に対する毒性反応と治療薬剤の代謝反応の測定に關連する。ステロイドで治療する疾患とステロイド治療への代謝反応によって引き起こされる病気には、腺腫症、胆汁鬱滞、肝硬変、血管腫、シェーンライン・ヘノッフ紫斑病、肝炎、肝細胞癌、転移性癌、特発性血小板減少性紫斑病、ポルフィリン症、サルコイドーシス、ウィルソン病が含まれる。反応は、ミフェプリストン、プロゲステロン、ベクロメタゾン、メドロキシプロゲステロン、ブデソニド、プレドニゾン、デキサメサゾン、ベタメタゾン、あるいはダナゾール等のステロイド化合物に曝露したあるいは治療した患者の組織で発現したレベルおよび配列の両者を、正常な治療されていない組織のレベルおよび配列と比較することによって測定され得る。

20

【0032】

ステロイドは、コレステロール、胆汁酸、ビタミンD、ホルモン等の脂質可溶性分子の1クラスであり、シクロペンタヒドロフェナントレン(cyclopentanoperhydrophenanthrene)に基づく共通な4リング構造を共有し、広範囲な機能を実施する。例えば、コレステロールは膜流動性を制御する細胞膜の1成分である。それは脂質を可溶して、消化中に小腸で吸収を促進する胆汁酸の前駆体でもある。ビタミンDは小腸におけるカルシウムの吸収を調節し、原形質におけるカルシウムの濃度を制御する。副腎皮質、卵巣、精巣によって生成されるステロイドホルモンには、グルココルチコイド、電解質コルチコイド、アンドロゲン、エストロゲンが含まれる。それらは核内の特定の遺伝子の転写を調節する細胞内受容体に結合することにより、多様な生体プロセスを制御する。例えばグルココルチコイドは、肝臓における糖新生の調節により血糖濃度、脂肪組織のリポリシスの促進により脂肪酸の血中濃度を上昇させ、中枢神経内のカテコールアミンへの感度を調節し、炎症を減少する。主要な電解質コルチコイドであるアルドステロンは副腎皮質により生成され、ナトリウムイオン再吸収を促進するために腎臓の遠位尿細管の細胞に作用する。精巣内のライディヒの間細胞により生成されるアンドロゲンには、思春期の変化すなわち精子の生成と第二次性徴の維持を誘発する男性ホルモンテストステロン等がある。女性ホルモンであるエストロゲンとプロゲステロンは、卵巣そしてさらに妊娠中に胎盤と胎児の副腎皮質により生成される。エストロゲンは、女性生殖プロセスと第二次性徴を調節する。プロゲステロンは、月経周期と妊娠中の子宮内膜における変化を調節する。

30

40

【0033】

ステロイドホルモンは、避妊法、また怪我や関節炎、喘息、自己免疫疾患等の疾患での抗炎症治療において広範囲に利用されている。天然のプロゲステチンであるプロゲステロンは、無月経、異常子宮出血を治療するために、または避妊薬として主に使用される。内因性的プロゲステロンは、受精卵着床と妊娠の維持の準備でエストロゲンに感作された子宮の子宮内膜における分泌作用に關与する。それは黄体形成ホルモン(LH)に応答して黄体から分泌される。外因性的プロゲステチンの主要な避妊効果には、LHのmidcycle surgeの抑制が

50

関係する。細胞レベルではプロゲステロンは自由に拡散して標的細胞に入っていく、プロゲステロン受容体に結合する。標的細胞には、女性の生殖管、乳腺、視床下部、下垂体がある。いったん受容体に結合すると、プロゲステロンは視床下部からのゴナドトロピン放出ホルモンの放出の頻度を遅くし、前排卵LHサージを鈍くする。それにより卵胞成熟と排卵を防ぐ。プロゲステロンには、最小限のエストロゲンとアンドロゲンの作用がある。プロゲステロンは、肝臓でプレグナジオールに代謝され、グルクロン酸と抱合する。

【0034】

6-メチル-17-ヒドロキシプロゲステロンとしても知られるメドロキシプロゲステロン(MAH)は、プロゲステロンよりも15倍も大きな薬理学的作用を有する合成プロゲステロンである。MAHは腎臓癌と子宮内膜癌、無月経、異常子宮出血、ホルモンの失調と関連する子宮内膜症の治療のために使用される。MAHは呼吸中枢に刺激的作用を有し、睡眠時無呼吸、慢性閉塞性肺疾患、あるいは炭酸過剰症によって引き起こされる血液の低酸素化の場合に使用されてきた。

10

【0035】

RU-486としても知られるミフェプリストンは、プロゲステロンの受容体を阻害する抗プロゲステロン剤である。それは、妊娠を維持するのに必要であるプロゲステロンの影響を打ち消す。ミフェプリストンは、妊娠初期に投与し、続いてプロスタグランジン、ミソプロストールで治療すると自然流産を誘発する。さらに研究によると、避妊手段を用いない性交の後、5日以内に投与するとミフェプリストンは非常に低量の投与で、性交後の避妊として非常に効果的になり得る。従って避妊失敗や性的暴行の場合に女性に「モーニングアフターピル」を提供する。ミフェプリストンにはまた、腫瘍がプロゲステロン依存性である場合の乳癌と卵巣癌の治療において潜在的な使用可能性がある。それは脳髄膜腫のステロイド依存性成長に干渉し、また子宮内膜組織のエストロゲン依存性成長を阻止して子宮内膜症の治療においても有用であろう。子宮類線維腫とクッシング症候群の治療においても有用であろう。ミフェプリストンは糖質コルチコイド受容体に結合し、コルチゾール結合に干渉する。ミフェプリストンは抗糖質コルチコイドとしても作用し、AIDS、神経性無食欲症、潰瘍、糖尿病、パーキンソン病、多発性硬化症、アルツハイマー病等の、コルチゾールレベルが上昇する症状を治療するのに効果がある。

20

【0036】

ダナゾールは、エチニルテストステロン由来の合成ステロイドである。ダナゾールは、卵胞刺激ホルモンとLHの下垂体合成を低下させることによりエストロゲン生成を間接的に減少する。ダナゾールはまた標的組織において性ホルモン受容体に結合する。従って同化作用、抗エストロゲン作用、弱いアンドロゲンの作用を示す。ダナゾールはいかなるプロゲステロン作用も有しておらず、コルチコトロピンの正常な下垂体放出あるいは副腎によるコルチゾールの放出を抑制しない。ダナゾールは、子宮内膜症の治療において、痛みを緩和し、子宮内膜細胞成長を阻害するために使用される。また乳房腺維嚢胞病と遺伝性血管浮腫を治療するためにも使用される。

30

【0037】

コルチコステロイドは炎症を緩和して、免疫応答を抑制するために使用される。炎症反応を仲介するサイトカインの調節により、好酸球、好塩基球、気道上皮性細胞機能を阻害する。炎症部位で白血球侵入を阻害し、炎症反応のメディエーターの機能において干渉し、さらに体液の免疫応答を抑制する。コルチコステロイドは、アレルギー、喘息、関節炎、皮膚病を治療するために使用される。ベクロメタゾンは、ステロイド依存性喘息の治療、アレルギーまたは非アレルギー(血管運動)鼻炎に付随する症状を緩和するため、あるいは外科的切除に続く再発性鼻ポリープを予防するために使用される合成グルココルチコイドである。鼻腔内ベクロメタゾンの抗炎症作用および血管収縮作用は、ヒドロコルチゾンによるものより5000倍強力である。ブデソニドは、アレルギー性鼻炎または喘息に付随する症状を制御するために使用されるコルチコステロイドである。ブデソニドは、全身作用が弱い強い局所抗炎症作用を有する。デキサメタゾンは、抗炎症組成物または免疫抑制組成物において使用される合成グルココルチコイドである。喘息の症状を予防するために

40

50

吸入薬においても使用される中枢神経系に到達する強力な能力のために、デキサメタゾン
は脳水腫を調節するために通常選択される治療である。デキサメタゾンは、ヒドロコルチ
ゾンより約20~30倍、そしてプレドニゾンより5~7倍強力である。プレドニゾンは、肝臓
で代謝されて、活性型であるプレドニゾロン（抗炎症特性を有するグルココルチコイド）
になる。プレドニゾンはヒドロコルチゾンより約4倍強力であり、そしてプレドニゾンの
作用持続時間はヒドロコルチゾンとデキサメタゾンの中間である。プレドニゾンは、同種
移植の拒絶反応、喘息、全身性紅斑性狼瘡、関節炎、潰瘍性大腸炎、その他の炎症症状を
治療するために使用される。ベタメタゾンは抗炎症作用と免疫抑制作用を有する合成グル
ココルチコイドであり、乾癬、水虫や白癬等の真菌感染症を治療するために使用される。

【0038】

コルチコステロイドの抗炎症作用は、リポコルチンと総称されるホスホリパーゼA₂抑制タン
パク質に関係すると考えられる。逆にリポコルチンは、前駆体分子アラキドン酸の放出
を阻害することによりプロスタグランジン、ロイコトリエン等炎症の強力な媒介物の生合
成を制御する。提案されている作用の機構には、減少したIgE合成、白血球上で増加した
-アドレナリン受容体の数、減少したアラキドン酸代謝が含まれる。慢性気管支喘息等
の即時アレルギー反応中、アレルゲンは肥満細胞の表面上のIgE抗体を架橋し、これらの
細胞の化学走化性物質の放出を誘発する。従って肥満細胞の流入と活性化が、喘息患者の
炎症と口腔粘膜の過剰刺激感受性に部分的に関与している。この炎症は、コルチコステロ
イドの投与によって遅れ得る。

【0039】

肝臓代謝とホルモン除去機構への影響は、薬剤の薬力学を理解するために重要である。ヒ
トC3A肝臓細胞株は、強力な接触阻害の成長に関して選択されたHepG2/C3（肝臓腫瘍を患
う15歳の男子から単離した肝臓癌細胞株）のクローン誘導体である。クローン集団の使用
は、細胞の再現性を強化する。C3A細胞は、培養中の主要なヒト肝細胞の多くの特徴を有
する。i)インシュリン受容体とインシュリン様成長因子II受容体の発現、ii)フェトプ
ロテインと比較した血清アルブミンの高率分泌、iii)アンモニアの尿素とグルタミンへの
転換、iv)芳香アミノ酸代謝、v)グルコースの無いまたインシュリンの無い培地での増殖
がそれらの特徴である。C3A細胞株は、成熟したヒト肝臓の *in vitro* モデルとして今や十
分に確立される(Mickelson 他 (1995) Hepatology 22:866-875、Nagendra 他 (1997) Am
J Physiol 272:G408-G416)。

【0040】

新規の分泌タンパク質およびそれらをコードするポリヌクレオチドの発見により、新規の
組成物を提供することで、当分野の要望に応えることができる。この新規の組成物は、肝
臓疾患、細胞増殖異常、自己免疫/炎症疾患、心血管障害、神経障害、および発生障害の
、診断・予防・治療において有用であり、また分泌タンパク質の核酸配列およびアミノ酸
配列の発現における外来性化合物の効果についての評価にも有用である。

【発明の開示】

【発明の効果】

【0041】

本発明は、総称して「SECP」、個別にはそれぞれ「SECP-1」、「SECP-2」、「SECP-3」、
「SECP-4」、「SECP-5」、「SECP-6」、「SECP-7」、「SECP-8」、「SECP-9」、「SECP-1
0」、「SECP-11」、「SECP-12」、「SECP-13」、「SECP-14」、「SECP-15」、「SECP-16
」、「SECP-17」、「SECP-18」、「SECP-19」、「SECP-20」、「SECP-21」、「SECP-22」
、「SECP-23」、「SECP-24」、「SECP-25」、「SECP-25」、「SECP-26」、「SECP-27」
、「SECP-28」、「SECP-29」、及び「SECP-30」と呼ぶ分泌タンパク質である精製されたポリ
ペプチドを提供する。或る実施態様において本発明は、(a)SEQ ID NO:1-30からなる群
から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-30からなる群から選択
したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド
、(c)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの生物
学的活性断片、または(d)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有する

10

20

30

40

50

ポリペプチドの免疫原性断片から成る群から選択した単離されたポリペプチドを提供する。一実施態様では、本発明はSEQ ID NO:1-30のアミノ酸配列を含む単離されたポリペプチドを提供する。

【0042】

また、本発明は(a)SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列と少なくとも90%同一である或る天然アミノ酸配列を有するポリペプチド、(c)SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片からなる群から選択した或るポリペプチドをコードする、単離されたポリヌクレオチドを提供する。一実施態様では、該ポリヌクレオチドは、SEQ ID NO:1-30からなる群から選択した或るポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO:31-60からなる群から選択される。

10

【0043】

本発明は更に、(a)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-30からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO:1-30を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択されたポリペプチドをコードするポリヌクレオチドと機能的に連結したプロモーター配列を含む組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の態様では、本発明は組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

20

【0044】

また、本発明は、(a)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-30からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択したポリペプチドを製造する方法を提供する。製造方法は、(a)組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、(b)そのように発現したポリペプチドを回収する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に連結したプロモーター配列を有する。

30

【0045】

本発明は更に、(a)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-30からなる群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d)SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択されたポリペプチドに特異結合するような単離された抗体を提供する。

40

【0046】

本発明は更に、(a)SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b)SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列と少なくとも90%が同一である天然のポリヌクレオチド配列を含むポリヌクレオチド、(c)(a)のポリヌクレオチドに相補的なポリヌクレオチド、(d)(b)のポリヌクレオチドに相補的なポリヌクレオチド、および(e)(a)~(d)のRNA等価物からなる群から選択された単離されたポリヌクレオチドを提供する。一態様では、ポリヌクレオチドは少なくとも60の連続したヌクレオチドを有する。

【0047】

50

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは (a) SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、 (b) SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列と少なくとも90%が同一である天然のポリヌクレオチド配列を含むポリヌクレオチド、 (c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、 (d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、および (e) (a) ~ (d) のRNA等価物からなる群から選択されたポリヌクレオチドの配列を有する。検出方法は、 (a) サンプル中の上記標的ポリヌクレオチドに相補的な或る配列からなる少なくとも20の連続したヌクレオチド群からなる或るプローブを用いて該サンプルをハイブリダイズする過程と、 (b) 該ハイブリダイゼーション複合体の有無を検出し、複合体が存在すればオプションでその量を検出する過程からなる。該プローブと該標的ポリヌクレオチドあるいはその断片との間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは、該標的ポリヌクレオチドに対し特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチドを含む。

10

【0048】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する一方法を提供する。ここで、標的ポリヌクレオチドは、 (a) SEQ ID NO:31-60からなる群から選択した或るポリヌクレオチド配列を含むポリヌクレオチド、 (b) SEQ ID NO:31-60からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を含むポリヌクレオチド、 (c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、 (d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、および (e) (a) ~ (d) のRNA等価物、からなる群から選択した、或るポリヌクレオチドの配列を有する。検出方法は、 (a) ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、 (b) 増幅した標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

20

【0049】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む組成物を提供する。有効量のポリペプチドは、 (a) SEQ ID NO:1-30からなる群から選択したアミノ酸配列を含むポリペプチド、 (b) SEQ ID NO:1-30からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、 (c) SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および (d) SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択される。一実施例では、SEQ ID NO:1-30からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的SECPの発現の低下に関連した疾患やその症状の治療方法を提供する。

30

【0050】

本発明はまた、 (a) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を含むポリペプチド、 (b) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然アミノ酸配列を含むポリペプチド、 (c) SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および (d) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を有するポリペプチドの免疫原性断片、からなる群から選択したポリペプチドのアゴニストとしての有効性を確認するために、或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、 (a) 該ポリペプチドを含むサンプルを或る化合物に曝す過程と、 (b) サンプル中のアゴニスト活性を検出する過程からなる。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、機能的SECPの発現の低下に関連した疾患やその症状の治療方法を提供し、その内にはそのような治療を必要とする患者にこの組成物を投与することが含まれる。

40

50

【0051】

本発明は更に、(a) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を含むポリペプチド、(b) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然アミノ酸配列を含むポリペプチド、(c) SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択したポリペプチドのアンタゴニストとしての有効性につき、或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを含むサンプルを或る化合物に曝す過程と、(b) サンプル中のアンタゴニスト活性を検出する過程からなる。別法で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む組成物を提供する。更なる別法では、本発明は、機能的SECPの過剰な発現に関連した疾患やその症状の治療方法を提供し、その内にはそのような治療を必要とする患者にこの組成物を投与することが含まれる。

10

【0052】

本発明は更に、(a) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を含むポリペプチド、(b) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を含むポリペプチド、(c) SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO:1-30からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドに特異結合する或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを適切な条件下で少なくとも1つの試験化合物に混合させる過程と、(b) この試験化合物と該ポリペプチドとの結合を検出し、それにより該ポリペプチドに特異結合する化合物を同定する過程からなる。

20

【0053】

本発明は更に、(a) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を含むポリペプチド、(b) SEQ ID NO:1-30からなる群から選択したアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を含むポリペプチド、(c) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-30からなる群から選択した或るアミノ酸配列を有するポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドの活性を調節する或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドの活性にとり許容し得る条件下で、該ポリペプチドを少なくとも1つの試験化合物と混合する過程と、(b) 該ポリペプチドの活性をこの試験化合物の存在下で算定する過程と、(c) この試験化合物の存在下での該ポリペプチドの活性をこの試験化合物の不存在下での該ポリペプチドの活性と比較する過程からなり、この試験化合物の存在下での該ポリペプチドの活性の変化は、該ポリペプチドの活性を調節する化合物を標示する。

30

【0054】

更に本発明は、SEQ ID NO:31-60からなる群から選択されたポリヌクレオチド配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、(a) この標的ポリヌクレオチドを含むサンプルを或る化合物に曝露する過程と、(b) この標的ポリヌクレオチドの発現の変化を検出する過程と、(c) 可変量の該化合物の存在下での標的ポリヌクレオチドの発現を該化合物の不存在下での発現と比較する過程を含む、該スクリーニング方法を提供する。

40

【0055】

本発明は更に、(a) 核酸を含む生物学的サンプルを試験化合物で処理する過程と、(b) (i) SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii) SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列と少なくとも90%が同一である天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii) (i) に相補的な配列を有するポリヌクレオチド、(iv) (ii) のポリヌクレオチドに相補的

50

なポリヌクレオチド、(v)(i)~(iv)のRNA等価物からなる群から選択したポリヌクレオチドの少なくとも20の連続したヌクレオチドを含むプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程とを含む試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、前記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、前記標的ポリヌクレオチドは、(i)SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii)SEQ ID NO:31-60からなる群から選択したポリヌクレオチド配列と少なくとも90%が同一である天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii)(i)のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、および(v)(i)~(iv)のRNA等価物からなる群から選択する。或いは、標的ポリヌクレオチドは、上記(i)~(v)からなる群から選択したポリヌクレオチド配列の断片を有する。毒性の算定方法には更に(c)ハイブリダイゼーション複合体の量を定量する過程と、(d)処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量の差は、試験化合物の毒性を示す。

10

【発明を実施するための最良の形態】

【0056】

(本発明の記載について)

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

20

【0057】

請求の範囲及び明細書中で用いている単数形の「或る」及び「その(この)」の表記は、文脈から明らかにそうでないとされる場合を除いて複数のものを指す場合もあることに注意しなければならない。従って、例えば「或る宿主細胞」は複数の宿主細胞を含み、その「抗体」は複数の抗体が含まれ、当業者には周知の等価物なども含まれる。

【0058】

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

30

【0059】

(定義)

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

40

【0060】

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

【0061】

用語「対立遺伝子変異配列」は、SECPをコードする遺伝子の別の形を指す。対立遺伝子変

50

異体は、核酸配列における少なくとも1つの突然変異から作製し得る。また、変異RNAまたはポリペプチドを作製し得る。ポリペプチドの構造または機能は、変異することもしないこともある。或る遺伝子は、その天然型の対立遺伝子変異配列を全く持たない場合もあり、1個以上持つこともある。対立遺伝子変異体を生じさせる通常の突然変異性変化は一般に、ヌクレオチドの自然な欠失、付加または置換に帰するものである。これら各変化は、単独で或いは他の変化と共に、所定の配列内で1回若しくは数回生じ得る。

【0062】

SECPをコードする「変異」核酸配列は、様々なヌクレオチドの欠失、挿入、或いは置換が起こっても、SECPと同じポリペプチド或いはSECPの機能特性の少なくとも1つを備えるポリペプチドを指す。この定義には、SECPをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置での対立遺伝子変異配列との不相当或いは予期しないハイブリダイゼーション、並びにSECPをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能な或いは検出困難な多型性を含む。コードされたタンパク質も「変異」し得るものであり、サイレント変化を生ぜしめて結果的に機能的に等価なSECPとなるようなアミノ酸残基の欠失、挿入または置換を含み得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にSECPの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び/または両親媒性についての類似性に基づいて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似した非荷電極性側鎖を持つアミノ酸には、アスパラギンとグルタミン、セリンとトレオニンがある。親水性値が近似した非荷電側鎖を持つアミノ酸には、ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

10

20

【0063】

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

【0064】

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

30

【0065】

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

【0066】

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

40

【0067】

用語「抗原決定基」は、特定の抗体と接触する分子の領域(即ちエピトープ)を指す。タンパク質またはタンパク質断片を用いて宿主動物を免疫化する場合、タンパク質の多数の領域が、抗原決定基(タンパク質の特定の領域または3次元構造)に特異結合する抗体の

50

産生を誘導し得る。抗原決定基は、抗体への結合について、無損傷抗原（即ち免疫応答を誘導するために用いられる免疫原）と競合し得る。

【0068】

用語「アプタマー」は、特定の分子標的に結合する核酸またはオリゴヌクレオチド分子を指す。アプタマーは *in vitro* の進化過程（例えば米国特許番号第5,270,163号に記載された SELEX (Systematic Evolution of Ligands by EXponential Enrichment)）から由来するもので、そのような過程は大きな組み合わせライブラリから標的特異的なアプタマー配列を選択する。アプタマー組成は、2本鎖または1本鎖であってもよく、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導體、または他のヌクレオチド様分子を含み得る。アプタマーのヌクレオチド成分は、修飾された糖基（例えばリボヌクレオチドの2'-OH基が2'-Fまたは2'-NH₂で置換し得る）を有することが可能で、そのような糖基はヌクレアーゼへの抵抗性または血液中でのより長い寿命などの望ましい性質に改善し得る。循環系からアプタマーが除去される速度を遅くするために、アプタマーを高分子量キャリア等の分子に抱合させることができる。アプタマーは、たとえば架橋剤の光活性化によって、各々の同種リガンドと特異的に架橋させることができる（Brody, E.N.及びL. Gold (2000) *J. Biotechnol.* 74:5-13等を参照）。

10

【0069】

「イントラマー (intramer)」の用語は *in vivo* で発現されるアプタマーを意味する。例えば、ワクシニアウイルスに基づく或るRNA発現系を用いて、白血球の細胞質内で特定のRNAアプタマー類が高レベルに発現されている（Blind, M.他 (1999) *Proc. Natl Acad. Sci. USA* 96:3606-3610）。

20

【0070】

「スピーゲルマー (spiegelmer)」の語はL-DNA、L-RNAその他の左旋性ヌクレオチド誘導體またはヌクレオチド様分子を含むアプタマーを指す。左旋性のヌクレオチドを含むアプタマーは右旋性ヌクレオチドに作用する天然の酵素による分解に対して耐性がある。

【0071】

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

30

【0072】

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

40

【0073】

「相補的」は、塩基対合によってアニールする2つの一本鎖核酸配列間の関係を指す。例えば、配列「5'A-G-T3'」は、相補配列「3'T-C-A5'」と対を形成する。

【0074】

「～のポリヌクレオチド配列を含む組成物」または「～のアミノ酸配列を含む(有する)組成物」は広い意味で、所定のポリヌクレオチド配列若しくはアミノ酸配列を含む任意の組

50

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

【0075】

「コンセンサス配列」は、不要な塩基を分離するためにDNA配列の解析を繰り返し行い、X L-PCRキット（Applied Biosystems, Foster City CA）を用いて5'及び/または3'の方向に伸長され、再度シーケンシングされた核酸配列、またはGELVIEW断片アセンブリシステム（GCG, Madison, WI）またはPhrap（University of Washington, Seattle WA）等の断片アセンブリ用のコンピュータプログラムを用いて1つ或いはそれ以上のオーバーラップするcDNAやEST、またはゲノムDNA断片からアセンブリされた核酸配列を指す。伸長及びアセンブリの両方を行ってコンセンサス配列を作製する配列もある。

10

【0076】

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

20

【0077】

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

30

40

【0078】

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

【0079】

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

50

【0080】

用語「誘導体」は、化学修飾されたポリヌクレオチドまたはポリペプチドを指す。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチド配列の化学修飾に含まれ得る。ポリヌクレオチド誘導体は、天然分子の生物学的または免疫学的機能を少なくとも1つは保持しているポリペプチドをコードする。ポリペプチド誘導体は、グリコシル化、ポリエチレングリコール化(pegylation)、或いは任意の同様なプロセスであって、誘導起源のポリペプチドの少なくとも1つの生物学的若しくは免疫学的機能を保持するプロセスによって、修飾されたポリペプチドである。

【0081】

「検出可能な標識」は、測定可能な信号を発生し得る、ポリヌクレオチドやポリペプチドに共有結合或いは非共有結合するレポーター分子または酵素を指す。 10

【0082】

「差次的発現」は、少なくとも2つの異なったサンプルを比較することによって決められる、増加(上方調節)、あるいは減少(下方調節)、または遺伝子発現の欠損またはタンパク発現の欠損を指す。このような比較は例えば、処理済サンプルと不処理サンプル、または病態のサンプルと正常サンプルの間で行われ得る。

【0083】

「エキソンシャフリング」は、異なるコード領域(エキソン)の組換えを意味する。1つのエキソンはコードされるタンパク質の1つの構造的または機能的ドメインを代表し得るため、安定した基礎構造の新たな再構築(reassortment)を介して新しいタンパク質がアセンブリされることが可能であり、これにより新しいタンパク質機能の進化を促進できる。 20

【0084】

用語「断片」は、SECPまたはSECPをコードするポリヌクレオチドの固有の部分であって、その親配列と同一であるがその配列より長さが短いものを指す。断片は、定義された配列の全長から1ヌクレオチド/アミノ酸残基を差し引いた長さよりも短い長さを有し得る。例えば或る断片は、5~1000の連続したヌクレオチドまたはアミノ酸残基を有し得る。プローブ、プライマー、抗原、治療用分子として、或いはその他の目的のために用いられる断片は、少なくとも5、10、15、16、20、25、30、40、50、60、75、100、150、250若しくは500の連続したヌクレオチド或いはアミノ酸残基の長さであり得る。断片は、或る分子の特定領域から優先的に選択し得る。例えば、ポリペプチド断片は、所定の配列に示すようなポリペプチドの最初の250または500アミノ酸(または最初の25%または50%)から選択された或る長さの連続したアミノ酸を有し得る。これらの長さは明らかに例として挙げているものであり、本発明の実施例では、配列表、表及び図面を含む明細書に裏付けされた任意の長さであってよい。 30

【0085】

SEQ ID NO:31-60の断片には、固有のポリヌクレオチド配列領域が含まれる。この領域は、SEQ ID NO:31-60を特異的に同定するものであり、例えば断片が得られる同一ゲノム中の他すべての配列とは異なるものである。SEQ ID NO:31-60のある断片は、例えば、ハイブリダイゼーションや増幅技術、またはSEQ ID NO:31-60を関連ポリヌクレオチド配列から区別する類似の方法に有用である。SEQ ID NO:31-60の断片の正確な長さ及び、断片が対応するSEQ ID NO:31-60の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。 40

【0086】

SEQ ID NO:1-30のある断片は、SEQ ID NO:31-60のある断片によってコードされる。SEQ ID NO:1-30のある断片には、SEQ ID NO:1-30を特異的に同定する固有のアミノ酸配列領域が含まれている。例えば、SEQ ID NO:1-30のある断片は、SEQ ID NO:1-30を特異認識する抗体を産出するための免疫原性ペプチドとして有用である。SEQ ID NO:1-30のある断片の正確な長さ、及びその断片に対応するSEQ ID NO:1-30の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0087】

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

【0088】

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

【0089】

ポリヌクレオチド配列に適用される「一致率」または「%一致」の語は、標準化されたアルゴリズムを用いてアラインメントされた、2つ以上のポリヌクレオチド配列間で一致する残基の割合を意味する。このようなアルゴリズムは、2配列間のアラインメントを最適化するために、比較する配列において、標準化された再現性のある方法でギャップを挿入しうるので、2つの配列をより有意に比較できる。

10

【0090】

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

20

【0091】

或いは、一般的に用いられ且つ自由に入手できる配列比較アルゴリズム一式が、国立バイオテクノロジー情報センター (NCBI) Basic Local Alignment Search Tool (BLAST) から提供されており (Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410)、これはメリーランド州ベセスダにあるNCBI及びインターネット (<http://www.ncbi.nlm.nih.gov/BLAST/>) を含む幾つかの情報源から入手可能である。このBLASTソフトウェア一式には、既知のポリヌクレオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールも入手可能であり、2つのヌクレオチド配列を直接にペアワイズで比較するために用いられる。「BLAST 2 Sequences」は、<http://www.ncbi.nlm.nih.gov/gorf/bl2.html> にアクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn及びblastp (以下に記載) の両方に用いることができる。BLASTプログラムは、一般的には、ギャップ及び他のパラメータをデフォルト設定に設定して用いる。例えば、2つのヌクレオチド配列を比較するために、デフォルトパラメータとして設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (2000年4月21日) を用いてblastnを実行してもよい。デフォルトパラメータの設定例を以下に示す。

30

【0092】

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

40

【0093】

一致率は、ある定義された配列の全長 (例えば特定のSEQ IDナンバーで定義された配列)

50

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

【0094】

高度の同一性を示さない核酸配列が、それにもかかわらず遺伝子コードの縮重が原因で類似のアミノ酸配列をコードする場合がある。この縮重を利用して核酸配列内で変化を生じさせて、全ての核酸配列が実質上同一のタンパク質をコードするような多数の核酸配列を生成し得ることを理解されたい。

10

【0095】

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

【0096】

ポリペプチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組み込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる（既に説明したのでそれを参照されたい）。CLUSTAL Vを用いて、ポリペプチド配列をペアワイズアラインメントする際のデフォルトパラメータは、Ktuple=1、gap penalty=3、window=5、「diagonals saved」=5と設定される。デフォルトの残基重み付け表としてPAM250マトリクスが選択される。ポリヌクレオチドアラインメントと同様に、CLUSTAL Vは、アラインメントされたポリペプチド配列対間の「percent similarity」として一致率を報告する。

20

【0097】

或いは、NCBI BLASTソフトウェア一式を用いてもよい。例えば、2つのポリペプチド配列をペアワイズで比較する場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツールVersion 2.0.12（2000年4月21日）でblastpを使用するであろう。デフォルトパラメータの設定例を以下に示す。

30

【0098】

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

【0099】

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

40

【0100】

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

50

【0101】

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

【0102】

「ハイブリダイゼーション」とは、所定のハイブリダイゼーション条件下で、ある一本鎖ポリヌクレオチドがある相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。特異的ハイブリダイゼーションは、2つの核酸配列が高い相補性を共有することの指標である。特異的ハイブリダイゼーション複合体は許容的アニーリング条件下で形成され、「洗浄」ステップ後もハイブリダイズされたままである。洗浄ステップは、ハイブリダイゼーションプロセスのストリンジェンシーを決定する際に特に重要であり、よりストリンジェントな条件では、非特異結合（即ち完全には一致しない核酸鎖間の対の結合）が減少する。核酸配列のアニーリングのための許容的條件は、当業者が慣例的に決定できる。許容的條件は、どのハイブリダイゼーション実験でも一定でありうるが、洗浄条件は所望のストリンジェンシーを得るように、従ってハイブリダイゼーション特異性も得るように実験によって変更されうる。アニーリングが許容される条件は、例えば、温度が68 で、約6×SSC、約1% (w/v) のSDS、並びに約100 μg/mlのせん断して変性したサケ精子DNAが含まれる。

10

【0103】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度及びpHにおける特異配列の融点 (T_m) より約5~20 低くなるように選択する。この T_m は、所定のイオン強度及びpHの条件下で、完全に一致するプローブに標的配列の50%がハイブリダイズする温度である。 T_m を計算する式及び核酸のハイブリダイゼーション条件はよく知られており、Sambrook J 他 (1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NYに記載されており、特に2巻の9章を参照されたい。

20

【0104】

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

30

【0105】

用語「ハイブリダイゼーション複合体」は、相補的な塩基間の水素結合によって形成された、2つの核酸配列の複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る (C_0t または R_0t 解析等)。或いは、一方の核酸配列が溶解状態で存在し、もう一方の核酸配列が固体支持体（例えば紙、膜、フィルタ、チップ、ピンまたはガラススライド、或いは他の適切な基板であって細胞若しくはその核酸が固定される基板）に固定されているような2つの核酸配列間に形成され得る。

40

【0106】

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

【0107】

「免疫応答」は、炎症、外傷、免疫異常症、あるいは伝染性疾患または遺伝性疾患に関連

50

する症状を指し得る。これらの症状は、細胞及び全身の防御系に作用し得る種々の因子、例えばサイトカイン、ケモカイン、その他のシグナル伝達分子の発現によって特徴づけることができる。

【0108】

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

【0109】

用語「マイクロアレイ」は、基板上の複数のポリヌクレオチド、ポリペプチドまたはその他の化合物の構成を指す。 10

【0110】

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

【0111】

用語「調節」は、SECPの活性の変化を指す。例えば、調節によって、SECPのタンパク質活性、或いは結合特性、またはその他の生物学的特性、機能的特性或いは免疫学的特性の変化が起こる。

【0112】

「核酸」及び「核酸配列」の語は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチドまたはこれらの断片を指す。「核酸」及び「核酸配列」の語は、ゲノム起源または合成起源のDNAまたはRNAであって一本鎖または二本鎖であるか或いはセンス鎖またはアンチセンス鎖を表し得るようなDNAまたはRNAや、ペプチド核酸(PNA)や、任意のDNA様またはRNA様物質を指すこともある。 20

【0113】

「機能的に連結した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に連結している。機能的に連結したDNA配列は非常に近接するか、或いは連続的に隣接し得る。また、2つのタンパク質コード領域を結合するために必要な場合は、同一のリーディングフレーム内に在り得る。 30

【0114】

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

【0115】

SECPの「翻訳後修飾」には、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、蛋白分解性切断及びその他の当分野で既知の修飾を含まれ得る。これらのプロセスは、合成或いは生化学的に生じ得る。生化学的修飾は、SECPの酵素環境に依存し、細胞の種類によって異なり得る。 40

【0116】

「プローブ」とは、同一配列或いは対立遺伝子核酸配列、関連する核酸配列の検出に用いる、SECPやそれらの相補体、またはそれらの断片をコードする核酸配列のことである。プローブは、単離されたオリゴヌクレオチドまたはポリヌクレオチドであって、検出可能な標識またはレポーター分子に結合したものである。典型的な標識には、放射性アイソトープ、リガンド、化学発光試薬及び酵素がある。「プライマー」とは、相補的な塩基対を形成して標的のポリヌクレオチドにアニーリング可能な、通常はDNAオリゴヌクレオチドである短い核酸である。プライマーは次に、DNAポリメラーゼ酵素によって標的DNA鎖に沿って延長し得る。プライマー対は、例えばポリメラーゼ連鎖反応(PCR)による核酸配列の 50

増幅（及び同定）に用い得る。

【0117】

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

【0118】

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

【0119】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野で既知のソフトウェアを用いて行う。例えばOLIGO 4.06ソフトウェアは、各100ヌクレオチドまでのPCRプライマー対の選択に有用であり、32キロボースまでのインプットポリヌクレオチド配列からオリゴヌクレオチド及び最大5,000ヌクレオチドまでの大きめのポリヌクレオチドとオリゴヌクレオチドを分析するのにも有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組込まれている。例えば、PrimOUプライマー選択プログラム（テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能）は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム（Whitehead Institute/MIT Center for Genome Research（マサチューセッツ州ケンブリッジ）より入手可能）によって、ユーザーは、プライマー結合部位として避けたい配列を指定できる「ミスプライミングライブラリ」を入力できる。Primer3は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である（後二者のプライマー選択プログラムのソースコードは、それぞれの情報源から得てユーザー固有のニーズを満たすように修正し得る）。PrimerGenプログラム（英国ヒトゲノムマッピングプロジェクト-リソースセンター（英国ケンブリッジ）から一般向けに入手可能）は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラムは、固有なもの、及び保存されたもの双方のオリゴヌクレオチドとポリヌクレオチド断片との同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチドおよびポリヌクレオチド断片は、ハイブリダイゼーション技術において、例えばPCRまたはシーケンシングプライマーとして、マイクロアレイエレメントとして、あるいは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

【0120】

「組換え核酸」は天然の配列ではなく、2つ以上の配列の離れたセグメントを人工的に組み合わせた配列である。この人為的組合せはしばしば化学合成によって達成するが、より一般的には核酸の単離セグメント群の人為的操作によって、例えばSambrookの文献（前出）に記載されているような遺伝子工学的的手法によって達成する。組換え核酸の語は、単に核酸の一部が付加、置換または欠失により改変された核酸も含む。しばしば組換え核酸に

10

20

30

40

50

は、プロモーター配列に機能的に連結した核酸配列が含まれる。このような組換え核酸は、例えば細胞を形質転換するために使用されるベクターの一部とすることが可能である。

【0121】

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

【0122】

「調節エレメント」は、通常は遺伝子の非翻訳領域に由来する核酸配列であり、エンハンサー、プロモーター、イントロン及び5'及び3'の非翻訳領域(UTR)を含む。調節エレメントは、転写、翻訳またはRNA安定性を制御する宿主タンパク質またはウイルスタンパク質と相互作用する。

10

【0123】

「レポーター分子」は、核酸、アミノ酸または抗体の標識化に用いられる化学的または生化学的部分である。レポーター分子としては、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、阻害因子、磁気粒子及びその他の当分野で既知の成分を含む。

【0124】

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

20

【0125】

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

【0126】

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

30

【0127】

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

【0128】

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

40

【0129】

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

【0130】

「転写物イメージ」あるいは「発現プロフィール」は、所定条件下での所定時間における

50

特定の細胞の種類または組織による集合的遺伝子発現のパターンを指す。

【0131】

「形質転換(transformation)」とは、外来DNAが受容細胞に導入されるプロセスのことである。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核宿主細胞または真核宿主細胞に挿入する任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。限定するものではないが形質転換方法には、バクテリオファージあるいはウイルス感染、電気穿孔法(エレクトロポレーション)、熱ショック、リポフェクション及び微粒子銃を用いる方法がある。用語「形質転換された細胞」には、導入されたDNAが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一過的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

10

【0132】

ここで用いる「遺伝形質転換体(transgenic organism)」とは任意の生物体であり、限定するものではないが動植物を含み、生物体の1個以上の細胞が、ヒトの関与によって、例えば本技術分野でよく知られているトランスジェニック(transgenic)技術によって導入された異種核酸を有する。核酸の細胞への導入は、直接または間接的に、細胞の前駆物質に導入することによって、計画的な遺伝子操作によって、例えば微量注射法によって或いは組換えウイルスでの感染によって行う。或いは核酸の導入は、組換えウイルスベクター、例えばレンチウイルスベクターを感染させて成し得る(Lois, C. 他(2002) Science 295 :868-872)。遺伝子操作の語は、古典的な交雑育種或いはin vitro受精を指すものではなく、組換えDNA分子の導入を指すものである。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌および動植物がある。本発明の単離されたDNAは、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合によって宿主に導入することができる。本発明のDNAをこのような生物体に移入する技術はよく知られており、前出のSambrook 他(1989)等の参考文献に記載されている。

20

【0133】

特定の核酸配列の「変異体/変異配列」は、核酸配列1本全部の長さに対して特定の核酸配列と少なくとも40%の配列同一性を有する核酸配列であると定義する。定義づけには、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9(1999年5月7日)を用いてblastnを実行する。このような核酸対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。或る変異体は、例えば「対立遺伝子」変異体(前述)、「スプライス」変異体、「種」変異体または「多型性」変異体として説明し得る。スプライス変異体は参照分子との有意な同一性を有し得るが、mRNAプロセッシング中のエキソンの選択的スプライシングによって通常、より多くまたはより少数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメイン群を有するか、あるいは参照分子には存在するドメイン群が欠落していることがある。種変異体は、種によって異なるポリヌクレオチド配列である。結果的に生じるポリペプチドは通常、相互に有意なアミノ酸同一性を有する。多型性変異体は、所与の種の個体間での特定遺伝子のポリヌクレオチド配列中での変異である。多型変異配列はまた、ポリヌクレオチド配列の1つのヌクレオチド塩基が異なる「1塩基多型性」(SNP)も含み得る。SNPの存在は、例えば特定の個体群、病状または病状性向を示し得る。

30

40

【0134】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の1本の長さ全体で特定のポリペプチド配列に対して少なくとも40%の配列同一性を有するポリペプチド配列として定義される。定義づけには、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9(1999年5月7日)を用いてblastpを実行する。このようなポリペプチド対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、90%、91%、92

50

%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。

【0135】

(発明)

本発明は、新規のヒト分泌タンパク質 (SECP) 群およびSECPをコードするポリヌクレオチド群の発見に基づき、また、これらの組成物を使用した、肝臓疾患、細胞増殖異常、自己免疫/炎症疾患、心血管障害、神経障害、および発生障害の診断、治療、ならびに予防の発見に基づく。

【0136】

表1は、本発明の完全長ポリヌクレオチド配列およびポリペプチド配列の命名の概略である。各ポリヌクレオチドおよびその対応するポリペプチドは、1つのIncyteプロジェクト識別番号 (IncyteプロジェクトID) に関連する。各ポリペプチド配列は、ポリペプチド配列識別番号 (ポリペプチドSEQ ID NO) とIncyteポリペプチド配列番号 (IncyteポリペプチドID) によって表示した。各ポリヌクレオチド配列は、ポリヌクレオチド配列識別番号 (ポリヌクレオチドSEQ ID NO) とIncyteポリヌクレオチドコンセンサス配列番号 (IncyteポリヌクレオチドID) によって表示した。列6は、本発明のポリペプチド配列とポリヌクレオチド配列とに対応する物理的完全長クローン群のIncyte ID番号を示す。完全長クローンは、列3に示すポリペプチド配列に対して少なくとも95%の配列同一性を持つポリペプチドをコードする。

10

【0137】

表2は、GenBankタンパク質 (genpept) データベースとPROTEOMEデータベースとに対するBLAST分析で同定した、本発明のポリペプチド群に相同な配列群を示す。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号 (ポリペプチドSEQ ID NO:) とそれに対応するIncyteポリペプチド配列番号 (IncyteポリペプチドID) を示す。列3は、GenBankの最も近い相同体のGenBank識別番号 (Genbank ID NO:) と最も近いPROTEOMEデータベース相同体のPROTEOMEデータベース識別番号 (PROTEOME ID NO:) を示す列4は、各ポリペプチドとその相同体1つ以上との間の一致に関する確率スコアを示す。列5は、GenBankとPROTEOMEデータベースの相同体の注釈を示し、更に該当箇所には関連する引用文献も示す。これらを引用することを以って本明細書の一部とする。

20

【0138】

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

30

40

【0139】

表2および3は共に、本発明の各々のポリペプチドの特性を要約しており、それら特性が、請求の範囲に記載されたポリペプチドが分泌タンパク質であることを確立している。

【0140】

例えば、SEQ ID NO:1はヒトのIgG Fc受容体I (GenBank ID g180279) と残基G18から残基T289まで99%の同一性を有することがBasic Local Alignment Search Tool (BLAST) によって示された (表2参照)。BLAST確率スコアは $5.4e-146$ であるが、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示す。SEQ ID NO:1はまた、免疫グロブリンドメインを有することが、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決

50

定された (表3参照)。BLIMPS分析から得たデータは、SEQ ID NO:1が分泌タンパク質であることを裏づける証拠を更に提供する (「免疫グロブリンドメイン」はマトリクスタンパク質の特徴であることに注意)。

【0141】

別の例として、SEQ ID NO:11は長さが150残基であり、M1残基からS105残基までヒトの膜貫通 α -カルボキシグルタミン酸タンパク質4(TGM4)(GenBank ID g12656635)に100%同一であることが、Basic Local Alignment Search Tool(BLAST)によって示された (表2参照)。BLAST確率スコアは $8.4e-54$ であり、これは観測されたポリペプチド配列を偶然に得る確率を示す。SEQ ID NO:11はまた、ビタミンK依存性カルボキシル化 α -カルボキシグルタミン酸(GLA)ドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリドメイン群のPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表3参照)。BLIMPS解析、MOTIFS解析、およびPROFILESCAN解析よりのデータは、SEQ ID NO:11がビタミンK依存性1回貫通型膜内在性タンパク質のファミリのメンバーである、更に実証的な証拠を提供する。

10

【0142】

別の例として、SEQ ID NO:16は長さが519アミノ酸であり、M274残基からE519残基までヒトのAD021タンパク質(GenBank ID g7578787)に100%同一であることが、Basic Local Alignment Search Tool(BLAST)によって示された (表2参照)。BLAST確率スコアは $8.3e-134$ であり、これは観測されたポリペプチド配列を偶然に得る確率を示す。また他の例として、SEQ ID NO:22はマウスの推定上のタンパク質配列(GenBank ID g12845943)とM1残基からQ138残基まで81%同一であることがBasic Local Alignment Search Tool (BLAST)によって示された (表2参照)。BLAST確率スコアは $9.1e-57$ であり、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:22はまたシグナル切断部位を有し、これはSPScan重み行列解析プログラムにおいて、統計的に有意な一致を検索して決定された (表3参照)。

20

【0143】

また他の例として、SEQ ID NO:23はヒトのMac-2 結合タンパク質(GenBank ID g307153)とS42残基からP345残基まで28%同一であることがBasic Local Alignment Search Tool (BLAST)によって示された (表2参照)。BLAST確率スコアは $9.3e-27$ であり、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:23はまた、BTB/POZドメインを有するが、これは隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリドメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表3参照)。

30

【0144】

SEQ ID NO:2-10、SEQ ID NO:12-15、SEQ ID NO:17-21、およびSEQ ID NO:24-30については、同様の方法で分析し、注釈を付けた。SEQ ID NO:1-30の解析用のアルゴリズム及びパラメータを表7に記載した。

【0145】

表4に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA配列またはゲノムDNA由来のコード(エキソン)配列を用いて、或いはこれら2種類の配列を任意に組み合わせてアセンブリした。列1は本発明の各ポリヌクレオチドに対するポリヌクレオチド配列識別番号(ポリヌクレオチドSEQ ID NO)および対応するIncyteポリヌクレオチドコンセンサス配列番号(Incyte ID)、および塩基対の各ポリヌクレオチド配列の長さを示している。列2は、本発明の完全長ポリヌクレオチド配列をアセンブリするのに用いたcDNA配列および/またはゲノム配列の、また、例えばSEQ ID NO:31-60を同定するため、或いはSEQ ID NO:31-60と関連するポリヌクレオチド配列群とを区別するためのハイブリダイゼーション技術または増幅技術に有用なポリヌクレオチド配列の断片の、開始ヌクレオチド(5')位置および終了ヌクレオチド(3')位置を示す。

40

【0146】

表4の列2で記述されたポリヌクレオチド断片は特に、例えば組織特異的cDNAライブラリ

50

あるいはプールしたcDNAライブラリに由来するIncyte cDNAを指す場合もある。或いは列2のポリヌクレオチド断片は、完全長ポリヌクレオチド配列のアセンブリに寄与したGenBank cDNAまたはESTを指す場合もある。さらに、列2のポリヌクレオチド断片は、ENSEMBL (The Sanger Centre、(英国ケンブリッジ)) データベースから由来した配列を同定し得る(即ち「ENST」命名を含む配列)。或いは、列2で記述されたポリヌクレオチド断片は、NCBI RefSeq Nucleotide Sequence Records データベースから由来する場合もあり(即ち「NM」または「NT」の命名を含む配列)、またNCBI RefSeq Protein Sequence Recordsから由来する場合もある(即ち「NP」の命名を含む配列)。または列2のポリヌクレオチド断片は、「エキソンスティッチング(exon-stitching)」アルゴリズムにより結び合わせたcDNA及びGenscan予想エキソンの両方のアセンブリ体を意味する場合がある。例えば、FL_XXXXXX_N₁_N₂_YYYYY_N₃_N₄ と同定されるポリヌクレオチドは、アルゴリズムが適用される配列のクラスターの識別番号がXXXXXXであり、アルゴリズムにより生成される予測の番号がYYYYYであり、(もし存在すれば)N_{1,2,3...}が解析中に手動で編集された可能性のある特定のエキソンであるような「スティッチされた(stitched)」配列である(実施例5参照)。または、列2のポリヌクレオチド断片は「エキソンストレッチング(exon-stretching)」アルゴリズムにより結び合わせたエキソンのアセンブリ体を指す場合もある。例えば、FLXXXXXX_gAAAAA_gBBBBB_1_Nとして同定されるポリヌクレオチド配列は、「ストレッチされた」配列である。XXXXXXはIncyteプロジェクト識別番号、gAAAAAは「エキソンストレッチング」アルゴリズムを適用したヒトゲノム配列のGenBank識別番号、gBBBBBは一番近いGenBankタンパク質相同体のGenBank識別番号またはNCBI RefSeq 識別番号であり、Nは特定のエキソンを指す(実施例5を参照)。あるRefSeq配列が「エキソンストレッチング」アルゴリズムのためのタンパク質相同体として使用された場合には、RefSeq識別子(「NM」、「NP」、または「NT」によって表される)が、GenBank識別子(即ち、gBBBBB)の代わりに使用される場合もある。

10

20

【0147】

或いは、接頭コードは、手動で編集された構成配列、ゲノムDNA配列から予測された構成配列、または組み合わされた配列解析方法から由来する構成配列を同定する。次の表は、構成配列の接頭コードと、接頭コードに対応する配列分析方法の例を列記する(実施例4と5を参照)。

30

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

40

【0148】

場合によっては、最終コンセンサスポリヌクレオチド配列を確認するために表4に示すような配列の適用範囲と重複するIncyte cDNAの適用範囲が得られたが、それに関連するIncyte cDNA識別番号は示さなかった。

【0149】

表5は、Incyte cDNA配列を用いてアセンブリされた完全長ポリヌクレオチド配列のため

50

の代表的なcDNAライブラリを示している。代表的なcDNAライブラリは、上記のポリヌクレオチド配列群をアセンブリ及び確認するために用いられたIncyte cDNA配列によって最も頻繁に代表される、Incyte cDNAライブラリである。cDNAライブラリを作製するために用いた組織およびベクターを表5に示し、表6で説明している。

【0150】

表8は、本発明のポリヌクレオチド配列に見られる一塩基多型(SNP)を、種々のヒト集団での対立遺伝子(アレル)頻度と共に示す。列1および列2はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号(SEQ ID NO:)およびそれに対応するIncyteプロジェクト識別番号(PID)を示す。列3および列4はそれぞれ、SNPが検出されたESTのIncyte識別番号(EST ID)およびSNPの識別番号(SNP ID)を示す。列5および列6はそれぞれ、SNPがあるEST配列内の位置(EST SNP)および全長ポリヌクレオチド配列内のSNPの位置(CB1 SNP)を示す。列7は、EST配列内に見られるアレルを示す。列8と9は、SNP部位に見られる2つのアレルを示す。列10は、ESTに見られるアレルに基づくSNP部位を含むコドンによってコードされたアミノ酸を示す。列11~14は4つの異なるヒトの集団のアレル1の頻度を示す。エントリn/d(検出せず)は、集団におけるアレル1の頻度が低すぎて検出できないことを示し、n/a(利用不可)はアレル頻度が集団について決定できなかったことを示す。

10

【0151】

本発明はまた、SECPの変異体も含む。好適なSECPの変異体は、SECPアミノ酸配列に対して、少なくとも約80%、あるいは少なくとも約90%、さらには少なくとも約95%のアミノ酸配列同一性を有し、SECPの機能的または構造的特徴を少なくとも1つ含む変異体である。

20

【0152】

本発明はまた、SECPをコードするポリヌクレオチドを含む。特定の実施例において、本発明は、SECPをコードするSEQ ID NO:31-60からなる群から選択した配列を含むポリヌクレオチド配列を含む。SEQ ID NO:31-60のポリヌクレオチド配列には、配列表に示したように等価RNA配列をも含むが、そこでは窒素塩基チミンの出現はウラシルで置換され、糖のバックボーンはデオキシリボースではなくリボースで構成される。

【0153】

本発明はまた、SECPをコードするポリヌクレオチド配列の変異配列を含む。詳細には、このようなポリヌクレオチド配列の変異配列は、SECPをコードするポリヌクレオチド配列と少なくとも約70%のポリヌクレオチド配列同一性、或いは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を有する。本発明の或る実施態様では、SEQ ID NO:31-60からなる群から選択されたアミノ酸配列と少なくとも約70%、或いは少なくとも約85%、または少なくとも約95%もの一致率を有するようなSEQ ID NO:31-60からなる群から選択された配列を有するポリヌクレオチド配列の変異配列を含む。上記の任意のポリヌクレオチドの変異体は、SECPの機能的若しくは構造的特徴を少なくとも1つ有するアミノ酸配列をコードし得る。

30

【0154】

さらに、或いは別法では、本発明のポリヌクレオチド変異体は、SECPをコードするポリヌクレオチド配列のスプライス変異体である。スプライス変異体は、SECPをコードするポリヌクレオチド配列と有意な配列同一性をもつ部分であり得るが、mRNAプロセッシング中のエキソンの選択的スプライシングによって生じるブロックの配列の追加または欠損のため、その変異体は通常、より多くまたはより少数のヌクレオチドを有することになる。スプライス変異体は、SECPをコードするポリヌクレオチド配列とその全長にわたって、約70%以下の、或いは約60%以下の、或いは約50%以下のポリヌクレオチド同一性を有することがありうるが、スプライス変異体の部分は、SECPをコードするポリヌクレオチド配列の部分に、少なくとも約70%の、或いは少なくとも約85%の、或いは100%のポリヌクレオチドの配列同一性を有することになる。上記したスプライス変異配列は何れも、SECPの機能的或いは構造的特徴の少なくとも1つを有する或るアミノ酸配列をコードし得る。

40

【0155】

遺伝暗号の縮重により作り出され得るSECPをコードする種々のポリヌクレオチド配列には

50

、自然発生する任意の既知の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得るあらゆる可能なポリヌクレオチド配列のバリエーションを網羅し得る。これらの組み合わせは、天然のSECPのポリヌクレオチド配列に適用される標準的なトリプレット遺伝暗号を基に作られ、全てのそのような変異が明確に開示されているとみなす。

【0156】

SECPをコードするヌクレオチド配列及びその変異配列は一般に、好適に選択されたストリンジেন্টな条件下で、天然のSECPのヌクレオチド配列とハイブリダイズ可能であるが、非天然のコドンを含めるなどの実質的に異なった使い方のコドンを有するSECP或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。宿主が特定のコドンを利用する頻度に基づいて、特定の真核宿主又は原核宿主に発生するペプチドの発現率を高めるようにコドンを選択することが可能である。コードされたアミノ酸配列を変えないで、SECP及びその誘導体をコードするヌクレオチド配列を実質的に変更する別の理由は、天然の配列から作られる転写物より例えば長い半減期など好ましい特性を備えるRNA転写物を作ることにある。

10

【0157】

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

20

【0158】

更に本発明には、種々のストリンジエンシー条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO:31-60及びそれらの断片群にハイブリダイズ可能なポリヌクレオチド配列群が含まれる(例えば、Wahl, G.M.及びS.L. Berger (1987) *Methods Enzymol.* 152:399-407、Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511等を参照)。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

30

【0159】

DNAシーケンシングの方法は当分野では公知であり、本発明のいずれの実施例もDNAシーケンシング方法を用いて実施可能である。DNAシーケンシング方法には酵素を用いることができ、例えばDNAポリメラーゼ1のクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech, Piscataway NJ)を用いることができる。或いは、例えばELONGASE増幅システム (Life Technologies, Gaithersburg MD)において見られるように、ポリメラーゼと校正エキソヌクレアーゼを併用することができる。好適には、MICROLAB2200液体転移システム (Hamilton, Reno, NV)、PTC200サーマルサイクラー (MJ Research, Watertown MA) 及びABI CATALYST 800サーマルサイクラー (Applied Biosystems) 等の装置を用いて配列の準備を自動化する。次に、ABI 373或いは377 DNAシーケンシングシステム (Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野でよく知られている他の方法を用いてシーケンシングを行う。結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する (Ausubel, F.M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7、Meyers, R.A. (1995) *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, 856-853ページ等を参照)。

40

【0160】

当分野で周知のPCR法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、SECPをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば、使用し得る方法の1つである制限部位PCR法は、ユニバーサ

50

ルプライマーおよびネステッドプライマーを用いてクローニングベクター内のゲノムDNAからの未知の配列を増幅する方法である（例えば、Sarkar, G. (1993) PCR Methods Appl. 2:318-322を参照）。別の方法に逆PCR法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知の配列を増幅する方法である。鋳型は、或る既知のゲノム遺伝子座およびその周辺の配列群からなる制限酵素断片群から得る（例えばTriglia, T. 他 (1988) Nucleic Acids Res. 16:8186を参照）。第3の方法としてキャプチャPCR法があり、これはヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片をPCR増幅する方法に参与している（例えばLagerstrom, M. 他 (1991) PCR Methods Applic. 1:111119を参照）。この方法では、PCRを行う前に複数の制限酵素の消化及びライゲーション反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。未知の配列群を検索するために用い得る他の複数の方法も当分野で既知である（例えばParker, J.D.他 (1991) Nucleic Acids Res. 19:30553060を参照）。更に、PCR、ネステッドプライマー及びPromoterFinder（商標）ライブラリ（Clontech, Palo Alto CA）を用いてゲノムDNAをウォーキングすることができる。この手順は、ライブラリ類をスクリーニングする必要がなく、イントロン/エキソン接合部の発見に有用である。全てのPCRベースの方法では、市販ソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア（National Biosciences, Plymouth MN）或いは別の好適なプログラムを用いて、長さが約22~30ヌクレオチド、GC含有率が約50%以上で、温度約68 ~ 72 で鋳型に対してアニーリングするように、プライマー群を設計し得る。

10

20

30

40

50

【0161】

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

【0162】

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

【0163】

本発明の別の実施例では、SECPをコードするポリヌクレオチド配列またはその断片を組換えDNA分子にクローニングして、適切な宿主細胞内にSECP、その断片または機能的等価物を発現させることが可能である。遺伝暗号に固有の縮重により、実質的に同じ或いは機能的に等価のアミノ酸配列をコードする別のDNA配列が作られ得り、これらの配列をSECPの発現に利用可能である。

【0164】

種々の目的でSECPをコードする配列を変えるために、当分野で一般的に知られている方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び/または発現の調節が含まれるが、これらに限定されるものではない。遺伝子断片及び合成オリゴヌクレオチドのランダムなフラグメンテーション及びPCR再アセンブリによるDNAシャッフリングを用い、ヌクレオチド配列を組み換えることが可能である。例えば、オリゴヌクレオチドを介した部位特異的変異誘導を利用して、新規な制限部位の作製、グリコシル化パターンの変更、コドン優先の変更、スプライス変異体の生成等を起こす突然変異を導入し得る。

【0165】

本発明のヌクレオチドを、MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号、Chang, C.-C. 他 (1999) Nat. Biotechnol. 17:793-797、Christians, F.C. 他 (1999) Nat. Biotechnol. 17:259-264、Cramer, A. 他 (1996) Nat. Biotechnol. 14:315-319)などのDNAシャッフリング技術を用いてシャッフリングして、SECPの生物学的または酵素的な活性、或いは他の分子や化合物と結合する能力などのSECPの生物学的特性を変更或いは改良することができる。DNAシャッフリングは、遺伝子断片のPCRを介する組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、所望の特性を持つ遺伝子変異体群を同定する、選択またはスクリーニングの手順を経る。続いて、これら好適な変異体をプールし、更に反復してDNAシャッフリングおよび選択/スクリーニングを行い得る。かくして、「人工的な」育種及び急速な分子進化によって遺伝的多様性が作られる。例えば、ランダムポイント突然変異を持つ単一の遺伝子の断片を組換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングし得る。或いは、所与遺伝子の断片を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子の断片と組み換え、それによって天然に存在する複数の遺伝子の遺伝的多様性を、定方向の、制御可能な方法で最大化させることができる。

10

【0166】

別の実施例によれば、SECPをコードする配列は、当分野で周知の化学的方法を用いて、全体或いは一部が合成可能である(例えば、Caruthers, M.H. 他 (1980) Nucl. Acids Res. Symp. Ser 7:215-223、Horn, T. 他 (1980) Nucl. Acids Res. Symp. Ser7.225-232を参照)。別法として、化学的方法を用いてSECP自体またはその断片を合成することが可能である。例えば、種々の液相または固相技術を用いてペプチド合成を行い得る(例えばCreighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, 55-60、Roberge, J.Y. 他 (1995) Science 269:202204を参照)。自動合成はABI 431Aペプチドシンセサイザ(Applied Biosystems)を用いて達成し得る。更にSECPのアミノ酸配列または任意のその一部は、直接的な合成の際の変更、及び/または他のタンパク質の配列または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

20

【0167】

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

30

【0168】

生物学的に活性なSECPを発現させるために、SECPをコードするヌクレオチド配列またはその誘導体を好適な発現ベクターに挿入する。この発現ベクターは、好適な宿主に挿入されたコーディング配列の転写及び翻訳の調節に必要な要素を含む。これらの要素には、ベクター及びSECPをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。このようなエレメントは、強度及び特異性が様々である。特定の開始シグナルによって、SECPをコードする配列のより効果的な翻訳を達成することが可能である。開始シグナルの例には、ATG開始コドンと、コザック配列など近傍の配列とが含まれる。SECPをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コード配列あるいはその断片のみが挿入された場合は、インフレームATG開始コドンなど外来性の翻訳制御シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳エレメント及び開始コドンは、様々な天然物及び合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である(例えばScharf, D. 他 (1994) Results Probl. Cell Differ. 20:125162を参照)。

40

【0169】

50

当業者に周知の方法を用いて、SECPをコードする配列、好適な転写及び翻訳調節エレメントを含む発現ベクターを作製することが可能である。これらの方法には、in vitro組換えDNA技術、合成技術、及びin vivo遺伝子組換え技術が含まれる(例えば、Sambrook, J. 他.(1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章、8章及び16-17章、Ausubel, F.M. 他.(1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, 9章、13章及び16章等を参照)。

【0170】

種々の発現ベクター/宿主系を利用して、SECPをコードする配列の保持及び発現が可能である。限定するものではないがこのような発現ベクター/宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌などの微生物等や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター(例えばバキュロウイルス)に感染した昆虫細胞系や、ウイルス発現ベクター(例えばカリフラワーモザイクウイルスCaMVまたはタバコモザイクウイルスTMV)または細菌発現ベクター(例えばTiプラスミドまたはpBR322プラスミド)で形質転換させた植物細胞系または動物細胞系がある(例えば前出Sambrook; 前出Ausubel; Van Heeke, GおよびS.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509、Engelhard, E.K. 他(1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227、Sandig, V. 他.(1996) *Hum. Gene Ther.* 7:1937-1945、Takamatsu, N. (1987) *EMBO J.* 6:307311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, 191196ページ、Logan, J.およびT. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:36553659、Harrington, J.J. 他.(1997) *Nat. Genet.* 15:345-355を参照)。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ送達することができる(Di Nicola, M. 他(1998) *Cancer Gen. Ther.* 5(6):350-356、Yu, M. 他(1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344、Buller, R.M. 他(1985) *Nature* 317(6040):813-815; McGregor, D.P. 他(1994) *Mol. Immunol.* 31(3):219-226、Verma, I.M.およびN. Somia (1997) *Nature* 389:239-242等を参照)。本発明は使用される宿主細胞によって限定されるものではない。

【0171】

細菌系では、多数のクローニングベクター及び発現ベクターが、SECPをコードするポリヌクレオチド配列の使用目的に応じて選択可能である。例えば、SECPをコードするポリヌクレオチド配列の日常的なクローニング、サブクローニング、増殖は、PBLUESCRIPT(Stratagene, La Jolla CA)またはPSPORT1 プラスミド(Life Technologies)などの多機能の大腸菌ベクターを用いて達成することができる。ベクターのマルチクローニング部位にSECPをコードする配列を連結反応するとlacZ 遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列におけるin vitro転写、ジデオキシのシーケンシング、ヘルパーファージによる一本鎖のレスキュー、入れ子状態の欠失の生成にも有用であろう(Van Heeke, G.及びS.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509等を参照)。例えば、抗体の産生のためなどに多量のSECPが必要な場合は、SECPの発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導SP6バクテリオファージプロモーターまたは誘導T7バクテリオファージプロモーターを有するベクターが使用できる。

【0172】

酵母の発現系を使用してSECPを産出し得る。因子、アルコールオキシダーゼ、PGHプロモーター等の構成型或いは誘導型のプロモーターを含む多数のベクターが、出芽酵母菌(Saccharomyces cerevisiae)またはピキア酵母(Pichia pastoris)に使用可能である。更に、このようなベクターは、発現したタンパク質の、分泌か細胞内での保持かのどちらかを指示するものであり、安定した増殖のために宿主ゲノムの中に外来配列群を組み込むことを可能にする(例えば前出のAusubel, 1995; Bitter, G.A. 他(1987) *Methods Enzymol.* 153:516-544、Scorer, C.A. 他(1994) *Bio/Technology* 12:181-184を参照)。

【0173】

植物系を使用してSECPを発現することも可能である。SECPをコードする配列の転写は、ウイルスプロモーター、例えば単独あるいはTMV由来のオメガリーダー配列と組み合わせて用いられるようなCaMV由来の35Sおよび19Sプロモーターによって促進される (Takamatsu, N. (1987) *EMBO J.* 6:307311)。あるいは、RuBisCOの小サブユニットなどの植物プロモーター、または熱ショックプロモーターを用い得る (例えばCoruzzi, G.他(1984) *EMBO J.* 3:1671-1680、Broglie, R.他(1984) *Science* 224:838843、Winter, J.他(1991) *Results Probl. Cell Differ.* 17:85105等を参照)。これらの作製物は、直接DNA形質転換にまたは病原体を媒介とする形質移入によって、植物細胞内に導入可能である (『マグローヒル科学技術年鑑』(*The McGraw Hill Yearbook of Science and Technology*) (1992) McGraw Hill New York NY, 191-196ページ等を参照)。

10

【0174】

哺乳類細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び3連リーダー配列からなるアデノウイルス転写物/翻訳複合体にSECPをコードする配列を連結し得る。アデノウイルスゲノムの非必須のE1またはE3領域へ挿入することにより、宿主細胞内でSECPを発現する感染ウイルスを得ることができる (例えば、Logan, J.及びT. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659等を参照)。更に、ラウス肉腫ウイルス(RSV)エンハンサー等の転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。SV40またはEBVをベースにしたベクターを用いてタンパク質を高レベルで発現させることも

20

【0175】

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

【0176】

哺乳動物系の組換えタンパク質の長期にわたる産生のためには、株化細胞におけるSECPの安定した発現が望ましい。例えば、発現ベクターを用いて、SECPをコードする配列を株化細胞に形質転換することが可能である。このような発現ベクターは、ウイルス起源の複製要素及び/または内因性の発現要素や、同じベクター上に或いは別のベクター上に選択マーカー遺伝子を含み得る。ベクターの導入後、選択培地に移す前に強化培地で約1~2日間、細胞を増殖させ得る。選択可能マーカーの目的は選択剤への抵抗性を与えることであり、選択可能マーカーの存在により、導入した配列をうまく発現するような細胞の成長および回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

30

【0177】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、 tk^{-} 単純細胞のために用いられる単純ヘルペスウイルスのチミジンキナーゼ遺伝子と、 apr^{-} 細胞のために用いられるアデニンホスホリボシルトランスフェラーゼ遺伝子がある (例えばWigler, M.他 (1977) *Cell* 11:223-232、Lowy, I. 他. (1980) *Cell* 22:817-823を参照)。また、選択の基礎として代謝拮抗物質、抗生物質あるいは除草剤への耐性を用いることができる。例えばdhfrはメトトレキセートに対する耐性を与え、neoはアミノグリコシッドネオマイシンおよびG-418に対する耐性を与え、alsはクロルスルフロンに対する耐性を、patはホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える (Wigler, M. 他. (1980) *Proc. Natl. Acad. Sci. USA* 77:35673570、ColbereGarapin, F. 他. (1981) *J. Mol. Biol.* 150:114等を参照)。この他の選択可能な遺伝子、例えば、代謝物のための細胞の必要条件を変えるtrpB及びhisDが、文献に記載されている (Hartman, S.C.及びR.C. Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:8047-8051等を参照)。可視マーカー、例えばアントシアニン、緑色蛍光タンパク質 (GF

40

50

P; Clontech)、グルクロニダーゼ及びその基質 グルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを同定するだけでなく、特定のベクター系に起因する一過性あるいは安定したタンパク質発現を定量し得る(Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131等を参照)。

【0178】

マーカー遺伝子発現の有無によって目的の遺伝子の存在が示唆されても、その遺伝子の存在および発現の確認が必要な場合もある。例えば、SECPをコードする配列がマーカー遺伝子配列の中に挿入された場合、SECPをコードする配列を含む形質転換された細胞は、マーカー遺伝子機能の欠落により同定可能である。または、1つのプロモーターの制御下でマーカー遺伝子がSECPをコードする配列とタンデムに配置することも可能である。誘導または選択に応答したマーカー遺伝子の発現は通常、タンデム遺伝子の発現も示す。

10

【0179】

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

【0180】

特異的ポリクローナル抗体または特異的モノクローナル抗体を用いてSECPの発現の検出及び計測を行うための免疫学的方法は、当分野で公知である。このような技術の例としては、酵素に結合した免疫吸着剤検定法(ELISA)、ラジオイムノアッセイ(RIA)、フローサイトメーター(FACS)などが挙げられる。SECP上の2つの非干渉エピトープに反応するモノクローナル抗体を用いた、2部位のモノクローナルベースイムノアッセイ(two-site, monoclonal-based immunoassay)が好ましいが、競合結合アッセイも用いることもできる。これらのアッセイおよび他のアッセイは、当分野で周知である(例えばHampton, R.他(1990) *Serological Methods, a Laboratory Manual*, APS Press, St. Paul MN, Sect. I V, Coligan, J.E. 他.(1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY, Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJを参照)。

20

30

【0181】

多岐にわたる標識方法及び抱合方法が、当業者に知られており、様々な核酸アッセイおよびアミノ酸アッセイにこれらの方法を用い得る。SECPをコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いはPCRプローブを生成する方法には、オリゴ標識化、ニックトランスレーション、末端標識化、または標識されたヌクレオチドを用いるPCR増幅が含まれる。別法として、SECPをコードする配列、またはその任意の断片をmRNAプローブの生成のためのベクターにクローニングすることも可能である。このようなベクターは、当分野において知られており、市販もされており、T7、T3またはSP6などの好適なRNAポリメラーゼおよび標識されたヌクレオチドを加えて、*in vitro*でRNAプローブの合成に用いることができる。このような方法は、例えばAmersham Pharmacia Biotech、Promega(Madison WI)、U.S. Biochemicalなどから市販されている種々のキットを用いて実行することができる。検出を容易にするために用い得る好適なレポーター分子あるいは標識には、基質、補助因子、インヒビター、磁気粒子のほか、放射性核種、酵素、蛍光剤、化学発光剤、発色剤などがある。

40

【0182】

SECPをコードするヌクレオチド配列で形質転換された宿主細胞は、細胞培地でのこのタンパク質の発現及び回収に好適な条件下で培養され得る。形質転換細胞から製造されたタンパク質が分泌されるか細胞内に留まるかは、使用される配列、ベクター、あるいはその両者に依存する。SECPをコードするポリヌクレオチドを含む発現ベクターは、原核細胞膜及

50

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

【0183】

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

10

【0184】

本発明の別の実施例では、SECPをコードする天然の核酸配列、変更された核酸配列、または組換えの核酸配列を上記した任意の宿主系の融合タンパク質の翻訳となる異種配列に連結させ得る。例えば、市販の抗体によって認識できる異種部分を含むキメラSECPタンパク質が、SECP活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分および異種ペプチド部分も、市販されている親和性マトリックスを用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ (GST)、マルトース結合タンパク質 (MBP)、チオレドキシン (Trx)、カルモジュリン結合ペプチド (CBP)、6-His、FLAG、c-myc、赤血球凝集素 (HA) がある。GSTは固定化グルタチオン上で、MBPはマルトース上で、Trxはフェニルアルシンオキシド上で、CBPはカルモジュリン上で、そして6-Hisは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。FLAG、c-mycおよび赤血球凝集素 (HA) は、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体およびポリクローナル抗体を用いた、融合タンパク質の免疫親和性精製を可能にする。また、SECPをコードする配列と異種タンパク質配列との間にあるタンパク質分解切断部位を融合タンパク質が含むように遺伝子操作すると、SECPが精製の後に異種部分から切断され得る。融合タンパク質の発現および精製方法は、前出のAusubel (1995) 10章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現および精製を促進することもできる。

20

30

【0185】

本発明の別の実施例では、TNTウサギ網状赤血球可溶化液またはコムギ胚芽抽出系 (Promega) を用いて *in vitro* で放射能標識したSECPの合成が可能である。これらの系は、T7、T3またはSP6プロモーターと機能的に連結したタンパク質コード配列の転写及び翻訳をカップルさせる。翻訳は、例えば³⁵Sメチオニンのような放射能標識したアミノ酸前駆体の存在下で起こる。

【0186】

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

40

【0187】

一実施例では、このように同定された化合物は、例えばリガンドやその断片などのSECPの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している (Coligan, J.E. 他 (1991) *Current Protocols in Immunology* 1(2)の5章等を参照)。同様に、化合物は、SECPが結合する天然受容体、あるいは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連する場合がある。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。ある実施態様では、このような化合物に対するスクリーニングには、分泌タンパク質あるいは細胞

50

膜上のタンパク質のいずれか一方としてSECPを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョウバエ、あるいは大腸菌からの細胞が含まれる。SECPを発現する細胞またはSECPを含有する細胞膜断片を試験化合物と接触させて、SECPまたは化合物のいずれかの結合、刺激または阻害を分析する。

【0188】

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

10

【0189】

本発明のSECPまたはその断片を用いて、SECPの活性を調整する化合物をスクリーニングすることが可能である。このような化合物としては、アゴニスト、アンタゴニスト、部分的アゴニストまたは逆アゴニストなどが含まれ得る。ある実施態様では、SECPの活性が許容される条件下でアッセイを実施し、そのアッセイでは少なくとも1つの試験化合物をSECPと混合し、試験化合物の存在下でSECPの活性を試験化合物不在下でのSECPの活性と比較する。試験化合物の存在下でのSECPの活性の変化は、SECPの活性を調整する化合物の存在を示唆する。別の実施態様において、試験化合物をSECPの活性に適した条件下でSECPを含む *in vitro* または無細胞再構成系と結合させてアッセイを実施する。これらアッセイのいずれかにおいて、SECPの活性を調整する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも1つ、または複数の試験化合物をスクリーニングすることができる。

20

【0190】

別の実施態様では、胚性幹細胞（ES細胞）における相同組み換えを用いて動物モデル系内で、SECPまたはその哺乳動物相同体をコードするポリヌクレオチドを「ロックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号および第5,767,337号などを参照）。例えば129/SvJ細胞系などのマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo: Capecchi, M.R. (1989) Science 244:1288-1292）などのマーカー遺伝子で破壊した、目的の遺伝子を持つベクターで形質転換する。このベクターは、相同組換えにより、宿主ゲノムの対応する領域に組込まれる。或いは、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をロックアウトする（Marth, J.D. (1996) Clin. Invest. 97:1999-2002、Wagner, K.U.他 (1997) Nucleic Acids Res. 25:4323-4330）。形質転換したES細胞を同定し、例えばC57BL/6マウス株などから採取したマウス細胞胚盤胞に微量注入する。これらの胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を特定し、これらを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、潜在的な治療薬や毒性薬物で検査されうる。

30

40

【0191】

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

【0192】

SECPをコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ロックイン」ヒト化動物（ブタ）または遺伝子組み換え動物（マウスまたはラット）を作製することが可

50

能である。ノックイン技術を用いて、SECPをコードするポリヌクレオチドのある領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について試験し、潜在的医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばSECPを乳汁内に分泌するなどSECPを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る (Janne, J. 他 (1998) *Biotechnol. Annu. Rev.* 4:55-74)。

【0193】

(治療)

SECPのいくつかの領域と分泌タンパク質群との間には、例えば配列およびモチーフの文脈における、化学的および構造的類似性が存在する。更にSECPの発現は、罹患した乳房、脳、脳腫瘍、病変した脳、神経、発達及び病変した子宮内膜組織、骨癌、副腎及び脳組織、また肺、卵巣、樹状細胞の腫瘍に随伴する隣接組織、腎癌、鼻、篩状、卵巣腫瘍組織に随伴する腎臓皮質組織と密接に関連する。また、SECPを発現する組織の数例は、表6を見られたい。このように、SECPは、肝臓疾患、細胞増殖異常、自己免疫/炎症疾患、心血管疾患、神経系疾患、および発達異常において、或る役割を果たすと考えられる。SECPの発現若しくは活性の増大に関連する疾患の治療においては、SECPの発現または活性を低下させることが望ましい。また、SECPの発現または活性の低下に関連する疾患の治療においては、SECPの発現または活性を増大させることが望ましい。

【0194】

したがって、一実施態様において、SECPの発現または活性の低下に関連した疾患の治療または予防のために、患者にSECPまたはその断片や誘導体を投与することが可能である。限定するものではないが、このような疾患には、肝臓疾患の中には、例えば腺腫症、胆汁鬱滞、肝硬変、血管腫、アレルギー性紫斑病、肝炎、肝細胞癌、転移性癌、特発性血小板減少性紫斑病、ポルフィリン症、サルコイドーシス、ウィルソン病、また肝臓の疾患の診断、ステロイドを使用した治療への代謝及び毒性反応の検出があり、細胞増殖異常の中には日光角化症、動脈硬化、アテローム硬化、滑液包炎、硬変、肝炎、混合性結合組織病 (MCTD)、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに、腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、および奇形癌など癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、および子宮の癌が含まれ、自己免疫/炎症性の疾患の中には、後天性免疫不全症候群 (AIDS)、アジソン病 (慢性原発性副腎機能不全)、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー (APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球傷害因子性偶発性リンパ球減少症、新生児溶血性疾患 (胎児赤芽球症)、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、血小板減少性紫斑病、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、心血管障害の中には、鬱血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪部石灰化 (mitral annular calcification)、僧帽弁逸脱、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性エリテマトーデスの心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心移植の合併症、動静脈瘤、アテローム硬化、高血圧、脈管炎、レイノー病、動脈瘤、動脈解離、静脈瘤、血栓静脈炎および静脈血栓、血管系腫瘍、血栓溶解の合併症、バルーン血管形成術、血管置換術 (vascular replaceme

10

20

30

40

50

nt)、冠動脈バイパス移植術が含まれ、神経障害の中には、癲癇、虚血性脳血管障害、脳卒中、脳腫瘍、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病およびその他の錐体外路障害、筋萎縮性側索硬化およびその他の運動ニューロン障害、進行性神経性筋萎縮症、網膜色素変性症(色素性網膜炎)、遺伝性運動失調、多発性硬化症および他の脱髄疾患、細菌性およびウイルス性髄膜炎、脳膿瘍、硬膜下膿瘍、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎および神経根炎、ウイルス性中枢神経系疾患と、プリオン病(クールー、クロイツフェルト ヤコブ病、およびGerstmann-Straussler-Scheinker症候群を含む)、致死性家族性不眠症、神経系性栄養病および代謝病、神経線維腫症、結節硬化症、小脳網膜血管腫症(cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、ダウン症を含む中枢神経系性精神遅滞および他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄疾患、筋ジストロフィー他の神経筋障害、末梢神経疾患、皮膚筋炎および多発性筋炎と、遺伝性、代謝性、内分泌性、および中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺、精神障害(気分性、不安性の障害、統合失調症/分裂病)、季節性感情障害(SAD)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、遅発性ジスキネジア、ジストニー、パラノイド精神病、帯状疱疹後神経痛、トゥレット病、進行性核上麻痺、大脳皮質基底核変性(corticobasal degeneration)、および家族性前頭側頭型痴呆とが含まれ、また発達障害も含まれ、その中には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ/ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神遅滞)、スミス マジェニス症候群(Smith-Magenis syndrome)、骨髄異形成症候群、遺伝性粘膜上皮異形成、遺伝性角皮症や、シャルコーマリーツース病および神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症や、Sydenham舞蹈病(Sydenham's chorea)および脳性小児麻痺などの発作障害、二分脊椎、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感音難聴が含まれる。

10

20

【0195】

別の実施態様では、SECPまたはその断片や誘導体を発現し得るベクターを患者に投与して、上記しだけに限られるものではないが疾患を含むSECPの発現または活性の低下に関連した疾患を治療または予防することも可能である。

【0196】

さらに別の実施態様では、実質的に精製されたSECPを含む組成物を好適な医薬用キャリアと共に患者に投与して、限定するものではないが上記した疾患を含むSECPの発現または活性の低下に関連した疾患を治療または予防することも可能である。

30

【0197】

更に別の実施例では、限定するものではないが上に列記した疾患を含むSECPの発現または活性の低下に関連した疾患の治療または予防のために、SECPの活性を調節するアゴニストを患者に投与することも可能である。

【0198】

更なる実施例では、SECPの発現または活性の増大に関連した疾患の治療または予防のために、患者にSECPのアンタゴニストを投与することが可能である。限定するものではないが、このような疾患の例には、上記した肝臓疾患、細胞増殖異常、自己免疫/炎症疾患、心血管障害、神経障害、および発生障害が含まれる。一実施態様では、SECPと特異的に結合する抗体が直接アンタゴニストとして、或いはSECPを発現する細胞または組織に薬剤を運ぶ標的化機構、或いは送達機構として間接的に用いられ得る。

40

【0199】

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

【0200】

別の実施態様では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補配列、またはベクターを、別の好適な治療薬と組合せて投与することもできる。併用療

50

法で用いる好適な治療薬は、当業者が従来の医薬原理に従って選択し得る。治療薬と組合せることにより、上記した種々の疾患の治療または予防に相乗効果をもたらす得る。この方法により、少量の各薬物で医薬効果をあげることが可能となり、それによって副作用の可能性を低減し得る。

【0201】

SECPのアンタゴニストは、当分野で一般的な方法を用いて製造することが可能である。詳しくは、精製されたSECPを用いて抗体を作ったり、治療薬のライブラリをスクリーニングしてSECPと特異的に結合するものの同定が可能である。SECPに対する抗体も、当分野で一般的な方法を用いて製造することが可能である。このような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖、Fab断片、及びFab発現ライブラリによって作られた断片が含まれる。但し、これらに限定されるものではない。中和抗体（すなわち二量体の形成を阻害する抗体）は通常、治療用に好適である。一本鎖抗体（例えば、ラクダまたはラマ）は有力な酵素阻害剤であり、またペプチド擬態物質の設計および免疫吸着剤やバイオセンサーの開発に利点があるであろう(Muyldermans, S. (2001) J. Biotechnol. 74:277-302)。

10

【0202】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ラクダ、ヒトコブラクダ、ラマ、ヒト及びその他のものを含む種々の宿主が、SECPまたは任意の断片、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。限定するものではないがこのようなアジュバントには、フロイントアジュバントと、水酸化アルミニウムなどのミネラルゲルアジュバントと、リゾレシチン、プルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、スカシガイのヘモシアニン（KLH）、ジニトロフェノールなどの界面活性剤とがある。ヒトに用いられるアジュバントの中では、BCG（カルメット ゲラン杆菌）及びコリネバクテリウム パルバム（*Corynebacterium parvum*）が特に好ましい。

20

【0203】

SECPに対する抗体を誘発するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約5個のアミノ酸からなり、一般的には約10個以上のアミノ酸からなるものが好ましい。これらのオリゴペプチド、ペプチドまたは断片は、天然のタンパク質のアミノ酸配列の一部と同一であることも望ましい。SECPのアミノ酸群の短い区間を、KLHなどの別のタンパク質の配列と融合し、キメラ分子に対する抗体を産生し得る。

30

【0204】

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

【0205】

更に、ヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの「キメラ抗体」作製のために開発した技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる（例えば、Morrison, S.L.他（1984）Proc. Natl. Acad. Sci. 81 - 4851 - 4855、Neuberger, M.S.他（1984）Nature 312:604-608、Takeda, S.他（1985）Nature 314:452,454を参照）。別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、SECP特異的一本鎖抗体を生成する。関連した特異性を有するがイデオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリ類からチェーンシャッフリングによって産生することもできる（Burton D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137等を参照）。

40

【0206】

抗体の産生は、リンパ球集団における *in vivo* 産生の誘導によって、或いは文献に開示さ

50

れているように非常に特異的な結合試薬の免疫グロブリンのライブラリまたはパネルのスクリーニングによっても行い得る（例えばOrlandi, R.他 (1989) Proc. Natl. Acad. Sci. USA 86:38333837、Winter, G. 他. (1991) Nature 349:293299を参照）。

【0207】

SECPに対する特異的な結合部位を含む抗体も得ることができる。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製される $F(ab')_2$ 断片と、 $F(ab')_2$ 断片のジスルフィド架橋を還元することによって作製されるFab断片とがある。或いは、Fab発現ライブラリを作製することによって、所望の特異性を持つモノクローナルFab断片を迅速且つ容易に同定することが可能となる(Huse, W.D他 (1989) Science 246:12751281等を参照)。

10

【0208】

種々の免疫学的検定(イムノアッセイ)を用いてスクリーニングすることにより、所望の特異性を有する抗体を同定し得る。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的結合アッセイ、または免疫放射定量測定法のための数々のプロトコルが、当分野では周知である。通常このような免疫測定法には、SECPとその特異性抗体との間の複合体の計測が含まれる。二つの非干渉性SECPエピトープに対して反応性のモノクローナル抗体を用いる、2部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合的結合アッセイも利用することができる(Pound、前出)。

【0209】

放射免疫測定法技術と共にScatchard分析などの様々な方法を用いて、SECPに対する抗体の親和性を評価する。親和性を結合定数 K_a で表すが、この K_a は、平衡状態の下でSECP抗体複合体のモル濃度を遊離抗体と遊離抗原のモル濃度で除して得られる値である。多数のSECPエピトープに対して親和性が不均一なポリクローナル抗体試薬の K_a は、SECPに対する抗体の平均親和性または結合活性を表す。特定のSECPエピトープに単一特異的なモノクローナル抗体試薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12}$ liter/molの高親和性抗体試薬は、SECP抗体複合体が過酷な処理に耐えなければならないイムノアッセイに用いるのが好ましい。 K_a 値が $10^6 \sim 10^7$ L/molの低親和性抗体試薬は、SECPが抗体から最終的に活性化状態で解離する必要がある免疫精製(immunopurification)及び類似の処理に用いるのが好ましい(Catty, D. (1988) Antibodies, Volume I: A Practical Approach. IRL Press, Washington, DC; Liddell, J. E.及びCryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

20

30

【0210】

ポリクローナル抗体製剤の抗体価及び結合活性を更に評価して、後に使う或る適用例に対するこのような製剤の品質及び適性を決定することができる。例えば、少なくとも $1 \sim 2$ mg/mlの特異的な抗体、好ましくは $5 \sim 10$ mg/mlの特異的な抗体を含むポリクローナル抗体試薬は一般に、SECP抗体複合体を沈殿させなければならない処理に用いられる。抗体の特異性、抗体価、結合活性、様々な適用例における抗体の品質や使用に対する指針については、一般に入手可能である(前出のCattyの文献、同Coligan他の文献等を参照)。

【0211】

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

40

【0212】

治療に用いる場合、アンチセンス配列を適切な標的細胞に導入するのに好適な、任意の遺

50

伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を作製する発現プラスミドの形で細胞内に送達することが可能である（例えばSlater, J.E. 他 (1998) J. Allergy Clin. Immunol. 102(3):469-475、Scanlon, K.J. 他 (1995) 9(13):1288-1296を参照）。アンチセンス配列はまた、例えばレトロウイルスやアデノ随伴ウイルスベクターなどのウイルスベクターを用いて細胞内に導入することもできる（Miller, A.D. (1990) Blood 76:271、前出のAusubel、Uckert, W.及びW. Walther (1994) Pharmacol. Ther. 63(3):323-347等を参照）。その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロープおよび当分野で公知のその他のシステムが含まれる（例えばRossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J.他 (1998) J. Pharm. Sci. 87(11):1308-1315、Morris, M.C. 他 (1997) Nucleic Acids Res. 25(14):2730-2736を参照）。

【0213】

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

【0214】

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

【0215】

限定するものではないがSECPの発現に有効であり得る発現ベクターには、PCDNA 3.1、EPI TAG、PRCCMV2、PREP、PVAX、PCR2-TOPOTAベクター（Invitrogen, Carlsbad CA）、PCMV-S CRIP、PCMV-TAG、PEGSH/PERV（Stratagene, La Jolla CA）及びPTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG（Clontech, Palo Alto CA）がある。SECPを発現させるために、(i) 恒常的に活性なプロモーター（例えば、サイトメガロウイルス（CMV）、ラウス肉腫ウイルス（RSV）、SV40ウイルス、チミジンキナーゼ（TK）、若しくは - アクチン遺伝子等）、(ii) 誘導性プロモーター（例えば、市販されているT-REXプラスミド（Invitrogen）に含まれている、テトラサイクリン調節性プロモーター（Gossen, M及びH. Bujard (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551、Gossen, M. 他 (1995) Science 26

8:1766-1769、Rossi, F.M.V.及びH.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456)、エクジソン誘導性プロモーター(市販されているプラスミドPVGRXR及びPINDに含まれている: Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストーン誘導性プロモーター(Rossi, F.M.V.及びH.M. Blau, 前出)、または(ii)正常な個体に由来するSECPをコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いることが可能である。

【0216】

市販のリポソーム形質転換キット(例えばInvitrogen社のPerFect Lipid Transfection Kit)を用いれば、当業者はポリヌクレオチドを培養中の標的細胞に送達することが可能になる。また、実験パラメータ群を最適化するのに必要な努力が最小限になる。別法では、リン酸カルシウム法(Graham, F.L.及びA.J. Eb (1973) *Virology* 52:456-467)若しくは電気穿孔法(Neumann, B. 他 (1982) *EMBO J.* 1:841-845)を用いて形質転換を行う。初代培養細胞にDNAを導入するためには、標準化された哺乳類の形質移入プロトコルの修正が必要である。

10

【0217】

本発明の別の実施例では、SECPの発現に関連する遺伝子欠損によって起こる疾患や異常症は、(i)レトロウイルス末端反復配列(LTR)プロモーター若しくは独立したプロモーターのコントロール下でSECPをコードするポリヌクレオチドと、(ii)好適なRNAパッケージングシグナルと、(iii)追加のレトロウイルス・シス作用性RNA配列及び効率的なベクターの増殖に必要なコーディング配列を伴うRev応答性エレメント(RRE)とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター(例えばPFB及びPFBNE0)はStratagene社から市販されており、刊行データ(Riviere, I. 他 (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92:6733-6737)に基づいている。上記データを引用することをもって本明細書の一部とする。このベクターは、好適なベクター産生細胞株(VPCL)において増殖される。VPCLは、各標的細胞上の受容体への親和性を持つエンベロープ遺伝子を、またはVSVgなど汎親和性エンベロープタンパク質を発現する(Armentano, D. 他 (1987) *J. Virol.* 61:1647-1650、Bender, M.A. 他 (1987) *J. Virol.* 61:1639-1646、Adam, M.A及びA.D. Miller (1988) *J. Virol.* 62:3802-3806、Dull, T. 他 (1998) *J. Virol.* 72:8463-8471、Zufferey, R. 他 (1998) *J. Virol.* 72:9873-9880)。Riggに付与された米国特許第5,910,434号(「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」)において、レトロウイルスパッケージング細胞株を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクター類の繁殖や、細胞集団(例えばCD4⁺T細胞群)の形質導入、および形質導入した細胞群の患者への戻しは、遺伝子治療の分野では当業者に周知の手法であり、多数の文献に記載がある(Ranga, U.他 (1997) *J. Virol.* 71:7020-7029、Bauer, G. 他 (1997) *Blood* 89:2259-2267、Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716、Ranga, U. 他 (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206、Su, L. (1997) *Blood* 89:2283-2290)。

20

30

【0218】

別法では、アデノウイルス系遺伝子治療の送達系を用いて、SECPの発現に関連する1つ或いは複数の遺伝子異常を有する細胞にSECPをコードするポリヌクレオチドを送達する。アデノウイルス系ベクター類の作製およびパッケージングについては、当業者に周知である。複製欠損型アデノウイルスベクター類は、種々の免疫調節タンパク質をコードする遺伝子群を、無損傷の臍島内に導入する目的で多様に利用し得ることが証明された(Csete, M.E.他 (1995) *Transplantation* 27:263268)。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第5,707,618号(「Adenovirus vectors for gene therapy」)に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについては、Antinozzi, P.A. 他 (1999) *Annu. Rev. Nutr.* 19:511-544、Verma, I.M.及びN. Somia (1997) *Nature* 18:389:239-242も参照されたい。両文献は、引用することをもって本明細書の一部とする。

40

50

【0219】

別法では、ヘルペス系遺伝子治療の送達系を用いて、SECPの発現に関連する1つ或いは複数の遺伝子異常を有する標的細胞にSECPをコードするポリヌクレオチドを送達する。単純ヘルペスウイルス(HSV)系のベクターは、HSV親和性の中枢神経細胞にSECPを導入する際に特に有用であり得る。ヘルペス系ベクター類の作製およびパッケージングは、当業者に公知である。或る複製適格性単純ヘルペスウイルス(HSV)1型系のベクターが、或るレポーター遺伝子の、霊長類の眼への送達に用いられている(Liu, X. 他(1999) *Exp. Eye Res.* 169:385395)。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号(「Herpes simplex virus strains for gene transfer」)に開示されており、該特許の引用をもって本明細書の一部とする。米国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも1つの外在性遺伝子を有するゲノムを含む組換えHSV d92の使用についての記載がある。上記特許はまた、ICP4、ICP27及びICP22を欠失した組換えHSV系統の作製及び使用について開示している。HSVベクターについては、Goins, W.F. 他(1999) *J. Virol.* 73:519-532、Xu, H. 他(1994) *Dev. Biol.* 163:152-161も参照されたい。両文献は、引用をもって本明細書の一部とする。クローン化ヘルペスウイルス配列の操作、巨大ヘルペスウイルスのゲノムの異なったセグメント群を含む多数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。

10

【0220】

別法では、ウイルス(正の一本鎖RNAウイルス)ベクターを用いてSECPをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのウイルスであるセムリキ森林熱ウイルス(SFV)の生物学的研究が広範に行われており、遺伝子導入ベクターはSFVゲノムに基づいている(Garoff, H.及びK.-J. Li(1998) *Cun. Opin. Biotech.* 9:464-469)。

20

ウイルスRNAの複製中に、通常はウイルスのカプシドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAは、完全長のゲノムRNAより高いレベルに複製されるため、酵素活性(例えばプロテアーゼ及びポリメラーゼ)を有するウイルスタンパク質に比べてカプシドタンパク質が過剰産生される。同様に、SECPをコードする配列をウイルスゲノムのキャプシドをコードする領域に導入することによって、ベクター導入細胞において多数のSECPをコードするRNAが産生され、高いレベルでSECPが合成される。通常、ウイルスの感染は、数日以内の細胞溶解を伴う。一方、シンドビスウイルス(SIN)の或る変異体を有するハムスター正常腎臓細胞(BHK-21)群が持続的な感染を確立する能力は、ウイルス類の溶解複製を、遺伝子治療に応用し得るように好適に改変可能であることを示唆する(Dryga, S.A.他(1997) *Virology* 228:74-83)。様々な宿主にウイルスを導入できることから、様々なタイプの細胞にSECPを導入することできる。或る集団における或るサブセットの細胞群の特異的形質導入は、形質導入前に細胞の選別を必要とし得る。ウイルスの感染性cDNAクローンの処置方法、ウイルスのcDNAおよびRNAの形質移入方法およびウイルスの感染方法は、当業者に公知である。

30

【0221】

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

40

【0222】

50

リボザイムは酵素的RNA分子であり、RNAの特異的切断を触媒するためにリボザイムを用いることもできる。リボザイム作用のメカニズムは、相補的標的RNAへのリボザイム分子の配列特異的ハイブリダイゼーションとその後に起こる内ヌクレオチド鎖切断に参与している。例えば、人工的に作製されたハンマーヘッド型リボザイム分子が、SECPをコードする配列の内ヌクレオチド鎖分解性の切断を特異的且つ効果的に触媒できる可能性がある。

【0223】

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

10

【0224】

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

20

【0225】

細胞内の安定性を高め、半減期を長くするためにRNA分子を修飾することができる。限定するものではないが可能な修飾としては、分子の5'末端、3'末端、あるいはその両方において隣接配列群を追加することや、分子の主鎖内においてホスホジエステラーゼ結合ではなくホスホロチオエートまたは2' O-メチルを使用することが含まれる。この概念は、PNAの産出に固有のものであるが、これら全ての分子に拡大することができる。そのためには、内因性エンドヌクレアーゼによって容易には認識されない、アデニン、シチジン、グアニン、チミン、及びウリジンにアセチル -、メチル -、チオ - 及び同様の修飾をしたものや、非従来型塩基、例えばイノシン、クエオシン (queosine)、ワイプトシン (wybutosine) 等を含める。

30

【0226】

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

40

【0227】

特異ポリヌクレオチドの発現改変における有効性に対して、少なくとも1個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を変異させる場合と、既に市販のまたは私的な、天然または非天然の化合物ライブラリから選択する場合と、標的ポリヌクレオチドの化学的及び/または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に

50

有効であることが知られているような化合物の化学修飾がある。SECPをコードするポリヌクレオチドを含むサンプルは、このようにして得られた試験化合物の少なくとも1つに曝露する。サンプルには例えば、無傷細胞、または透過処理した細胞、あるいは *in vitro* 無細胞系すなわち再構成生化学系があり得る。SECPをコードするポリヌクレオチドの発現における変化は、当分野で通常知られている任意の方法でアッセイする。通常、SECPをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特異ヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量し、それによって1つ以上の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を改変する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの発現改変に有効な化合物に対して、例えば分裂酵母 (*Schizosaccharomyces pombe*) 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5,932,435号、Arndt, G.M. 他 (2000) *Nucleic Acids Res.* 28:E15) またはHeLa細胞等のヒト細胞系 (Clarke, M.L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268:8-13) を用いてスクリーニングを実行する。本発明の或る特定の実施態様は、或る特定ポリヌクレオチド配列に対するアンチセンス活性について、オリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾したオリゴヌクレオチド) の組合せライブラリをスクリーニングする過程に関する (Bruce, T.W. 他 (1997) U.S. Patent No. 5,686,242、Bruce, T.W. 他 (2000) U.S. Patent No. 6,022,691)。

10

20

30

40

50

【0228】

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

【0229】

上記の治療方法はいずれも、例えば、ヒト、イヌ、ネコ、ウシ、ウマ、ウサギ、サルなどの哺乳類を含めて治療が必要な全ての被験体に適用できる。

【0230】

本発明のさらなる実施態様は、通常薬剤として許容できる賦形剤で処方される活性成分を有する組成物の投与に関連する。賦形剤には例えば、糖、でんぷん、セルロース、ゴムおよびタンパク質がある。様々な剤型が広く知られており、詳細は *Remington's Pharmaceutical Sciences* (Maack Publishing, Easton PA) の最新版に記載されている。このような組成物は、SECP、SECPの抗体、擬態、アゴニスト、アンタゴニスト、またはSECPのインヒビターなどからなる。

【0231】

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

【0232】

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

【0233】

本発明での使用に適した組成物には、所定の目的を達成するために必要なだけの量の活性

成分を含有する組成物が含まれる。有効投与量の決定は、当業者の能力の範囲内で行う。

【0234】

SECPまたはその断片を含む高分子を直接細胞内に送達するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、その高分子の細胞融合と細胞内送達とを促進し得る。別法では、SECPまたはその断片をHIV Tat-1タンパク質の陽イオン性N末端部に結合することもできる。このようにして生成された融合タンパク質類は、或るマウスモデル系の、脳を含む全ての組織の細胞群に形質導入することがわかっている (Schwarze, S.R. 他 (1999) Science 285:1569-1572)。

【0235】

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

【0236】

治療有効量は、症状や容態を回復させる活性処方成分 (例えば、SECPまたはその断片、SECPの抗体、SECPのアゴニストまたはアンタゴニスト、インヒビターなど) 量を意味する。治療有効度及び毒性は、細胞培養または動物実験における標準的な薬剤手法によって、例えばED₅₀ (集団の50%の治療有効量) またはLD₅₀ (集団の50%の致死量) を測定するなどして決定することができる。毒性効果の治療効果に対する投与量比が治療指数であり、LD₅₀/ED₅₀比として表すことができる。高い治療指数を示すような組成物が望ましい。細胞培養アッセイ及び動物実験から得られたデータは、ヒトに用いるための投与量の範囲を策定するのに用いられる。このような組成物に含まれる投与量は、毒性を殆どあるいは全く持たず、ED₅₀を含むような血中濃度の範囲にあることが好ましい。用いられる投与形態、患者の感受性および投与の経路によって、投与量はこの範囲内で様々に変わる。

【0237】

正確な投与量は、治療が必要な被験者に関する要素を考慮して、現場の医者が決定することになる。十分なレベルの活性成分を与え、あるいは所望の効果を維持すべく、用法および用量を調整する。被験者に関する要素としては、疾患の重症度、被験者の全身健康状態、被験者の年齢、体重及び性別 (ジェンダー)、投与の時間及び頻度、薬剤の併用、反応感受性及び治療に対する応答等を考慮しうる。作用期間が長い組成物は、特定の薬剤の半減期及びクリアランス率によって、3~4日毎に1度、1週間に1度、或いは2週間に1度の間隔で投与し得る。

【0238】

通常の投与量は、投与の経路にもよるが約0.1~100.000 µgであり、合計で約1gまでとする。特定の投与量及び送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。当業者は、タンパク質またはインヒビターに対する処方とは異なる、ヌクレオチドに対する処方を利用することになる。同様に、ポリヌクレオチドまたはポリペプチドの送達は、特定の細胞、症状、部位などに特異的なものとなる。

【0239】

(診断)

別の実施例では、SECPに特異的に結合する抗体が、SECPの発現によって特徴付けられる疾患の診断、またはSECPやSECPのアゴニストまたはアンタゴニスト、インヒビターで治療を受けている患者をモニターするためのアッセイに用いられる。診断目的に有用な抗体は、上記の治療の箇所に記載した方法と同じ方法で調合される。SECPの診断アッセイには、抗体及び標識を用いてヒトの体液或いは細胞や組織から採取されたものからSECPを検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記

で説明されている。

【0240】

SECPを測定するためのELISA, RIA, 及びFACSを含む種々のプロトコルは、当分野では周知であり、変わった或いは異常なレベルのSECPの発現を診断する元となるものを提供する。正常或いは標準的なSECPの発現の値は、複合体の形成に適した条件下で、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞とSECPに対する抗体とを混合させることによって決定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者、対照、及び疾患生検組織からの各サンプルのSECPの発現の量が基準値と比較される。基準値と被験体との偏差が、疾患を診断するパラメータを確定する。

【0241】

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

10

【0242】

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

20

【0243】

プローブはまた、関連する配列の検出に利用され、SECPをコードする任意の配列と少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO:31-60の配列、或いはSECP遺伝子のプロモーター、エンハンサー、イントロンを含むゲノム配列に由来し得る。

30

【0244】

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

40

【0245】

SECPをコードするポリヌクレオチド配列を用いて、SECPの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患には、肝臓疾患の中には、例えば腺腫症、胆汁鬱滞、肝硬変、血管腫、アレルギー性紫斑病、肝炎、肝細胞癌、転移性癌、特発性血小板減少性紫斑病、ポルフィリン症、サルコイドーシス、ウィルソン病、また肝臓の疾患の診断、ステロイドを使用した治療への代謝及び毒性反応の検出があり、細胞増殖異常の中には日光角化症、動脈硬化、アテローム硬化、滑液包炎、硬変、肝炎、混合性結合組織病（MCTD）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに、腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、および奇形癌など癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、

50

消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、および子宮の癌が含まれ、自己免疫/炎症性の疾患の中には、後天性免疫不全症候群(AIDS)、アジソン病(慢性原発性副腎機能不全)、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球傷害因子性偶発性リンパ球減少症、新生児溶血性疾患(胎児赤芽球症)、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、血小板減少性紫斑病、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、心血管障害の中には、鬱血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪部石灰化(mitral annular calcification)、僧帽弁逸脱、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性エリテマトーデスの心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心移植の合併症、動静脈瘻、アテローム硬化、高血圧、脈管炎、レイノー病、動脈瘤、動脈解離、静脈瘤、血栓静脈炎および静脈血栓、血管系腫瘍、血栓溶解の合併症、バルーン血管形成術、血管置換術(vascular replacement)、冠動脈バイパス移植術が含まれ、神経障害の中には、癲癇、虚血性脳血管障害、脳卒中、脳腫瘍、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病およびその他の錐体外路障害、筋萎縮性側索硬化およびその他の運動ニューロン障害、進行性神経性筋萎縮症、網膜色素変性症(色素性網膜炎)、遺伝性運動失調、多発性硬化症および他の脱髄疾患、細菌性およびウイルス性髄膜炎、脳膿瘍、硬膜下膿瘍、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎および神経根炎、ウイルス性中枢神経系疾患と、プリオン病(クールー、クロイツフェルト ヤコブ病、およびGerstman-Straussler-Scheinker症候群を含む)、致死性家族性不眠症、神経系性栄養病および代謝病、神経線維腫症、結節硬化症、小脳網膜血管腫症(cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、ダウン症を含む中枢神経系性精神遅滞および他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄疾患、筋ジストロフィー他の神経筋障害、末梢神経疾患、皮膚筋炎および多発性筋炎と、遺伝性、代謝性、内分泌性、および中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺、精神障害(気分性、不安性の障害、統合失調症/分裂病)、季節性感情障害(SAD)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、遅発性ジスキネジア、ジストニー、パラノイド精神病、帯状疱疹後神経痛、トゥレット病、進行性核上麻痺、大脳皮質基底核変性(corticobasal degeneration)、および家族性前頭側頭型痴呆とが含まれ、また発達障害も含まれ、その中には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ/ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神遅滞)、スミス マジェニス症候群(Smith-Magenis syndrome)、骨髄異形成症候群、遺伝性粘膜上皮異形成、遺伝性角皮症や、シャルコーマリーツス病および神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症や、Sydenham舞踏病(Sydenham's chorea)および脳性小児麻痺などの発作障害、二分脊椎、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感音難聴が含まれる。変異SECPの発現を検出するために、患者から採取した体液或いは組織を利用して、SECPをコードするポリヌクレオチド配列を、サザン法やノーザン法、ドットプロット法、或いはその他の膜系の技術、PCR法と、ディップスティック(dipstick)、ピン(pin)、およびマルチフォーマットのELISA様アッセイ、及びマイクロアレイに使用することが可能である。このような定性方法または定量方法は、当分野で公知である。

10

20

30

40

50

【0246】

ある実施態様では、SECPをコードするヌクレオチド配列は、関連する疾患、特に上記した疾患を検出するアッセイにおいて有用であろう。SECPをコードするヌクレオチド配列は、標準的な方法で標識化され、ハイブリダイゼーション複合体の形成に好適な条件下で、患者から採取した体液或いは組織のサンプルに加えることができるであろう。好適なインキュベーション期間が経過したらサンプルを洗浄し、シグナルを定量して標準値と比較する。患者のサンプルのシグナルの量が、対照サンプルと較べて著しく変わっている場合は、サンプル内のSECPをコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイは、動物実験、臨床試験における特定の治療効果を評価するため、あるいは個々の患者の治療をモニターするために用いることもできる。

10

【0247】

SECPの発現に関連する疾患の診断の基準となるものを提供するために、発現の正常すなわち標準的なプロファイルが確立される。これは、ハイブリダイゼーション或いは増幅に好適な条件下で、動物或いはヒトの何れかの正常な被験者から抽出された体液或いは細胞と、SECPをコードする配列或いはその断片とを結合させることにより達成され得る。実質的に精製されたポリヌクレオチドを既知量で用いて行った実験から得た値を正常な被験者から得た値と比較することにより、標準ハイブリダイゼーションを定量することができる。このようにして得た標準値は、疾患の徴候を示す患者から得たサンプルから得た値と比較することができる。標準値からの偏差を用いて疾患の存在を確定する。

【0248】

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

20

【0249】

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

30

【0250】

SECPをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、PCRの利用が含まれ得る。これらのオリゴマーは、化学的に合成するか、酵素により生産するか、あるいは*in vitro*で産出し得る。オリゴマーは、好ましくはSECPをコードするポリヌクレオチドの断片、或いはSECPをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件下で、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェント条件下で、近縁のDNA或いはRNA配列の検出、定量、或いはその両方のため用いることが可能である。

【0251】

或る実施態様において、SECPをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、一塩基多型性（SNP）を検出し得る。SNPは、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入及び欠失である。限定するものではないがSNPの検出方法には、SSCP(single-stranded conformation polymorphism)及び蛍光SSCP(fSSCP)法がある。SSCPでは、SECPをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応(PCR)でDNAを増幅する。DNAは例えば、病変組織または正常組織、生検サンプル、体液その他に由来し得る。DNA内のSNPは、一本鎖形状のPCR生成物の2次及び3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSSCPでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってDNAシーケンシング機などの高処理機器でアンプリマー(amplimer)の検出が可能になる。更に、インシリコSNP(in silico)

40

50

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

【0252】

ヒト疾患の遺伝的根拠を研究するためにSNPを用いることができる。例えば、少なくとも16の一般的SNPが、非インスリン依存型真性糖尿病と関連がある。SNPは、嚢胞性線維症、鎌状赤血球貧血あるいは慢性肉芽腫症等の単一遺伝子病の転帰の差異を研究する面でも有用である。例えば、マンノース結合レクチンの変異体であるMBL2は、嚢胞性線維症の肺での有害な転帰と関連することが示されてきた。SNPはまた、生命を脅かす毒性等の薬剤への患者の反応に影響する遺伝変異体の同定という薬理ゲノミクスにおいても有用性がある。例えば、ALOX5遺伝子のコア・プロモーターにおける或る変異は5-リポキシゲナーゼ経路を標的とする抗喘息剤を用いた治療への臨床反応の減少につながるが、Nアセチルトランスフェラーゼの或る変異は抗結核薬剤イソニアジドに反応する末梢神経障害の高発症率と関連する。異なる集団におけるSNP分布の分析は、集団の起源と移動の追跡以外にも遺伝的浮動、突然変異、組換え、選択の調査において有用である(Taylor, J.G. 他 (2001) Trends Mol. Med. 7:507-512、Kwok, P.-Y及びZ. Gu (1999) Mol. Med. Today 5:538-543、Nowotny, P. 他 (2001) Curr. Opin. Neurobiol. 11:637-641)。

10

20

【0253】

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

【0254】

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

30

40

【0255】

別の実施例では、SECP、SECPの断片、SECPに特異的な抗体をマイクロアレイ上のエレメントとして用いることができる。マイクロアレイを用いて、上記のようなタンパク質-タンパク質相互作用、薬物-標的相互作用および遺伝子発現プロファイルをモニターまたは測定することが可能である。

【0256】

或る実施態様は、或る組織または細胞タイプの転写イメージを作製する、本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプによる遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所与の条件下で所与の時

50

間に発現した遺伝子の数および相対存在量を定量することにより分析し得る (Seilhamer 他の特許第5,840,484号「Comparative Gene Transcript Analysis」を参照。当該文献は特に引用することを以って本明細書の一部となす)。従って、特定の組織または細胞タイプの転写または逆転写全体に本発明のポリヌクレオチドまたはその相補体をハイブリダイズすることにより、転写イメージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはその相補体がマイクロアレイ上のエレメントのサブセットを複数含むような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロファイルを提供し得る。

【0257】

転写イメージは、組織、細胞株、生検またはその生体サンプルから単離した転写物を用いて作製し得る。転写イメージはしたがって、組織または生検サンプルの場合には in vivo、細胞株の場合には in vitro での遺伝子発現を反映する。 10

【0258】

本発明のポリヌクレオチドの発現のプロフィールを作製する転写物イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、in vitroモデル系及び薬剤の前臨床評価に関連して使用し得る。全ての化合物は、作用および毒性のメカニズムを標示し、しばしば分子フィンガープリントまたは毒性シグネチャ (toxicant signatures) と称される、特徴的な遺伝子発現パターンを惹起する (Nuwaysir, E.F. 他 (1999) Mol. Carcinog. 24:153-159、Steiner, S及びN.L. Anderson (2000) Toxicol. Lett. 112-113:467-471、これらを引用することを以って本明細書の一部とする)。試験化合物が既知の毒性を有する化合物のシグネチャと類似のシグネチャを有する場合には、毒性特性を共有している可能性 20
がある。フィンガープリントまたはシグネチャは、多数の遺伝子および遺伝子ファミリーからの発現情報を含んでいる場合に、最も有用且つ正確である。理想的には、ゲノム全域にわたる発現の測定が、最高品質のシグネチャを提供する。たとえ、発現が任意の試験された化合物によって変容しない遺伝子があったとしても、それらの発現レベルを残りの発現データをノーマライズするために使用できるため、それらの遺伝子は重要である。ノーマライズ手順は、異なる化合物で処理した後の発現データの比較に有用である。毒物シグネチャの要素に遺伝子機能を割り当てるのが毒性メカニズムの解釈に役立つが、毒性の予測につながる、シグネチャを統計的に一致させる過程に、遺伝子機能の知識は必要とされない (例えば2000年2月29日に米国国立環境健康科学研究所 (National Institute of Environmental Health Sciences) より発行されたPress Release 00-02を参照されたい。 30
これについては<http://www.niehs.nih.gov/oc/news/toxchip.htm>で入手可能である)。従って、毒性シグネチャを用いる中毒学的スクリーニングの際に、全ての発現した遺伝子配列を含めることは重要且つ望ましいことである。

【0259】

或る実施例では、核酸を有する生体サンプルを試験化合物で処理することにより、この試験化合物の毒性を算定する。処理した生物学的サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1つ若しくは複数のプローブでハイブリダイズし、それによって本発明のポリヌクレオチドに対応する転写物レベルを定量し得る。処理した生物学的サンプル中の転写レベルを、未処理生物学的サンプル中のレベルと比較する。両サンプルの転写物レベルの差は、処理されたサンプル中で試験化合物が引き起こす毒性反応を示す。 40

【0260】

別の実施態様は、本発明のポリペプチド配列群を用いて或る組織または細胞タイプのプロテオームを分析することに関する。プロテオームの語は、特定の組織または細胞タイプでのタンパク質発現の包括的パターンを指す。プロテオームの各タンパク質成分は、個々に更なる分析の対象とすることができる。プロテオーム発現パターンすなわちプロファイルは、所与の条件下で所与の時間に発現したタンパク質の数および相対存在量を定量することにより分析し得る。したがって、或る細胞のプロテオームのプロファイルは、特定の組織または細胞タイプのポリペプチドを分離および分析することにより作成し得る。或る実施態様では、1次元等電点電気泳動によりサンプルからタンパク質を分離し、次に2次元 50

ドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に応じて分離するような2次元ゲル電気泳動により分離が達成される(前出のSteinerおよびAnderson)。タンパク質は、通常はクーマシーブルー、あるいは銀染色液または蛍光染色液などの物質を用いてゲルを染色することにより、分散した、独自の位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常、サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または未処理のいずれかの生物学的サンプルから得られる同等に位置するタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシーケンシングする。或るスポット内のタンパク質の同一性は、その部分配列を、好適には少なくとも5個の連続するアミノ酸残基を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列データが得られる。

10

【0261】

プロテオームのプロファイルは、SECPに特異的な抗体を用いてSECP発現レベルを定量することによっても作成可能である。或る実施態様では、マイクロアレイ上のエレメントとしてこれら抗体を用い、マイクロアレイをサンプルに曝して各アレイエレメントへのタンパク質結合レベルを検出することにより、タンパク質発現レベルを定量する(Lueking, A. 他(1999) Anal. Biochem. 270:103-111、Mendoze, L.G. 他(1999) Biotechniques 27:778-788)。検出は当分野で既知の様々な方法で行うことができ、例えば、チオール反応性またはアミノ反応性蛍光化合物とサンプル中のタンパク質を反応させ、各アレイエレメントにおける蛍光結合の量を検出し得る。

20

【0262】

プロテオームレベルでの毒性シグネチャも中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行に分析するべきである。数種の組織の数種のタンパク質に対しては、転写物とタンパク質との存在量の相関が乏しいので(Anderson, N.L.及びJ. Seilhamer(1997) Electrophoresis 18:533-537)、転写物イメージには有意に影響しないがプロテオームのプロフィールを変化させるような化合物の分析においてプロテオーム毒物シグネチャは有用たり得る。更に、体液中の転写物の分析はmRNAの急速な分解のために困難なので、プロテオームのプロファイル作成はこのような場合により信頼し得、情報価値があり得る。

30

【0263】

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

【0264】

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

40

【0265】

マイクロアレイは、本技術分野で既知の方法を用いて調製し、使用し、そして分析する(Brennan, T.M. 他(1995)の米国特許第5,474,796号、Schena, M. 他(1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler 他(1995) PCT出願第W095/251116

50

号、Shalon, D.他 (1995) PCT出願第W095/35505号、Heller, R.A. 他(1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M.J. 他 (1997) 米国特許第5,605,662号等を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, 編集. (1999) Oxford University Press, Londonに記載されている。該文献は、特に引用することを以って本明細書の一部となす。

【0266】

本発明の別の実施例ではまた、SECPをコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コード配列または非コード配列のいずれかを用いることができ、或る例では、コード配列よりも非コード配列が好ましい。例えば、多重遺伝子族のメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。配列は、或る特定の染色体に、または或る染色体の或る特定領域に、または人為形成の染色体、例えば、ヒト人工染色体 (HAC)、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、細菌P1産物、あるいは単一染色体cDNAライブラリ群に対してマッピングされる (例えばHarrington, J.J. 他 (1997) Nat. Genet. 15:345-355、Price, C.M. (1993) Blood Rev. 7:127-134、Trask, B.J. (1991) Trends Genet. 7:149-154を参照)。一度マッピングすると、本発明の核酸配列を用いて例えば或る病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型 (RFLP) と関連させるような遺伝子連鎖地図を発生させ得る。(例えば、Lander, E.S及びD. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

10

20

【0267】

蛍光原位置ハイブリッド形成法 (FISH) は、他の物理的及び遺伝地図データと関連し得る (例えばHeinz-Ulrich,他 (1995) in Meyers, 前出 965-968ページを参照)。遺伝地図データの例は、種々の科学雑誌あるいはOnline Mendelian Inheritance in Man (OMIM) のウェブサイトに見ることができる。物理的な染色体地図上のSECPをコードする遺伝子の位置と、特定の疾患との相関性、あるいは特定の疾患に対する素因との相関性は、この疾患と関連するDNAの領域の決定に役立ち得るため、位置を決定するクローニングの作業を促進し得る。

【0268】

確定した染色体マーカーを用いた結合分析等の物理的マッピング技術及び染色体標本原位置ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳動物の染色体上に遺伝子を配置することにより、正確な染色体上の遺伝子座が未知でも、関連するマーカー類をしばしば明らかにし得る。この情報は、位置クローニングその他の遺伝子発見技術を用いて疾患遺伝子を探す研究者にとって価値がある。疾患または症候群に關与する遺伝子が、血管拡張性失調症の11q22-23領域のように、特定のゲノム領域への遺伝的連関によって大まかに位置決めがなされると、該領域にマップされた任意の配列は、更なる調査のための関連遺伝子或いは調節遺伝子を提示している可能性がある (例えばGatti, R.A. 他 (1988) Nature 336:577580を参照)。転座、反転などに起因する、健常者、保有者、罹病者の三者間における染色体位置の相違を検出する場合にも、本発明のヌクレオチド配列を用い得る。

30

40

【0269】

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

【0270】

別の薬物スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化合物を高い処理能力でスクリーニングするために用いられる (Geysen, 他 (1984) PCT 出願番号 W084/03564等を参照)。この方法においては、多数の異なる小さな試験用化合物を

50

固体基板上で合成する。試験用化合物は、SECP、或いはその断片と反応してから洗浄される。次に、本技術分野でよく知られている方法で、結合したSECPを検出する。精製したSECPはまた、上記した薬剤のスクリーニング技術において用いるプレート上で直接コーティングすることもできる。別法では、非中和抗体を用いてペプチドを捕捉し、ペプチドを固体支持物に固定することもできる。

【0271】

別の実施例では、SECPと特異結合可能な中和抗体がSECPとの結合について試験用化合物と競合する、競合的薬剤スクリーニングアッセイを用いることができる。このようにして、SECPと1つ以上の抗原決定因子を共有するどのペプチドの存在をも、抗体を使って検出できる。

10

【0272】

別の実施例では、将来に開発される分子生物学技術が、現在知られているヌクレオチド配列の特性(限定はされないが、トリプレット遺伝コード、特異的な塩基対相互作用等を含む)に依存しているならば、SECPをコードするヌクレオチド配列をその新技術に用い得る。

【0273】

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

【0274】

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第60/285、207号、第60/287、114号、第60/288、640号、第60/290、516号、第60/292、184号、第60/343、553号、第60/358、279号、第60/366、041号、及び第60/357、002号は、言及することをもって本明細書の一部となす。

20

【実施例】

【0275】

1 cDNAライブラリの作製

Incyte cDNA群の由来は、LIFSEQ GOLDデータベース (Incyte Genomics, Palo Alto CA) に記載されたcDNAライブラリ群である。幾つかの組織はホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解し、他の組織はホモジナイズしてフェノールにまたは変性剤群の好適な混合液に溶解した。混合液の1例であるTRIZOL (Life Technologies) は、フェノールとグアニジンイソチオシアネートとの単相溶液である。結果として得られた溶解物は、塩化セシウムクッション上で遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、或いは別の方法を用いて、溶解物からRNAを沈殿させた。

30

【0276】

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

40

【0277】

場合によってはStratagene社へのRNA提供を行い、対応するcDNAライブラリをStratagene社が作製することもあった。そうでない場合は、UNIZAPベクターシステム (Stratagene) またはSUPERScriptプラスミドシステム (Life Technologies) を用いて本技術分野で既知の推奨方法または類似の方法でcDNAを合成し、cDNAライブラリを作製した(前出のAusubel, 1997, 5.1-6.6ユニット等を参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素または酵素でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズ(

50

300～1000bp) 選択は、SEPHACRYL S1000、SEPHAROSE CL2BまたはSEPHAROSE CL4Bカラムクロマトグラフィー (Amersham Pharmacia Biotech)、或いは分取アガロースゲル電気泳動法を用いて行った。cDNAは好適なプラスミドのポリリンカーの適合する制限酵素部位へ連結された。好適なプラスミドは、例えばPBLUESCRIPTプラスミド (Stratagene)、PSPORT1プラスミド (Life Technologies)、PCDNA2.1 プラスミド (Invitrogen, Carlsbad CA)、PBK-CMV プラスミド (Stratagene)、PCR2-TOPOTAプラスミド (Invitrogen)、PCMV-ICISプラスミド (Stratagene)、pIGEN (Incyte Genomics, Palo Alto CA)、pRARE (Incyte Genomics) またはpINCY (Incyte Genomics) 等およびその誘導体である。組換えプラスミドは、Stratagene社のXL1-Blue、XL1-BlueMRFまたはSOLR、或いはLife Technologies社のDH5、DH10BまたはELECTROMAX DH10Bを含むコンピテント大腸菌細胞に形質転換した。

10

【0278】

2 cDNAクローンの単離

UNIZAPベクターシステム (Stratagene) を用いた *in vivo* 切除によって、或いは細胞溶解によって、実施例1のようにして得たプラスミドを宿主細胞から回収した。プラスミドの精製には、下記の少なくとも1つを用いた。すなわちMagicまたはWIZARD Minipreps DNA精製システム (Promega)、AGTC Miniprep精製キット (Edge Biosystems, Gaithersburg MD)、QIAGEN社のQIAWELL 8 Plasmid、QIAWELL 8 Plus PlasmidおよびQIAWELL 8 Ultra Plasmid精製システム、R.E.A.L. Prep 96プラスミド精製キットのいずれかである。プラスミドは、沈殿させた後、0.1mlの蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥しないで4℃で保管した。

20

【0279】

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

【0280】

3 シークエンシング及び分析

実施例2に記載したようにプラスミドから回収したIncyte cDNAを、以下に示すようにシークエンシングした。cDNAのシークエンス反応は、標準的方法或いは高処理装置、例えばABI CATALYST 800 サーマルサイクラー (Applied Biosystems) またはPTC-200 サーマルサイクラー (MJ Research) をHYDRAマイクロディスペンサー (Robbins Scientific) またはMICROLAB 2200 (Hamilton) 液体転移システムと併用して処理した。cDNAのシークエンス反応は、Amersham Pharmacia Biotech社が提供する試薬、またはABIシークエンシングキット、例えばABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems) の試薬を用いて準備した。cDNAのシークエンス反応の電気泳動的分離及び標識したポリヌクレオチドの検出には、MEGABACE 1000 DNAシークエンシングシステム (Molecular Dynamics) が、標準ABIプロトコル及び塩基呼び出し (base calling) ソフトウェアを用いるABI PRISM 373または377シークエンシングシステム (Applied Biosystems) が、或いはその他の本技術分野で既知の配列解析システムを用いた。cDNA配列内のリーディングフレームは、標準的方法 (前出のAusubel, 1997, 7.7ユニットに概説) を用いて決定した。cDNA配列の幾つかを選択して、実施例8に記載した方法で配列を伸長させた。

30

40

【0281】

IncyteのcDNA配列に由来するポリヌクレオチド配列は、ベクター、リンカー及びポリ(A)配列を除去し、あいまいな塩基をマスクすることによって有効性を確認した。その際、BLAST、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びプログラムを用いた。次に、Incyte cDNA配列またはそれらの翻訳の問い合わせを、以下のデータベース群に対して行った。すなわち、選抜した公共のデータベース群 (例えば

50

GenBankの霊長類、げっ歯類、哺乳類、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM)と、ヒト、ラット、マウス、線虫 (*Caenorhabditis elegans*)、出芽酵母 (*Saccharomyces cerevisiae*)、分裂酵母 (*Schizosaccharomyces pombe*) および *Candida albicans* からの配列群を持つPROTEOMEデータベース群 (Incyte Genomics, Palo Alto CA)、および、隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベース群、例えばPFAM、INCY、およびTIGRFAM (Haft, D.H. 他(2001) *Nucleic Acids Res.* 29:41-43); 並びにHMMベースのタンパク質ドメインデータベースたとえばSMART (Schultz他(1998) *Proc. Natl. Acad. Sci. USA* 95:5857-5864; Letunic, I. 他(2002) *Nucleic Acids Res.* 30:242-244). (HMMは、遺伝子ファミリーのコンセンサス1次構造を分析する確率的アプローチである。例えばEddy, S.R. (1996) *Cuff. Opin. Struct. Biol.* 6:361-365等を参照)。問合せは、BLAST、FASTA、BLIMPS及びHMMERに基づくプログラムを用いて行った。Incyte cDNA配列は、完全長のポリヌクレオチド配列を産出するようにアセンブリされた。あるいは、GenBank cDNA群、GenBank EST群、スティッチされた配列群、ストレッチされた配列群、またはGenscan予測コード配列群 (実施例4および5を参照) を用い、Incyte cDNAのアセンブリ体を完全長まで伸長させた。PhredとPhrapとConsedとに基づくプログラムを用いて構築し、GenMarkとBLASTとFASTAとに基づくプログラムを用いて、cDNAのアセンブリ体を、オープンリーディングフレームについてスクリーニングした。完全長ポリヌクレオチド配列を翻訳し、対応する完全長ポリペプチド配列を誘導した。あるいは、本発明のポリペプチドは、完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。完全長ポリペプチド配列群の続いたの分析としての問い合わせを、GenBankタンパク質データベース群 (genpept)、SwissProt、PROTEOMEデータベース群、BLOCKS、PRINTS、DOMO、PRODOMおよびProsites等のデータベースや、PFAM、INCY、およびTIGRFAM等の隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベース群、並びにSMART等のHMMベースのタンパク質ドメインデータベース群に対し行った。完全長ポリヌクレオチド配列はまた、MACDNASIS PROソフトウェア (日立ソフトウェアエンジニアリング, South San Francisco CA) およびLASERGENEソフトウェア (DNASTAR) を用いて分析する。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算するMEGALIGNマルチシーケンシングアラインメントプログラム (DNASTAR) に組み込まれているようなCLUSTALアルゴリズムによって特定されるデフォルトパラメータを用いて生成する。

【0282】

Incyte cDNA及び完全長配列の分析及びアセンブリに利用したツール、プログラム及びアルゴリズムの概略と、適用可能な説明、参照文献、閾値パラメータを表7に示す。用いたツール、プログラム及びアルゴリズムを表7の列1に、それらの簡単な説明を列2に示す。列3は適切な参照文献であり、全ての文献は全体を引用を以って本明細書の一部となす。適用可能な場合には、列4は2つの配列が一致する強さを評価するために用いたスコア、確率値その他のパラメータを示す (スコアが高いほど、また確率値が低いほど、2配列間の同一性が高くなる)。

【0283】

完全長ポリヌクレオチド配列およびポリペプチド配列のアセンブリ及び分析に用いる上記のプログラムは、SEQ ID NO:31-60のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20~約4000ヌクレオチドの断片を表4の列2に示した。

【0284】

4 ゲノムDNAからのコード配列の同定及び編集

推定上の分泌タンパク質は、公共のゲノム配列データベース (例えば、gbpriやgbhtg) においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物からゲノムDNA配列を分析する汎用遺伝子同定プログラムである (Burge, C及びS. Karlin (1997) *J. Mol. Biol.* 268:78-94、Burge, C及びS. Karlin (1998) *Cuff. Opin. Struct. Biol.* 8:346-354参照)。プログラムは予測エキソンを連結し、メチオニンから停止

コドンに及ぶアセンブリされたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチドおよびポリペプチド配列のFASTAデータベースである。Genscanが一度に分析する配列の最大範囲は、30kbに設定した。これらのGenscan推定cDNA配列の内、どの配列が分泌タンパク質をコードするかを決定するために、コードされたポリペプチドをPFAMモデルにおいて分泌タンパク質について問合せて分析した。潜在的な分泌タンパク質はまた、分泌タンパク質として注釈が付けられていたIncyte cDNA配列への相同性を基に同定された。こうして選択されたGenscan予測配列は、次にBLAST分析により公共データベースgenpept及びgbpriと比較した。必要であれば、genpeptからのトップBLASTヒットと比較することによりGenscan予測配列を編集し、余分なまたは省略されたエキソンなど、Genscanが予測した配列におけるエラーを補正した。BLAST分析はまた、Genscan予測配列の、いかなるIncyte cDNAまたは公共cDNAカバレッジ (coverage) の発見にも用いられ、したがって転写の証拠を提供した。Incyte cDNAカバレッジが利用できた場合には、この情報を用いてGenscan予測配列を補正または確認した。完全長ポリヌクレオチド配列は、実施例 3に記載したアセンブリプロセスを用いて、Incyte cDNA配列および/または公共cDNA配列でGenscan予測コード配列をアセンブリして得た。或いは、完全長ポリヌクレオチド配列は、編集した、または非編集のGenscan予測コード配列に完全に由来する。

10

【0285】

5 cDNA配列データとのゲノム配列データのアセンブリ

ステッチ配列 (Stitched Sequence)

部分cDNA配列は、実施例 4に記載のGenscan遺伝子同定プログラムにより予測されたエキソンを用いて伸長させた。実施例 3に記載されたようにアセンブリされた部分cDNAは、ゲノムDNAにマッピングし、関連するcDNA及び1つ若しくは複数のゲノム配列から予測されたGenscanエキソンを含むクラスタに分解した。cDNA及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の2以上の配列に存在するような配列を同定し、そのように同定された間隔は推移性により等しいと考えられた。例えば、1つのcDNA及び2つのゲノム配列に間隔が存在する場合、3つの間隔は全て等しいと考えられた。このプロセスは、無関係であるが連続したゲノム配列をcDNA配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列 (parent sequence) に沿って現われる順に

20

30

【0286】

ストレッチ配列 (Stretched Sequence)

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

40

50

れるストレッチ配列を検査し、完全遺伝子を含んでいるか否かを判定した。

【0287】

6 SECPをコードするポリヌクレオチドの染色体マッピング

SEQ ID NO:31-60をアセンブリするために用いた配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、Incyte LIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO:31-60と一致するこれらのデータベースの配列を、Phrapなどの構築アルゴリズム(表7)を使用して、連続及びオーバーラップした配列のクラスターにアセンブリした。スタンフォード・ヒトゲノムセンター(SHGC)、ホワイトヘッド・ゲノム研究所(WIGR)、Genethon等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスタ化された配列が既にマッピングされていたかを決定した。マッピングされた配列が或るクラスターに含まれている場合、そのクラスターの全配列が、個々の配列番号と共に、地図上の位置に割り当てられた。

10

【0288】

地図上の位置は、ヒト染色体の範囲または区間として表される。センチモルガン間隔の地図上の位置は、染色体のpアームの末端に関連して測定する。(センチモルガン(cM)は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1cMは、ヒト中のDNAの1メガベース(Mb)にほぼ等しい。尤も、この値は、組換えのホットスポット及びコールドスポットに起因して広範囲に変化する。)cM距離は、各クラスター内に配列が含まれる放射線ハイブリッドマーカー類に対して境界を提供するGenethonによってマッピングされた遺伝マーカー群に基づく。NCBI「GeneMap'99」(<http://www.ncbi.nlm.nih.gov/genemap/>)などの一般個人が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、既に同定されている疾患遺伝子群が、上記した区間内若しくは近傍に位置するかを決定できる。

20

【0289】

7 ポリヌクレオチド発現の分析

ノーザン分析は、転写された遺伝情報の存在を検出するために用いられる実験技術であり、特定の細胞種または組織からのRNAが結合される膜への標識されたヌクレオチド配列のハイブリダイゼーションに関与している(例えば前出のSambrook, 7章、同Ausubel (1995) 4章および16章を参照)。

【0290】

BLASTを適用する類似のコンピュータ技術を用いて、GenBankやLIFESEQ(Incyte Genomics)等のcDNAデータベースにおいて同一または関連分子を検索した。ノーザン分析は、多数の膜系ハイブリダイゼーションよりも非常に速い。更に、特定の一致を厳密な或いは相同的なものとして分類するか否かを決定するため、コンピュータ検索の感度を変更することができる。検索の基準は積スコアであり、次式で定義される。

30

【0291】

【数1】

(BLASTスコア×配列一致率)

$$\frac{\text{BLASTスコア} \times \text{配列一致率}}{5 \times (\text{長さ(配列1)}, \text{長さ(配列2)}) \text{の最小値}}$$

40

【0292】

積スコアは、2つの配列間の類似度と、配列が一致する長さとの両方を考慮している。積スコアは、0~100のノーマライズされた値であり、次のようにして求める。BLASTスコアにヌクレオチドの配列一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。高スコアリングセグメント対(HSP)に一致する各塩基に+5のスコアを割り当て、各不一致塩基対に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る(ギャップにより隔離される)。2以上のHSPがある場合には、最高BLASTスコアのセグメント対を用いて積スコアを計算する。積スコアは、断片的オーバーラップとBLASTアラインメントの質とのバランスを表す。例えば積スコア100は、比較した

50

2つの配列の短い方の長さ全体にわたって100%一致する場合のみ得られる。積スコア70は、一端が100%一致し、70%オーバーラップしているか、他端が88%一致し、100%オーバーラップしているかのいずれかの場合に得られる。積スコア50は、一端が100%一致し、50%オーバーラップしているか、79%一致し、100%オーバーラップしているかのいずれかの場合に得られる。

【0293】

或いは、SECPをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば幾つかの完全長配列は、少なくとも一部は、オーバーラップするIncyte cDNA配列群を用いてアセンブリされる(実施例3を参照)。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、以下の臓器/組織カテゴリーの1つに分類される。即ち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液及び免疫系、肝、筋骨格系、神経系、膵臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性/混合性または尿路である。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患/条件カテゴリー即ち癌、細胞株、発達、炎症、神経性、外傷、心血管、プール、その他の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、SECPをコードするcDNAの組織特異的および疾患特異的な発現を反映する。cDNA配列およびcDNAライブラリ/組織の情報は、LIFESEQ GOLD データベース(Incyte Genomics, Palo Alto CA)から得ることができる。

10

【0294】

8 ポリヌクレオチドをコードするSECPの伸長

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

20

【0295】

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

30

【0296】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた。PCRは、PTC-200サーマルサイクラー(MJ Research, Inc.)を用いて96ウェルプレート内で行った。反応混合液は、鋳型DNA及び200nmolの各プライマー、 Mg^{2+} と $(NH_4)_2SO_4$ と2-メルカプトエタノールを含む反応バッファー、Taq DNAポリメラーゼ(Amersham Pharmacia Biotech)、ELONGASE酵素(Life Technologies)、Pfu DNAポリメラーゼ(Stratagene)を含む。プライマー対、PCI AとPCI Bに対して以下のパラメータで増幅を行った。ステップ1: 94で3分間、ステップ2: 94で15秒間、ステップ3: 60で1分間、ステップ4: 68で2分間、ステップ5: ステップ2、3および4を20回繰り返す、ステップ6: 68で5分間、ステップ7: 4で保存。別法では、プライマー対であるT7とSK+とに対して以下のパラメータで増幅を行った。ステップ1: 94で3分間、ステップ2: 94で15秒間、ステップ3: 57で1分間、ステップ4: 68で2分間、ステップ5: ステップ2、3および4を20回繰り返す、ステップ6: 68で5分間、ステップ7: 4で保存。

40

【0297】

各ウェルのDNA濃度は、1X TE及び0.5µlの希釈していないPCR産物に溶解した100µlのPICOGREEN定量試薬(0.25%(v/v) PICOGREEN; Molecular Probes, Eugene OR)を不透明な蛍光光度計プレート(Corning Costar, Acton MA)の各ウェルに分配してDNAが試薬と結合でき

50

るようにして測定する。サンプルの蛍光を計測してDNAの濃度を定量すべく、プレートにFluoroskan II (Labsystems Oy, Helsinki, Finland) でスキャンした。反応混合物のアリコート5~10 μ lを1%アガロースゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを判定した。

【0298】

伸長させたヌクレオチドは、脱塩及び濃縮して384穴プレートに移し、CviJIコレラウイルスエンドヌクレアーゼ (Molecular Biology Research, Madison WI) を用いて消化し、pUC 18ベクター (Amersham Pharmacia Biotech) への再連結反応前に音波処理またはせん断した。ショットガン・シーケンシングのために、消化したヌクレオチドを低濃度(0.6~0.8%)のアガロースゲル上で分離し、断片を切除し、寒天をAgar ACE (Promega) で消化した。伸長させたクローンをT4リガーゼ (New England Biolabs, Beverly MA) を用いてpUC 18ベクター (Amersham Pharmacia Biotech) に再連結し、Pfu DNAポリメラーゼ (Stratagene) で処理して制限部位のオーバーハングを満たし、大腸菌細胞に形質移入した。形質移入した細胞群の選択を、抗生物質を含む培地で行い、それぞれのコロニーを採取し、LB/2Xカルベニシリン培養液中の384ウェルプレート群に37 $^{\circ}$ Cで一晩培養した。

10

【0299】

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

20

【0300】

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

30

【0301】

9 SECPをコードするポリヌクレオチドの一塩基多型の同定
一塩基多型性 (SNP) として知られる一般的なDNA配列変異体は、LIFESEQデータベース (Incyte Genomics) を用いてSEQ ID NO:31-60において同定された。実施例3に記述されているように同じ遺伝子からの配列は共にクラスター化され、アセンブリされ、遺伝子内の全ての配列変異体を同定することができた。一連のフィルタから成るアルゴリズムは、SNPを他の配列変異体から区別するために用いられる。前段フィルタ群が、最小限のPhredクオリティスコア15を要求することにより大多数のベースコールのエラーを除去し、また、配列アラインメントエラーと、ベクター配列、キメラ、スプライス変異体の、不適切なトリミングに起因するエラーを除去した。先進の染色体分析の自動化した手順により、推定上のSNPの近傍の本来のクロマトグラムファイル进行分析した。クローンエラー-フィルタ群は、統計的に生み出されたアルゴリズムを用いて、逆転写酵素、ポリメラーゼあるいは体細胞性突然変異によって引き起こされる等の、実験プロセッシング中に導入されたエラーを同定した。クラスタリングエラー-フィルタ群は、統計的に生み出されたアルゴリズムを用いて、近縁の相同体あるいは偽遺伝子のクラスタリングに起因するエラー、または非ヒト配列によるコンタミネーションによるエラーを同定した。フィルタの最終セットは、免疫グロブリンまたはT細胞受容体に見出される重複とSNPを除去した。

40

【0302】

異なる4つのヒト集団のSNP部位における対立遺伝子頻度を分析するために、高処理MAS

50

SARRAYシステム(Sequenom, Inc.)を用いる質量分析によって、更なる特徴付けのためにいくつかのSNPが選択された。白人集団は92人(男性46人、女性46人)から成り、そのうち83人はユタ州、4人はフランス、3人はベネズエラ、2人はアーミッシュの出身である。アフリカ系集団は194人(男性97人、女性97人)から成り、全てアフリカ系米国人である。ラテンアメリカ系集団は324人(男性162人、女性162人)から成り、全てメキシコ出身である。アジア系の集団は126人(男性64人、女性62人)からなり、中国人43%、日本人31%、コリアン13%、ベトナム人5%、他のアジア系8%の親の構成が報告されている。対立遺伝子頻度は最初に白人集団で分析された。この集団で対立遺伝子変異を示さないSNPの時には他の3つの集団で更に試験されない場合もあった。

【0303】

10

10 個々のハイブリダイゼーションプローブの標識化及び使用

SEQ ID NO:31-60から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても本質的に同一の手順が用いられる。オリゴヌクレオチドを、OLIG04.06ソフトウェア(National Bioscience)のような最新式のソフトウェアを用いてデザインし、50pmolの各オリゴマーと、250 μ Ciの[³²P]アデノシン三リン酸(Amersham Pharmacia Biotech)及びT4ポリヌクレオチドキナーゼ(DuPont NEN, Boston MA)とを組み合わせるにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25超細繊維分子サイズ排除デキストラン ビードカラム(Amersham Pharmacia Biotech)を用いて実質的に精製する。下記のいずれか1つのエンドヌクレアーゼで消化されたヒトゲノムDNAの、典型的な膜ベースのハイブリダイゼーション解析において、毎分10⁷カウントの標識されたプローブを含むアリコットを用いる。すなわちAse I、Bgl II、Eco RI、Pst I、Xba I、またはPvu II(DuPont NEN)である。

20

【0304】

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

30

【0305】

11 マイクロアレイ

マイクロアレイの表面上でアレイエレメントの結合または合成は、フォトリソグラフィ、ピエゾ式印刷(インクジェット印刷、前出のBaldeschweiler等を参照)、機械的マイクロスポッティング技術及びこれらから派生したものをを用いて達成することが可能である。上記各技術において基板は、均一な、無孔の表面を持つ固体とすべきである(Schena(1999)前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。或いは、ドットプロット法またはスロットプロット法に類似した手順を利用して、熱的、紫外線的、化学的または機械的結合手順を用いて基板の表面にエレメントを配置及び結合させてもよい。通常のアレイは、利用可能な、当業者に公知の方法と機械とを用いて作製でき、任意の適正数のエレメントを有し得る(例えばSchena, M. 他.(1995) Science 270:467-470、Shalon, D.他.(1996) Genome Res. 6:639-645、Marshall, A.及びJ. Hodgson(1998) Nat. Biotechnol. 16:27-31等を参照)。

40

【0306】

完全長cDNA、発現配列タグ(EST)、またはその断片またはオリゴマーが、マイクロアレイのエレメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、LASERGENEソフトウェア(DNASTAR)等の本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメント群を、生体サンプル中のポリヌクレオチド群とハイブリダイズする。生体サンプル中のポリヌクレオチドは、検出を容易にするために蛍光

50

標識などの分子タグに抱合させる。ハイブリダイゼーション後、生体サンプルからのハイブリダイズされていないヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントでのハイブリダイゼーションを検出する。あるいは、レーザー脱離および質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上の或るエレメントにハイブリダイズする各ポリヌクレオチドの、相補性の度合と相対存在度とを算定し得る。一実施例におけるマイクロアレイの調製及び使用について、以下に詳述する。

【0307】

組織または細胞サンプルの調製

グアニジニウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ(dT)セルソース法を用いてポリ(A)⁺RNAを精製する。各ポリ(A)⁺RNAサンプルを、MMLV逆転写酵素、0.05pg/ μ lのオリゴ(dT)プライマー(21mer)、1×第一鎖合成バッファー、0.03unit/ μ lのRNアーゼ阻害因子、500 μ MのdATP、500 μ MのdGTP、500 μ MのdTTP、40 μ MのdCTP、40 μ MのdCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。逆転写反応は、GEMBRIGHTキット(Incyte)を用い、200ngのポリ(A)⁺RNAを含む体積25mlで行う。特異的対照ポリ(A)⁺RNAは、非コード酵母ゲノムDNAから *in vitro* 転写により合成する。37℃で2時間インキュベートした後、各反応サンプル(1つはCy3、もう1つはCy5標識)は、2.5mlの0.5M水酸化ナトリウムで処理し、85℃で20分間インキュベートし、反応を停止させてRNAを分解させる。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピナラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。結合後、2つの反応サンプルは、1mlのグリコーゲン(1mg/ml)を用いて析出させたエタノール、60mlの酢酸ナトリウム及び300mlの100%エタノールである。サンプルは次に、SpeedVAC(Savant Instruments Inc., Holbrook NY)を用いて完全に乾燥させ、14 μ lの5×SSC/0.2%SDS中で再懸濁する。

【0308】

マイクロアレイの調製

本発明の配列を用いて、アレイエレメントを生成する。各アレイエレメントは、クローン化したcDNAインサートを有するベクターを含有する細菌細胞から増幅する。PCR増幅は、cDNAインサートの側面に位置するベクター配列に相補的なプライマーを用いる。30サイクルのPCRで1~2ngの初期量から5 μ gより大きい最終量までアレイエレメントを増幅する。増幅したアレイエレメントは、SEPHACRYL-400(Amersham Pharmacia Biotech)を用いて精製する。

【0309】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス(Corning)は、処理の間及び処理後に0.1%のSDS及びアセトン中で超音波処理をかけ、蒸留水で十分に洗浄する。スライドガラスは、4%フッ化水素酸(VWR Scientific Products Corporation (VWR), West Chester PA)中でエッチングし、蒸留水中で十分に洗浄し、95%エタノール中で0.05%アミノプロピルシラン(Sigma)を用いてコーティングする。コーティングしたスライドガラスは、110℃のオーブンで硬化させる。

【0310】

米国特許第5,807,522号で説明されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。該特許は、引用を以って本明細書の一部となす。平均濃度が100ng/ μ lのアレイエレメントDNA 1 μ lを高速機械装置により開放型キャピラリープリンティングエレメント(open capillary printing element)に充填する。装置はここで、スライド毎に約5nlのアレイエレメントサンプルを加える。

【0311】

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

10

20

30

40

50

洗浄することにより、非特異結合部位をブロックする。

【0312】

ハイブリダイゼーション

ハイブリダイゼーション反応は、5×SSC、0.2% SDSハイブリダイゼーション緩衝液中のCy3及びCy5標識したcDNA合成生成物を各0.2µg含む9µlのサンプル混合体を有する。サンプル混合体は、65℃まで5分間加熱し、マイクロアレイ表面上で等分して1.8cm²のカバーガラスで覆う。アレイは、顕微鏡用スライドより僅かに大きい空洞を有する防水チェンバーに移す。チェンバーのコーナーに140µlの5×SSCを加えることにより、チェンバー内部を湿度100%に保持する。アレイを含むチェンバーは、60℃で約6.5時間インキュベートする。アレイは、第1洗浄緩衝液中(1×SSC、0.1% SDS)において45℃で10分間洗浄し、第2洗浄緩衝液中(0.1×SSC)において各々45℃で10分間、3度洗浄して乾燥させる。

10

【0313】

検出

レポーター標識ハイブリダイゼーション複合体は、Cy3の励起のためには488nm、Cy5の励起のためには632nmでスペクトル線が発生し得るInnova 70混合ガス10 Wレーザ(Coherent, Inc., Santa Clara CA)を備えた顕微鏡で検出する。励起レーザー光の焦点をアレイ上に置くため、20×顕微鏡対物レンズ(Nikon, Inc., Melville NY)を用いる。アレイを含むスライドを、顕微鏡の、コンピュータ制御のX-Yステージに置き、対物レンズを通してラスタースキャンする。本実施例で用いる1.8cm×1.8cmのアレイは、解像度20µmでスキャンする。

20

【0314】

2回の異なるスキャンで、混合ガスマルチラインレーザは2つの蛍光色素を連続的に励起する。発光された光は、波長に基づき分離され、2つのフルオロフォアに対応する2つの光電子増倍管検出器(PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ)に送られる。好適なフィルタ群をアレイと光電子増倍管との間に設置して、シグナルをフィルタする。用いるフルオロフォアの最大発光の波長は、Cy3では565nm、Cy5では650nmである。装置は両方の蛍光色素からのスペクトルを同時に記録し得るが、レーザ源において好適なフィルタを用いて、蛍光色素1つにつき1度スキャンし、各アレイを通常2度スキャンする。

30

【0315】

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

【0316】

光電子増倍管の出力は、IBMコンパチブルPCコンピュータにインストールされた12ビットRTI-835Hアナログ-デジタル(A/D)変換ボード(Analog Devices, Inc., Norwood MA)を用いてデジタル化される。デジタル化されたデータは、青色(低シグナル)から赤色(高シグナル)までの擬似カラー範囲へのリニア20色変換を用いてシグナル強度がマッピングされたようなイメージとして表示される。データは、定量的にも分析される。2つの異なる蛍光色素を同時に励起及び測定する場合には、各蛍光色素の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク(発光スペクトルの重なりに起因する)を補正する。

40

【0317】

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

50

00LS遺伝子発現分析プログラム (Incyte) である。

【0318】

ヒトのTHP-1細胞 (American Type Culture Collection, Manassas VA)は、10%の胎児血清 (v/v)、0.45%のグルコース (w/v)、10mMのHepes、1mMのピルビン酸ナトリウム、 1×10^{-5} Mの -メルカプトエタノール、ペニシリン (100units/ml)、ストレプトマイシン (100mg/ml) 含有のRPMI1640培地で成長した。酸化LDLの取り込み実験のために、 1×10^{-7} Mの12-0-テトラデカノイル-ホルボール-13-アセテート (Research Biochemical International, Natick MA)を含む培地において 1×10^6 細胞/mlの密度で、細胞を撒いて、24時間置く。次に培地は、Hammer 他.(1995; Arterio Thromb Vasc Biol 15:704-713)の方法に基づき、 CuSO_4 で“十分に”酸化した、 $100 \mu\text{g/ml}$ のLDL (Intracel, Rockville MD)を入れた培地、あるいは入れない培地と交換した。培地は培養中に2日おきに交換した。細胞は30分から4日間の範囲の時点で渡ってOx-LDLで処理した。この期間の間、細胞は接着性を維持し、通常のまだらのある Nile redの染色パターンを有する。0、0.5、2.5、8時間と1、2、4日間のOx-LDL曝露の時点での発現プロファイル作成のために、RNAを調製した。

10

【0319】

ヒト肝臓のC3A細胞のステロイド処理

コンフルエント前期のC3A細胞は、 $1 \mu\text{M}$ 、 $10 \mu\text{M}$ 、 $100 \mu\text{M}$ の各濃度で、1、3、6時間の間、ミフェプリストン、プロゲステロン、ベクロメタゾン、メドロキシプロゲステロン、プレドニド、プレドニゾン、デキサメサゾン、ベタメタゾン、あるいはダナゾールで処理された。全ての場合において、未処理のコンフルエント前期のC3A細胞からのmRNAが並行して調製された。

20

【0320】

このようにして、SEQ ID NO:31に関して、SECPは試験された6つの肺腫瘍組織の内の3つで少なくとも2倍下方調節されたことが示された。さらに、泡沫細胞形成のTHP-1モデルにおいて少なくとも3倍下方調節された。SEQ ID NO:34に関して、下記のステロイド化合物で処理された時、SECP発現はC3A肝臓細胞株において少なくとも2.5倍上方調節されいたことが示された。ベクロメタゾン (Beclio)、メドロキシプロゲステロン (MAH)、プレドニゾン (Prdsne)、デキサメサゾン (Dex)、ベタメタゾン (Betam)がそれらのステロイドである。

【0321】

1.2 相補的ポリヌクレオチド

SECPをコードする配列或いはその任意の一部に対して相補的な配列は、天然のSECPの発現を検出、低下、または阻害するために用いられる。約15~30塩基対を含むオリゴヌクレオチドの使用について記すが、これより小さなあるいは大きな配列の断片の場合でも、本質的に同じ手順を用いる。Oligo4.06ソフトウェア (National Biosciences) 及びSECPのコーディング配列を用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な5'配列から相補的オリゴヌクレオチドを設計し、これを用いて、プロモーターがコーディング配列に結合するのを防止する。翻訳を阻害するためには、相補的なオリゴヌクレオチドを設計して、リボソームがSECPをコードする転写物に結合するのを阻害する。

30

【0322】

1.3 SECPの発現

SECPの発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌でSECPが発現するために、抗生物質耐性遺伝子及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクローニングする。このようなプロモーターの例には、lacオペレーター調節エレメントと併用するT5またはT7バクテリオファージプロモーター及び、trp-lac (tac)ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、BL21 (DE3)等の好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル -Dチオガラクトピラノシド (IPTG)で誘発されるとSECPを発現する。真核細胞でのSECPの発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイルスとして知られている *Autographica californica*核

40

50

多角体病ウイルス(AcMNPV)の組換え型を感染させて行う。バキュロウイルスの非必須の多角体遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、SECPをコードするcDNAと置換する。ウイルスの感染力は維持され、強力なポリヘドリンプロモーターによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合は*Spodoptera frugiperda* (Sf9) 昆虫細胞への感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる (Engelhard, E. K.他 (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. 他 (1996) Hum. Gene Ther. 7:1937-1945.等を参照)。

【0323】

殆どの発現系では、SECPが、例えばグルタチオンSトランスフェラーゼ(GST)、またはFLAGや6-Hisなどのペプチドエピトープ標識で合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製が素早く1回で行うことができる。GSTは日本住血吸虫からの26kDaの酵素であり、タンパク質の活性および抗原性を維持した状態で、固定化したグルタチオン上での融合タンパク質の精製を可能とする (Amersham Pharmacia Biotech)。精製の後、GST部分を特定の操作部位でSECPからタンパク質的に切断できる。FLAGは8アミノ酸のペプチドであり、市販されているモノクローナル及びポリクローナル抗FLAG抗体 (Eastman Kodak) を用いた免疫親和性精製を可能にする。6ヒスチジン残基が連続して伸長した6-Hisは、金属キレート樹脂上での精製を可能にする (QIAGEN)。タンパク質の発現および精製の方法は、前出のAusubel (1995) 10章、16章に記載されている。これらの方法で精製したSECPを直接用いて以下の実施例17、18および19のアッセイを行うことができる。

【0324】

1.4 機能的アッセイ

SECP機能は、哺乳動物細胞培養系において生理学的に高められたレベルでのSECPをコードする配列の発現によって評価する。cDNAを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにcDNAをサブクローニングする。選択されるベクターには、pCMV SPORTプラスミド (Life Technologies) 及びpCR 3.1プラスミド (Invitrogen, Carlsbad CA) が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤あるいは電気穿孔法を用いて、5~10 μ gの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に、一過的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2 μ gのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質 (GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザ光学に基づく技術であるフローサイトメトリー (FCM) を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するか或いは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、ヨウ化プロピジウムによるDNA染色によって計測される核DNA含量の変化、前方光散乱と90°側方光散乱によって計測される細胞サイズと粒度の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞内におけるタンパク質の発現の変化、及びフルオレセイン抱合したアネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。フローサイトメトリー法については、Ormerod, M.G. (1994) *Flow Cytometry*, Oxford, New York NYに記述がある。

【0325】

遺伝子発現におけるSECPの影響は、SECPをコードする配列とCD64またはCD64-GFPのどちらかが形質移入された高度に精製された細胞集団を用いて評価することができる。CD64またはCD64-GFPは、形質移入された細胞表面で発現し、ヒト免疫グロブリンG (IgG) の保存領

10

20

30

40

50

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

【0326】

15 SECPに特異的な抗体の作製

ポリアクリルアミドゲル電気泳動法(PAGE; Harrington, M.G. (1990) Methods Enzymol. 182:488-495等を参照)または他の精製技術を用いて実質上精製されたSECPを用いて、標準プロトコルでウサギ、マウス等の動物を免疫化して抗体を産出する。

10

【0327】

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

【0328】

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

20

【0329】

16 特異的な抗体を用いる天然SECPの精製

天然SECP或いは組換えSECPを、SECPに特異的な抗体を用いるイムノアフィニティークロマトグラフィにより実質的に精製する。イムノアフィニティークラムは、CNBr-活性化SEPHA ROSE(Amersham Pharmacia Biotech)のような活性化クロマトグラフィー用樹脂と抗SECP抗体とを共有結合させることにより形成する。結合後に、製造者の使用説明書に従ってこのレジンをブロックし、洗浄する。

30

【0330】

SECPを含む培養液をイムノアフィニティークラムに通し、SECPを選択的に吸着できる条件で(例えば、界面活性剤の存在下において高イオン強度のバッファーで)そのカラムを洗浄する。そのカラムを、抗体とSECPとの結合を切るような条件で(例えば、pH 2~3のバッファー、或いは高濃度の尿素またはチオシアン酸塩イオンのようなカオトロピックイオンで)溶出させ、SECPを回収する。

【0331】

17 SECPと相互作用する分子の同定

SECPまたは生物学的に活性なその断片を、¹²⁵Iボルトンハンター試薬で標識する(例えばBolton A.E.およびW.M. Hunter (1973) Biochem. J. 133:529-539を参照)。マルチウェルプレートに予め配列しておいた候補の分子を、標識したSECPと共にインキュベートし、洗浄して、標識したSECP複合体を有する全てのウェルをアッセイする。様々なSECP濃度で得られたデータを用いて、候補分子と結合したSECPの数量及び親和性、会合についての値を計算する。

40

【0332】

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

50

【0333】

SECPはまた、ハイスループットな方法で酵母2ハイブリッドシステムを使用するPATHCALLINGプロセス(CuraGen Corp., New Haven CT)に用いて、遺伝子の2大ライブラリにコードされる遺伝子間の全ての相互作用を決定することができる(Nandabalan, K. 他(2000) 米国特許第6,057,101号)。

【0334】

18 SECP活性の実証

SECPの成長刺激活性若しくは阻害活性の為のアッセイは、スイスマウスの3T3細胞におけるDNA合成の量を測定する(McKay, I.及びLeigh, I., 編集.(1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY)。このアッセイにおいては、様々な量のSECPが、放射性DNA前駆物質である[³H]チミジンの存在中で静止状態3T3培養細胞へと加えられる。このアッセイのためのSECPを得る手段は、組換えでも良く、生化学的な調製より得ても良い。酸沈殿し得るDNAへの[³H]チミジンの組み込みが、適切な時間間隔で測定され、組み込まれた量は新規に合成されたDNAの量に直接比例する。少なくとも100倍のSECP濃度範囲にわたる線形の用量 応答曲線は、成長の調節活性を示す。ミリリットルあたりの1単位の活性は、50%の応答レベルを提供するSECPの濃度として定義される。ここで、100%の応答レベルは、酸沈殿し得るDNAへの[³H]チミジンの最大の組み込みを表す。

10

【0335】

別法では、SECP活性のためのアッセイで、培養細胞中の神経伝達の刺激若しくは抑制を測定する。培養CHO繊維芽細胞をSECPに曝す。エンドサイトーシスによるSECP取込後に、これらの細胞を新鮮培養液で洗浄し、細胞全膜電位固定したアフリカツメガエル筋細胞を操作して、SECPを含まない媒質中で前記の繊維芽細胞の1つと接触させる。膜電流を、この筋細胞から記録する。対照値に対し増加若しくは減少した電流は、SECPの神経修飾性効果を示している(Morimoto, T.他(1995) Neuron 15: 689-696)。

20

【0336】

別法では、SECP活性のアッセイで、分泌性の膜で囲まれた細胞小器官におけるSECPの量を測定する。上述したように形質移入された細胞を採取し、溶解する。ライセートは、当業者に既知の、ショ糖密度勾配超遠心法などの方法で分画する。そのような方法は、ゴルジ体、ER、小膜結合小胞、およびその他の分泌細胞小器官のような、細胞内要素の隔離を可能とする。分割された細胞ライセートおよび全体の細胞ライセートよりの免疫沈降は、SECP特定抗体を用いて実行され、また免疫沈降サンプルはSDS-PAGEおよび免疫プロット技術を用いて解析される。全細胞ライセート中のSECPに対する分泌性細胞小器官中のSECP濃度は、分泌経路を通るSECPの量に比例する。

30

【0337】

別法では、AMP結合活性の測定を、SECPと³²P標識したAMPとを混合させて行う。この反応溶液を37℃でインキュベートし、反応はトリクロロ酢酸の添加で終了させる。酸抽出物を中和し、ゲル電気泳動にかけて非結合標識を除去する。ゲル内に留まる放射活性が、SECPの活性に比例する。

【0338】

19 免疫グロブリン活性の実証

SECP活性の或るアッセイでは、SECPが血清からの抗原類を認識し沈殿させる能力を測定する。この活性の測定は、定量沈降反応で成し得る(Golub, E. S. 他(1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, 113-115ページ)。SECPを、当分野で既知の方法で同位体標識する。一定量の標識SECPに種々の血清濃度を加える。SECP-抗原複合体を溶液から沈殿させ、遠心分離で収集する。沈殿性SECP-抗原複合体の量は、沈殿物中に検出される放射性同位元素の量に比例する。沈殿性SECP-抗原複合体の量を、血清濃度に対してプロットする。多様な血清中濃度に対しての特徴的沈降素曲線が得られ、沈殿性SECP-抗原複合体の量は、初め血清中濃度の増加に比例して増加し、当量点を頂点とし、その後は血清中濃度の増加に比例して減少する。このように、沈殿性SECP-抗原

40

50

複合体の量は、抗原の制限量と過剰量の両方に対する感受性によって特徴付けられるSECP活性の測定量である。

【0339】

別法として、SECP活性の或るアッセイは、細胞表面におけるSECPの発現を測定する。SECPをコードするcDNAを、非白血球細胞株に形質移入する。細胞表面タンパク質は、ビオチンで標識する(de la Fuente, M.A. 他(1997) Blood 90:2398-2405)。SECP特異的抗体群を用いて免疫沈降を行い、免疫沈降したサンプルをSDS-PAGEと免疫ブロット法で分析する。標識した免疫沈降剤と未標識免疫沈降剤の比は、細胞表面に発現したSECPの量に比例する。

【0340】

別法で、SECP活性の或るアッセイは、SECPの過剰発現が誘発する、細胞凝集の量を測定する。このアッセイにおいては、NIH3T3などの培養細胞にSECPをコードするcDNAで形質移入する。このcDNAは、或る強力なプロモーターの制御下にある或る適切な哺乳類発現ベクター内に含まれるようにする。緑色蛍光タンパク質(CLONTECH)などの蛍光標識タンパク質をコードするcDNAとの共形質移入を行うと、安定な形質移入体を同定するのに役立つ。形質移入された細胞と形質移入されない細胞で細胞の凝集(塊化)量を比較する。細胞の凝集量が、SECPの活性の直接の測定値となる。

【0341】

当業者には、本発明の範囲及び精神から逸脱しない、本発明の記載した方法及びシステムの種々の修正および変更は自明であろう。本発明について説明するにあたり幾つかの実施例に関連して説明を行ったが、本発明の請求の範囲が、そのような特定の実施例に不当に制限されるべきではないことを理解されたい。実際に、分子生物学または関連分野の専門家には明らかな、本明細書に記載されている本発明の実施方法の様々な修正は、下記の特許請求の範囲内にあるものとする。

【0342】

(表の簡単な説明)

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

【0343】

表2は、本発明のポリペプチド群のGenBank識別番号と、最も近いGenBank相同体の注釈(annotation)と、PROTEOMEデータベース識別番号と、PROTEOMEデータベース相同体群の注釈とを示す。各ポリペプチドとその相同体が一致する確率スコアも併せて示す。

【0344】

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

【0345】

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

【0346】

表5は、本発明のポリヌクレオチドの代表的なcDNAライブラリを示す。

【0347】

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

【0348】

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

【0349】

表8は、本発明のポリヌクレオチド配列から得られる一塩基多型を異なるヒト集団の対立遺伝子頻度と共に示す。

10

20

30

40

50

【 0 3 5 0 】
 【 表 1 - 1 】

表 1 - 1

Incyte プロジェクト ID	ポリペプチド SEQ ID NO:	Incyte ポリペプ チド ID	ポリヌクレオチド SEQ ID NO:	Incyte ポリヌク レオチド ID	Incyte 全長クローン
1895273	1	1895273CD1	31	1895273CB1	1895273CA2
7007222	2	7007222CD1	32	7007222CB1	813820CA2
3559223	3	3559223CD1	33	3559223CB1	
3441255	4	3441255CD1	34	3441255CB1	
1958917	5	1958917CD1	35	1958917CB1	90067353CA2
6219465	6	6219465CD1	36	6219465CB1	
3576625	7	3576625CD1	37	3576625CB1	
4765758	8	4765758CD1	38	4765758CB1	4765758CA2
7236661	9	7236661CD1	39	7236661CB1	6764860CA2, 6772618CA2, 7236661CA2
7714187	10	7714187CD1	40	7714187CB1	
5136540	11	5136540CD1	41	5136540CB1	5136540CA2
3277403	12	3277403CD1	42	3277403CB1	90093289CA2, 90097965CA2
1517569	13	1517569CD1	43	1517569CB1	
2415991	14	2415991CD1	44	2415991CB1	4031777CA2, 7383971CA2

10

20

30

40

【 0 3 5 1 】
 【 表 1 - 2 】

表1-2

Incyte プロジェクトID	ポリペプチド SEQ ID NO:	Incyte ポリペプ チドID	ポリヌクレオチド SEQ ID NO:	Incyte ポリヌク レオチドID	Incyte 全長クローン
2735742	15	2735742CD1	45	2735742CB1	1842245CA2, 1862749CA2, 2765810CA2, 3174585CA2, 6034213CA2, 6361652CA2, 7638315CA2, 7977239CA2, 8087951CA2, 8094152CA2, 8250338CA2, 8531020CA2, 8743705CA2
2768535	16	2768535CD1	46	2768535CB1	4239666CA2
6848851	17	6848851CD1	47	6848851CB1	
7040722	18	7040722CD1	48	7040722CB1	56022479CA2
6430290	19	6430290CD1	49	6430290CB1	90076008CA2, 90076371CA2, 90076387CA2, 90076395CA2
2640251	20	2640251CD1	50	2640251CB1	6452666CA2, 8516219CA2
3839350	21	3839350CD1	51	3839350CB1	3839350CA2

【 0 3 5 2 】
【 表 1 - 3 】

10

20

30

表1-3

Incyte プロジェクトID	ポリペプチド SEQ ID NO:	Incyte ポリペプ チドID	ポリヌクレオチド SEQ ID NO:	Incyte ポリヌク レオチドID	Incyte 全長クローン
6393813	22	6393813CD1	52	6393813CB1	1615668CA2, 2193894CA2, 6557563CA2, 6557564CA2, 7057161CA2
5685755	23	5685755CD1	53	5685755CB1	
71728459	24	71728459CD1	54	71728459CB1	5405225CA2
1904303	25	1904303CD1	55	1904303CB1	
2911343	26	2911343CD1	56	2911343CB1	90094777CA2, 90094885CA2
7500308	27	7500308CD1	57	7500308CB1	2343465CA2
7501098	28	7501098CD1	58	7501098CB1	4660380CA2
7503839	29	7503839CD1	59	7503839CB1	
7503698	30	7503698CD1	60	7503698CB1	

【 0 3 5 3 】

【 表 2 - 1 】

10

20

30

40

表2-1

ポリペプチド SEQ ID NO:	Incyte ポリペプ チド ID	GenBank ID NO.またはプロテ オーム ID NO:	確率スコア	注釈
1	1895273CD1	g180279	5.40E-146	[ヒト] IgG Fc 受容体 I van de Winkel, J.G.J. 他, (1991) J. Biol. Chem. 266:13449-13455
6	6219465CD1	g6649221	2.40E-184	[ヒト] セレン含有タンパク質 N Lescure, A. 他, (1999) J. Biol. Chem. 274:38147-38154
7	3576625CD1	g4406679	1.10E-54	[ヒト] ヒト神経細胞オルファクトメジン関連 ER 局在性タンパク質 Danielson, P.E. 他, (1994) J. Neurosci. Res. 38:468-78
10	7714187CD1	g3041877	1.30E-222	[ヒト] IB3089A Nishiyama, H. 他, (1999) Genes Chromosomes Cancer 26:171-5
11	5136540CD1	g12656635	8.40E-54	[ヒト] (AF326351) 膜貫通 γ -カルボキシグルタミン酸タンパク質 4 TMG4 Kulman, J.D. 他, (2001) Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375
12	3277403CD1	g9864185	8.10E-104	[シロウジ] ジョウバヒ [Crossveinless 2 Olson, D.J. 他, (2000) Development 127:3947-3959
16	2768535CD1	g7578787	8.30E-134	AD021 タンパク質 [ヒト]

10

20

30

40

【 0 3 5 4 】

【 表 2 - 2 】

表2-2

ポリペプチド SEQ ID NO:	Incyte ポリペ プチド ID	GenBank ID NO:またはプロテ オーム ID NO:	確率スコア	注釈
23	5685755CD1	g307153	9.30E-27	[ヒト]Mac-2 結合タンパク質 Koths, K., Taylor, E., Halenbeck, R., Caspit, C. 及び Wang, A. (1993) Cloning and characterization of a human Mac-2 binding protein, a new member of the superfamily defined by the macrophage scavenger receptor cysteine-rich domain J. Biol. Chem. 268:14245-14249
26	2911343CD1	g11225366	4.10E-13	[ヒト]精巢特異的ロイシニツチピレートタンパク質
28	7501098CD1	g12656635	2.6E-13	[ヒト]膜貫通 γ -カルボキシグルタミン酸タンパク質 4TMG4 Kulman, J. D. 他. (2001) Identification of two novel transmembrane gamma-carboxyglutamic acid proteins expressed broadly in fetal and adult tissues. Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375
		685373 TMG4	2.2E-14	[ヒト][原形質膜]Gla ファミリーのメンバーである推定上の膜貫通タンパク質で ある膜貫通 γ -カルボキシグルタミン酸タンパク質 4 は、WW ドメイン含有タン パク質との相互作用を仲介し得る、1 つの細胞質内 PPXY モチーフを含む Kulman, J. D. 他. (2001) Identification of two novel transmembrane gamma-carboxyglutamic acid proteins expressed broadly in fetal and adult tissues. Proc. Natl. Acad. Sci. U. S. A. 98:1370-1375.

10

20

30

40

【 0 3 5 5 】

【 表 3 - 1 】

表3-1

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
1	1895273CD1	289	Signal_cleavage:M1-G15 シグナルペプチド:M1-G15 免疫グロブリンドメイン:G120-A177, G32-G85 膜貫通ドメイン:L4-L20, P203-K230N-末端は細胞質内である。 受容体 FC 免疫グロブリン PD01270:E33-F61, L131-R166, M1-S26 γ FC 受容体 I 高親和性 免疫グロブリン前駆体 FC γ RI PD013195:W238-T289 受容体 FC γ 親和性 免疫グロブリン 前駆体 タンパク質 膜貫通 糖タンパク質 シグナル PD002534:W19-W64, L4-G15 受容体 FC γ I 高親和性 免疫グロブリン 前駆体 タンパク質 膜貫通 PD004578:SI57-K237, N65-V108 γ FC 受容体 I 高親和性 免疫グロブリン前駆体 FC γ RI PD013193:P107-T156	SPSCAN HMMER HMMER_PFAM TMAP BLIMPS_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM
2	7007222CD1	159	IG-様 C2-タイプドメイン DM03427P12314 189-331:F104-G247, P26151 198-339:F104-S246, I48471 199-336:E102-K235 ミエリン関連糖タンパク質 DM00682 P12314 1-187:G18-L103 潜在的リン酸化部位: S25 S246 S255 T30 T99 T150 T156 T164 T227 Y173 潜在的グリコシル化部位: N67 N74 N78 N110 N155 Signal_cleavage:M1-A26 シグナルペプチド:M1-A26 膜貫通ドメイン:P55-C83 N-末端は細胞質内 潜在的リン酸化部位: S141 T90 T132	BLAST_DOMO MOTIFS MOTIFS SPSCAN HMMER TMAP MOTIFS SPSCAN
3	3559223CD1	559	Signal_cleavage:M1-S19	SPSCAN

10

20

30

40

【 0 3 5 6 】

【 表 3 - 2 】

表3-2

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
4	3441255CD1	1222	潜在的リン酸化部位: S17 S31 S66 S93 S187 S188 S242 S260 S267 S318 S331 S412 S442 S446 S525 S549 T141 T148 T160 T200 T245 T426 T467 T480 T488 潜在的グリコシル化部位: N248 N386 Signal_cleavage: M1-F32 シグナルペプチド: M11-S30 シグナルペプチド: M11-E33 シグナルペプチド: M11-S35 膜貫通ドメイン: P8-E34, P179-N196, R773-S793, W801-L821 N 末端は細胞質内でない。	MOTIFS MOTIFS SPSCAN HMMER HMMER HMMER TMAP
5	1958917CD1	696	潜在的リン酸化部位: S30 S282 S444 S447 S465 S657 S701 S748 S806 S849 S873 S909 S976 S1007 S1033 S1060 T31 T54 T189 T249 T293 T416 T539 T672 T842 T844 T903 T1020 T1148 T1210 潜在的リン酸化部位: N38 N676 N698 N719 N884 Signal_cleavage: M1-A27 シグナルペプチド: M3-A27 シグナルペプチド: M3-A34 膜貫通ドメイン: S8-W36 N: 末端はサイトゾル内でない do ムチン; MUG5; 気管支; DM05454 S55316 1-317:Q355-S581 潜在的リン酸化部位: S72 S99 S122 S184 S383 S473 S475 S499 S506 S511 S549 S612 S649 S662 T235 T248 T512 T542 T618 潜在的グリコシル化部位: N483 Signal_cleavage: M1-A43 膜貫通ドメイン: R25-A53, V231-I258 N 末端はサイトゾル内でない EF-ハンドカルシウム-結合ドメイン BL00018:D80-F92 潜在的リン酸化部位: S77 S88 S107 S319 S332 S359 S385 S404 T81 Y398	MOTIFS SPSCAN HMMER HMMER TMAP BLAST_DOMO MOTIFS
6	6219465CD1	436	潜在的リン酸化部位: N483 Signal_cleavage: M1-A43 膜貫通ドメイン: R25-A53, V231-I258 N 末端はサイトゾル内でない EF-ハンドカルシウム-結合ドメイン BL00018:D80-F92 潜在的リン酸化部位: S77 S88 S107 S319 S332 S359 S385 S404 T81 Y398	MOTIFS SPSCAN TMAP BLIMPS_PRINTS MOTIFS

【 0 3 5 7 】
【 表 3 - 3 】

10

20

30

40

表3-3

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
7	3576625CD1	652	潜在的グリコシル化部位: N156 Signal_cleavage: M1-A27 シグナルペプチド: M1-A27 オルファクトメジン様ドメイン: G397-V652 膜貫通ドメイン: A4-S21 N 末端は細胞質内にならない タンパク質 前駆体 シグナル ミオシリン 小柱 (Trabecular) 糖質コルチコイド 糖タンパク質 オルファクトメジン 網目誘発応答 PD006897: L440-T645	MOTIFS SPSCAN HMMER HMMER_PFAM TMAP BLAST_PRODUM
8	4765758CD1	91	潜在的リン酸化部位: S21 S66 S82 S121 S155 S170 S242 S266 S314 S345 S382 S439 S519 S551 S574 T25 T151 T225 T275 T398 T565 T606 T624 Y75 Y236 潜在的グリコシル化部位: N159 N183 Signal_cleavage: M1-A42 シグナルペプチド: M20-A42 シグナルペプチド: M20-G44 クエン酸シントナーゼシグネチャ: E6-D67 リボソームタンパク質 S24e シグネチャ: K3-A57 潜在的リン酸化部位: S77 S82 T22 T70 T73	MOTIFS SPSCAN HMMER HMMER PROFILES SCAN PROFILES SCAN MOTIFS
9	7236661CD1	155	Signal_cleavage: M1-G17 シグナルペプチド: M1-G17 シグナルペプチド: M1-A20 潜在的リン酸化部位: S38 S43 S91 S95 S135 T47 T90 T115	SPSCAN HMMER HMMER MOTIFS
10	7714187CD1	765	Signal_cleavage: M1-A33 シグナルペプチド: M15-A33 シグナルペプチド: M15-A34 シグナルペプチド: M15-D37	SPSCAN HMMER HMMER MOTIFS SPSCAN HMMER HMMER HMMER

10

20

30

40

【 0 3 5 8 】
【 表 3 - 4 】

表3-4

SEQ ID NO:	Incyte ポリペプアミノ酸残基数	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
11	5136540CD1 150	シグナルペプチド:M15-Q38	HMMER
		シグナルペプチド:M15-A40	HMMER
		シグナルペプチド:M15-A33	HMMER
		膜貫通ドメイン:G8-A33	TMAP
		細胞膜傷害複合体構成物/パーフォリンシグネチャ L103-K155	PROFILES SCAN
		IB3089A PDI47692:M1-C765	BLAST_PROD OM
		潜在的リン酸化部位:S42 S49 S72 S157 S218 S267 S315 S361 S380 S418 S523 S528 S539 S563 S579 S592 S735 T119 T191 T309 T442 T504 T536 T599 T620 T761	MOTIFS
		潜在的リン酸化部位:N168 N337 N456 N562 N609 N641	MOTIFS
		Signal_cleavage:M1-G17	SPSCAN
		シグナルペプチド:M1-G17	HMMER
		シグナルペプチド:M1-P19	HMMER
12	3277403CD1 685	ピタミン K 依存性カルボキシシ化ドメイン:L57-A98	HMMER_P F AM
		ピタミン K 依存性カルボキシシ化ドメイン:V35-S110	PROFILES SCAN
		凝固因子 GLA ドメインシグネチャ PR00001:V84-A98, D56-C69, N70-F83	BLIMPS_PRINTS
		GLA ドメイン DM00454 P08709 22-100:K29-W93	BLAST_DOMO
		GLA ドメイン DM00454:P00742 2-80:G32-W93	BLAST_DOMO
		GLA ドメイン DM00454 P25155 2-80:C21-W93	BLAST_DOMO
		GLA ドメイン DM00454 S49075 2-80:G32-W93	BLAST_DOMO
		ピタミン K 依存性カルボキシシ化ドメイン:D56-W93	MOTIFS
		潜在的リン酸化部位:S38 S97 S110 S120 T37 T102	MOTIFS
		Signal_cleavage:M1-G33	SPSCAN
		シグナルペプチド:R16-A39	HMMER
シグナルペプチド:L10-A39	HMMER		
トリプシンインヒビター様 システインリッチドメイン:C629-C682	HMMER_P F AM		

10

20

30

40

【 0 3 5 9 】
【 表 3 - 5 】

表3-5

SEQ ID NO:	Incyte ポリペプ チドID	アミノ酸残 基数	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータ ベース
13	1517569CD1	126	フォンウイルブランド因子タイプCドメイン:C108-C163, C166-C224, C301-C357, C238-C289, C50-C105 フォンウイルブランド因子タイプDドメイン:C364-N514 C末端システインクノート BL01185:C238-C286, K351-V389 シグナル前駆体 糖タンパク質 ピテロガニン 細胞 反復 ウイルブランドフォン α TECTORIN PD001080:C331-C663 do ムチン、フォン、ウイルブランド、ヘモシチン、DM01378 P04275 512-692:A486-L669 do ムチン、フォン、ウイルブランド、ヘモシチン、DM01378 P04275 160-33:G495-P666 do ムチン、フォン、ウイルブランド、ヘモシチン、DM01378 P04275 986-1176:K492-H664 do ムチン、フォン、ウイルブランド、ヘモシチン、DM01378 P98092 370-549:K490-C663 ピタミンK依存性カルボキシ化ドメイン:K271-F308 VWFCドメインシグネチャ:C67-C105, C186-C224, C319-C357 潜在的リン酸化部位:S118 S123 S273 S479 S525 S573 T78 T260 T358 T437 T466 T549	HMMER_PFAM HMMER_PFAM BLIMPS_BLOCKS BLAST_PRODROM BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO MOTIFS MOTIFS MOTIFS MOTIFS SPSCAN SPSCAN TMAP MOTIFS MOTIFS HMMER MOTIFS MOTIFS SPSCAN MOTIFS SPSCAN
14	2415991CD1	149	潜在的グリコシル化部位:N30 N116 N247 N255 N318 N441 Signal_cleavage:M1-G33 膜貫通領域:I58-S86 潜在的リン酸化部位:S56 T13 T27 T52 Y71 潜在的グリコシル化部位:N11 シグナルペプチド:M1-G21 潜在的リン酸化部位:S122 潜在的グリコシル化部位:N33	MOTIFS SPSCAN TMAP MOTIFS MOTIFS HMMER MOTIFS MOTIFS SPSCAN MOTIFS SPSCAN
15	2735742CD1	114	Signal_cleavage:M1-A18 S63 S72 S105	SPSCAN MOTIFS SPSCAN
16	2768535CD1	519	Signal_cleavage:M1-L60	SPSCAN

10
20
30
40

【 0 3 6 0 】
【 冊 3 - 6 】

表3-6

SEQ ID	Incyte ポリペプチド ID	アミノ酸残基数	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
17	6848851CD1	1164	膜貫通領域:L38-G56 N-末端は細胞質内 潜在的リン酸化部位:S28 S29 S80 S113 S199 S217 S239 S296 S327 T2 T34 T110 T124 潜在的グリコシル化部位:N98 N289 N322 Signal_cleavage:M1-A45 膜貫通領域:S901-A923 N-末端はサイトゾル内 潜在的リン酸化部位:S18 S19 S65 S79 S107 S108 S110 S116 S128 S132 S216 S240 S241 S358 S379 S486 S506 S628 S665 S725 S956 S1027 T71 T78 T98 T112 T114 T162 T165 T209 T211 T225 T600 T604 T843 T1017 T1062 T1145 Y859 潜在的リン酸化部位:N377 N413 N935 N1009 Signal_cleavage:M1-E19 シグナルペプチド:M2-G23 潜在的リン酸化部位:S102 T86 潜在的グリコシル化部位:N103 Signal_cleavage:M1-G54	MOTIFS SPSCAN TMAP MOTIFS MOTIFS SPSCAN
18	7040722CD1	112	膜貫通領域:C23-S51 N-末端はサイトゾル内 潜在的リン酸化部位:S9 S51 S143 S144 T66 T121 Signal_cleavage:M1-A42, M18-A42, M14-A42 シグナルペプチド:M18-S44 シグナルペプチド:M18-A42 潜在的リン酸化部位:S20 T4 T53 Signal_cleavage:M1-A18 シグナルペプチド:M1-A18 潜在的リン酸化部位:S20 S50 T6 Signal_cleavage:M1-A42, M16-A42	MOTIFS SPSCAN HMMER MOTIFS MOTIFS SPSCAN TMAP MOTIFS SPSCAN HMMER HMMER MOTIFS SPSCAN HMMER MOTIFS SPSCAN
19	6430290CD1	170		
20	2640251CD1	80		
21	3839350CD1	118		
22	6393813CD1	140		

【 0 3 6 1 】
【 排 3 - 7 】

10

20

30

40

表3-7

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
23	5685755CD1	478	潜在的リン酸化部位:S23 S33 S61 S72 S122 S124 Signal_cleavage:M1-R31 シグナルペプチド:M1-A32 BTB/POZ ドメイン:Q47-L162 膜貫通ドメイン:G10-T26 N-末端はサイトソル内 タンパク質 C 関連 MAMA シグロフィリン ペプチジルロリル イソララゼ 腫瘍関連 MAC2 結合 PD014408:L53-A323, S338-P362 SPERACT 受容体アミノ末端 DM05438/A47161 126-585:S42-P345 潜在的リン酸化部位:S346 S361 S368 S400 S434 S452 T179 T355 T476 潜在的グリコシル化部位:N44 N61 N100 N195 N307 Signal_cleavage:M1-A15	MOTIFS SPSCAN HMNER HMNER_PFAM TMAP BLAST_PRODUM BLAST_DOMO MOTIFS MOTIFS SPSCAN
24	71728459CD1	80	シグナルペプチド:M1-A15 シグナルペプチド:M1-G28 膜貫通ドメイン:T4-W23 潜在的リン酸化部位:S26 S35 S54 S68 Signal_cleavage:M1-A20	HMNER HMNER TMAP MOTIFS SPSCAN
25	1904303CD1	505	シグナルペプチド:M1-A20 シグナルペプチド:M1-G24 膜貫通ドメイン:V312-L340, P467-L494 N-末端はサイトソル内 潜在的リン酸化部位:S21 S174 S237 S364 S368 S413 S426 T46 T176 T203 T248 T288 T338 T411	HMNER HMNER TMAP MOTIFS SPSCAN
26	2911343CD1	321	潜在的グリコシル化部位:N220 N229 N278 N336 Signal_cleavage:M39-A96 ロイシンリッチドメイン:N67-P88, S89-P110, C111-P131, S133-L162, E45-Q66	MOTIFS SPSCAN HMNER_PFAM

10

20

30

40

【 0 3 6 2 】
【 排 3 - 8 】

表3-8

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
27	7500308CD1	60	潜在的リン酸化部位: S36 S89 S175 S182 S183 S196 S304 S319 T23 T28 T42 T48 T58 T124 T145 T232 T260 T274 T311 T315 潜在的グリコシル化部位: N26 N138 Signal_cleavage: M1-P21 シグナルペプチド: M1-I25, M1-L26 潜在的リン酸化部位: S42 T33 Signal_cleavage: M1-G17 シグナルペプチド: M1-G17 シグナルペプチド: M1-G24 シグナルペプチド: M1-P19 ヒスチジン 酸性ホスファターゼシグネチャ: M1-A44	MOTIFS SPSCAN HMMER MOTIFS SPSCAN HMMER HMMER HMMER HMMER PROFILESCAN
28	7501098CD1	45	潜在的リン酸化部位: S18 S19 T40 Signal_cleavage: M1-A20 シグナルペプチド: M1-A19, M1-S21, M1-A20, M1-G24	SPSCAN HMMER MOTIFS SPSCAN HMMER
29	7503839CD1	43	潜在的リン酸化部位: S21 S174 S315 S319 S364 S377 T46 T176 T203 T239 T289 T362	MOTIFS SPSCAN HMMER MOTIFS
30	7503698CD1	456	潜在的グリコシル化部位: N220 N229 N287	SPSCAN HMMER MOTIFS

10

20

30

40

【 0 3 6 3 】

【 表 4 - 1 】

表 4-1

ポリヌクレオチド SEQ ID NO/ Incyte ID/配列長	配列断片
31/1895273CB1/ 1053	1-211, 1-320, 67-288, 67-335, 67-644, 69-216, 69-230, 69-286, 69-293, 69-331, 69-344, 69-358, 69-363, 69-481, 69-495, 69-591, 69-607, 69-617, 69-636, 112-506, 132-985, 150-394, 186-786, 188-967, 190-641, 190-726, 190-727, 241-713, 243-711, 258-1000, 264-950, 268-877, 276-991, 286-712, 287-995, 310-953, 326-614, 367-1027, 423-606, 428-1036, 429-679, 434-724, 455-709, 459-653, 459-676, 459-998, 471-1003, 483-697, 488-1032, 489-971, 503-1050, 507-1023, 508-1053, 598-645, 743-1051
32/70072222CB1/ 1579	1-550, 1-613, 4-295, 28-284, 360-627, 362-623, 372-623, 377-664, 406-930, 430-830, 451-607, 459-1027, 459-1048, 544-1022, 564-1027, 590-1146, 609-1134, 665-1320, 676-1263, 691-1053, 694-1146, 709-1320, 730-1209, 740-1263, 746-1061, 751-1015, 751-1223, 772-1205, 784-1270, 792-1332, 805-1234, 805-1246, 815-1322, 816-1460, 843-1362, 860-1263, 876-1362, 904-1160, 925-1492, 943-1575, 976-1544, 1024-1307, 1033-1579, 1041-1069, 1041-1339, 1041-1341, 1041-1342, 1041-1343, 1041-1346, 1041-1347, 1045-1172, 1046-1069, 1054-1579, 1057-1347, 1084-1544, 1104-1579, 1120-1579
33/3559223CB1/ 2440	1-471, 51-284, 128-825, 129-705, 222-334, 253-785, 282-593, 283-699, 283-786, 357-400, 408-1003, 457-699, 457-826, 467-1050, 469-993, 520-1132, 534-752, 585-866, 585-1061, 635-1291, 703-966, 747-1291, 781-1350, 813-1417, 820-1034, 906-1151, 925-1291, 942-1341, 1021-1294, 1064-1326, 1083-1591, 1110-1380, 1110-1673, 1132-1482, 1164-1340, 1169-1705, 1237-1490, 1237-1553, 1237-1607, 1246-1793, 1253-1758, 1253-1843, 1270-1876, 1348-1997, 1418-1729, 1474-2026, 1486-1705, 1495-1683, 1495-1716, 1517-2020, 1551-2206, 1608-1890, 1611-1754, 1653-2268, 1665-2002, 1671-1947, 1681-1942, 1691-1801, 1731-2269, 1741-2278, 1752-1821, 1759-2418, 1780-2270, 1793-2408, 1819-2366, 1827-2321, 1829-2018, 1829-2358, 1839-2404, 1846-2362, 1883-2345, 1889-2337, 1901-2118, 1908-2061, 1908-2365, 1913-2171, 1913-2435, 1935-2420, 1936-2174, 1943-2314, 1957-2423, 1963-2440, 1968-2424, 1969-2423, 1970-2422, 1971-2411, 1975-2423, 1986-2423, 1997-2440, 2000-2266, 2009-2420, 2018-2440, 2019-2277, 2020-2414, 2023-2440, 2036-2432, 2038-2440, 2043-2424, 2044-2434, 2053-2424, 2066-2420, 2070-2420, 2071-2420, 2095-2395, 2141-2440, 2146-2420, 2147-2420, 2157-2416, 2157-2430, 2158-2334, 2158-2376, 2158-2390, 2201-2418, 2279-2420, 2311-2420, 2066-2420, 2070-2420, 2071-2420, 2095-2395, 2141-2440, 2146-2420, 2147-2420, 2157-2416, 2157-2430, 2158-2334, 2158-2376, 2158-2390, 2201-2418, 2279-2420, 2311-2420

【 0 3 6 4 】
【 表 4 - 2 】

10

20

30

40

表4-2

34/3441255CBI/ 4133	1-349, 1-651, 243-430, 489-4133, 655-1312, 665-1263, 925-1207, 925-1545, 932-1543, 945-1617, 1040-1585, 1260-1709, 1293-1597, 1388-1675, 1388-2004, 1681-1927, 1700-1936, 1700-2164, 1725-2284, 1796-2295, 1809-2215, 2074-2304, 2074-2336, 2074-2465, 2127-2774, 2148-2409, 2205-2442, 2515-2735, 2515-2918, 2515-3045, 2517-2784, 2640-2908, 2659-3187, 2725-3340, 2743-2990, 2776-3233, 2777-3023, 2790-3014, 2794-3321, 2832-3337, 2832-3649, 2841-3160, 2872-3447, 2948-3199, 2988-3368, 3027-3368, 3105-3321, 3237-3507, 3327-3983, 3327-4004, 3381-4029, 3406-3904, 3427-4003, 3514-4077, 3594-4128, 3632-4077, 3653-3886, 3709-4018
35/1958917CBI/ 4689	1-70, 1-623, 2-516, 185-463, 356-989, 464-660, 464-846, 609-803, 609-896, 609-1018, 609-1133, 609-1137, 617-1058, 617-1201, 660-1445, 676-1349, 677-904, 679-1420, 687-1445, 706-935, 709-1445, 725-1137, 726-1443, 779-1137, 782-904, 852-1137, 936-1137, 977-1137, 1000-1137, 1029-1202, 1044-1137, 1089-1137, 1097-1137, 1113-1137, 1138-1327, 1203-1327, 1203-1502, 1213-1744, 1328-1502, 1328-1616, 1375-1766, 1394-1654, 1394-1806, 1394-1827, 1489-2117, 1503-1744, 1617-1744, 1617-2054, 1655-2312, 1745-2054, 1777-2316, 1848-2366, 2001-2435, 2047-2477, 2055-2275, 2092-2575, 2132-2409, 2180-2748, 2273-2762, 2283-2724, 2313-2555, 2355-2802, 2369-2917, 2467-3119, 2530-2979, 2547-3208, 2572-3124, 2576-2805, 2591-2896, 2603-2875, 2670-3140, 2671-2971, 2683-3224, 2683-3269, 2695-3261, 2711-3335, 2821-3340, 2821-3354, 3048-3544, 3122-3694, 3160-3603, 3172-3376, 3217-3410, 3217-3752, 3245-3782, 3332-3875, 3349-3924, 3435-3924, 3503-4186, 3503-4244, 3504-3748, 3551-3840, 3551-4036, 3611-4005, 3631-3845, 3714-3745, 3714-3750, 3714-3751, 3714-3753, 3714-3780, 3714-3781, 3714-3792, 3714-3797, 3714-3798, 3714-3801, 3714-3828, 3714-3829, 3714-3846, 3714-3848, 3714-3876, 3714-3896, 3717-3755, 3718-3753, 3718-3800, 3719-4008, 3721-4145, 3735-3893, 3735-3896, 3761-3795, 3761-3801, 3761-3830, 3761-3887, 3761-3892, 3761-3923, 3761-3944, 3766-3944, 3768-3944, 3777-3850, 3782-4032, 3782-4301, 3809-3843, 3809-3848, 3809-3877, 3809-3935, 3809-3940, 3809-3943, 3813-3847, 3813-3895, 3814-3943, 3816-3943, 3830-3944, 3856-3890, 3856-3896, 3856-3925, 3856-3944, 3856-3945, 3860-3894, 3860-4689, 3861-3944, 3863-3944, 3877-3943, 3904-3938, 3904-3943, 3904-3949, 3904-4188, 3908-3942, 3909-3943, 3911-3943, 3930-4402, 3976-4498, 3979-4543, 3982-4587, 3993-4262, 3993-4480, 4083-4304, 4108-4359, 4136-4328, 4136-4416, 4246-4537, 4360-4645, 4491-4547, 4628-4651

表4-3

36/6219465CBI/ 3294	<p>1-1311, 168-355, 168-483, 168-548, 168-573, 168-588, 168-592, 168-601, 168-643, 168-646, 168-679, 168-682, 168-686, 168-689, 168-696, 168-715, 168-745, 168-765, 168-797, 168-819, 168-822, 168-835, 168-843, 169-745, 171-551, 172-577, 172-587, 172-608, 172-745, 174-745, 175-771, 178-678, 178-733, 178-745, 187-642, 188-526, 189-526, 202-498, 211-798, 216-745, 224-349, 224-532, 224-535, 224-745, 226-721, 229-745, 230-452, 230-525, 237-415, 237-596, 238-526, 251-436, 263-772, 278-813, 306-951, 325-967, 326-800, 338-510, 341-510, 341-868, 347-745, 382-790, 397-970, 513-745, 514-977, 561-1309, 567-800, 568-804, 570-772, 570-843, 593-1149, 641-1123, 641-1266, 663-1116, 698-1255, 703-1188, 727-1173, 740-1313, 755-1176, 763-1333, 765-1384, 772-1298, 785-1417, 788-1148, 788-1366, 788-1381, 788-1389, 788-1467, 820-1413, 829-1319, 835-1313, 836-1437, 837-1511, 854-1422, 870-1516, 873-1212, 882-1101, 918-986, 936-986, 940-986, 951-986, 962-986, 1014-1226, 1014-1443, 1014-1456, 1014-1485, 1014-1544, 1014-1548, 1014-1685, 1015-1476, 1016-1105, 1031-1561, 1031-1588, 1046-1528, 1049-1458, 1049-1694, 1079-1590, 1086-1604, 1089-1469, 1096-1288, 1108-1730, 1109-1626, 1110-1532, 1113-1384, 1113-1564, 1129-1341, 1136-1783, 1142-1640, 1144-1760, 1172-1692, 1177-1767, 1183-1790, 1186-1773, 1190-1524, 1194-1743, 1194-1790, 1200-1642, 1202-1790, 1207-1790, 1211-1783, 1215-1712, 1215-1805, 1217-1747, 1223-1776, 1224-1790, 1236-1790, 1236-1864, 1237-1827, 1244-1682, 1244-1790, 1257-1905, 1259-1835, 1289-1790, 1292-1873, 1294-1648, 1295-1541, 1302-1790, 1335-1773, 1337-1790, 1344-1978, 1352-1790, 1358-1663, 1370-2054, 1378-1888, 1411-1922, 1434-1596, 1434-1927, 1439-2074, 1445-1868, 1461-1868, 1477-1790, 1505-1958, 1506-1913, 1516-1807, 1529-1790, 1549-1793, 1587-1818, 1587-1889, 1587-1998, 1587-2083, 1587-2139, 1587-2224, 1587-2254, 1589-2051, 1595-1840, 1640-2133, 1644-1960, 1646-2251, 1647-2275, 1648-2280, 1667-1970, 1670-1995, 1671-2250, 1672-2225, 1673-1939, 1677-2245, 1715-2331, 1720-1986, 1730-2238, 1736-2233, 1753-1928, 1766-2356, 1770-2293, 1777-2448, 1792-2284, 1792-2380, 1793-2244, 1797-2050, 1799-2441, 1804-2342, 1818-2225, 1822-2434, 1829-2135, 1846-2088, 1862-2540, 1885-2470, 1890-2483, 1892-2145, 1909-2535, 1912-2526, 1933-2130, 1943-2516, 1954-2582, 1963-2492, 1971-2559, 1971-2614, 1980-2524, 1983-2503, 1989-2438, 1993-2464, 1999-2547, 1999-2733, 2008-2492, 2009-2420, 2009-2535, 2010-2614, 2017-2464, 2051-2332, 2065-2669, 2071-2535, 2104-2689, 2109-2649, 2117-2669, 2118-2379, 2125-2748, 2131-2558, 2136-2639, 2146-2739, 2146-2802, 2155-2686, 2159-2821, 2161-2685, 2162-2748, 2170-2692, 2189-2702, 2200-2740, 2222-2723, 2250-2764, 2264-2780, 2293-2720, 2296-2537, 2299-2748, 2328-2575, 2359-2608, 2359-2929, 2363-2599, 2381-2679, 2385-2729, 2393-2693, 2476-2690, 2493-3030, 2494-2712, 2494-2721, 2530-2776, 2608-2872, 2618-2850, 2667-2925, 2667-2948, 2698-2949, 2700-2947, 2701-2920, 2706-2989, 2763-2964, 2763-2970, 2776-3020, 2781-2976, 2793-3038, 2859-3020, 2869-3140, 2967-3239, 3040-3186, 3040-3294</p>
------------------------	--

10

20

30

40

【 0 3 6 6 】
【 表 4 - 4 】

表4-4

37/3576625CBI/ 3635	1-90, 1-588, 62-157, 62-354, 62-464, 62-551, 62-564, 62-582, 62-686, 67-282, 67-571, 68-623, 71-657, 75-618, 78-571, 78-582, 80-493, 91-1959, 98-332, 98-354, 98-543, 98-549, 108-677, 125-756, 130-815, 175-814, 227-722, 245-892, 258-622, 337-1004, 348-416, 355-433, 396-779, 423-661, 461-613, 677-1098, 792-1075, 798-1386, 810-1458, 818-1106, 866-1113, 876-1458, 935-1098, 1133-1568, 1141-1713, 1153-1445, 1167-1598, 1169-1383, 1355-1735, 1355-1745, 1355-1794, 1355-1812, 1355-1836, 1355-1852, 1355-1856, 1355-1885, 1355-1887, 1355-1891, 1355-1918, 1355-1926, 1355-1962, 1355-1976, 1356-1891, 1389-1695, 1389-1900, 1389-1965, 1389-2010, 1389-2017, 1414-1891, 1556-2122, 1556-2177, 1556-2185, 1556-2202, 1556-2216, 1567-1937, 1575-1988, 1578-2212, 1579-2035, 1579-2081, 1579-2118, 1579-2215, 1579-2216, 1580-2190, 1581-2212, 1584-2216, 1588-1891, 1588-2216, 1641-2212, 1654-2216, 1658-1963, 1660-2216, 1674-2243, 1680-2212, 1714-2230, 1719-2295, 1726-2121, 1734-1869, 1748-2305, 1752-2334, 1753-2203, 1760-2419, 1783-2305, 1806-2367, 1817-2393, 1866-2332, 1868-2527, 1877-2525, 1883-2554, 1884-2071, 1910-2141, 1919-2565, 1930-2345, 1934-2595, 1952-2422, 1996-2712, 2040-2525, 2044-2726, 2053-2630, 2058-2754, 2092-2616, 2107-2707, 2114-2736, 2127-2713, 2128-2659, 2159-2740, 2168-2510, 2231-2749, 2240-2707, 2245-2491, 2246-2885, 2285-2961, 2320-2972, 2330-2895, 2350-2806, 2380-2903, 2393-3096, 2495-2705, 2496-3107, 2511-3203, 2547-2927, 2553-3249, 2573-3263, 2602-3263, 2607-3232, 2623-3100, 2626-2882, 2643-3294, 2688-2943, 2689-3407, 2699-3154, 2699-3249, 2712-3274, 2719-3350, 2734-3203, 2752-2993, 2759-3413, 2770-3455, 2773-3326, 2773-3437, 2779-3440, 2783-3146, 2796-2825, 2856-3402, 2863-3154, 2869-3322, 2887-3435, 2932-3635, 2947-3523, 2951-3473, 3002-3497, 3127-3635, 3129-3522, 3131-3352, 3195-3431, 3200-3460, 3216-3483
38/4765758CBI/ 635	1-280, 1-611, 23-635
39/7236661CBI/ 793	1-56, 1-210, 1-379, 1-414, 1-503, 1-565, 1-793, 43-679, 63-679, 130-565, 202-793
40/7714187CBI/ 2741	1-140, 15-532, 15-874, 158-262, 189-784, 379-662, 485-1033, 680-1171, 980-1068, 1062-1612, 1069-1599, 1292-1588, 1373-1629, 1374-1843, 1404-2741, 1585-1828, 1585-2271, 1587-1822, 1682-2250, 2065-2734
41/5136540CBI/ 1074	1-252, 137-209, 137-412, 137-556, 137-575, 137-583, 137-597, 137-707, 137-738, 137-767, 145-191, 169-699, 290-675, 298-791, 315-574, 417-901, 426-901, 428-1071, 428-1074, 431-518, 452-1026, 468-1062, 488-1048, 518-771, 518-795, 518-920, 518-1062, 521-1054, 523-773, 523-1068, 527-974, 536-980, 561-785, 574-1068, 595-1070, 595-1071, 626-1071, 637-1071, 647-1058, 811-1071, 984-1068
42/3277403CBI/ 2560	1-76, 1-632, 7-76, 19-71, 21-76, 119-426, 189-458, 234-750, 269-526, 381-2519, 472-1015, 667-1197, 994-1487, 1224-1446, 1281-1385, 1453-1603, 1516-1790, 1516-1967, 1516-2010, 1620-2113, 1621-2213, 1677-1962, 1677-2201, 1711-1976, 1711-1999, 1743-2271, 1858-2560, 1863-2066, 1865-2303, 2112-2358, 2528-2559

【 0 3 6 7 】
【 表 4 - 5 】

表4-5

<p>43/1517569CBI/ 5833</p> <p>1-541, 1-604, 12-686, 47-797, 124-507, 179-765, 198-661, 234-822, 237-811, 275-947, 377-1021, 414-1020, 430-909, 472-1022, 517-986, 547-1067, 640-1021, 713-1086, 790-1268, 793-1366, 793-1446, 797-1333, 797-1443, 797-1453, 799-1225, 799-1333, 850-1493, 949-1544, 1064-1706, 1073-1570, 1122-1724, 1131-1704, 1137-1707, 1218-1916, 1279-1535, 1392-1617, 1501-1559, 1505-1755, 1505-1971, 1522-1749, 1609-1871, 1609-1877, 1609-2052, 1824-2236, 1930-2502, 1963-2559, 1980-2641, 1981-2454, 2019-2590, 2038-2544, 2038-2672, 2038-2706, 2149-2767, 2175-2494, 2200-2428, 2200-2692, 2201-2738, 2211-2858, 2220-2931, 2265-2968, 2267-2463, 2268-2480, 2296-2888, 2333-2588, 2352-2971, 2355-2614, 2478-3083, 2479-3121, 2487-3119, 2493-3096, 2523-3136, 2552-3021, 2589-3194, 2591-3249, 2598-3150, 2608-3176, 2619-3303, 2627-3219, 2635-3357, 2638-3040, 2639-3049, 2639-3211, 2640-3052, 2643-3321, 2648-3147, 2677-3184, 2702-3275, 2712-2972, 2715-2986, 2715-3188, 2720-3099, 2733-3428, 2789-3021, 2872-3480, 2882-3443, 2882-3571, 2884-3350, 2888-3388, 2888-3591, 2889-3167, 2890-3360, 2948-3681, 2954-3443, 2960-3664, 2975-3571, 2976-3689, 2986-3652, 2991-3692, 2997-3685, 3026-3579, 3032-3667, 3048-3713, 3064-3530, 3080-3375, 3085-3729, 3109-3362, 3114-3548, 3118-3722, 3118-3726, 3119-3399, 3119-3722, 3120-3693, 3135-3738, 3136-3736, 3138-3736, 3143-3377, 3148-3739, 3153-3415, 3160-3435, 3160-3670, 3160-3734, 3162-3713, 3162-3738, 3164-3422, 3173-3424, 3179-3688, 3183-3517, 3187-3739, 3188-3738, 3193-3728, 3206-3482, 3207-3741, 3209-3692, 3213-3420, 3231-3525, 3231-3725, 3231-3743, 3232-3738, 3234-3738, 3243-3508, 3252-3741, 3292-3858, 3299-3858, 3699-4031, 3699-4088, 3718-4036, 3749-3968, 3749-4255, 3749-4285, 3814-4000, 3926-4281, 4087-4111, 4087-4115, 4087-4165, 4087-4196, 4087-4200, 4087-4250, 4089-4170, 4094-4250, 4126-4250, 4128-4717, 4137-4250, 4141-4165, 4172-4316, 4172-4333, 4172-4335, 4179-4431, 4211-4385, 4211-4394, 4213-4335, 4222-4491, 4226-4250, 4257-4335, 4264-4335, 4296-4335, 4298-4335, 4307-4335, 4311-4536, 4408-4674, 4408-4963, 4487-4797, 4664-4956, 4671-5297, 4871-5193, 4871-5396, 4957-5255, 5024-5292, 5024-5586, 5033-5297, 5085-5238, 5098-5365, 5158-5369, 5391-5630, 5406-5684, 5406-5699, 5406-5700, 5406-5714, 5408-5691, 5442-5718, 5513-5740, 5542-5819, 5549-5698, 5584-5833, 5607-5833, 5665-5727</p>	<p>44/2415991CBI/ 2264</p> <p>1-707, 230-774, 245-538, 308-596, 351-1022, 419-681, 419-1022, 451-911, 451-959, 487-707, 487-727, 487-989, 598-906, 599-933, 611-858, 620-833, 661-1022, 713-1308, 786-1423, 798-1084, 875-1107, 890-1681, 892-1474, 903-1172, 927-1197, 927-1227, 947-1173, 1009-1271, 1022-1291, 1171-1756, 1174-1435, 1175-1471, 1206-1423, 1244-1760, 1274-1463, 1298-1559, 1308-1539, 1308-1722, 1308-1884, 1310-1590, 1311-1521, 1321-1591, 1344-1623, 1344-1997, 1359-1565, 1395-1654, 1453-1744, 1510-2164, 1528-2178, 1536-1819, 1539-1827, 1570-2221, 1592-2216, 1644-1875, 1644-2155, 1648-2225, 1698-1959, 1730-1959, 1747-2264, 1748-1914, 1757-2019, 1812-2061, 1817-2043, 1893-2108, 1895-2166, 2037-2248, 2054-2264, 2063-2134</p>
---	---

10

20

30

40

【 0 3 6 8 】
【 表 4 - 6 】

表4-6

<p>45/2735742CBI/ 1638</p>	<p>1-625, 326-449, 326-502, 326-532, 326-533, 326-534, 326-544, 326-548, 326-549, 326-552, 326-557, 326-561, 326-568, 326-569, 326-571, 326-572, 326-573, 326-574, 326-575, 326-578, 326-579, 326-584, 326-587, 326-590, 326-599, 326-601, 326-610, 326-662, 326-747, 326-771, 326-836, 326-848, 326-894, 345-566, 362-586, 362-595, 372-624, 418-618, 460-719, 482-725, 516-675, 582-847, 588-715, 599-885, 601-803, 601-814, 627-986, 633-884, 646-900, 653-972, 680-895, 680-904, 684-917, 711-1000, 738-945, 757-993, 763-1019, 764-995, 798-1042, 805-1053, 830-1101, 878-1102, 889-1286, 914-1526, 931-1157, 937-1152, 942-1201, 942-1519, 947-1156, 947-1321, 958-1183, 965-1188, 965-1195, 965-1514, 997-1566, 1019-1292, 1033-1269, 1050-1262, 1085-1319, 1112-1389, 1133-1369, 1146-1386, 1146-1475, 1146-1587, 1191-1381, 1191-1549, 1203-1330, 1350-1563, 1352-1612, 1363-1638, 1395-1582, 1395-1594</p>
<p>46/2768535CBI/ 4790</p>	<p>1-520, 46-518, 46-520, 88-707, 129-472, 186-883, 187-765, 206-353, 421-1034, 505-1159, 623-908, 644-1251, 652-1211, 654-1195, 678-1221, 690-1269, 703-1254, 704-1393, 706-989, 762-1001, 770-1352, 781-1389, 797-1335, 813-1420, 870-1413, 945-1256, 1050-1703, 1108-1788, 1114-1738, 1132-1698, 1173-1815, 1189-1796, 1195-1801, 1204-1798, 1220-1719, 1248-1823, 1262-1921, 1266-1924, 1270-1883, 1272-1807, 1283-1664, 1283-1870, 1336-2021, 1425-2056, 1438-2076, 1527-2070, 1621-2136, 1676-2224, 1696-2084, 1750-2313, 1863-2252, 1918-2152, 1923-2186, 1926-2568, 1964-2230, 1964-2464, 2072-2606, 2275-2500, 2401-3007, 2464-2726, 2470-2676, 2494-2785, 2533-2819, 2540-2809, 2595-3128, 2621-2839, 2640-2882, 2649-2884, 2654-2925, 2656-2930, 2698-2958, 2727-2928, 2727-3213, 2782-2953, 2782-3400, 2853-3148, 2858-3045, 2889-3189, 2891-3338, 2891-3379, 2894-3205, 2934-3187, 2948-3209, 2972-3515, 2986-3244, 3000-3209, 3003-3246, 3027-3291, 3028-3230, 3028-3300, 3035-3298, 3074-3268, 3074-3298, 3074-3666, 3106-3350, 3192-3404, 3192-3456, 3221-3449, 3223-3640, 3266-3531, 3266-3756, 3289-3860, 3292-3804, 3305-3643, 3308-3608, 3309-3598, 3361-3649, 3364-3659, 3400-3643, 3406-3654, 3407-3672, 3411-3811, 3420-3867, 3421-3695, 3424-3703, 3428-3705, 3450-3787, 3454-3696, 3460-3739, 3473-3868, 3473-3875, 3487-3648, 3498-3745, 3503-3767, 3516-3753, 3533-3728, 3539-3762, 3599-3809, 3601-3733, 3620-4123, 3638-3878, 3647-3669, 3647-3884, 3647-3919, 3647-4186, 3716-3988, 3724-3996, 3737-4008, 3746-3998, 3752-3983, 3755-4415, 3777-4010, 3777-4404, 3809-4098, 3823-4098, 3836-4109, 3857-4095, 3858-4065, 3863-4367, 3875-4093, 3882-4128, 3900-4146, 3909-4147, 3984-4413, 3995-4242, 4025-4670, 4043-4296, 4056-4349, 4086-4123, 4100-4398, 4100-4593, 4100-4688, 4115-4333, 4116-4594, 4143-4696, 4147-4677, 4152-4410, 4191-4680, 4202-4429, 4215-4494, 4221-4479, 4222-4509, 4277-4532, 4318-4682, 4322-4659, 4328-4543, 4380-4630, 4394-4669, 4422-4661, 4425-4667, 4439-4696, 4444-4696, 4450-4665, 4459-4696, 4472-4676, 4494-4696, 4494-4790</p>

【 0 3 6 9 】
【 表 4 - 7 】

表4-7

47/6848851CBI/ 3916	1-293, 136-477, 137-447, 170-482, 171-429, 171-477, 172-432, 172-446, 173-303, 173-433, 173-447, 174-416, 175-443, 180-393, 193-365, 193-447, 197-447, 202-447, 204-476, 206-476, 257-599, 366-599, 366-749, 386-429, 502-684, 509-743, 600-749, 600-1197, 604-1073, 604-1184, 627-1215, 627-1308, 750-1197, 750-1603, 775-951, 921-1562, 1143-1704, 1198-1603, 1198-1737, 1244-1601, 1293-1712, 1365-1613, 1596-1748, 1596-2025, 1604-1737, 1604-1858, 1605-1997, 1680-2413, 1728-2255, 1732-2340, 1749-1997, 1782-2357, 2040-2689, 2040-2690, 2059-2331, 2254-2690, 2421-2777, 2421-2794, 2424-2824, 2435-3135, 2465-2742, 2480-2802, 2551-3204, 2571-2852, 2609-3171, 2637-2912, 2732-3429, 2946-3149, 2946-3219, 2989-3258, 2994-3238, 3162-3406, 3186-3436, 3187-3610, 3214-3590, 3214-3620, 3226-3767, 3232-3512, 3235-3493, 3235-3616, 3254-3916, 3280-3916, 3378-3607
48/7040722CBI/ 1702	1-1702, 5-561, 20-561, 107-167, 107-194, 107-240, 264-546, 420-852, 420-878, 420-1108, 420-1120, 424-1094, 947-1702, 952-988, 953-1702, 956-1702, 965-1702
49/6430290CBI/ 1462	1-274, 27-274, 28-274, 29-274, 29-583, 40-682, 567-1426, 567-1429, 567-1448, 567-1458, 567-1462, 931-1211
50/2640251CBI/ 3958	1-282, 19-539, 73-658, 78-392, 83-598, 97-759, 106-624, 236-700, 335-611, 337-611, 375-634, 384-987, 404-759, 489-700, 508-825, 523-790, 528-947, 594-899, 600-899, 600-902, 605-899, 668-931, 678-925, 685-899, 703-947, 746-983, 780-1061, 843-1073, 881-1389, 977-1275, 1043-1287, 1046-1323, 1111-1692, 1231-1506, 1322-1572, 1379-1630, 1408-1814, 1408-1873, 1433-2210, 1440-1692, 1563-2005, 1614-2034, 1626-2024, 1732-2034, 1750-1972, 1820-2378, 1853-2417, 1975-2249, 2009-2210, 2043-2287, 2084-2626, 2091-2619, 2454-3008, 2468-3078, 2480-3019, 2584-3143, 2678-3169, 2687-3270, 2754-3231, 2901-3169, 2902-3169, 2936-3597, 3261-3958, 3382-3686, 3382-3779
51/3839350CBI/ 826	1-296, 1-432, 1-450, 1-549, 1-563, 1-796, 3-360, 3-826, 102-826, 161-430, 190-798, 199-401, 221-736, 246-564, 254-773, 272-489, 299-436
52/6393813CBI/ 729	1-501, 18-719, 32-407, 32-493, 42-536, 51-302, 57-630, 58-305, 68-351, 68-608, 124-411, 173-710, 193-644, 208-232, 216-516, 230-463, 230-716, 243-706, 266-706, 278-536, 283-729, 321-729, 334-717, 538-729, 594-655
53/5685755CBI/ 1610	1-201, 76-697, 79-316, 197-478, 197-1448, 349-828, 351-803, 351-845, 353-846, 354-843, 358-845, 362-820, 363-828, 365-846, 434-846, 459-846, 476-787, 479-1448, 480-1553, 530-816, 546-846, 606-846, 609-828, 839-1279, 839-1333, 840-1331, 843-1210, 865-1323, 874-1051, 882-1321, 1252-1610
54/71728459CBI/ 869	1-337, 1-576, 1-617, 1-869, 120-487, 283-749, 283-750, 285-696, 286-739, 326-829, 327-828, 383-774, 420-819, 456-869

10

20

30

40

【 0 3 7 0 】
【 表 4 - 8 】

表4-8

55/1904303CB1/ 2209	1-515, 290-1414, 404-692, 416-828, 422-698, 429-676, 429-980, 530-747, 530-765, 530-1056, 586-858, 624-835, 744-1052, 761-1021, 780-1070, 793-1288, 877-1193, 976-1439, 978-1330, 1018-1382, 1057-1255, 1057-1397, 1096-1378, 1097-1412, 1097-1414, 1173-1451, 1203-1325, 1247-1687, 1247-1690, 1248-2190, 1254-1539, 1258-1414, 1298-1773, 1339-1918, 1345-1614, 1345-1880, 1345-1918, 1399-1618, 1434-1836, 1459-2128, 1500-2115, 1517-1747, 1523-1786, 1582-2168, 1587-2209, 1590-2179, 1693-2200, 1719-2165, 1722-2208, 1725-2190, 1761-2184, 1763-2188, 1772-2189, 1774-2182, 1779-2187, 1788-2156, 1796-2187, 1812-2208, 1818-2182, 1820-2208, 1822-2182, 1823-2192, 1825-2187, 1827-2182, 1833-2182, 1834-2182, 1846-2190, 1849-2192, 1858-2190, 1861-2184, 1863-2192, 1870-2182, 1871-2182, 1878-2180, 1879-2160, 1879-2180, 1897-2182, 1907-2187, 1928-2208, 2004-2167, 2095-2182
56/2911343CB1/ 1520	1-582, 1-1513, 43-138, 43-315, 43-573, 86-573, 101-347, 202-456, 465-569, 466-569, 523-996, 525-772, 528-1013, 571-831, 571-838, 573-1124, 578-1373, 620-849, 658-806, 666-931, 666-1181, 732-1273, 732-1368, 735-1345, 774-1377, 800-944, 829-1410, 830-1400, 834-1104, 885-1146, 933-1210, 939-1520, 942-1460, 970-1191, 999-1274, 1000-1411, 1051-506, 1053-1514, 1053-1520, 1066-1520, 1067-1520, 1068-1514, 1075-1470, 1083-1448, 1098-1520, 1108-1423, 1111-1493, 1138-1514, 1152-1437, 1163-1520, 1180-1520, 1187-1514, 1260-1520, 1274-1514
57/7500308CB1/ 1282	1-1282, 4-295, 363-612, 363-888, 402-684, 402-707, 402-1050, 406-1001, 407-1018, 439-1006, 452-764, 454-859, 463-718, 474-1001, 495-1035, 516-976, 580-1045, 646-1278, 748-1046, 748-1050, 835-1116, 887-1282, 1011-1243, 1060-1282
58/7501098CB1/ 1228	1-252, 1-394, 1-397, 1-399, 1-1014, 87-193, 87-197, 87-198, 359-843, 363-843, 370-1010, 370-1016, 394-968, 410-1004, 430-990, 460-713, 460-862, 460-1024, 463-996, 465-715, 465-1228, 469-916, 537-1016, 568-1019, 580-1018
59/7503839CB1/ 3582	1-128, 10-3582, 215-936, 270-739, 270-987, 293-883, 293-933, 293-1125, 441-617, 587-1256, 602-1405, 803-1370, 909-1423, 959-1382, 1014-1803, 1031-1279, 1081-1663, 1262-1414, 1262-1691, 1279-2186, 1346-1720, 1347-1654, 1347-1700, 1347-1705, 1347-1720, 1347-1868, 1347-2098, 1394-1922, 1394-2009, 1415-1663, 1446-2023, 1706-2355, 1706-2356, 1710-2355, 1717-2355, 1725-1999, 1779-2355, 1888-2355, 1909-2355, 1920-2355, 1923-2355, 1976-2284, 2002-2355, 2034-2947, 2087-2443, 2087-2460, 2090-2490, 2101-2801, 2137-2265, 2137-2408, 2147-2468, 2217-2870, 2237-2518, 2275-2837, 2303-2578, 2334-2745, 2340-3263, 2341-2615, 2341-2879, 2341-2920, 2341-2927, 2341-3008, 2341-3037, 2341-3051, 2341-3082, 2341-3088, 2341-3134, 2341-3233, 2346-3199, 2352-3252, 2384-2822, 2398-3095, 2423-2696, 2488-2721, 2489-3363, 2562-3289, 2612-2815, 2612-2885, 2614-3559, 2632-3296, 2660-2904, 2661-2924, 2674-3520, 2710-3484, 2828-3072, 2853-3102, 2853-3275, 2874-3555, 2880-3256, 2898-3178, 2901-3159, 2901-3282, 2920-3582, 2946-3582, 3049-3273, 3142-3546, 3258-3449, 3258-3561, 3373-3501

【 0 3 7 1 】
【 表 4 - 9 】

表4-9

60/7503698CB1/ 2063	1-2043, 404-692, 416-828, 422-698, 429-676, 530-747, 530-765, 530-1056, 586-858, 624-835, 744-1052, 761-1021, 780-1070, 897-1056, 963-1267, 1100-1540, 1100-1543, 1107-1392, 1111-1267, 1147-1267, 1151-1626, 1192-1771, 1198-1352, 1198-1467, 1198-1733, 1198-1800, 1236-1984, 1252-1471, 1312-1606, 1312-1981, 1370-1600, 1384-1609, 1435-2021, 1440-2043, 1443-2032, 1517-1809, 1518-1743, 1546-2043, 1572-2018, 1575-2043, 1578-2063, 1580-2037, 1584-2039, 1614-2037, 1616-2041, 1617-1993, 1625-2042, 1627-2035, 1632-2040, 1649-2040, 1665-2043, 1671-2035, 1673-2043, 1675-2035, 1676-2035, 1678-2040, 1680-2035, 1681-2040, 1686-2035, 1687-2035, 1699-2046, 1711-2047, 1714-2037, 1716-2034, 1720-2031, 1720-2033, 1720-2039, 1720-2063, 1723-2035, 1724-2035, 1724-2039, 1729-2039, 1731-2013, 1731-2033, 1731-2039, 1732-2013, 1732-2031, 1732-2033, 1732-2039, 1750-2035, 1760-2040, 1781-2043, 1857-2020, 1871-2063, 1948-2035, 1980-2043
------------------------	---

10

20

30

40

【 0 3 7 2 】

【 表 5 - 1 】

表5-1

ポリクレオチド SEQ ID NO:	Incyte プロジェク ト ID	代表的ライブラリ
31	1895273CB1	THPINOT03
32	70072222CB1	BRAIFEN05
33	3559223CB1	BRAHTDK01
34	3441255CB1	PLACNOB01
35	1958917CB1	BRSTNOT23
36	6219465CB1	BRAUNOR01
37	3576625CB1	NERDTDN03
38	4765758CB1	PLACNOT05
39	7236661CB1	BRAUNOR01
40	7714187CB1	BRAIHCT01
41	5136540CB1	UTREDIT07
42	3277403CB1	FIBRUNT02
43	1517569CB1	LUNGNOT18
44	2415991CB1	BRAINOT11
45	2735742CB1	LUNGFET03
46	2768535CB1	ADRENOT07
47	6848851CB1	BRAIFER05
48	7040722CB1	UTRSTMR02
49	6430290CB1	LUNGNON07
50	2640251CB1	BRAITUT03
51	3839350CB1	DENDTNT01
52	6393813CB1	KIDCTMT01
53	5685755CB1	BRAIUNT01
54	71728459CB1	BRAHNOT01

10

20

30

40

【 0 3 7 3 】

【 表 5 - 2 】

表5-2

ポリヌクレオチド SEQ ID NO:	Incyte プロジェクト ト ID	代表的ライブラリ
55	1904303CB1	OVARUT10
56	2911343CB1	BSCNNOT03
57	7500308CB1	BRAYDIN03
58	7501098CB1	UTREDIT07
59	7503839CB1	BRAIFER05
60	7503698CB1	LUNGNOT02

10

20

30

40

【 0 3 7 4 】

【 表 6 - 1 】

表6-1

ライブラリ	ベクター	ライブラリの説明
ADRENOT07	pINCY	ライブラリは61才の女性の両側副腎摘出時に採取した副腎組織から単離したRNAを用いて作製した。患者の病歴には、詳細不明の副腎障害がある。
BRAHNOT01	pINCY	ライブラリは、心臓麻痺で死亡した35歳の白人男性から採取した後部海馬組織から単離したRNAを使用して作製された。病理検査では、軽度の軟膜線維症と大脳新皮質の複数微小梗塞が見られた。顕微的に、大脳半球は、焦点石灰沈着を伴う、軽度の軟膜線維症と判明した。大脳半球全体に、縮んだわずかに好酸性の錐体ニューロンの形跡があった。さらに、大脳皮質全体に散在して、周辺神経膠性を伴う複数の小さな空洞発生があった。当患者の病歴には、拡張型心筋症、鬱血性心不全、心肥大、及び肥大した脾臓と肝臓が含まれる。
BRAHTDK01	PSPORT1	この増幅され標準化されたライブラリは、胆管癌で死亡した55才白人女性から切除した旧皮質、前海馬組織、および後海馬組織から単離したプールされたRNAを用いて作製された。病理は、円蓋部全体に優勢な軽度の軟膜線維形成があり、帯状皮質白質および視床に軸索球状態が散在し、内嗅皮質および中脳水道灰白質領域に、少数の神経原線維変化が散在した。随伴する腫瘍組織の病理は、残存腫瘍または再発性腫瘍を持つ高分化型の肝臓胆管癌を示した。患者の病歴には、胆管癌、術後パッド・キアリー症候群、胆汁性腹水、水胸症、脱水症、栄養不良、乏尿、および急性腎不全がある。過去の手術には、胆嚢切除術(48時間/一回)の再アノールングによるハイブリダイゼーションを除いたことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を適応した条件を用いて1回で標準化した。
BRAIFEN05	pINCY	この標準化胎児脳組織ライブラリは胎児脳組織ライブラリの326万の独立したクローンから作製した。RNA源は妊娠23週目で左側発育不全心臓のため死産の白人男子胎児から採取した脳組織から作製した。このライブラリは、極めて長時間(48時間/1回)の再アノールングによるハイブリダイゼーションを除いたことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を適応した条件を用いて2回にわたり標準化した。
BRAIFER05	pINCY	ライブラリは妊娠23週目で死産の左心低形成症胎児(白人男子)から採取した脳組織から単離したRNAを用いて作製した。

【 0 3 7 5 】
【 冊 6 - 2 】

表6-2

BRAIHCT01	pINCY	ライブラリは脳血管事故死亡の57才の白人男性の脳から除去した罹病後頭葉から単離したRNAを用いて作製した。患者の病歴にはハンチントン病および肺炎腫がある。
BRAINOT11	pINCY	ライブラリは、5才の白人男子の大腦半球切除術時に右側頭葉から抽出された脳組織から単離されたRNAを用いて作製された。病理は、慢性癲癇発作に一致する広範囲の多小脳回症 (polymicrogyria) と軽度から中等度の神経膠症(主に、軟膜下、皮質下)を示した。家族歴には子宮頸癌が含まれる。
BRAITUT03	PSPORT1	ライブラリは、17才の白人女性の脳髄膜病変の切除時に、左前頭葉から採取した脳腫瘍組織から単離したRNAを用いて作製した。病理はグレード4の繊維状の巨大および小細胞の星状細胞腫を示した。家族歴には良性高血圧および心臓血管系疾患および脳血管系疾患がある。
BRAIUNT01	pINCY	ライブラリは、眼窩上領域への転移を伴う神経上皮腫を患う14才の白人女子に由来する神経上皮腫細胞株(ATCC HTB-10)であるSK-N-MCから単離したRNAを用いて作製した。
BRAUNOR01	pINCY	このランダムプライムされたライブラリは、アテローム性動脈硬化症により胸部大動脈破裂と大動脈出血で死亡の81才白人女性から得た線条体、淡蒼球および被殻組織から単離したRNAを用いて作製した。病理学検査では、双方の内部頸動脈に中程度のアテローム性動脈硬化、前頭皮質および海馬に微視的な梗塞、および散在広汎性のアミロイド斑、および神経原線維変化が見られたが、これは年齢と矛盾しない。肉眼的には、軟膜は上矢状静脈洞に沿った軽度の肥厚と硝子質化とを示すのみであった。残りの軟膜は薄かった。また、充血した血管がいくらか見られた。主に前頭極および前頭葉、及び側頭葉(両側)に軽度の萎縮が見られた。微視的には、アルツハイマータイプII星状細胞のペアが新皮質の深層にあった。中央前頭皮質の灰白質の奥深くのニューロンの周りでサテライトシス (satellitosis) が増大していた。扁桃はまれな広汎性斑と神経原線維変性を含んでいた。後部海馬には、反応性グリオシスに囲まれたヘモジリン含有マクロファージを伴う、微視的な囊胞性の空洞化領域を含んでいた。患者の病歴は、敗血症、胆管炎、術後無気肺、肺炎冠動脈疾患(CAD)、左心室肥大に起因する心肥大、巨脾症、細動脈性腎硬化、結節性コロイド甲状腺腫、気腫、鬱血心不全(CHF)、甲状腺機能低下症および抹消血管疾患を含んでいた。

10
20
30
40

【 0 3 7 6 】
【 表 6 - 3 】

表6-3

BRAYDIN03	pINCY	標準化したライブラリは、脳神経ライブラリの670万個の独立性クローンから作製した。開始RNAは脳血管事故から死亡した57才の白人男性から採取された罹患視床下部組織から単離したRNAから作製した。患者の病歴にはハンチントン病および肺気腫がある。このライブラリは、極めて長時間(48時間/1回)の再アニーリングによるハイブリダイゼーションを用いたことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 から適応した条件を用いて2回にわたり標準化した。このライブラリはクローンを含む挿入断片用を選択するために直線化し、再度環状にした。
BRSTNOT23	pINCY	ライブラリは、35才の白人女性の両側乳房縮小時に採取した罹患乳房組織から単離したRNAを用いて作製した。病理は非増殖性繊維嚢胞性疾患(乳腺症)を示した。家族歴には、2型糖尿病、アテローム性冠動脈疾患、急性心筋梗塞、高脂血症、および冠動脈ハイパスがある。
BSCNNOT03	pINCY	ライブラリは92才の男性の脳から採取した尾状核組織から単離したRNAを用いて作製した。病理検査では、いくつかの小さい脳梗塞が見られたが、老人斑や神経原線維変性は見られなかった。患者の病歴には、放射線治療を受けた咽喉癌がある。
DENDTNT01	pINCY	ライブラリは、抹消血液の処理済樹状細胞から単離されたRNAを用いて作製した。
FIBRUNT02	pINCY	ライブラリは14才の白人男子から採取した骨肉腫に由来するMG-63細胞株から単離したRNAを用いて作製した。
KIDCTMT01	pINCY	ライブラリは65才の男性の腎尿管切除術時に採取した腎臓皮質組織から単離したRNAを用いて作製した。随伴する腫瘍組織の病理学検査では、グレード3の腎細胞癌が腎臓と腎臓周囲の中間部位内に見られた。
LUNGFET03	pINCY	ライブラリは、妊娠20週で死亡した白人女子胎児から採取した肺組織から単離したRNAを用いて作製された。
LUNGNON07	pINCY	この標準化した肺組織ライブラリは、或る肺組織ライブラリの510万個の独立クローンから作製した。開始RNAの作製は、肺組織から単離したRNAから行った。このライブラリは、極めて長時間(48時間/1回)の再アニーリングによるハイブリダイゼーションを用いた他は、Soares 他, PNAS (1994) 91:9228-9232 及び Bonaldo 他, Genome Research (1996) 6:791 を適用した条件を用いて2回にわたりノープライズした。
LUNGNOT02	PBLUESCRIPT	ライブラリは、くも膜下出血で死亡した47才白人男子の肺組織から単離されたRNAを用いて作製された。

10

20

30

40

表6-4

LUNGNOT18	pINCY	ライブラリは66才の白人女性から採取した左上葉肺組織から単離したRNAを用いて作製した。随伴する腫瘍組織の病理学検査では、グレード2の腺癌を示した。患者の病歴には脳血管疾患、アテローム硬化型冠動脈疾患および肺不全が含まれる。家族の病歴には、心筋梗塞およびアテローム硬化型冠動脈疾患が含まれる。
NERDNDN03	pINCY	この標準化された後根神経節組織のライブラリは、後根神経節組織ライブラリの105万の独立性クローンから作製した。開始RNAは、急性肺水腫、急性気管支肺炎、双方胸膜滲出、心膜滲出、及び悪性リンパ腫(ナチュラルキラー細胞タイプ)のため死亡した32才の白人男性の頸部脊椎から取り除かれた後根神経節組織から作製された。当患者には、原因不明の発熱、倦怠感、疲労、及び胃腸出血があった。患者の病歴には、推定サイトメガロウイルス感染、肝臓鬱血、脂肪肝、脾腫、出血性膀胱炎、甲状腺出血、呼吸不全、左肺の肺炎、咽頭のナチュラルキラー細胞リンパ腫、ヘルペス、及びタバコとアルコールの濫用等がある。以前に受けた手術には、結腸内視鏡、結腸生検、アデノイド切除術、鼻咽腔、内視鏡検査、生検等がある。当患者の薬物療法には、Diflucan (fluconazole)、Deltasone (prednisone)、hydrocodone、Lortab、アルプラゾラム(Alprazolam)、Reazodone、ProMace-Cytabom、Etoposide、シタラビン(Cisplatin)、シタラビン(Cytarabine)、及びデキサメタゾン(dexamethasone)等がある。当患者は放射線療法及び複数の輸血を受けた。このライブラリは、極めて長時間(48時間/一回)の再アニーリングによるハイブリダイゼーションを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を適応した条件を用いて2回にわたり標準化した。
OVARTUT10	pINCY	ライブラリは58才白人女性の腹式子宮全摘出、孤立卵巣の除去、鼠径ヘルニアの修復時に左卵巣から除去された卵巣腫瘍組織から単離されたRNAを使って作製された。病理は、左卵巣に部分的に嚢胞性および壊死性腫瘍塊を形成する、結腸起源の転移性グレード3の腺癌、左卵巣間膜に小瘤を形成する結腸起源の腺癌を示した。子宮筋層において1つの壁内平滑筋腫が同定された。頸部は軽度の慢性嚢胞性子宮頸管炎を示した。患者の病歴には、良性高血圧、卵巣の卵胞嚢胞、結腸癌、良性の結腸腫瘍、変形性関節症がある。家族の病歴には気腫、心筋梗塞、アテローム硬化型冠動脈疾患、良性高血圧、高脂血症がある。
PLACNOB01	PBLUESCRIPT	ライブラリは、胎盤から単離されたRNAを用いて作製された。
PLACNOT05	pINCY	ライブラリは胎児性死亡で妊娠18週目に死亡した白人男性胎児から採取した胎盤組織から単離したRNAを用いて作製した。

【 0 3 7 8 】
【 表 6 - 5 】

表6-5

THPINOT03	pINCY	ライブラリは、非処理の THP-1 細胞から単離した RNA を使用して作製された。THP-1 は、急性単球性白血病の 1 才の白人男児の末梢血に由来するヒト前単球株である(引用文献: Int. J. Cancer (1980) 26:171)。”
UTREDIT07	pINCY	ライブラリは子宮内膜生検時に女性から採取した病変した子宮内膜組織から単離した RNA を用いて作製した。病理学検査は、欠損したベータ3、タイプII 欠損を伴う子宮内膜示した。
UTRSTMR02	PCDNA2.1	このランダムブライムされたライブラリは、二人の異なる提供者からプールしたcDNA を使って作製された。cDNA は、32 才の白人女性(提供者A)から採取した子宮内膜組織から単離した mRNA、及び腔式子宮摘出術と両側性卵巣摘出時に 45 才の女性(提供者B)から子宮筋層から単離した mRNA を用いて作製された。提供者Aに関しては、病理学検査では子宮内膜は分泌相であった。頸部は、時計にあたる 1 時、6 時、および 7 時の位置で限局的に扁平円柱上皮境界を伴う重症の子宮頸部形成異常(CIN III)を示した。軽度のコイロサイト異形成(koilocytotic dysplasia)も子宮頸管内で確認された。提供者Bに関しては、一致する腫瘍組織の病理学検査では、複数(23)の漿膜、壁内、粘膜下の平滑筋腫が見られた。患者の病歴には、提供者Bでは、緊張性尿失禁、喘息発作重症状態のない外因性喘息、正常分娩が含まれる。提供者Bの家族の病歴には、脳血管疾患、うつ病、アテローム硬化型冠動脈疾患が見られた。

10

20

30

40

【 0 3 7 9 】

【 表 7 - 1 】

表7-1

【 0 3 8 0 】
【 表 7 - 2 】

プログラム	説明	参考文献	パラメータ閾値
ABI FACTURA	ベクター配列を除き、核酸配列内のまぎらわしい塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	1種のFast Data Finderであり、アミノ酸配列または核酸配列の比較、注釈付けに有用である。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列をアセンブリするプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool。アミノ酸配列および核酸配列の配列類似性検索に有用であり、BLASTには5つの機能がある:blastp、blastn、blastx、tblastnおよびtblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410, Altschul, S.F. 他 (1997) Nucleic Acids Res. 25:3389-3402.	EST確率値=1.0E-8 以下 完全長配列確率値=1.0E-10以下
FASTA	問合せ配列と同種の配列群との類似性を検索するPearson およびLipman アルゴリズム。FASTAは最低5つの機能からなる:fasta、tfasta、fastx、tfastxおよびssearch。	Pearson, W.R.及びD.J. Lipman (1988) Proc. Natl. Acad. Sci.USA 85:2444-2448, Pearson, W.R. (1990) Methods Enzymol.183:63-98, Smith, T.F.及びM.S. Waterman (1981) Adv. Appl. Math. 2:482-489	ESTfasta E 値=1.0E-6 アセンブリされた ESTfasta 同一性=95%以上 一致した長さ=200塩基以上 fasta E 値=1.0E-8以下 完全長配列: fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOMおよびPFAM データベースの配列に対応させて遺伝子ファミリー、配列相溶性および構造的フィンガープリント領域を検索するBLOCKS IMPROVED サーチャー。	Henikoff, S.及びJ.G. Henikoff (1991) Nucleic Acids Res.19:6565-6572, Henikoff, J.G.及びS. Henikoff (1996) Methods Enzymol.266:88-105, Attwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値=1.0E-3以下
HMMER	PFAMのようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol. 235:1501-1531, Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res.26:320-322, Durbin, R. 他(1998) Our World View, in a Nutshell, Cambridge Univ.Press, 1-350ページ。	PFAMヒット確率値=1.0E-3 以下 シングルペプチドヒット:スコア=0以上

10

20

30

40

表7-2

【 0 3 8 1 】
【 表 8 1 1 】

プログラム	説明	参照文献	パラメータ閾値
ProfileScan	Prositesで定義された配列パターンと一致するタンパク質配列内の構造的モチーフおよび配列モチーフを検索するプログラム。	Gribskov, M. 他 (1988) CABIOS 4:61-66, Gribskov, M. 他 (1989) Methods Enzymol.183:146-159, Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221.	ノーマライズされた質のスコアとその特 定のPrositesモチーフに対するGCC 指定 "HIGH" 値 通常、スコア=1.4~2.1.
Phred	高い感度と確率で自動配列決定機のトレースを調べるベ- スコーリングアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185, Ewing, B及びP. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的実装に基づいたプロ グラムである、SWATとCrossMatchを含むPhils 改訂アセンブ リプログラムで、配列相同性検索やDNA配列のアセンブリに 有用である。	Smith, T.F.およびM.S. Waterman (1981) Adv. Appl. Math. 2:482-489, Smith, T.F.およびM.S. Waterman (1981) J. Mol. Biol. 147:195-197, Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap でアセンブリしたものの表示、編集用グラフィカルツ- ル。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンして、分泌シグナルペプチドの存在を 調べる重み行列解析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6, Clavere, J.M.及びS. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	重み行列を用いてタンパク質配列上の膜貫通セグメントを描 写し、配向を決定するプログラム。	Persson, B.およびP.Argos (1994) J. Mol. Biol. 237:182-192, Persson, B.およびP. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れMarkov モデル(HMM) を使ってタンパク質配列上の膜貫 通セグメントを描写し、配向を決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc.Sixth Intl. Conf. on Intelligent Systems for Mol.Biol., Glasgow 他, 編集, The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182ページ。	
Motifs	Prositesで定義された配列と一致したパターンのアミノ酸配列 を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, 第9版, M51-59ページ, Genetics Computer Group, Madison, WI.	

10
20
30
40

表8-1

SEQ ID NO.	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST アレル	アレル 1	アレル 2	アミノ酸	白人アレル 1 頻度	アフリカ系アレル 1 頻度	アジア系アレル 1 頻度	ヒスパニック系アレル 1 頻度
57	7500308	6824818H1	SNP0005839 ₅	72	952	A	A	G	非コード	n/d	n/a	n/a	n/a
57	7500308	6824818H1	SNP0014250 ₆	233	791	A	A	G	非コード	n/a	n/a	n/a	n/a
57	7500308	6824818J1	SNP0005839 ₅	580	952	A	A	G	非コード	n/d	n/a	n/a	n/a
57	7500308	6824818J1	SNP0014250 ₆	419	791	A	A	G	非コード	n/a	n/a	n/a	n/a
59	7503839	1233581H1	SNP0010566 ₀	163	3212	A	A	G	非コード	n/a	n/a	n/a	n/a
59	7503839	1960883H1	SNP0005153 ₈	159	3011	C	C	T	非コード	n/a	n/a	n/a	n/a
59	7503839	2396122H2	SNP0005153 ₈	111	3011	C	C	T	非コード	n/a	n/a	n/a	n/a
59	7503839	2439231H1	SNP0005153 ₈	184	3011	C	C	T	非コード	n/a	n/a	n/a	n/a
59	7503839	3525859H1	SNP0005153 ₈	114	3011	C	C	T	非コード	n/a	n/a	n/a	n/a
59	7503839	5676172H1	SNP0002172 ₅	41	1071	C	C	A	非コード	n/d	n/d	n/d	n/d
59	7503839	6367830H1	SNP0000353 ₃	249	3361	C	C	T	非コード	n/a	n/a	n/a	n/a
59	7503839	7579662H1	SNP0002172 ₅	113	1071	C	C	A	非コード	n/d	n/d	n/d	n/d
59	7503839	7653343H1	SNP0002172 ₅	485	1071	C	C	A	非コード	n/d	n/d	n/d	n/d
59	7503839	7763335J1	SNP0005153 ₈₅	85	3011	C	C	T	非コード	n/a	n/a	n/a	n/a

【 0 3 8 2 】
【 表 8 - 2 】

表8-2

59	7503839	8450847J1	8	SNP0000353	257	3361	C	C	T	非コード	n/a	n/a	n/a	n/a
59	7503839	8454173J1	3	SNP0000353	257	3361	C	C	T	非コード	n/a	n/a	n/a	n/a
59	7503839	8454612J1	3	SNP0000353	255	3361	C	C	T	非コード	n/a	n/a	n/a	n/a
60	7503698	1336111H1	8	SNP0015123	61	1312	A	A	G	E297	n/a	n/a	n/a	n/a
60	7503698	1904303H1	8	SNP0015123	115	1312	A	A	G	E297	n/a	n/a	n/a	n/a
60	7503698	3417040H1	9	S.NP0015123	194	1563	C	C	T	P381	n/a	n/a	n/a	n/a
60	7503698	5943105H1	9	SNP0015123	45	1561	C	C	T	S380	n/a	n/a	n/a	n/a

【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
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 02/086069 A2

- (51) International Patent Classification: C12N
- (21) International Application Number: PCT/US02/12464
- (22) International Filing Date: 19 April 2002 (19.04.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
 - 60/285,207 20 April 2001 (20.04.2001) US
 - 60/287,114 27 April 2001 (27.04.2001) US
 - 60/288,640 3 May 2001 (03.05.2001) US
 - 60/290,516 11 May 2001 (11.05.2001) US
 - 60/292,184 18 May 2001 (18.05.2001) US
 - 60/343,553 21 December 2001 (21.12.2001) US
 - 60/357,002 13 February 2002 (13.02.2002) US
 - 60/358,279 20 February 2002 (20.02.2002) US
 - 60/366,041 19 March 2002 (19.03.2002) US

Circle, Fremont, CA 94555 (US). **LI, Joana, X.** [US/US]; 1264 Geneva Avenue, San Francisco, CA 94112 (US). **THANGAVELU, Kavitha** [IN/US]; 1950 Montecito Avenue #23, Mountain View, CA 94034 (US). **GETZEN, Kimberly, J.** [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). **DING, Li** [CN/US]; 3353 Alma Street, #146, Palo Alto, CA 94306 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **YAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US). **WALIA, Narinder, K.** [US/US]; 890 Davis Street A #205, San Leandro, CA 94577 (US). **MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95037 (US). **LAL, Preeti, G.** [US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **GRAUL, Richard, C.** [US/US]; 682-29th Avenue, San Francisco, CA 94121 (US). **REDDY, Roopa** [IN/US]; 1233 W. McKinley Drive #3, Sunnyvale, CA 94086 (US). **BECHA, Shanya, D.** [US/US]; 21062 Gary Drive # 117, Castro Valley, CA 94546 (US). **SAPPERSTEIN, Stephanie, K.** [US/US]; 3531 Highland Avenue, Redwood City, CA 94062 (US). **RICHARDSON, Thomas, W.** [US/US]; 616 Canyon Road #107, Redwood City, CA 94062 (US). **TRAN, Uyen, K.** [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Pollen Place Way, San Jose, CA 95121 (US). **TANG, Y., Tom** [US/US]; 4230 Ramwick Court, San Jose, CA 95118 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **YAN, Lu** [CA/US]; 3885 Corrina Way, Palo Alto, CA 94303 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US).

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

(72) Inventors; and

(75) Inventors/Applicants (for US only): **KLAMMER, Aaron, A.** [US/US]; 565 Ortega Avenue #1, Mountain View, CA 94040 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **DUGGAN, Brendan, M.** [AU/US]; 243 Buena Vista Avenue #306, Sunnyvale, CA 94086 (US). **WARREN, Bridget, A.** [US/US]; 2250 Homestead Court #2, Los Altos, CA 94024 (US). **EMERLING, Brooke, M.** [US/US]; 1735 Woodland Avenue #71, Palo Alto, CA 94303 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **ARVIZU, Chandra, S.** [US/US]; 1706 Morocco Drive, San Jose, CA 95125 (US). **HONCHELL, Cynthia, D.** [US/US]; 400 Laurel Street #203, San Carlos, CA 94070 (US). **NGUYEN, Danielle, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **KALLICK, Deborah, A.** [US/US]; 58 Linda Avenue, Alhambra, CA 94027 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **AU-YOUNG, Janice, K.** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **RAMKUMAR, Jayatxami** [IN/US]; 34359 Maybird

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

(81) Designated States (national): AI, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GI, GM, GR, GU, HK, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TH, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,

[Continued on next page]



WO 02/086069 A2

(54) Title: SECRETED PROTEINS

(57) Abstract: The invention provides human secreted proteins (SLCP) and polynucleotides which identify and encode SLCP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of SLCP.

WO 02/086069 A2 

GB, GR, IL, IT, LU, MC, NL, PT, SI, TR), OAPI patent (BI, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NT, SN, TD, TG). *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.*

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

WO 02/086069

PCT/US02/12464

SECRETED PROTEINS**TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of secreted proteins and to the use of these sequences in the diagnosis, treatment, and prevention of liver, cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

BACKGROUND OF THE INVENTION

Protein transport and secretion are essential for cellular function. Protein transport is mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is composed of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Proteins that are retained in the plasma membrane contain one or more transmembrane domains, each comprised of about 20 hydrophobic amino acid residues. Secreted proteins are generally synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include proteins with important roles in cell-to-cell signaling. Such proteins include transmembrane receptors and cell surface markers, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, enzymes, neuropeptides, and vasomediators. (Reviewed in Alberts, B. et al. (1994) *Molecular Biology of The Cell*, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

Cell surface markers include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shotgun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular

WO 02/086069

PCT/US02/12464

biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

5 Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently
10 characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, MPs may be heavily glycosylated and may contain an Arginine-Glycine-Aspartate (RGD) tripeptide motif which may play a role in adhesive interactions. MPs include extracellular proteins such as fibronectin, collagen, galectin, vitronectin and its proteolytic derivative somatomedin B; and cell adhesion receptors such as
15 cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) *Kidney Int.* 51:1413-1417; Sjaastad, M.D. and Nelson, W.J. (1997) *BioEssays* 19:47-55.)

Mucins are highly glycosylated glycoproteins that are the major structural component of the mucus gel. The physiological functions of mucins are cytoprotection, mechanical protection,
20 maintenance of viscosity in secretions, and cellular recognition. MUC6 is a human gastric mucin that is also found in gall bladder, pancreas, seminal vesicles, and female reproductive tract (Toribara, N.W. et al. (1997) *J. Biol. Chem.* 272:16398-16403). The MUC6 gene has been mapped to human chromosome 11 (Toribara, N.W. et al. (1993) *J. Biol. Chem.* 268:5879-5885). Hemomucin is a novel Drosophila surface mucin that may be involved in the induction of antibacterial effector molecules
25 (Theopold, U. et al. (1996) *J. Biol. Chem.* 271:12708-12715).

Tuftelins are one of four different enamel matrix proteins that have been identified so far. The other three known enamel matrix proteins are the amelogenins, enamelin and ameloblastin. Assembly of the enamel extracellular matrix from these component proteins is believed to be critical in producing a matrix competent to undergo mineral replacement. (Paine C.T. et al. (1998) *Connect Tissue Res.* 38:257-267). Tuftelin mRNA has been found to be expressed in human ameloblastoma tumor, a non-mineralized odontogenic tumor (Deutsch D. et al. (1998) *Connect Tissue Res.* 39:177-184)

Olfactomedin-related proteins are extracellular matrix, secreted glycoproteins with conserved C-terminal motifs. They are expressed in a wide variety of tissues and in broad range of species, from *Caenorhabditis elegans* to *Homo sapiens*. Olfactomedin-related proteins comprise a gene family with
35 at least 5 family members in humans. One of the five, TIGR/myocilin protein, is expressed in the eye and is associated with the pathogenesis of glaucoma (Kulkarni, N.H. et al., (2000) *Genet. Res.* 76:41-

WO 02/086069

PCT/US02/12464

50) Research by Yokoyama et al. (1996) found a 135-amino acid protein, termed AMY, having 96% sequence identity with rat neuronal olfactomedin-related ER localized protein in a neuroblastoma cell line cDNA library, suggesting an essential role for AMY in nerve tissue (Yokoyama, M. et al., (1996) DNA Res. 3:311-320). Neuron-specific olfactomedin-related glycoproteins isolated from rat brain
5 cDNA libraries show strong sequence similarity with olfactomedin. This similarity is suggestive of a matrix-related function of these glycoproteins in neurons and neurosecretory cells (Danielson, P.E. et al., (1994) J. Neurosci. Res. 38:468-478).

Mac-2 binding protein is a 90-kD serum protein (90K) and another secreted glycoprotein, isolated from both the human breast carcinoma cell line SK-BR-3, and human breast milk. It
10 specifically binds to a human macrophage-associated lectin, Mac-2. Structurally, the mature protein is 567 amino acids in length and is preceded by an 18-amino acid leader. There are 16 cysteines and seven potential N-linked glycosylation sites. The first 106 amino acids represent a domain very similar to an ancient protein superfamily defined by a macrophage scavenger receptor cysteine-rich domain (Koths, K. et al., (1993) J. Biol. Chem. 268:14245-14249). 90K is elevated in the serum of
15 subpopulations of AIDS patients and is expressed at varying levels in primary tumor samples and tumor cell lines. Ullrich et al. (1994) have demonstrated that 90K stimulates host defense systems and can induce interleukin-2 secretion. This immune stimulation is proposed to be a result of oncogenic transformation, viral infection or pathogenic invasion (Ullrich, A., et al. (1994) J. Biol. Chem. 269:18401-18407).

20 Semaphorins are a large group of axonal guidance molecules consisting of at least 30 different members and are found in vertebrates, invertebrates, and even certain viruses. All semaphorins contain the sema domain which is approximately 500 amino acids in length. Neuropilin, a semaphorin receptor has been shown to promote neurite outgrowth *in vitro*. The extracellular region of neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of
25 neuropilin have been suggested as having roles in protein-protein interactions and are suggested to be involved in the binding of semaphorins through the sema and the C-terminal domains (reviewed in Raper, J.A. (2000) Curr. Opin. Neurobiol. 10:88-94). Plexins are neuronal cell surface molecules that mediate cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Plexins have been shown to be expressed in the receptors and neurons of particular sensory systems (Ohta, K. et al. (1995) Cell 14:1189-1199). There is evidence that suggests that some plexins function to control motor and CNS axon guidance in the developing nervous system. Plexins, which themselves contain complete semaphorin domains, may be both the ancestors of classical semaphorins and binding partners for semaphorins (Winberg, M.L. et al (1998) Cell 95:903-916).

35 Human pregnancy-specific beta 1-glycoprotein (PSG) is a family of closely related glycoproteins of molecular weights of 72 KDa, 64KDa, 62KDa, and 54KDa. Together with the carcinoembryonic antigen, they comprise a subfamily within the immunoglobulin superfamily (Plouzek C.A. and Chou

WO 02/086069

PCT/US02/12464

J.Y.. Endocrinology 129:950-958) Different subpopulations of PSG have been found to be produced by the trophoblasts of the human placenta, and the amnionic, and chorionic membranes (Plouzek C.A. et al. (1993) Placenta 14:277-285).

5 Autocrine motility factor (AMF) is one of the motility cytokines regulating tumor cell migration, therefore identification of the signaling pathway coupled with it has critical importance. Autocrine motility factor receptor (AMFR) expression has been found to be associated with tumor progression in thymoma (Ohta Y. et al. (2000) Int J Oncol. 17:259-264). AMFR is a cell surface glycoprotein of molecular weight 78KDa.

10 Hormones are secreted molecules that travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical compositions and mechanisms of action, hormones can be grouped into two categories. One category includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes hydrophilic hormones that
15 function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines (epinephrine, norepinephrine) and histamine, and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotrophic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864.)

20 Pro-opiomelanocortin (POMC) is the precursor peptide of corticotropin (ACTH) a hormone synthesized by the anterior pituitary gland, which functions in the stimulation of the adrenal cortex. POMC is also the precursor peptide of the hormone, beta-lipotropin (beta-LPH). Each hormone includes smaller peptides with distinct biological activities: alpha-melanotropin (alpha-MSH) and corticotropin-like intermediate lobe peptide (CLIP) are formed from ACTH; gamma-lipotropin (gamma-LPH) and beta-endorphin are peptide components of beta-LPH, while beta-MSH is contained within gamma-LPH. Adrenal insufficiency due to ACTH deficiency, resulting from a genetic mutation in exons 2 and 3 of POMC has defined an endocrine disorder characterized by early-onset obesity,
30 adrenal insufficiency, and red hair pigmentation (Chretien, M. et al., (1979) *Canad. J. Biochem.* 57:1111-1121, Krude, H. et al., (1998) *Nature Genet.* 19:155-157, Online Mendelian Inheritance in Man, OMIM. Johns Hopkins University, Baltimore, MD. MIM Number: 176830: August 1, 2000. World Wide Web URL: www.ncbi.nlm.nih.gov/omim/).

35 Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with membrane proteins for activity. Complex interactions among these factors and their receptors trigger intracellular signal

WO 02/086069

PCT/US02/12464

transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth factors such as epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, and platelet-derived growth factor. The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs). Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin which function as hormones to regulate cellular functions other than proliferation.

Growth and differentiation factors play critical roles in neoplastic transformation of cells in vitro and in tumor progression in vivo. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both in vivo and in vitro. Moreover, some growth factors and growth factor receptors are related both structurally and functionally to oncoproteins. In addition, growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)

The Slit protein, first identified in *Drosophila*, is critical in central nervous system midline formation and potentially in nervous tissue histogenesis and axonal pathfinding. Itoh et al. have identified mammalian homologues of the slit gene (human Slit-1, Slit-2, Slit-3 and rat Slit-1). The encoded proteins are putative secreted proteins containing EFG-like motifs and leucine-rich repeats, both are conserved protein-protein interaction domains. Slit-1, -2, and -3 mRNAs are expressed in the brain, spinal cord, and thyroid, respectively (Itoh, A. et al., (1998) *Brain Res. Mol. Brain Res.* 62:175-186). The Slit family of proteins are indicated to be functional ligands of glypican-1 in nervous tissue and suggests that their interactions may be critical in certain stages during central nervous system histogenesis (Liang, Y. et al., (1999) *J. Biol. Chem.* 274:17885-17892).

Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting.

WO 02/086069

PCT/US02/12464

(Reviewed in Martin, C.R. et al. (1985) Endocrine Physiology, Oxford University Press, New York, NY, pp. 57-62.)

- NP/VMs are involved in numerous neurological and cardiovascular disorders. For example, neuropeptide Y is involved in hypertension, congestive heart failure, affective disorders, and appetite regulation. Somatostatin inhibits secretion of growth hormone and prolactin in the anterior pituitary, as well as inhibiting secretion in intestine, pancreatic acinar cells, and pancreatic beta-cells. A reduction in somatostatin levels has been reported in Alzheimer's disease and Parkinson's disease. Vasopressin acts in the kidney to increase water and sodium absorption, and in higher concentrations stimulates contraction of vascular smooth muscle, platelet activation, and glycogen breakdown in the liver.
- 10 Vasopressin and its analogues are used clinically to treat diabetes insipidus. Endothelin and angiotensin are involved in hypertension, and drugs, such as captopril, which reduce plasma levels of angiotensin, are used to reduce blood pressure (Watson, S. and S. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 194; 252; 284; 55; 111).
- 15 Neuropeptides have also been shown to have roles in nociception (pain). Vasoactive intestinal peptide appears to play an important role in chronic neuropathic pain. Nociceptin, an endogenous ligand for the opioid receptor-like 1 receptor, is thought to have a predominantly anti-nociceptive effect, and has been shown to have analgesic properties in different animal models of tonic or chronic pain (Dickinson, T. and Fleetwood-Walker, S.M. (1998) Trends Pharmacol. Sci. 19:346-348).
- 20 Other proteins that contain signal peptides include secreted proteins with enzymatic activity. Such activity includes, for example, oxidoreductase/dehydrogenase activity, transferase activity, hydrolase activity, lyase activity, isomerase activity, or ligase activity. For example, matrix metalloproteinases are secreted hydrolytic enzymes that degrade the extracellular matrix and thus play an important role in tumor metastasis, tissue morphogenesis, and arthritis (Reponen, P. et al. (1995) Dev. Dyn. 202:388-396; Firestein, G.S. (1992) Curr. Opin. Rheumatol. 4:348-354; Ray, J.M. and Stetler-Stevenson, W.G. (1994) Eur. Respir. J. 7:2062-2072; and Mignatti, P. and Rifkin, D.B. (1993) Physiol. Rev. 73:161-195). Additional examples are the acetyl-CoA synthetases which activate acetate for use in lipid synthesis or energy generation (Luong, A. et al. (2000) J. Biol. Chem. 275:26458-264566). The result of acetyl-CoA synthetase activity is the formation of acetyl-CoA from acetate and CoA. Acetyl-CoA synthetases share a region of sequence similarity identified as the AMP-binding domain signature. Acetyl-CoA synthetase has been shown to be associated with hypertension (H. Toh, (1991) Protein Seq. Data Anal. 4:111-117, and Iwai, N. et al., (1994) Hypertension 23:375-380).
- 35 A number of isomerases catalyze steps in protein folding, phototransduction, and various anabolic and catabolic pathways. One class of isomerases is known as peptidyl-prolyl *cis-trans* isomerases (PPIases). PPIases catalyze the *cis* to *trans* isomerization of certain proline imidic bonds in proteins. Two families of PPIases are the FK506 binding proteins (FKBPs), and cyclophilins (CyPs).

WO 02/086069

PCT/US02/12464

FKBPs bind the potent immunosuppressants FK506 and rapamycin, thereby inhibiting signaling pathways in T-cells. Specifically, the PPIase activity of FKBPs is inhibited by binding of FK506 or rapamycin. There are five members of the FKBP family which are named according to their calculated molecular masses (FKBP12, FKBP13, FKBP25, FKBP52, and FKBP65), and localized to different regions of the cell where they associate with different protein complexes (Coss, M. et al. (1995) J. Biol. Chem. 270:29336 - 29341; Schreiber, S.L. (1991) Science 251:283 - 287).

The peptidyl-prolyl isomerase activity of CyP may be part of the signaling pathway that leads to T-cell activation. CyP isomerase activity is associated with protein folding and protein trafficking, and may also be involved in assembly/disassembly of protein complexes and regulation of protein activity. For example, in *Drosophila*, the CyP NinaA is required for correct localization of rhodopsins, while a mammalian CyP (Cyp40) is part of the Hsp90/Hsc70 complex that binds steroid receptors. The mammalian CypA has been shown to bind the gag protein from human immunodeficiency virus 1 (HIV-1), an interaction that can be inhibited by cyclosporin. Since cyclosporin has potent anti-HIV-1 activity, CypA may play an essential function in HIV-1 replication. Finally, Cyp40 has been shown to bind and inactivate the transcription factor c-Myb, an effect that is reversed by cyclosporin. This effect implicates CyPs in the regulation of transcription, transformation, and differentiation (Bergsma, D.J. et al (1991) J. Biol. Chem. 266:23204 - 23214; Hunter, T. (1998) Cell 92: 141-143; and Levenson, J.D. and Ness, S.A. (1998) Mol. Cell. 1:203-211).

Gamma-carboxyglutamic acid (Gla) proteins rich in proline (PRGPs) are members of a family of vitamin K-dependent single-pass integral membrane proteins. These proteins are characterized by an extracellular amino terminal domain of approximately 45 amino acids rich in Gla. The intracellular carboxyl terminal region contains one or two copies of the sequence PPXY, a motif present in a variety of proteins involved in such diverse cellular functions as signal transduction, cell cycle progression, and protein turnover (Kulman, J.D. et al., (2001) Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375). The process of post-translational modification of glutamic residues to form Gla is Vitamin K-dependent carboxylation. Proteins which contain Gla include plasma proteins involved in blood coagulation. These proteins are prothrombin, proteins C, S, and Z, and coagulation factors VII, IX, and X. Osteocalcin (bone-Gla protein, BGP) and matrix Gla-protein (MGP) also contain Gla (Friedman, P.A., and C.T. Przysiecki (1987) Int. J. Biochem. 19:1-7; C. Vermeer (1990) Biochem. J. 266:625-636).

The *Drosophila* sp. gene *crossveinless 2* is characterized as having a putative signal or transmembrane sequence, and a partial Von Willebrand Factor D domain similar to those domains known to regulate the formation of intramolecular and intermolecular bonds and five cysteine-rich domains, known to bind BMP-like (bone morphogenetic proteins) ligands. These features suggest that *crossveinless 2* may act extracellularly or in the secretory pathway to directly potentiate ligand signaling and hence, involvement in the BMP-like signaling pathway known to play a role in vein specification (Conley, C.A. et al., (2000) Development 127:3947-3959). The dorsal-ventral patterning in both

WO 02/086069

PCT/US02/12464

vertebrate and *Drosophila* embryos requires a conserved system of extracellular proteins to generate a positional information gradient.

The presence of signal peptides has also been seen in proteins associated with prostate cancer. The development of androgen-independent cells in prostate tumors is useful as a method of determining disease progression. As with most tumors, prostate cancer develops through a multistage progression ultimately resulting in an aggressive tumor phenotype. The initial step in tumor progression involves the hyperproliferation of normal luminal and/or basal epithelial cells. Androgen responsive cells become hyperplastic and evolve into early-stage tumors. Although early-stage tumors are often androgen-sensitive and respond to androgen ablation, a population of androgen-independent cells evolve from the hyperplastic population. These cells represent a more advanced form of prostate tumor that may become invasive and potentially become metastatic to the bone, brain, or lung.

Expression profiling

Array technology can provide a simple way to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, condition, disease, or disorder.

20 Foam cell formation and Secreted Proteins

Atherosclerosis and the associated coronary artery disease and cerebral stroke represent the most common cause of death in industrialized nations. Although certain key risk factors have been identified, a full molecular characterization that elucidates the causes and provides care for this complex disease has not been achieved. Molecular characterization of growth and regression of atherosclerotic vascular lesions requires identification of the genes that contribute to features of the lesion including growth, stability, dissolution, rupture and, most lethally, induction of occlusive vessel thrombus.

An early step in the development of atherosclerosis is formation of the "fatty streak". Lipoproteins, such as the cholesterol-rich low-density lipoprotein (LDL), accumulate in the extracellular space of the vascular intima, and undergo modification. Oxidation of LDL occurs most avidly in the sub-endothelial space where circulating antioxidant defenses are less effective. The degree of LDL oxidation affects its interaction with target cells. "Minimally oxidized" LDL (MM-LDL) is able to bind to LDL receptor but not to the oxidized LDL (Ox-LDL) or "scavenger" receptors that have been identified, including scavenger receptor types A and B, CD36, CD68/macrosialin and LOX-1 (Navab *et al.* (1994) *Arterioscler Thromb Vasc Biol* 16:831-842; Kodama *et al.* (1990) *Nature* 343:531-535; Acton *et al.* (1994) *J Biol Chem* 269:21003-21009; Endemann *et al.* (1993) *J Biol Chem* 268:11811-11816; Ramprasad *et al.* (1996) *Proc Natl Acad Sci* 92:14833-14838; Kataoka *et al.* (1999) *Circulation*

WO 02/086069

PCT/US02/12464

99:3110-3117). MM-LDL can increase the adherence and penetration of monocytes, stimulate the release of monocyte chemoattractant protein 1 (MCP-1) by endothelial cells, and induce scavenger receptor A (SRA) and CD36 expression in macrophages (Cushing *et al.* (1990) *Proc Natl Acad Sci* 87:5134-5138; Yoshida *et al.* (1998) *Arterioscler Thromb Vasc Biol* 18:794-802; Steinberg (1997) *J Biol Chem* 272:20963-20966). SRA and the other scavenger receptors can bind Ox-LDL and enhance uptake of lipoprotein particles.

Mononuclear phagocytes enter the intima, differentiate into macrophages, and ingest modified lipids including Ox-LDL. In most cell types, cholesterol content is tightly controlled by feedback regulation of LDL receptors and biosynthetic enzymes (Brown and Goldstein (1986) *Science* 232:34-47). In macrophages, however, the additional scavenger receptors lead to unregulated uptake of cholesterol (Brown and Goldstein (1983) *Annu Rev Biochem* 52:223-261) and accumulation of multiple intracellular lipid droplets producing a "foam cell" phenotype. Cholesterol-engorged and dead macrophages contribute most of the mass of early "fatty streak" plaques and typical "advanced" lesions of diseased arteries. Numerous studies have described a variety of foam cell responses that contribute to growth and rupture of atherosclerotic vessel wall plaques. These responses include production of multiple growth factors and cytokines, which promote proliferation and adherence of neighboring cells; chemokines, which further attract circulating monocytes into the growing plaque; proteins, which cause remodeling of the extracellular matrix; and tissue factor, which can trigger thrombosis (Ross (1993) *Nature* 362:801-809; Quin *et al.* (1987) *Proc Natl Acad Sci* 84:2995-2998). Thus, cholesterol-loaded macrophages which occur in abundance in most stages of the atherosclerotic plaque formation contribute to inception of the atherosclerotic process and to eventual plaque rupture and occlusive thrombus.

During Ox-LDL uptake, macrophages produce cytokines and growth factors that elicit further cellular events that modulate atherogenesis such as smooth muscle cell proliferation and production of extracellular matrix. Additionally, these macrophages may activate genes involved in inflammation including inducible nitric oxide synthase. Thus, genes differentially expressed during foam cell formation may reasonably be expected to be markers of the atherosclerotic process.

Steroid molecules and Secreted Proteins

The potential application of gene expression profiling is particularly relevant to measuring the toxic response to potential therapeutic compounds and of the metabolic response to therapeutic agents. Diseases treated with steroids and disorders caused by the metabolic response to treatment with steroids include adenomatosis, cholestasis, cirrhosis, hemangioma, Henoch-Schonlein purpura, hepatitis, hepatocellular and metastatic carcinomas, idiopathic thrombocytopenic purpura, porphyria, sarcoidosis, and Wilson disease. Response may be measured by comparing both the levels and sequences expressed in tissues from subjects exposed to or treated with steroid compounds such as mifepristone, progesterone, beclomethasone, medroxyprogesterone, budesonide, prednisone, dexamethasone, betamethasone, or danazol with the levels and sequences expressed in normal untreated tissue.

WO 02/086069

PCT/US02/12464

Steroids are a class of lipid-soluble molecules, including cholesterol, bile acids, vitamin D, and hormones, that share a common four-ring structure based on cyclopentanoperhydrophenanthrene and that carry out a wide variety of functions. Cholesterol, for example, is a component of cell membranes that controls membrane fluidity. It is also a precursor for bile acids which solubilize lipids and facilitate absorption in the small intestine during digestion. Vitamin D regulates the absorption of calcium in the small intestine and controls the concentration of calcium in plasma. Steroid hormones, produced by the adrenal cortex, ovaries, and testes, include glucocorticoids, mineralocorticoids, androgens, and estrogens. They control various biological processes by binding to intracellular receptors that regulate transcription of specific genes in the nucleus. Glucocorticoids, for example, increase blood glucose concentrations by regulation of gluconeogenesis in the liver, increase blood concentrations of fatty acids by promoting lipolysis in adipose tissues, modulate sensitivity to catecholamines in the central nervous system, and reduce inflammation. The principal mineralocorticoid, aldosterone, is produced by the adrenal cortex and acts on cells of the distal tubules of the kidney to enhance sodium ion reabsorption. Androgens, produced by the interstitial cells of Leydig in the testis, include the male sex hormone testosterone, which triggers changes at puberty, the production of sperm and maintenance of secondary sexual characteristics. Female sex hormones, estrogen and progesterone, are produced by the ovaries and also by the placenta and adrenal cortex of the fetus during pregnancy. Estrogen regulates female reproductive processes and secondary sexual characteristics. Progesterone regulates changes in the endometrium during the menstrual cycle and pregnancy.

Steroid hormones are widely used for fertility control and in anti-inflammatory treatments for physical injuries and diseases such as arthritis, asthma, and auto-immune disorders. Progesterone, a naturally occurring progestin, is primarily used to treat amenorrhea, abnormal uterine bleeding, or as a contraceptive. Endogenous progesterone is responsible for inducing secretory activity in the endometrium of the estrogen-primed uterus in preparation for the implantation of a fertilized egg and for the maintenance of pregnancy. It is secreted from the corpus luteum in response to luteinizing hormone (LH). The primary contraceptive effect of exogenous progestins involves the suppression of the midcycle surge of LH. At the cellular level, progestins diffuse freely into target cells and bind to the progesterone receptor. Target cells include the female reproductive tract, the mammary gland, the hypothalamus, and the pituitary. Once bound to the receptor, progestins slow the frequency of release of gonadotropin releasing hormone from the hypothalamus and blunt the pre-ovulatory LH surge, thereby preventing follicular maturation and ovulation. Progesterone has minimal estrogenic and androgenic activity. Progesterone is metabolized hepatically to pregnenediol and conjugated with glucuronic acid.

Medroxyprogesterone (MAH), also known as 6 α -methyl-17-hydroxyprogesterone, is a synthetic progestin with a pharmacological activity about 15 times greater than progesterone. MAH is used for the

WO 02/086069

PCT/US02/12464

treatment of renal and endometrial carcinomas, amenorrhea, abnormal uterine bleeding, and endometriosis associated with hormonal imbalance. MAH has a stimulatory effect on respiratory centers and has been used in cases of low blood oxygenation caused by sleep apnea, chronic obstructive pulmonary disease, or hypercapnia.

5 Mifepristone, also known as RU-486, is an antiprogesterone drug that blocks receptors of progesterone. It counteracts the effects of progesterone, which is needed to sustain pregnancy. Mifepristone induces spontaneous abortion when administered in early pregnancy followed by treatment with the prostaglandin, misoprostol. Further, studies show that mifepristone at a substantially lower dose can be highly effective as a postcoital contraceptive when administered within five days after
10 unprotected intercourse, thus providing women with a "morning-after pill" in case of contraceptive failure or sexual assault. Mifepristone also has potential uses in the treatment of breast and ovarian cancers in cases in which tumors are progesterone-dependent. It interferes with steroid-dependent growth of brain meningiomas, and may be useful in treatment of endometriosis where it blocks the estrogen-dependent growth of endometrial tissues. It may also be useful in treatment of uterine fibroid
15 tumors and Cushing's Syndrome. Mifepristone binds to glucocorticoid receptors and interferes with cortisol binding. Mifepristone also may act as an anti-glucocorticoid and be effective for treating conditions where cortisol levels are elevated such as AIDS, anorexia nervosa, ulcers, diabetes, Parkinson's disease, multiple sclerosis, and Alzheimer's disease.

Danazol is a synthetic steroid derived from ethinyl testosterone. Danazol indirectly reduces
20 estrogen production by lowering pituitary synthesis of follicle-stimulating hormone and LH. Danazol also binds to sex hormone receptors in target tissues, thereby exhibiting anabolic, antiestrogenic, and weakly androgenic activity. Danazol does not possess any progestogenic activity, and does not suppress normal pituitary release of corticotropin or release of cortisol by the adrenal glands. Danazol is used in the treatment of endometriosis to relieve pain and inhibit endometrial cell growth. It is also used to treat
25 fibrocystic breast disease and hereditary angioedema.

Corticosteroids are used to relieve inflammation and to suppress the immune response. They inhibit eosinophil, basophil, and airway epithelial cell function by regulation of cytokines that mediate the inflammatory response. They inhibit leukocyte infiltration at the site of inflammation, interfere in the function of mediators of the inflammatory response, and suppress the humoral immune response.

30 Corticosteroids are used to treat allergies, asthma, arthritis, and skin conditions. Beclomethasone is a synthetic glucocorticoid that is used to treat steroid-dependent asthma, to relieve symptoms associated with allergic or nonallergic (vasomotor) rhinitis, or to prevent recurrent nasal polyps following surgical removal. The anti-inflammatory and vasoconstrictive effects of intranasal beclomethasone are 5000 times greater than those produced by hydrocortisone. Budesonide is a corticosteroid used to control
35 symptoms associated with allergic rhinitis or asthma. Budesonide has high topical anti-inflammatory activity but low systemic activity. Dexamethasone is a synthetic glucocorticoid used in anti-

WO 02/086069

PCT/US02/12464

inflammatory or immunosuppressive compositions. It is also used in inhalants to prevent symptoms of asthma. Due to its greater ability to reach the central nervous system, dexamethasone is usually the treatment of choice to control cerebral edema. Dexamethasone is approximately 20-30 times more potent than hydrocortisone and 5-7 times more potent than prednisone. Prednisone is metabolized in the liver to its active form, prednisolone, a glucocorticoid with anti-inflammatory properties. Prednisone is approximately 4 times more potent than hydrocortisone and the duration of action of prednisone is intermediate between hydrocortisone and dexamethasone. Prednisone is used to treat allograft rejection, asthma, systemic lupus erythematosus, arthritis, ulcerative colitis, and other inflammatory conditions. Betamethasone is a synthetic glucocorticoid with antiinflammatory and immunosuppressive activity and is used to treat psoriasis and fungal infections, such as athlete's foot and ringworm.

The anti-inflammatory actions of corticosteroids are thought to involve phospholipase A₂ inhibitory proteins, collectively called lipocortins. Lipocortins, in turn, control the biosynthesis of potent mediators of inflammation such as prostaglandins and leukotrienes by inhibiting the release of the precursor molecule arachidonic acid. Proposed mechanisms of action include decreased IgE synthesis, increased number of β -adrenergic receptors on leukocytes, and decreased arachidonic acid metabolism. During an immediate allergic reaction, such as in chronic bronchial asthma, allergens bridge the IgE antibodies on the surface of mast cells, which triggers these cells to release chemotactic substances. Mast cell influx and activation, therefore, is partially responsible for the inflammation and hyperirritability of the oral mucosa in asthmatic patients. This inflammation can be retarded by administration of corticosteroids.

The effects upon liver metabolism and hormone clearance mechanisms are important to understand the pharmacodynamics of a drug. The human C3A cell line is a clonal derivative of HepG2/C3 (hepatoma cell line, isolated from a 15-year-old male with liver tumor), which was selected for strong contact inhibition of growth. The use of a clonal population enhances the reproducibility of the cells. C3A cells have many characteristics of primary human hepatocytes in culture: i) expression of insulin receptor and insulin-like growth factor II receptor; ii) secretion of a high ratio of serum albumin compared with α -fetoprotein; iii) conversion of ammonia to urea and glutamine; iv) metabolize aromatic amino acids; and v) proliferate in glucose-free and insulin-free medium. The C3A cell line is now well established as an *in vitro* model of the mature human liver (Mickelson et al. (1995) Hepatology 22:866-875; Nagendra et al. (1997) Am J Physiol 272:G408-G416).

The discovery of new secreted proteins, and the polynucleotides encoding them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of liver, cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

WO 02/086069

PCT/US02/12464

SUMMARY OF THE INVENTION

The invention features purified polypeptides, secreted proteins, referred to collectively as "SECP" and individually as "SECP-1," "SECP-2," "SECP-3," "SECP-4," "SECP-5," "SECP-6," "SECP-7," "SECP-8," "SECP-9," "SECP-10," "SECP-11," "SECP-12," "SECP-13," "SECP-14," "SECP-15," "SECP-16," "SECP-17," "SECP-18," "SECP-19," "SECP-20," "SECP-21," "SECP-22," "SECP-23," "SECP-24," "SECP-25," "SECP-26," "SECP-27," "SECP-28," "SECP-29," and "SECP-30." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-30.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-30. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:31-60.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90%

WO 02/086069

PCT/US02/12464

identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, e) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a

WO 02/086069

PCT/US02/12464

polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from

WO 02/086069

PCT/US02/12464

the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. The method comprises a) exposing a sample comprising the polypeptide to a compound, and
5 b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment the composition.

10 The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an amino acid
15 sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

20 The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, c) a biologically active fragment of a polypeptide having an
25 amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of
30 the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide
35 sequence selected from the group consisting of SEQ ID NO:31-60, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target

WO 02/086069

PCT/US02/12464

polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing
5 the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, iii) a polynucleotide having a
10 sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, ii) a polynucleotide comprising a
15 naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization
20 complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog, and the PROTEOME database identification numbers and annotations of PROTEOME database homologs, for polypeptides of the invention. The probability scores for the matches between
30 each polypeptide and its homolog(s) are also shown.

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

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

WO 02/086069

PCT/US02/12464

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

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

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

Table 8 shows single nucleotide polymorphisms found in polynucleotide sequences of the invention, along with allele frequencies in different human populations.

DESCRIPTION OF THE INVENTION

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

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

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

DEFINITIONS

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

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

An "allelic variant" is an alternative form of the gene encoding SECP. Allelic variants may

WO 02/086069

PCT/US02/12464

result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each
5 of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

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

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

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

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

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof,

WO 02/086069

PCT/US02/12464

such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind SBCP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

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

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH₂), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

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

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

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

WO 02/086069

PCT/US02/12464

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

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

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

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

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

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

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys

WO 02/086069

PCT/US02/12464

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

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

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

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

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

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

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of SECP or the polynucleotide encoding SECP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to

WO 02/086069

PCT/US02/12464

the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

10 A fragment of SEQ ID NO:31-60 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:31-60, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:31-60 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:31-60 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:31-60 and the region of SEQ ID NO:31-60 to which the fragment corresponds are routinely determinable by one of 15 ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-30 is encoded by a fragment of SEQ ID NO:31-60. A fragment of SEQ ID NO:1-30 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-30. For example, a fragment of SEQ ID NO:1-30 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-30. The precise length of a 20 fragment of SEQ ID NO:1-30 and the region of SEQ ID NO:1-30 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon 25 (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

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

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to 30 the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default 35 parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular

WO 02/086069

PCT/US02/12464

biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bi2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

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

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in

WO 02/086069

PCT/US02/12464

a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

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

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

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

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

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

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles

WO 02/086069

PCT/US02/12464

a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity.

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

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

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may
25 be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily
30 apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization
35 complex may be formed in solution (e.g., C_{4t} or R_{4t} analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g.,

WO 02/086069

PCT/US02/12464

paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

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

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

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

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

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

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

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

25 "Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

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

35 "Post-translational modification" of an SECP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by

WO 02/086069

PCT/US02/12464

cell type depending on the enzymatic milieu of SECP.

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

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

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

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public

WO 02/086069

PCT/US02/12464

from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

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

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

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

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

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

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

The terms "specific binding" and "specifically binding" refer to that interaction between a

WO 02/086069

PCT/US02/12464

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

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

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

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

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

"Transformation" describes a process by which exogenous DNA is introduced into a recipient
20 cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells"
25 includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid
30 introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. In one alternative, the nucleic acid can be introduced by infection with a recombinant viral vector, such as a lentiviral vector (Lois, C. et al. (2002) Science 295:868-872). The term genetic manipulation
35 does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the

WO 02/086069

PCT/US02/12464

present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al.

5 (1989), *supra*.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 10 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides 15 due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. 20 Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) 25 set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

30 THE INVENTION

The invention is based on the discovery of new human secreted proteins (SECP), the polynucleotides encoding SECP, and the use of these compositions for the diagnosis, treatment, or prevention of liver, cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders.

35 Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a

WO 02/086069

PCT/US02/12464

single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO.) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO.) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown. Column 6 shows the Incyte ID numbers of physical, full length clones corresponding to the polypeptide and polynucleotide sequences of the invention. The full length clones encode polypeptides which have at least 95% sequence identity to the polypeptide sequences shown in column 3.

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

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO.) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs including the locations of signal peptides (as indicated by "Signal Peptide" and/or "signal_cleavage"). Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are secreted proteins.

For example, SEQ ID NO:1 is 99% identical, from residue G18 to residue T289, to human IgG Fc receptor I (GenBank ID g180279) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.4e-146, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains immunoglobulin domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.)

WO 02/086069

PCT/US02/12464

Data from BLIMPS analyses provide further corroborative evidence that SEQ ID NO:1 is a secreted protein (note that "immunoglobulin domains" are characteristic of matrix proteins).

As another example, SEQ ID NO:11 is 150 residues in length and is 100% identical, from residue M1 to residue S105, to human transmembrane gamma-carboxyglutamic acid protein 4 (TGM4) (GenBank ID g12656635) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $8.4e-54$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:11 also contains a Vitamin K-dependent carboxylation/gamma-carboxyglutamic (GLA) domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:11 is a member of a family of vitamin K-dependent single-pass integral membrane proteins.

As another example, SEQ ID NO:16 is 519 amino acids in length and is 100% identical, from residue M274 to residue E519 to human AD021 protein (GenBank ID g7578787) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $8.3e-134$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance.

As another example, SEQ ID NO:22 is 81% identical, from residue M1 to residue Q138, to murine putative protein sequence (GenBank ID g12845943) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $9.1e-57$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:22 also contains signal cleavage sites as determined by searching for statistically significant matches in the SPScan weight matrix analysis program. (See Table 3.)

As another example, SEQ ID NO:23 is 28% identical, from residue S42 to residue P345, to human Mac-2 binding protein (GenBank ID g307153) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $9.3e-27$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:23 also contains a BTB/POZ domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.)

SEQ ID NO:2-10, SEQ ID NO:12-15, SEQ ID NO:17-21, and SEQ ID NO:24-30 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-30 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification

WO 02/086069

PCT/US02/12464

number (Polynucleotide SEQ ID NO.), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide sequences of the invention, and of fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:31-60 or that distinguish between SEQ ID NO:31-60 and related polynucleotide sequences.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL_XXXXXX_N₁_N₂_YYYY_N₃_N₄ represents a "stitched" sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYY is the number of the prediction generated by the algorithm, and N_{1,2,3,4}, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FLXXXXXX_gAAAAA_gBBBBB_i_N is a "stretched" sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

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

WO 02/086069

PCT/US02/12464

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

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

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

Table 8 shows single nucleotide polymorphisms (SNPs) found in polynucleotide sequences of the invention, along with allele frequencies in different human populations. Columns 1 and 2 show the polynucleotide sequence identification number (SEQ ID NO.) and the corresponding Incyte project identification number (PID) for polynucleotides of the invention. Column 3 shows the Incyte identification number for the EST in which the SNP was detected (EST ID), and column 4 shows the identification number for the SNP (SNP ID). Column 5 shows the position within the EST sequence at which the SNP is located (EST SNP), and column 6 shows the position of the SNP within the full-length polynucleotide sequence (CB1 SNP). Column 7 shows the allele found in the EST sequence. Columns 8 and 9 show the two alleles found at the SNP site. Column 10 shows the amino acid encoded by the codon including the SNP site, based upon the allele found in the EST. Columns 11-14 show the frequency of allele 1 in four different human populations. An entry of n/d (not detected) indicates that the frequency of allele 1 in the population was too low to be detected, while n/a (not available) indicates that the allele frequency was not determined for the population.

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

The invention also encompasses polynucleotides which encode SECP. In a particular

WO 02/086069

PCT/US02/12464

embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:31-60, which encodes SECP. The polynucleotide sequences of SEQ ID NO:31-60, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding SECP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding SECP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:31-60 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:31-60. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SECP.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide sequence encoding SECP. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding SECP, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to the polynucleotide sequence encoding SECP over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding SECP. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SECP.

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

Although nucleotide sequences which encode SECP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring SECP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding SECP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring

WO 02/086069

PCT/US02/12464

codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SECP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

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

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:31-60 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

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

WO 02/086069

PCT/US02/12464

from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:3186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

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

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

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

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter SECP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA

WO 02/086069

PCT/US02/12464

shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

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

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

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

35 In order to express a biologically active SECP, the nucleotide sequences encoding SECP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains

WO 02/086069

PCT/US02/12464

the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding SECP. Such elements may vary in their strength and specificity. Specific initiation signals

5 may also be used to achieve more efficient translation of sequences encoding SECP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding SECP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous

10 translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

15 Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding SECP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

20 A variety of expression vector/host systems may be utilized to contain and express sequences encoding SECP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus);

25 plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery

30 of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol.*

35

WO 02/086069

PCT/US02/12464

Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

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

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

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

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

WO 02/086069

PCT/US02/12464

virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

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

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

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apv* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be
20 used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorosulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements
25 for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A.
30 (1995) Methods Mol. Biol. 55:121-131.)

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

WO 02/086069

PCT/US02/12464

expression of the tandem gene as well.

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

Immunological methods for detecting and measuring the expression of SECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated
10 cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SECP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York
15 NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

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

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

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted
35 sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation,

WO 02/086069

PCT/US02/12464

lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and W138) are available from the American Type Culture Collection
5 (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding SECP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric SECP protein containing a
10 heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of SECP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA).
15 GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoprecipitation of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SECP encoding
20 sequence and the heterologous protein sequence, so that SECP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

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

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

In one embodiment, the compound thus identified is closely related to the natural ligand of SECP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a
35 natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which SECP

WO 02/086069

PCT/US02/12464

binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express SECP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells
5 expressing SECP or cell membrane fractions which contain SECP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either SECP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the
10 assay may comprise the steps of combining at least one test compound with SECP, either in solution or affixed to a solid support, and detecting the binding of SECP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

SECP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of SECP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for SECP activity, wherein SECP is combined with at least one test compound, and the activity of SECP in the presence of
15 a test compound is compared with the activity of SECP in the absence of the test compound. A change in the activity of SECP in the presence of the test compound is indicative of a compound that modulates the activity of SECP. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising SECP under conditions suitable for SECP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of SECP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds
20 may be screened.

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

WO 02/086069

PCT/US02/12464

cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

5 Polynucleotides encoding SECP may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) *Science* 282:1145-1147).

10 Polynucleotides encoding SECP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding SECP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential

15 pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress SECP, e.g., by secreting SECP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between
20 regions of SECP and secreted proteins. In addition, the expression of SECP is closely associated with diseased breast, brain, brain tumor, diseased brain, nervous, developmental, and diseased endometrial tissues, bone cancer, adrenal and brain tissues, and with neighboring tissues associated with tumors of the lung and ovary, dendritic cells, and kidney cortex tissue associated with a renal carcinoma, nasal, cribriform and ovarian tumor tissues. In addition, examples of tissues expressing SECP can be found in
25 Table 6. Therefore, SECP appears to play a role in liver, cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders. In the treatment of disorders associated with increased SECP expression or activity, it is desirable to decrease the expression or activity of SECP. In the treatment of disorders associated with decreased SECP expression or activity, it is desirable to

30 increase the expression or activity of SECP. Therefore, in one embodiment, SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP. Examples of such disorders include, but are not limited to, a liver disorder such as adenomatosis, cholestasis, cirrhosis, hemangioma, Henoch-Schonlein purpura, hepatitis, hepatocellular and metastatic carcinomas, idiopathic thrombocytopenic purpura, porphyria, sarcoidosis, and Wilson disease and for
35 diagnosis of liver disorders and detecting metabolic and toxicological responses to treatment with steroids; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,

WO 02/086069

PCT/US02/12464

cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxicity, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebotrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis,

WO 02/086069

PCT/US02/12464

encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, 5 inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a developmental disorder such as renal tubular acidosis, 10 anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and 15 cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss.

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

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

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

In a further embodiment, an antagonist of SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP. Examples of such disorders include, but are not limited to, those liver, cell proliferative, autoimmune/inflammatory, cardiovascular, 30 neurological, and developmental disorders described above. In one aspect, an antibody which specifically binds SECP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express SECP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding SECP may be administered to a subject to treat or prevent a disorder associated with increased 35 expression or activity of SECP including, but not limited to, those described above.

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

WO 02/086069

PCT/US02/12464

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

An antagonist of SECP may be produced using methods which are generally known in the art. In particular, purified SECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SECP. Antibodies to SECP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use. Single chain antibodies (e.g., from camels or llamas) may be potent enzyme inhibitors and may have advantages in the design of peptide mimetics, and in the development of immuno-adsorbents and biosensors (Muyldermans, S. (2001) *J. Biotechnol.* 74:277-302).

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

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

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

In addition, techniques developed for the production of "chimeric antibodies," such as the

WO 02/086069

PCT/US02/12464

splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SECP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

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

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

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

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

WO 02/086069

PCT/US02/12464

dissociation of SECP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

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

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

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

In another embodiment of the invention, polynucleotides encoding SECP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial

WO 02/086069

PCT/US02/12464

hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Sonnia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in SECP expression or regulation causes disease, the expression of SECP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

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

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

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

WO 02/086069

PCT/US02/12464

1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

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

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SECP to cells which have one or more genetic abnormalities with respect to the expression of SECP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby
30 incorporated by reference. For adenoviral vectors, see also Antinuzzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding SECP to target cells which have one or more genetic abnormalities with
35 respect to the expression of SECP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing SECP to cells of the central nervous system, for which HSV has a

WO 02/086069

PCT/US02/12464

tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395).

The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding SECP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for SECP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of SECP-coding RNAs and the synthesis of high levels of SECP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of SECP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases,

WO 02/086069

PCT/US02/12464

transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the
5 transcript from binding to ribosomes.

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

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

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

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

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding SECP. Compounds
35 which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which

WO 02/086069

PCT/US02/12464

are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased SECP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding SECP may be therapeutically useful, and in the treatment of disorders associated with decreased SECP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding SECP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding SECP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding SECP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding SECP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) *Nucleic Acids Res.* 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) *Biochem. Biophys. Res. Commun.* 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

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

WO 02/086069

PCT/US02/12464

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

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

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

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

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

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

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

WO 02/086069

PCT/US02/12464

A therapeutically effective dose refers to that amount of active ingredient, for example SECP or fragments thereof, antibodies of SECP, and agonists, antagonists or inhibitors of SECP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

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

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

25 DIAGNOSTICS

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

A variety of protocols for measuring SECP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of SECP expression. Normal or standard values for SECP expression are established by combining body fluids or cell extracts taken

WO 02/086069

PCT/US02/12464

from normal mammalian subjects, for example, human subjects, with antibodies to SECP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of SECP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation
5 between standard and subject values establishes the parameters for diagnosing disease.

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

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

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the SECP encoding sequences. The hybridization probes of the subject
20 invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:31-60 or from genomic sequences including promoters, enhancers, and introns of the SECP gene.

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

Polynucleotide sequences encoding SECP may be used for the diagnosis of disorders associated
30 with expression of SECP. Examples of such disorders include, but are not limited to, a liver disorder such as adenomatosis, cholestasis, cirrhosis, hemangioma, Henoch-Schonlein purpura, hepatitis, hepatocellular and metastatic carcinomas, idiopathic thrombocytopenic purpura, porphyria, sarcoidosis, and Wilson disease and for diagnosis of liver disorders and detecting metabolic and toxicological responses to treatment with steroids; a cell proliferative disorder such as actinic keratosis,
35 arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary

WO 02/086069

PCT/US02/12464

thrombocytopenia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus;

5 an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic

10 lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner

15 syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective

20 endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins,

25 thrombophlebitis and phlebotrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy,

30 retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis,

35 cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy,

WO 02/086069

PCT/US02/12464

neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polyomyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, 5 seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, 10 genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucocutaneous dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss. The polynucleotide sequences encoding SECP may 15 be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding SECP may be useful in assays that 20 detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding SECP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample 25 then the presence of altered levels of nucleotide sequences encoding SECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of SECP, a 30 normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding SECP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. 35 Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the

WO 02/086069

PCT/US02/12464

presence of a disorder.

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

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

Additional diagnostic uses for oligonucleotides designed from the sequences encoding SECP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding SECP,
15 or a fragment of a polynucleotide complementary to the polynucleotide encoding SECP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences
20 encoding SECP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding SECP are used to amplify DNA using the polymerase chain reaction
25 (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally,
30 sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass
35 spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

WO 02/086069

PCT/US02/12464

SNPs may be used to study the genetic basis of human disease. For example, at least 16 common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase pathway. Analysis of the distribution of SNPs in different populations is useful for investigating genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations and their migrations. (Taylor, J.G. et al. (2001) Trends Mol. Med. 7:507-512; Kwok, P.-Y. and Z. Gu (1999) Mol. Med. Today 5:538-543; Nowotny, P. et al. (2001) Curr. Opin. Neurobiol. 11:637-641.)

Methods which may also be used to quantify the expression of SECP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, SECP, fragments of SECP, or antibodies specific for SECP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of

WO 02/086069

PCT/US02/12464

gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,340,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing
5 the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

10 Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of
15 pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to
20 that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The
25 normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at
30 <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated
biological sample are hybridized with one or more probes specific to the polynucleotides of the present
35 invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated

WO 02/086069

PCT/US02/12464

biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global
5 pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the
10 polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally
15 proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the
20 protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for SECP to quantify the levels of SECP expression. In one embodiment, the antibodies are used as elements on a microarray,
25 and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendozze, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should
30 be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the
35 proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

WO 02/086069

PCT/US02/12464

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in *DNA Microarrays: A Practical Approach*, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding SECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968.) Examples of genetic map

WO 02/086069

PCT/US02/12464

data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding SECP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

5 In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery
10 techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc.,
15 among normal, carrier, or affected individuals.

In another embodiment of the invention, SECP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes
20 between SECP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with SECP, or fragments thereof, and
25 washed. Bound SECP is then detected by methods well known in the art. Purified SECP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding SECP specifically compete with a test compound for binding SECP. In
30 this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with SECP.

In additional embodiments, the nucleotide sequences which encode SECP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such
35 properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the

WO 02/086069

PCT/US02/12464

preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/285,207, U.S. Ser. No. 60/287,114, U.S. Ser. No. 60/288,640, U.S. Ser. No. 60/290,516, U.S. Ser. No. 60/292,184, U.S. Ser. No. 60/343,553, U.S. Ser. No. 60/358,279, U.S. Ser. No. 60/366,041, and U.S. Ser. No. 60/357,002, are hereby expressly incorporated by reference.

EXAMPLES

10 I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPIT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPIT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), 35 PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), piGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or piNCY (Incyte

WO 02/086069

PCT/US02/12464

Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

5 Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid
10 purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well
15 plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows.

20 Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM
25 BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA
30 sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA
35 sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS,

WO 02/086069

PCT/US02/12464

DOMO, PRODOM; PROTEOME databases with sequences from *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Candida albicans* (Incyte Genomics, Palo Alto CA); hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM (Haft, D.H. et al. (2001) *Nucleic Acids Res.* 29:41-43); and HMM-based protein domain databases such as SMART (Schultz et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:5857-5864; Letunic, I. et al. (2002) *Nucleic Acids Res.* 30:242-244). (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Gensean-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM; and HMM-based protein domain databases such as SMART. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:31-60. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and

WO 02/086069

PCT/US02/12464

amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative secreted proteins were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr1 and gbhtg). Genscan is a general-
5 purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a
10 FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode secreted proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted proteins. Potential secreted proteins were also identified by homology to Incyte
15 cDNA sequences that had been annotated as secreted proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr1 public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide
20 sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

25 "Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory
30 and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals
35 were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together

WO 02/086069

PCT/US02/12464

by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpr public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Ineye cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:31-60 were compared with sequences from the Ineye LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:31-60 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances

WO 02/086069

PCT/US02/12464

are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel (1995) *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

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

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding SECP are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic

WO 02/086069

PCT/US02/12464

structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding SECP. cDNA sequences and cDNA library/tissue information are found in the LIFEBSEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of SECP Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the

WO 02/086069

PCT/US02/12464

concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

25 IX. Identification of Single Nucleotide Polymorphisms in SECP Encoding Polynucleotides

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:31-60 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants. An automated procedure of advanced chromosome analysis analysed the original chromatogram files in the vicinity of the putative SNP. Clone error filters used statistically generated algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from

WO 02/086069

PCT/US02/12464

clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

Certain SNPs were selected for further characterization by mass spectrometry using the high throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezuelan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed no allelic variance in this population were not further tested in the other three populations.

X. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:31-60 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

XI. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schemm (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure

WO 02/086069

PCT/US02/12464

analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

5 Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the
10 biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be
15 assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand
20 buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast
25 genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a
30 SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses
primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified
35 in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

WO 02/086069

PCT/US02/12464

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

15 Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

25 Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is

WO 02/086069

PCT/US02/12464

typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array
5 contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are
10 hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a
15 linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra)
20 between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to
25 obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

Human THP-1 cells (American Type Culture Collection, Manassas VA) were grown in RPMI1640 medium containing 10% fetal serum (v/v), 0.45% glucose (w/v), 10mM Hepes, 1mM sodium pyruvate, 1×10^{-5} M β -mercaptoethanol, penicillin (100 units/ml) and streptomycin (100 mg/ml).
25 For oxidized-LDL loading experiments, cells were seeded at a density of 1×10^6 cells/ml in medium containing 12-O-tetradecanoyl-phorbol-13-acetate (Research Biochemical International, Natick MA) at 1×10^{-7} M for 24 hr. The medium was then replaced by culture medium with or without 100 μ g/ml of CuSO_4 "fully" oxidized LDL (Intracel, Rockville MD) according to the method of Hammer *et al.* (1995; Arterio Thromb Vasc Biol 15:704-713). Medium was replaced every two days during the time of
30 culture. Cells were treated with Ox-LDL over time points ranging from 30 minutes to 4 days. During this period, cells remained adherent and had a typical speckled Nile red staining pattern. RNA was prepared for expression profiling at 0, 0.5, 2.5, and 8 hours, and 1, 2, and 4 days of Ox-LDL exposure.

Steroid treatment of human liver C3A cells

35 Early confluent C3A cells were treated with mifepristone, progesterone, beclomethasone, medroxyprogesterone, budesonide, prednisone, dexamethasone, betamethasone, or danazol at concentrations of 1 μ M, 10 μ M, and 100 μ M for 1, 3, and 6 hours. In all cases mRNA from untreated

WO 02/086069

PCT/US02/12464

early confluent C3A cells were prepared in parallel.

In this way it was shown that for SEQ ID NO:31, SECP was downregulated at least 2-fold in three of six lung tumor tissues tested. In addition it was down-regulated at least 3-fold in a THP-1 model for foam cell formation. For SEQ ID NO:34, it was shown that SECP expression was up-regulated at least 2.5-fold in a C3A liver cell line when treated with the following steroid compounds: beclomethasone (Becl), medroxyprogesterone (MAH), prednisone (Prdsne), dexamethasone (Dex), and with betamethasone (Betam).

XII. Complementary Polynucleotides

Sequences complementary to the SECP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring SECP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of SECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the SECP-encoding transcript.

XIII. Expression of SECP

Expression and purification of SECP is achieved using bacterial or virus-based expression systems. For expression of SECP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express SECP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of SECP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding SECP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, SECP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton

WO 02/086069

PCT/US02/12464

enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from SECP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified SECP obtained by these methods can be used directly in the assays shown in Examples XVII, XVIII, and XIX, where applicable.

10 XIV. Functional Assays

SECP function is assessed by expressing the sequences encoding SECP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector.

20 Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) *Flow Cytometry*, Oxford, New York NY.

The influence of SECP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding SECP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY).

WO 02/086069

PCT/US02/12464

mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding SECP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XV. Production of SECP Specific Antibodies

5 SECP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize animals (e.g., rabbits, mice, etc.) and to produce antibodies using standard protocols.

Alternatively, the SECP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-SECP activity by, for example, binding the peptide or SECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

20 **XVI. Purification of Naturally Occurring SECP Using Specific Antibodies**

Naturally occurring or recombinant SECP is substantially purified by immunoaffinity chromatography using antibodies specific for SECP. An immunoaffinity column is constructed by covalently coupling anti-SECP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing SECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SECP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SECP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and SECP is collected.

XVII. Identification of Molecules Which Interact with SECP

SECP, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled SECP, washed, and any wells with labeled SECP complex are assayed. Data obtained using different concentrations of SECP are used to calculate values for the number, affinity, and association of SECP with the candidate

WO 02/086069

PCT/US02/12464

molecules.

Alternatively, molecules interacting with SECP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) Nature 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

5 SECP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVIII. Demonstration of SECP activity

10 An assay for growth stimulating or inhibiting activity of SECP measures the amount of DNA synthesis in Swiss mouse 3T3 cells (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY). In this assay, varying amounts of SECP are added to quiescent 3T3 cultured cells in the presence of [³H]thymidine, a radioactive DNA precursor. SECP for this assay can be obtained by recombinant means or from biochemical preparations.

15 Incorporation of [³H]thymidine into acid-precipitable DNA is measured over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold SECP concentration range is indicative of growth modulating activity. One unit of activity per milliliter is defined as the concentration of SECP producing a 50% response level, where 100% represents maximal incorporation of [³H]thymidine into
20 acid-precipitable DNA.

Alternatively, an assay for SECP activity measures the stimulation or inhibition of neurotransmission in cultured cells. Cultured CHO fibroblasts are exposed to SECP. Following endocytic uptake of SECP, the cells are washed with fresh culture medium, and a whole cell voltage-clamped Xenopus myocyte is manipulated into contact with one of the fibroblasts in SECP-free
25 medium. Membrane currents are recorded from the myocyte. Increased or decreased current relative to control values are indicative of neuromodulatory effects of SECP (Morimoto, T. et al. (1995) Neuron 15:689-696).

Alternatively, an assay for SECP activity measures the amount of SECP in secretory, membrane-bound organelles. Transfected cells as described above are harvested and lysed. The lysate
30 is fractionated using methods known to those of skill in the art, for example, sucrose gradient ultracentrifugation. Such methods allow the isolation of subcellular components such as the Golgi apparatus, ER, small membrane-bound vesicles, and other secretory organelles. Immunoprecipitations from fractionated and total cell lysates are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The
35 concentration of SECP in secretory organelles relative to SECP in total cell lysate is proportional to the amount of SECP in transit through the secretory pathway.

WO 02/086069

PCT/US02/12464

Alternatively, AMP binding activity is measured by combining SECP with ³²P-labeled AMP. The reaction is incubated at 37°C and terminated by addition of trichloroacetic acid. The acid extract is neutralized and subjected to gel electrophoresis to remove unbound label. The radioactivity retained in the gel is proportional to SECP activity.

5 **XIX. Demonstration of Immunoglobulin Activity**

An assay for SECP activity measures the ability of SECP to recognize and precipitate antigens from serum. This activity can be measured by the quantitative precipitin reaction. (Golub, E.S. et al. (1987) *Immunology: A Synthesis*, Sinauer Associates, Sunderland, MA, pages 113-115.) SECP is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled SECP. SECP-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable SECP-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable SECP-antigen complex is plotted against the serum concentration. For various serum concentrations, a characteristic precipitin curve is obtained, in which the amount of precipitable SECP-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable SECP-antigen complex is a measure of SECP activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

Alternatively, an assay for SECP activity measures the expression of SECP on the cell surface. cDNA encoding SECP is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et al. (1997) *Blood* 90:2398-2405). Immunoprecipitations are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of SECP expressed on the cell surface.

Alternatively, an assay for SECP activity measures the amount of cell aggregation induced by overexpression of SECP. In this assay, cultured cells such as NIH3T3 are transfected with cDNA encoding SECP contained within a suitable mammalian expression vector under control of a strong promoter. Cotransfection with cDNA encoding a fluorescent marker protein, such as Green Fluorescent Protein (CLONTECH), is useful for identifying stable transfectants. The amount of cell agglutination, or clumping, associated with transfected cells is compared with that associated with untransfected cells. The amount of cell agglutination is a direct measure of SECP activity.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments.

WO 02/086069

PCT/US02/12464

Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

WO 02/086069

PCT/US02/12464

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Incyte Full Length Clones
1895273	1	1895273CD1	31	1895273CB1	1895273CA2
70072222	2	70072222CD1	32	70072222CB1	813820CA2
3559223	3	3559223CD1	33	3559223CB1	
3441255	4	3441255CD1	34	3441255CB1	
1958917	5	1958917CD1	35	1958917CB1	90067153CA2
6219465	6	6219465CD1	36	6219465CB1	
3576625	7	3576625CD1	37	3576625CB1	
4765758	8	4765758CD1	38	4765758CB1	4765758CA2
7236661	9	7236661CD1	39	7236661CB1	6772618CA2, 7236661CA2
7714187	10	7714187CD1	40	7714187CB1	
5136540	11	5136540CD1	41	5136540CB1	5136540CA2
3277403	12	3277403CD1	42	3277403CB1	9009298CA2, 9009296CA2
1517569	13	1517569CD1	43	1517569CB1	4031777CA2
2415991	14	2415991CD1	44	2415991CB1	7383971CA2

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polymerside SEQ ID NO:	Incyte Polymerside ID	Incyte Full Length Clones
2735742	15	2735742CD1	45	2735742CB1	1842745CA2, 1862749CA2, 2765810CA2, 3174585CA2, 6094213CA2, 6361652CA2, 7638315CA2, 7977239CA2, 8087951CA2, 8094152CA2, 8250338CA2, 8531020CA2, 8743705CA2
2768535	16	2768535CD1	46	2768535CB1	4239666CA2
6848551	17	6848551CD1	47	6848551CB1	
7040722	18	7040722CD1	48	7040722CB1	56022479CA2
6430290	19	6430290CD1	49	6430290CB1	9007608CA2, 9007637CA2, 9007638CA2, 9007639CA2
2640251	20	2640251CD1	50	2640251CB1	6452666CA2, 8516219CA2
3839550	21	3839550CD1	51	3839550CB1	3839550CA2
6393813	22	6393813CD1	52	6393813CB1	1615668CA2, 2193894CA2, 6657565CA2, 6657564CA2, 7057161CA2
5685755	23	5685755CD1	53	5685755CB1	

WO 02/086069

PCT/US02/12464

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polymaleotide SEQ ID NO:	Incyte Polymaleotide ID	Incyte Clone	Incyte Full Length
71728459	24	71728459CD1	54	71728459CB1	5406225CA2	
1904303	25	1904303CD1	55	1904303CB1	9009477CA2,	
2911343	26	2911343CD1	56	2911343CB1	90094885CA2	
7500308	27	7500308CD1	57	7500308CB1	2343465CA2	
7501098	28	7501098CD1	58	7501098CB1	4660380CA2	
7503839	29	7503839CD1	59	7503839CB1		
7503698	30	7503698CD1	60	7503698CB1		

WO 02/086069

PCT/US02/12464

Table 2

Polypeptide SFC ID NO.	Incyte Polypeptide ID	GenBank ID NO. or PROTEOME ID NO.	Probability Score	Annotation
1	1895273CD1	g380279	5.40E-146	[Homo sapiens] IgG Fc receptor I van de Winkel, J.G.J. et al., (1991) J. Biol. Chem. 266:13449-13455
6	6219465CD1	g6649221	2.40E-184	[Homo sapiens] selenoprotein N Lescure, A. et al., (1999) J. Biol. Chem. 274:38147-38154
7	3576623CD1	g4406679	1.10E-54	[Homo sapiens] Human neuronal (beta)-tubulin related ER localized protein Danielsen, P.E. et al., (1994) J. Neurosci. Res. 38:468-78
10	7714187CD1	g3041877	1.30E-222	[Homo sapiens] JB3489A Nishiyama, H. et al., (1999) Genes Chromosomes Cancer 26:171-5
11	5136540CD1	g12656635	8.40E-54	[Homo sapiens] (A1536351) transmembrane gamma-carboxyglutamic acid protein 4, TM64 Kulman, J.D. et al., (2001) Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375
12	3277403CD1	g9864185	8.10E-104	[Drosophila melanogaster] Crossveinless 2 Olson, D.J. et al., (2000) Development 127:3947-3959
16	2768535CD1	g7578787	8.30E-134	[Homo sapiens] AD021 protein [Homo sapiens]
23	5685755CD1	g407153	9.30E-27	[Homo sapiens] Mac-2 binding protein Kodis, K., Taylor, E., Halebeck, R., Caspi, C. and Wang, A. (1993) Cloning and characterization of a human Mac-2 binding protein, a new member of the superfamily defined by the macrophage scavenger receptor cysteine-rich domain J. Biol. Chem. 268:14245-14249
26	2911343CD1	g11225366	4.10E-13	[Homo sapiens] Ixstis specific leucine rich repeat protein

WO 02/086069

PCT/US02/12464

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO. or PROTEOME ID NO.	Probability Score	Annotation
28	7301098CD1	g12656635	2.6E-13	[Homo sapiens] transmembrane gamma-carboxyglutamic acid protein 4 TMG4 Kulman, J. D. et al. (2001) Identification of two novel transmembrane gamma-carboxyglutamic acid proteins expressed broadly in fetal and adult tissues. Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375
		68537 TMG4	2.2E-14	[Homo sapiens][Plasma membrane] Transmembrane gamma-carboxyglutamic acid protein 4, a putative transmembrane spanning protein which is a member of the Gla family, contains one cytoplasmic PFX Y motif which may mediate interaction with WW domain-containing proteins Kulman, J. D. et al. (2001) Identification of two novel transmembrane gamma-carboxyglutamic acid proteins expressed broadly in fetal and adult tissues. Proc. Natl. Acad. Sci. U. S. A. 98:1370-1375.

WO 02/086069

PCT/US02/12464

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Signal Peptide: M1-E33 Signal Peptide: M1-E35 Transmembrane domain: P8-E34, P179-N196, R773-S793, W801-L821N-Terminus is non-cytosolic	HMMER HMMER TMAP
			Potential Phosphorylation Sites: S30 S282 S444 S447 S465 S657 S701 S748 S806 S849 S873 S909 S976 S1007 S1053 S1060 T31 T54 T189 T249 T293 T416 T539 T672 T842 T844 T903 T1020 T1148 T1210	MOTIFS
			Potential Glycosylation Sites: N38 N676 N698 N719 N884	MOTIFS
5	1958917CD1	696	Signal_cleavage: M1-A27 Signal Peptide: M3-A27 Signal Peptide: M3-A34 Transmembrane domain: S8 W36N-Terminus is non-cytosolic do MUCIN; MUC5: TRACHEOBROCHIAL; DM05454555516H1-317; Q355-S581 Potential Phosphorylation Sites: S72 S99 S122 S184 S385 S473 S475 S499 S506 S511 S549 S612 S649 S662 T235 T248 T512 T542 T618	SFSCAN HMMER HMMER TMAP BLAST_DOMO MOTIFS
6	6219463CD1	436	Signal_cleavage: M1-A43 Transmembrane domain: R25-A53, Y231-F238 N-Terminus is non-cytosolic EF-hand calcium-binding domain: BL00018; D380-F392 Potential Phosphorylation Sites: S77 S88 S107 S319 S332 S339 S385 S404 T81 Y398	MOTIFS SFSCAN TMAP BLIMPS_PRINTS MOTIFS
7	3576023CD1	652	Potential Glycosylation Sites: N156 Signal_cleavage: M1-A27 Signal Peptide: M1-A27 Olfactomedin-like domain: G397-V652 Transmembrane domain: A4-S21N-Terminus is non-cytosolic PROTEIN PRECURSOR SIGNAL MYOCYCLIN TRABECULAR GLIUCOCORTICOID GLYCOPROTEIN OLFACTOMEDIN MESH-WORK INDUCED RESPONSE PD0006897; L440-T645	SFSCAN HMMER HMMER TMAP BLAST_PRODUM

WO 02/086069

PCT/US02/12464

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Potential Phosphorylation Sites: S21 S66 S82 S121 S155 S170 S242 S266 S314 S345 S382 S439 S519 S551 S574 T25 T151 T223 T273 T398 T563 T606 T624 Y75 Y236	MOTIFS
			Potential Glycosylation Sites: N159 N183	MOTIFS
8	4765758CD1	91	Signal_cleavage: M1-A42 Signal Peptide: M20-A42 Signal Peptide: M20-GH4 Citrate synthase signature: E6-D67 Ribosomal protein S24e signature: K3-A57 Potential Phosphorylation Sites: S77 S82 T22 T70 T73	SPSCAN HMMER HMMER PROFLESCAN PROFLESCAN MOTIFS
9	7256661CD1	155	Signal_cleavage: M1-G17 Signal Peptide: M1-G17 Signal Peptide: M1-A20 Potential Phosphorylation Sites: S58 S45 S91 S95 S135 T47 T90 T115	SPSCAN HMMER HMMER MOTIFS
10	7714187CD1	765	Signal_cleavage: M1-A33 Signal Peptide: M15-A33 Signal Peptide: M15-A34 Signal Peptide: M15-D37 Signal Peptide: M15-Q38 Signal Peptide: M15-A40 Signal Peptide: M15-A33 Transmembrane domain: O8-A33 Membrane attack complex components/ perform signature: L103-K155 IB3089A_PD147692: M1-C765	SPSCAN HMMER HMMER HMMER HMMER HMMER HMMER TMAP PROFLESCAN
			Potential Phosphorylation Sites: S45 S49 S72 S157 S218 S267 S315 S361 S380 S418 S523 S528 S559 S563 S579 S592 S735 T119 T309 T442 T504 T536 T599 T620 T761	BLAST_PRODOM MOTIFS
11	5136540CD1	150	Potential Glycosylation Sites: N168 N937 N656 N562 N609 N641 Signal_cleavage: M1-G17 Signal Peptide: M1-G17 Signal Peptide: M1-P19	MOTIFS SPSCAN HMMER HMMER

WO 02/086069

PCT/US02/12464

Table 3

SEQ ID NO	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
12	3277402CD1	685	Vitamin K-dependent carboxylation/gamma-carboxyglutamic (GLA) domain: L57-A98 Vitamin K-dependent carboxylation domain: V55-S110 Coagulation factor GLA domain signature PR00001: Y84-A98, D56-C69, N70-F83 GLA DOMAIN DM00454P070922-100: K29-W93 GLA DOMAIN DM00454P07422-80: G32-W93 GLA DOMAIN DM00454P251552-80: C21-W93 GLA DOMAIN DM00454P251552-80: G32-W93 Vitamin K-dependent carboxylation domain: D56-W93 Potential Phosphorylation Sites: S38 S97 S110 S120 T37 T102 Signal cleavage: M1-G33 Signal Peptide: R16-A39 Signal Peptide: L10-A39 Tyrasin inhibitor like cysteine-rich domain: C629-C682 von Willebrand factor type C domain: C108-C163, C166-C224, C301-C357, C238-C289, C50-C105 von Willebrand factor type D domain: C364-N314 C-terminal cysteine knot BL01185, C238-C286, K351-Y389 SIGNAL PRECURSOR GLYCOPROTEIN VITELLOGENIN CELL REPEAT WILLEBRAND VON ALPHA RECTORIN PD001080: C331-C663 do MUCIN; VON; WILLEBRAND; HEMOCYTIN; DM01378P0427512-692: A486-L669 do MUCIN; VON; WILLEBRAND; HEMOCYTIN; DM01378P04275160-333: G495-P666 do MUCIN; VON; WILLEBRAND; HEMOCYTIN; DM01378P04275186-1176: K492-H664 do MUCIN; VON; WILLEBRAND; HEMOCYTIN; DM01378P980225703-549: K490-C663 Vitamin K-dependent carboxylation domain: K271-F508 VWFC domain signature: C67-C105, C186-C224, C319-C357 Potential Phosphorylation Sites: S118 S123 S273 S479 S525 S573 T78 T260 T358 T437 T466 T549 Signal cleavage: M1-G33 Transmembrane Region: L58-S86	HMMER_PFAM PROFILINSCAN BLIMPS_PRINTS BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO MOTIFS MOTIFS SFSCAN HMMER HMMER_PFAM HMMER_PFAM BLIMPS_BLOCKS BLAST_PRODOM BLAST_DOMO BLAST_DOMO BLAST_DOMO MOTIFS MOTIFS MOTIFS SFSCAN TM/AP
13	1517560CD1	126	Signal cleavage: M1-G33 Transmembrane Region: L58-S86	SFSCAN TM/AP

WO 02/086069

PCT/US02/12464

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Potential Phosphorylation Sites: S66 T13 T27 T52 Y71	MOTIFS
			Potential Glycosylation Sites: N11	MOTIFS
14	2445991CD1	149	Signal Peptide: M1-G21	EMBLER
			Potential Phosphorylation Sites: S122	MOTIFS
			Potential Glycosylation Sites: N33	MOTIFS
15	2735742CD1	114	Signal_cleavage: M1-A18	SPSCAN
			S63 S72 S105	MOTIFS
16	2766533CD1	519	Signal_cleavage: M1-L60	SPSCAN
			Transmembrane Region: L38-G56 N-terminus is cytosolic	TMAP
			Potential Phosphorylation Sites: S28 S29 S80 S113 S199 S217 S239 S296 S327 T2 T34 T110 T124	MOTIFS
			Potential Glycosylation Sites: N98 N289 N922	MOTIFS
17	6848851CD1	1164	Signal_cleavage: M1-A45	SPSCAN
			Transmembrane Region: S901-A923 N-terminus is cytosolic	TMAP
			Potential Phosphorylation Sites: S18 S19 S65 S79 S107 S108 S110 S116 S128 S132 S216 S240 S241 S358 S379 S486 S506 S628 S665 S725 S956 S1027 T71 T78 T98 T112 T114 T162 T165 T209 T211 T225 T600 T604 T843 T1017 T1062 T1145 Y859	MOTIFS
			Potential Glycosylation Sites: N377 N413 N935 N1009	MOTIFS
18	7040722CD1	112	Signal_cleavage: M1-E19	SPSCAN
			Signal Peptide: M2-G23	EMBLER
			Potential Phosphorylation Sites: S102 T86	MOTIFS
			Potential Glycosylation Sites: N103	MOTIFS
19	6430290CD1	170	Signal_cleavage: M1-G54	SPSCAN
			Transmembrane Region: C23-S51 N-terminus is non-cytosolic	TMAP
			Potential Phosphorylation Sites: S9 S51 S143 S144 T66 T121	MOTIFS
20	2640251CD1	80	Signal_cleavage: M1-A42, M1B-A42, M1-A42	SPSCAN
			Signal Peptide: M1B-S44	EMBLER
			Signal Peptide: M1B-A42	EMBLER
			Potential Phosphorylation Sites: S20 T4 T53	MOTIFS

WO 02/086069

PCT/US02/12464

Table 3

SIFQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
21	383955CD1	118	Signal_cleavage: M1-A18 Signal Peptide: M1-A18 Potential Phosphorylation Sites: S20 S50 T6	SFSCAN HMIMER MGITFS
22	6393813CD1	140	Signal_cleavage: M1-A42, M1G-A42 Potential Phosphorylation Sites: S23 S33 S61 S72 S122 S124	SFSCAN MGITFS SFSCAN
23	568755CD1	478	Signal_cleavage: M1-R31 Signal Peptide: M1-A32 BTP/POZ domain: Q47-L162 Transmembrane domain: G10-T26N-terminus is non-cytosolic PROTEIN C ASSOCIATED MAMA CYCLOPHILIN PHPTIDYLPROLYL ISOMERASE PANCREAS CANCER ASSOCIATED MAC2 BINDING PD014408: L53-A324, S338 P362 SPERACT RECEPTOR AMINO-TERMINAL DM05438(A471G1126-585: S42-P445 Potential Phosphorylation Sites: S346 S561 S568 S460 S434 S452 T179 T355 T476 Potential Glycosylation Sites: N44 N61 N100 N195 N307	SFSCAN HMIMER HMIMER_Pfam TMAP BLAST_PRODOM BLAST_DOMO MGITFS MGITFS SFSCAN
24	71728459CD1	80	Signal_cleavage: M1-A15 Signal Peptide: M1-A15 Signal Peptide: M1-G28 Transmembrane domain: T4-W23 Potential Phosphorylation Sites: S26 S35 S54 S68	SFSCAN HMIMER HMIMER TMAP MGITFS
25	1904303CD1	503	Signal_cleavage: M1-A20 Signal Peptide: M1-A20 Signal Peptide: M1-C24 Transmembrane domain: V312-L340, P467-L494N-terminus is non-cytosolic Potential Phosphorylation Sites: S21 S174 S237 S364 S368 S413 S426 T46 T176 T203 T248 T288 T338 T411 Potential Glycosylation Sites: N220 N229 N278 N346 Signal_cleavage: M39-A06	SFSCAN HMIMER HMIMER TMAP MGITFS SFSCAN MGITFS SFSCAN
26	2911343CD1	321	Signal_cleavage: M39-A06	MGITFS SFSCAN

WO 02/086069

PCT/US02/12464

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
			Leucine Rich Repeat: N67-P88, S89-P110, C111-P131, S133-L162, E45-Q66 Potential Phosphorylation Sites: S36 S89 S175 S182 S183 S196 S304 S319 T23 T28 T42 T48 T38 T124 T145 T232 T260 T274 T311 T315 Potential Glycosylation Sites: N26 N138	HMMER_PFAM MOTIFS
27	7500308CD1	60	Signal_cleavage: M1-P21 Signal Peptide: M1-L25, M1-L26 Potential Phosphorylation Sites: S42 T33	SPSCAN HMMER MOTIFS
28	7501098CD1	45	Signal_cleavage: M1-G17 Signal Peptide: M1-G17 Signal Peptide: M1-G24 Signal Peptide: M1-P19 Histidine acid phosphatases signatures: M1-A44 Signal_cleavage: M1-A15	SPSCAN HMMER HMMER HMMER PROSITE/ESCAN SPSCAN
29	7503839CD1	43	Signal Peptide: M1-S18 Potential Phosphorylation Sites: S18 S19 T40	HMMER MOTIFS
30	7503698CD1	456	Signal_cleavage: M1-A20 Signal Peptide: M1-A19, M1-S21, M1-A20, M1-G24 Potential Phosphorylation Sites: S21 S176 S315 S319 S364 S377 T46 T176 T203 T239 T289 T362 Potential Glycosylation Sites: N220 N229 N287	SPSCAN HMMER HMMER MOTIFS MOTIFS

Table 4

Polymer SEQ ID NO./ Incid ID / Sequence Length	Sequence Fragments
31/1895273CB1/ 1053	1-211, 1-320, 67-288, 67-535, 67-644, 69-216, 69-230, 69-286, 69-293, 69-331, 69-344, 69-338, 69-363, 69-481, 69-495, 69-591, 69-607, 69-617, 69-636, 112-506, 132-985, 150-394, 186-786, 188-967, 190-641, 190-726, 190-727, 241-713, 243-711, 258-1000, 264-950, 268-877, 276-991, 286-712, 287-995, 310-953, 326-614, 367-1027, 423-606, 428-1036, 429-679, 434-724, 435-709, 459-653, 459-676, 459-998, 471-1003, 483-697, 488-1032, 489-971, 503-1050, 507-1023, 508-1053, 508-645, 743-1051
32/7007222CB1/ 1579	1-550, 1-613, 4-295, 28-284, 300-627, 362-623, 372-623, 377-664, 406-930, 430-830, 451-607, 459-1027, 459-1048, 544-1022, 564-1027, 590-1146, 609-1134, 665-1320, 676-1263, 691-1053, 694-1146, 709-1320, 730-1209, 740-1263, 746-1061, 751-1015, 751-1223, 772-1205, 784-1270, 792-1332, 805-1234, 805-1246, 815-1322, 816-1460, 843-1562, 860-1263, 876-1362, 904-1160, 925-1492, 943-1575, 976-1544, 1024-1307, 1033-1579, 1041-1069, 1041-1339, 1041-1341, 1041-1342, 1041-1343, 1041-1346, 1041-1347, 1045-1172, 1046-1069, 1054-1579, 1057-1347, 1084-1544, 1104-1579, 1120-1579
33/4559223CB1/ 2440	1-471, 51-284, 128-825, 129-705, 222-334, 253-785, 282-593, 283-609, 283-786, 357-400, 408-1003, 457-699, 457-826, 467-1050, 469-993, 520-1132, 534-752, 585-866, 585-1061, 635-1291, 709-966, 747-1291, 781-1350, 813-1417, 820-1034, 906-1151, 925-1291, 942-1341, 1021-1294, 1064-1326, 1083-1591, 1110-1380, 1110-1673, 1132-1482, 1164-1340, 1169-1705, 1237-1490, 1237-1607, 1246-1793, 1253-1738, 1253-1843, 1270-1876, 1348-1997, 1418-1729, 1474-2026, 1486-1765, 1495-1683, 1495-1716, 1517-2020, 1551-2206, 1608-1890, 1611-1754, 1653-2268, 1665-2002, 1671-1947, 1681-1942, 1691-1801, 1731-2269, 1741-2278, 1752-1821, 1759-2418, 1780-2270, 1793-2408, 1819-2366, 1827-2321, 1829-2018, 1829-2358, 1839-2404, 1846-2362, 1883-2345, 1889-2337, 1901-2118, 1908-2061, 1908-2365, 1913-2171, 1913-2435, 1935-2420, 1936-2174, 1943-2314, 1957-2423, 1963-2440, 1968-2424, 1969-2423, 1970-2422, 1971-2411, 1975-2423, 1986-2423, 1997-2440, 2000-2266, 2009-2420, 2018-2440, 2019-2277, 2020-2414, 2023-2440, 2036-2432, 2038-2440, 2043-2424, 2044-2434, 2053-2424, 2066-2420, 2070-2420, 2071-2420, 2095-2395, 2141-2440, 2146-2420, 2147-2420, 2157-2416, 2157-2430, 2159-2334, 2158-2376, 2158-2390, 2201-2418, 2279-2420, 2311-2420

Table 4

Polymucleotide SEQ ID NO./ Incr/ID/ Sequence Length	Sequence Fragments
34/34/1255CB1/ 4133	1-349, 1-651, 243-430, 489-4133, 655-1312, 665-1263, 925-1207, 925-1543, 932-1543, 945-1617, 1040-1585, 1260-1709, 1293-1597, 1388-1675, 1388-2004, 1681-1927, 1700-1956, 1700-2164, 1725-2284, 1796-2295, 1809-2215, 2074-2304, 2074-2336, 2074-2465, 2127-2774, 2148-2409, 2205-2442, 2515-2918, 2515-3045, 2517-2784, 2640-2908, 2659-3187, 2725-3340, 2743-2990, 2777-3023, 2790-3014, 2794-3321, 2832-3337, 2832-3649, 2841-3160, 2872-3447, 2948-3109, 2988-3368, 3027-3368, 3105-3321, 3277-3507, 3327-3983, 3327-4004, 3381-4029, 3406-3904, 3427-4003, 3514-4077, 3594-4128, 3632-4077, 3653-3886, 3709-4018
35/1958017CB1/ 4689	1-70, 1-623, 2-516, 185-463, 356-989, 464-660, 464-846, 609-803, 609-896, 609-1018, 609-1133, 609-1137, 617-1058, 617-1201, 660-1445, 676-1349, 677-904, 679-1420, 687-1445, 706-935, 709-1445, 725-1137, 726-1443, 779-1137, 782-904, 852-1137, 936-1137, 977-1137, 1000-1137, 1020-1202, 1044-1137, 1089-1137, 1097-1137, 1113-1137, 1138-1327, 1203-1327, 1203-1502, 1213-1744, 1328-1502, 1328-1616, 1375-1766, 1394-1654, 1394-1806, 1394-1827, 1489-2117, 1503-1744, 1617-1744, 1617-2054, 1655-2312, 1745-2054, 1777-2316, 1848-2366, 2001-2435, 2047-2477, 2065-2275, 2092-2575, 2132-2409, 2180-2748, 2073-2762, 2283-2794, 2313-2555, 2355-2802, 2369-2917, 2467-3119, 2530-2979, 2547-3208, 2572-3124, 2576-2805, 2591-2896, 2603-2875, 2670-3140, 2671-2071, 2683-3224, 2683-3269, 2695-3261, 2711-3335, 2821-3340, 2821-3554, 3048-3544, 3172-3694, 3169-3603, 3172-3376, 3217-3410, 3217-3752, 3245-3782, 3332-3875, 3349-3924, 3435-3924, 3503-4186, 3503-4244, 3504-3748, 3551-3840, 3551-4036, 3611-4005, 3631-3845, 3714-3745, 3714-3750, 3714-3751, 3714-3753, 3714-3780, 3714-3781, 3714-3792, 3714-3797, 3714-3798, 3714-3801, 3714-3828, 3714-3829, 3714-3846, 3714-3848, 3714-3876, 3714-3896, 3717-3755, 3718-3753, 3718-3800, 3719-4008, 3721-4145, 3735-3893, 3735-3896, 3761-3795, 3761-3801, 3761-3830, 3761-3887, 3761-3892, 3761-3944, 3766-3944, 3768-3944, 3777-3850, 3782-4032, 3782-4301, 3809-3843, 3809-3848, 3809-3877, 3809-3955, 3809-3940, 3809-3943, 3813-3847, 3813-3895, 3814-3943, 3816-3943, 3830-3944, 3856-3890, 3856-3896, 3856-3925, 3856-3944, 3856-3945, 3860-3894, 3860-4689, 3861-3944, 3863-3944, 3877-3943, 3904-3938, 3904-3943, 3904-3949, 3904-4188, 3908-3942, 3909-3943, 3911-3943, 3930-4402, 3976-4498, 3979-4343, 3982-4387, 3993-4262, 3993-4480, 4083-4304, 4108-4339, 4136-4328, 4136-4416, 4226-4537, 4360-4645, 4491-4547, 4628-4651

Table 4

Polynucleotide SEQ ID NO./ Inverse ID/ Sequence Length	Sequence Fragments
36/6219465CBI/ 3294	<p>1-1311, 168-355, 168-483, 168-548, 168-573, 168-588, 168-592, 168-601, 168-643, 168-646, 168-679, 168-682, 168-686, 168-689, 168-696, 168-715, 168-745, 168-765, 168-797, 168-819, 168-822, 168-835, 168-843, 169-745, 171-551, 172-577, 172-587, 172-608, 172-745, 174-745, 175-771, 178-678, 178-733, 178-745, 187-642, 188-526, 189-526, 202-498, 211-798, 216-745, 224-548, 224-532, 224-535, 224-745, 226-721, 229-745, 230-452, 230-525, 237-415, 237-596, 238-526, 251-436, 263-772, 278-813, 306-951, 325-967, 326-800, 338-510, 341-510, 341-868, 347-745, 382-790, 397-670, 515-745, 514-577, 561-1309, 567-800, 368-804, 570-772, 570-843, 593-1169, 611-1123, 641-1266, 663-1116, 698-1255, 703-1188, 727-1173, 740-1313, 755-1176, 765-1333, 765-1384, 772-1298, 785-1417, 788-1448, 788-1366, 788-1381, 788-1389, 788-1467, 820-1413, 829-1319, 835-1313, 856-1437, 837-1511, 854-1422, 870-1516, 873-1212, 882-1101, 918-986, 936-986, 940-986, 951-986, 962-986, 1014-1226, 1014-1443, 1014-1456, 1014-1485, 1014-1544, 1014-1548, 1014-1655, 1015-1476, 1016-1105, 1031-1561, 1031-1588, 1046-1528, 1049-1458, 1049-1694, 1079-1590, 1086-1604, 1089-1469, 1096-1288, 1108-1730, 1109-1636, 1110-1532, 1113-1384, 1113-1564, 1129-1341, 1156-1783, 1142-1640, 1144-1760, 1172-1692, 1177-1767, 1183-1790, 1186-1773, 1190-1524, 1194-1743, 1194-1790, 1200-1642, 1202-1790, 1207-1790, 1211-1783, 1215-1712, 1215-1808, 1217-1747, 1223-1776, 1224-1790, 1236-1790, 1236-1864, 1237-1827, 1244-1662, 1244-1790, 1257-1905, 1259-1835, 1289-1790, 1292-1873, 1294-1648, 1295-1541, 1302-1790, 1335-1773, 1337-1790, 1344-1978, 1352-1790, 1358-1663, 1370-2054, 1378-1888, 1411-1922, 1434-1596, 1434-1927, 1439-2074, 1445-1868, 1461-1868, 1477-1790, 1505-1958, 1506-1913, 1516-1807, 1529-1790, 1549-1793, 1587-1818, 1587-1889, 1587-1998, 1587-2083, 1587-2139, 1587-2224, 1587-2254, 1589-2051, 1595-1840, 1640-2133, 1644-1960, 1646-2251, 1647-2275, 1648-2280, 1667-1970, 1670-1995, 1671-2250, 1672-2225, 1673-1999, 1677-2245, 1715-2331, 1720-1986, 1730-2238, 1736-2233, 1753-1928, 1766-2356, 1770-2448, 1777-2448, 1792-2384, 1792-2380, 1793-2244, 1797-2650,</p>

Table 4

Polynucleotide SEQ ID NO./ Incid ID / Sequence Length	Sequence Fragments
37/5776625CB1/ 3635	1799-2441, 1804-2342, 1818-2225, 1822-2434, 1829-2135, 1846-2088, 1862-2540, 1885-2470, 1890-2483, 1892-2145, 1909-2535, 1912-2526, 1933-2130, 1943-2516, 1954-2582, 1963-2492, 1971-2559, 1971-2614, 1980-2524, 1983-2503, 1989-2438, 1993-2464, 1999-2547, 1999-2733, 2008-2492, 2009-2420, 2009-2535, 2010-2614, 2017-2464, 2061-2332, 2065-2669, 2071-2535, 2104-2639, 2109-2649, 2117-2669, 2118-2379, 2125-2748, 2131-2558, 2136-2639, 2146-2739, 2146-2802, 2155-2686, 2159-2821, 2161-2685, 2169-2748, 2170-2692, 2189-2702, 2200-2740, 2222-2723, 2230-2764, 2264-2780, 2293-2720, 2296-2537, 2299-2748, 2328-2575, 2359-2608, 2359-2929, 2363-2599, 2381-2679, 2385-2729, 2393-2693, 2476-2690, 2493-3000, 2494-2712, 2494-2721, 2530-2776, 2668-2872, 2618-2850, 2667-2925, 2667-2948, 2698-2949, 2700-2947, 2701-2920, 2706-2989, 2763-2964, 2763-2970, 2776-3020, 2781-2976, 2793-3038, 2859-3020, 2869-3140, 2967-3239, 3040-3186, 3040-3294
	1-90, 1-588, 62-157, 62-354, 62-464, 62-551, 62-584, 62-582, 62-686, 67-282, 67-571, 68-623, 71-657, 75-618, 78-892, 238-622, 80-493, 91-1959, 98-332, 98-354, 98-549, 108-677, 125-756, 130-815, 175-814, 227-722, 245-818, 1106, 866-1113, 876-1458, 935-1098, 1133-1568, 1141-1713, 1153-1445, 1167-1598, 1169-1383, 1355-1735, 1355-1739, 1355-1745, 1355-1794, 1355-1812, 1355-1856, 1355-1852, 1355-1856, 1355-1885, 1355-1887, 1355-1891, 1355-1918, 1355-1926, 1355-1962, 1355-1976, 1356-1891, 1389-1695, 1389-1900, 1389-1905, 1389-2010, 1389-2017, 1414-1891, 1556-2122, 1536-2177, 1536-2185, 1556-2202, 1556-2216, 1567-1937, 1575-1988, 1578-2212, 1579-2035, 1579-2081, 1579-2118, 1579-2215, 1579-2216, 1580-2190, 1581-2212, 1584-2216, 1588-1891, 1588-2216, 1641-2212, 1654-2216, 1658-1963, 1660-2216, 1674-2243, 1680-2212, 1714-2230, 1719-2295, 1726-2121, 1734-1869, 1748-2305, 1752-2334, 1753-2203, 1760-2419, 1783-2305, 1806-2367, 1817-2393, 1866-2332, 1868-2327, 1877-2525, 1883-2554, 1884-2071, 1910-2141, 1919-2565, 1930-2345, 1934-2595, 1952-2422, 1996-2712, 2040-3525, 2044-2736, 2053-2630, 2058-2754, 2092-2616, 2107-2707, 2114-2736, 2127-2713, 2128-2659, 2159-2740, 2168-2510, 2231-2749, 2240-2707, 2245-2491, 2246-2885, 2285-2961, 2320-2972, 2330-2895, 2350-2806, 2380-2905, 2393-3096, 2495-2705, 2496-3107, 2511-3203, 2547-2927, 2553-3249, 2573-3283, 2602-3263, 2607-3232, 2623-3100, 2626-2882, 2643-3294, 2688-2943, 2689-3407, 2699-3154, 2699-3249, 2712-3274, 2719-3350, 2734-3203, 2752-2993, 2759-3413, 2770-3455, 2773-3326, 2773-3440, 2783-3446, 2783-3146, 2796-2825, 2856-3402, 2863-3154, 2869-3322, 2887-3435, 2932-3635, 2947-3523, 2951-3473, 3002-3497, 3127-3635, 3129-3522, 3131-3352, 3195-3431, 3200-3460, 3216-3483

Table 4

Polynucleotide SEQ ID NO./ Inye ID/ Sequence Length	Sequence Fragments
38/4765758CB1/ 635	1-280, 1-611, 23-635
39/7236661CB1/ 793	1-56, 1-210, 1-379, 1-414, 1-503, 1-565, 1-793, 43-679, 130-565, 202-793
40/7714187CB1/ 2741	1-140, 15-532, 15-874, 158-262, 189-784, 379-662, 485-1033, 680-1171, 986-1068, 1062-1612, 1069-1599, 1292-1588, 1373-1629, 1374-1843, 1404-2741, 1585-1828, 1585-2271, 1587-1822, 1682-2250, 2065-2734
41/5136540CB1/ 1074	1-252, 137-209, 137-412, 137-556, 137-575, 137-583, 137-597, 137-707, 137-738, 137-767, 145-191, 169-699, 290-675, 298-791, 315-574, 417-591, 426-901, 428-1071, 428-1074, 431-518, 452-1026, 468-1062, 488-1048, 518-771, 518-795, 518-920, 518-1062, 521-1054, 523-773, 523-1068, 527-974, 536-980, 561-785, 574-1068, 595-1070, 595-1071, 626-1071, 637-1071, 647-1088, 811-1071, 984-1068
42/3277403CB1/ 2560	1-76, 1-632, 7-76, 19-71, 21-76, 119-426, 189-458, 234-750, 269-526, 381-2519, 472-1015, 667-1197, 994-1487, 1224-1446, 1281-1385, 1453-1603, 1516-1790, 1516-1967, 1516-2010, 1620-2113, 1621-2213, 1677-1962, 1677-2201, 1711-1976, 1711-1999, 1743-2271, 1858-2500, 1863-2066, 1865-2303, 2112-2358, 2528-2559
43/1517569CB1/ 5833	1-541, 1-604, 12-686, 47-707, 124-507, 179-765, 198-661, 234-822, 237-811, 275-947, 377-1021, 414-1020, 430-909, 472-1022, 517-986, 547-1067, 640-1021, 713-1086, 790-1268, 793-1366, 793-1446, 793-1533, 797-1443, 797-1453, 799-1225, 799-1333, 850-1493, 949-1544, 1064-1706, 1073-1570, 1122-1724, 1131-1704, 1137-1707, 1218-1916, 1279-1535, 1392-1617, 1501-1559, 1505-1753, 1505-1971, 1522-1749, 1609-1871, 1609-1877, 1609-2052, 1824-2236, 1930-2502, 1963-2559, 1980-2641, 1981-2454, 2019-2590, 2038-2544, 2038-2672, 2038-2706, 2149-2767, 2175-2494, 2200-2428, 2200-2692, 2201-2738, 2214-2858, 2220-2931, 2265-2968, 2267-2463, 2268-2480, 2296-2888, 2333-2588, 2332-2971, 2355-2614, 2478-3083, 2479-3121, 2487-3119, 2493-3096, 2523-3136, 2552-3021, 2589-3194, 2591-3249, 2598-3176, 2608-3176, 2619-3303, 2627-3219, 2635-3357, 2638-3040, 2639-3049, 2659-3211, 2640-3052, 2643-3321, 2648-3147, 2677-3184, 2702-3275, 2712-2972, 2715-2986, 2715-3188, 2720-3099, 2733-3428, 2789-3021, 2872-3380, 2882-3443, 2882-3571, 2884-3350, 2888-3388, 2888-3591, 2889-3167,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
	2890-3360, 2948-3681, 2954-3443, 2960-3664, 2975-3571, 2976-3689, 2986-3652, 2991-3692, 2997-3685, 3026-3579, 3032-3667, 3048-3713, 3064-3530, 3080-3375, 3085-3729, 3109-3362, 3114-3548, 3118-3722, 3118-3726, 3119-3399, 3119-3722, 3120-3693, 3135-3738, 3136-3736, 3138-3736, 3143-3377, 3148-3739, 3153-3415, 3160-3435, 3160-3670, 3160-3734, 3162-3713, 3162-3738, 3164-3422, 3172-3424, 3179-3688, 3183-3517, 3187-3739, 3188-3738, 3193-3728, 3206-3482, 3207-3741, 3209-3692, 3213-3420, 3231-3525, 3231-3725, 3231-3743, 3232-3738, 3234-3738, 3243-3508, 3252-3741, 3292-3858, 3299-3858, 3699-4031, 3699-4088, 3718-4036, 3749-3948, 3749-4253, 3749-4285, 3814-4000, 3926-4281, 4087-4111, 4087-4115, 4087-4165, 4087-4196, 4087-4200, 4087-4250, 4089-4170, 4094-4250, 4126-4250, 4128-4717, 4137-4250, 4141-4165, 4172-4316, 4172-4333, 4172-4335, 4179-4431, 4211-4385, 4311-4536, 4408-4674, 4408-4963, 4487-4797, 4664-4956, 4671-5297, 4871-5193, 4871-5396, 4871-5255, 5024-5292, 5024-5586, 5033-5297, 5085-5238, 5098-5365, 5158-5369, 5391-5630, 5406-5684, 5406-5699, 5406-5700, 5406-5714, 5408-5691, 5442-5718, 5513-5740, 5542-5819, 5549-5698, 5584-5833, 5607-5833, 5665-5727
442415991CIB1/ 2264	1-707, 230-714, 245-538, 308-596, 351-1022, 419-681, 419-1022, 451-911, 451-959, 487-707, 487-989, 598-906, 599-933, 611-858, 620-833, 661-1022, 713-1308, 786-1423, 798-1084, 875-1107, 890-1681, 892-1474, 903-1172, 927-1197, 927-1227, 947-1173, 1009-1271, 1022-1291, 1171-1756, 1174-1435, 1175-1471, 1206-1423, 1244-1760, 1274-1463, 1298-1539, 1308-1539, 1308-1722, 1308-1884, 1310-1890, 1311-1521, 1321-1591, 1344-1623, 1344-1997, 1359-1565, 1395-1654, 1453-1744, 1510-2164, 1528-2178, 1536-1819, 1539-1827, 1570-2221, 1592-2216, 1644-1875, 1644-2155, 1648-2225, 1698-1959, 1730-1959, 1747-2264, 1748-1914, 1757-2019, 1812-2061, 1817-2043, 1893-2108, 1895-2166, 2037-2248, 2054-2264, 2063-2134

WO 02/086069

PCT/US02/12464

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
45/273742CB1/ 1638	1-625, 326-449, 326-502, 326-532, 326-533, 326-534, 326-544, 326-548, 326-549, 326-552, 326-557, 326-561, 326-568, 326-569, 326-571, 326-572, 326-573, 326-574, 326-575, 326-578, 326-579, 326-584, 326-587, 326-590, 326-599, 326-601, 326-610, 326-662, 326-747, 326-771, 326-836, 326-848, 326-884, 326-894, 345-566, 362-586, 362-595, 372-624, 418-618, 460-719, 482-725, 516-675, 582-847, 588-715, 599-885, 601-814, 627-986, 633-884, 646-900, 653-972, 680-895, 680-904, 684-917, 711-1000, 738-945, 757-993, 763-1019, 764-995, 798-1042, 805-1053, 830-1101, 878-1102, 889-1286, 914-1526, 931-1157, 937-1152, 942-1201, 942-1519, 947-1156, 947-1321, 958-1183, 965-1188, 965-1195, 965-1514, 997-1566, 1019-1292, 1033-1269, 1080-1262, 1083-1319, 1112-1389, 1133-1369, 1146-1386, 1146-1475, 1146-1587, 1191-1381, 1191-1549, 1203-1390, 1330-1563, 1352-1612, 1363-1638, 1395-1582, 1395-1594
46/2768535CB1/ 4790	1-520, 46-518, 46-520, 88-707, 129-472, 186-883, 187-765, 206-353, 421-1034, 505-1159, 623-908, 644-1251, 652-1211, 654-1195, 678-1221, 690-1269, 703-1254, 704-1393, 706-989, 762-1001, 770-1352, 781-1389, 797-1335, 813-1420, 870-1413, 945-1256, 1050-1703, 1108-1788, 1114-1738, 1132-1698, 1173-1815, 1189-1796, 1195-1801, 1204-1798, 1320-1719, 1248-1823, 1262-1921, 1266-1924, 1270-1883, 1272-1807, 1283-1664, 1283-1870, 1336-2021, 1425-2056, 1438-2076, 1527-2070, 1621-2136, 1676-2224, 1696-2084, 1750-2313, 1865-2252, 1918-2152, 1923-2186, 1926-2568, 1964-2230, 1964-2464, 2072-2606, 2275-2500, 2401-3007, 2464-2726, 2470-2676, 2494-2785, 2533-2819, 2540-2809, 2595-3128, 2621-2839, 2640-2882, 2649-2884, 2654-2925, 2656-2930, 2698-2958, 2727-2928, 2727-3213, 2782-2953, 2782-3400, 2853-3148, 2858-3045, 2889-3189, 2891-3338, 2891-3379, 2894-3205, 2934-3187, 2948-3209, 2972-3515, 2986-3244, 3000-3209, 3003-3246, 3027-3291, 3028-3230, 3028-3300, 3035-3298, 3074-3268, 3074-3298, 3074-3666, 3106-3350, 3192-3404, 3192-3456, 3221-3449, 3223-3640,

Table 4

Polynucleotide SEQ ID NO/ Incyte ID/ Sequence Length	Sequence Fragments
	3266-3531, 3266-3756, 3289-3860, 3292-3804, 3303-3643, 3308-3608, 3309-3598, 3361-3649, 3364-3659, 3400-3643, 3406-3654, 3407-3672, 3411-3811, 3420-3867, 3421-3695, 3424-3703, 3428-3705, 3450-3787, 3454-3696, 3460-3739, 3473-3868, 3473-3875, 3487-3648, 3498-3745, 3503-3767, 3516-3753, 3533-3728, 3539-3762, 3599-3809, 3601-3733, 3620-4103, 3638-3678, 3647-3884, 3647-3919, 3647-4186, 3716-3988, 3724-3996, 3737-4008, 3746-3998, 3752-3983, 3765-4415, 3771-4010, 3771-4004, 3809-4098, 3823-4098, 3836-4109, 3857-4095, 3858-4063, 3863-4367, 3873-4093, 3882-4128, 3900-4146, 3909-4147, 3984-4413, 3995-4242, 4025-4670, 4043-4296, 4056-4349, 4066-4123, 4100-4393, 4100-4688, 4115-4333, 4116-4594, 4143-4696, 4147-4677, 4152-4410, 4191-4680, 4202-4429, 4215-4494, 4221-4479, 4222-4509, 4277-4532, 4318-4682, 4322-4659, 4323-4543, 4380-4630, 4394-4609, 4423-4661, 4423-4667, 4439-4696, 4444-4696, 4450-4665, 4459-4696, 4472-4676, 4494-4696, 4494-4790
476848851CBI/ 3916	1-293, 136-477, 137-447, 170-482, 171-477, 172-432, 172-446, 173-303, 173-433, 173-447, 174-416, 175-443, 180-393, 193-365, 193-447, 197-447, 202-447, 204-476, 206-476, 257-599, 366-599, 366-749, 386-429, 502-684, 509-743, 600-749, 600-1197, 604-1073, 604-1184, 627-1215, 637-1308, 750-1197, 750-1603, 775-951, 921-1562, 1143-1704, 1198-1603, 1198-1737, 1244-1601, 1293-1712, 1365-1613, 1596-1748, 1596-2025, 1604-1737, 1604-1838, 1605-1997, 1680-2413, 1738-2255, 1732-2340, 1749-1997, 1782-2357, 2040-2689, 2040-2690, 2059-2331, 2254-2690, 2421-2777, 2421-2794, 2424-2824, 2435-3135, 2465-2742, 2480-2802, 2551-3204, 2571-2852, 2609-3171, 2637-2912, 2732-3429, 2946-3149, 2946-3219, 2980-3258, 2994-3238, 3162-3406, 3186-3436, 3187-3610, 3214-3590, 3214-3620, 3226-3767, 3232-3512, 3235-3493, 3235-3616, 3254-3916, 3280-3916, 3378-3607
487040722CBI/ 1702	1-1702, 5-561, 20-561, 107-167, 107-194, 107-240, 264-546, 420-852, 420-878, 420-1108, 420-1120, 424-1094, 947-1702, 952-988, 953-1702, 956-1702, 965-1702
496430290CBI/ 1462	1-274, 27-274, 28-274, 29-274, 29-583, 40-683, 567-1462, 567-1429, 567-1448, 567-1458, 567-1462, 931-1211

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
5072640251CB1/ 3958	1-282, 19-539, 73-658, 78-592, 83-598, 97-759, 106-624, 236-700, 335-611, 337-611, 375-634, 384-987, 404-759, 489-700, 508-825, 523-790, 528-947, 594-899, 606-899, 600-902, 605-899, 668-931, 678-923, 685-899, 705-947, 746-983, 780-1061, 843-1073, 881-1389, 977-1275, 1043-1287, 1046-1323, 1111-1692, 1231-1506, 1322-1572, 1379-1630, 1408-1814, 1408-1873, 1433-2210, 1440-1692, 1563-2005, 1614-2034, 1626-2024, 1732-2034, 1750-1972, 1820-2378, 1853-2417, 1975-2249, 2009-2210, 2043-2287, 2084-2626, 2091-2619, 2454-3008, 2468-3078, 2480-3019, 2584-3143, 2678-3169, 2687-3270, 2754-3251, 2901-3169, 2902-3169, 2936-3597, 3261-3938, 3382-3686, 3382-3779
51/3839350CB1/ 826	1-296, 1-432, 1-450, 1-549, 1-796, 3-826, 102-826, 161-430, 190-798, 199-401, 221-736, 246-564, 254-773, 272-489, 299-456
52/6393813CB1/ 729	1-501, 18-719, 32-407, 32-493, 42-536, 51-302, 57-630, 58-505, 68-351, 68-608, 124-411, 173-710, 193-644, 208-232, 216-516, 230-463, 230-716, 243-706, 256-706, 278-536, 283-729, 321-729, 334-717, 538-729, 594-655
53/688755CB1/ 1610	1-201, 16-697, 79-316, 197-478, 197-1448, 249-828, 351-803, 351-843, 353-846, 354-843, 358-845, 362-820, 363-828, 365-846, 366-846, 434-846, 459-846, 476-787, 479-1448, 480-1553, 530-816, 546-846, 606-846, 609-828, 839-1279, 839-1333, 840-1331, 843-1210, 865-1323, 874-1051, 882-1321, 1252-1610
54/7128459CB1/ 869	1-337, 1-576, 1-617, 1-869, 120-487, 283-749, 283-750, 285-696, 286-739, 326-829, 327-828, 383-774, 420-819, 456-869
55/1904303CB1/ 2209	1-515, 290-1414, 404-692, 416-828, 422-698, 429-676, 429-980, 530-747, 530-765, 530-1056, 586-838, 624-835, 744-1052, 761-1021, 780-1070, 793-1288, 877-1193, 976-1439, 978-1350, 1018-1382, 1057-1255, 1057-1397, 1096-1378, 1097-1412, 1097-1414, 1173-1451, 1203-1325, 1247-1687, 1247-1690, 1248-2190, 1254-1539, 1258-1414, 1298-1773, 1339-1918, 1345-1614, 1345-1830, 1345-1918, 1399-1618, 1434-1836, 1459-2128, 1500-2115, 1517-1747, 1523-1786, 1582-2168, 1587-2209, 1590-2179, 1693-2200, 1719-2165, 1722-2208, 1725-2190, 1761-2184, 1765-2188, 1772-2189, 1774-2182, 1779-2187, 1788-2156, 1796-2187, 1812-2208, 1818-2182, 1820-2208, 1822-2182, 1835-2192, 1827-2182, 1833-2182, 1846-2190, 1849-2192, 1838-2190, 1861-2184, 1865-2192, 1870-2182, 1871-2182, 1878-2180, 1879-2160, 1879-2180, 1897-2182, 1907-2187, 1928-2208, 2004-2167, 2095-2182

Table 4

Polymer SEQ ID NO./ Incid ID / Sequence Length	Sequence Fragments
56/2911343CB1/ 1520	1-582, 1-1513, 43-138, 43-315, 43-573, 86-573, 101-347, 202-436, 465-569, 466-569, 523-996, 525-772, 528-1013, 571-831, 571-838, 573-1124, 578-1373, 620-849, 658-806, 666-931, 666-1181, 732-1368, 732-1368, 735-1343, 774-1377, 800-944, 829-1410, 830-1400, 834-1104, 885-1146, 933-1210, 939-1520, 942-1460, 970-1191, 999-1274, 1000-1411, 1051-1506, 1053-1514, 1053-1520, 1066-1520, 1067-1520, 1068-1514, 1075-1470, 1083-1448, 1098-1520, 1108-1423, 1111-1493, 1138-1514, 1152-1487, 1163-1520, 1180-1520, 1187-1514, 1260-1520, 1274-1514
57750308CB1/ 1282	1-1282, 4-295, 363-612, 363-888, 402-684, 402-707, 402-1050, 406-1001, 407-1018, 459-1006, 452-764, 454-859, 463-718, 474-1001, 495-1053, 516-976, 580-1045, 646-1278, 748-1046, 748-1046, 748-1050, 835-1116, 887-1282, 1011-1243, 1060-1282
587501098CB1/ 1228	1-252, 1-394, 1-397, 1-399, 1-1014, 87-193, 87-197, 87-198, 359-843, 363-843, 370-1010, 370-1010, 370-1016, 394-968, 410-1004, 430-990, 460-713, 460-862, 460-1024, 463-996, 465-715, 465-1228, 469-916, 537-1016, 568-1019, 580-1018
597503839CB1/ 3582	1-128, 10-3582, 215-946, 270-739, 270-987, 293-883, 293-933, 293-1125, 441-617, 587-1256, 602-1405, 803-1370, 909-1423, 959-1382, 1014-1803, 1031-1279, 1081-1663, 1262-1414, 1262-1691, 1279-2186, 1346-1720, 1347-1654, 1347-1700, 1347-1705, 1347-1720, 1347-2098, 1394-1922, 1394-2098, 1415-1663, 1446-2023, 1706-2355, 1706-2356, 1710-2385, 1717-2355, 1725-1999, 1779-2355, 1888-2355, 1909-2355, 1920-2355, 1923-2355, 1976-2284, 2002-2355, 2034-2947, 2087-2443, 2087-2460, 2090-2490, 2101-2801, 2137-2265, 2137-2408, 2147-2468, 2217-2870, 2237-2518, 2275-2837, 2303-2578, 2334-2745, 2340-3263, 2341-2615, 2341-2879, 2341-2920, 2341-2927, 2341-3008, 2341-3037, 2341-3051, 2341-3088, 2341-3134, 2341-3233, 2346-3199, 2352-3252, 2384-2822, 2398-3095, 2423-2696, 2488-2721, 2489-3363, 2562-3289, 2612-2815, 2612-2885, 2614-3559, 3632-3296, 3660-2904, 3661-2924, 3674-3520, 3710-3484, 3828-3072, 3853-3102, 3853-3275, 3874-3555, 2880-3256, 2898-3178, 2901-3159, 2901-3282, 2920-3582, 2946-3582, 3049-3273, 3162-3546, 3258-3449, 3288-3561, 3373-3501

Table 4

Polymers SEQ ID NO./ Amino ID/ Sequence Length	Sequence Fragments
607503698CB1/ 2063	1-2043, 404-692, 416-828, 422-698, 429-676, 530-747, 530-765, 530-1056, 586-838, 624-835, 744-1052, 761-1021, 780-1070, 897-1056, 963-1267, 1100-1540, 1100-1543, 1107-1392, 1111-1267, 1147-1267, 1151-1626, 1192-1771, 1198-1352, 1198-1467, 1198-1733, 1198-1800, 1236-1984, 1252-1471, 1312-1606, 1312-1981, 1370-1600, 1384-1609, 1435-2021, 1440-2043, 1443-2032, 1517-1809, 1518-1743, 1546-2043, 1572-2018, 1575-2043, 1578-2063, 1580-2037, 1584-2039, 1614-2037, 1616-2041, 1617-1993, 1625-2042, 1627-2033, 1632-2040, 1669-2040, 1665-2043, 1671-2035, 1673-2043, 1675-2035, 1676-2035, 1678-2040, 1680-2035, 1681-2040, 1686-2035, 1687-2035, 1699-2046, 1711-2047, 1714-2037, 1716-2034, 1720-2031, 1720-2033, 1720-2039, 1720-2063, 1723-2035, 1724-2035, 1724-2039, 1729-2039, 1731-2013, 1731-2033, 1731-2039, 1732-2013, 1732-2033, 1732-2033, 1732-2039, 1750-2035, 1760-2040, 1781-2043, 1857-2020, 1871-2063, 1948-2033, 1980-2043

WO 02/086069

PCT/US02/12464

Table 5

Polynucleotide SEQ ID NO:	Incye Project ID:	Representative Library
31	1895273CB1	THPINOT03
32	70072222CB1	BRAIFEN05
33	3559223CB1	BRAHIDK01
34	3441255CB1	PLACNOB01
35	1958917CB1	BRSTNOT23
36	6219465CB1	BRAUNOR01
37	3576625CB1	NERDTDN03
38	4765758CB1	PLACNOT05
39	7236661CB1	BRAUNOR01
40	7714187CB1	BRAIHCT01
41	5136540CB1	UTREDIT07
42	3277403CB1	FIBRUNT02
43	1517569CB1	LUNGNOT18
44	2415991CB1	BRAINOT11
45	2735742CB1	LUNGPET03
46	2768535CB1	ADRENOT07
47	6848851CB1	BRAIFER05
48	7040722CB1	UTRSTMR02
49	6430290CB1	LUNGNON07
50	2640251CB1	BRAITUT03
51	3839350CB1	DENDTNT01
52	6393813CB1	KIDCTMT01
53	5685755CB1	BRAIUNT01
54	71728459CB1	BRAHNOT01
55	1904303CB1	OVARTUT10
56	2911343CB1	BSCNNOT03
57	7500308CB1	BRAYDIN03
58	7501098CB1	UTREDIT07
59	7503839CB1	BRAIFER05
60	7503698CB1	LUNGNOT02

WO 02/086069

PCT/US02/12464

Table 6

Library	Vector	Library Description
ADREN0707	pINCY	Library was constructed using RNA isolated from adrenal tissue removed from a 61-year-old female during a bilateral adrenalectomy. Patient history included an unspecified disorder of the adrenal glands.
BRAHN0701	pINCY	Library was constructed using RNA isolated from posterior hippocampus tissue removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Microscopically, the cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. In addition, scattered throughout the cerebral cortex, there were multiple small microscopic areas of cavitation with surrounding gliosis. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly and an enlarged spleen and liver.
BRAHDK01	PSFORT1	This amplified and normalized library was constructed using pooled RNA isolated from archicortex, anterior and posterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominantly over the convexitates, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver. 7.6x10 ⁶ independent clones from this amplified library were normalized in 1 round using conditions adapted Soares et al., PNAS (1994) 91:9228-9232 and Donaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAJEN05	pINCY	This normalized fetal brain tissue library was constructed from 3.26 million independent clones from a fetal brain library. Starting RNA was made from brain tissue removed from a Caucasian male fetus, who was stillborn with a hypoplastic left heart at 23 weeks' gestation. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Donaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAJFER05	pINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
BRAJHC701	pINCY	Library was constructed using RNA isolated from diseased occipital lobe tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema.

WO 02/086069

PCT/US02/12464

Table 6

Library	Vector	Library Description
BRAINOT11	pNCY	Library was constructed using RNA isolated from brain tissue removed from the right temporal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), consistent with chronic seizure disorder. Family history included a cervical neoplasm.
BRAITUD3	PSPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 17-year-old Caucasian female during excision of a cerebral meningioma. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
BRAUNTO1	pNCY	Library was constructed using RNA isolated from SK-N-MC, a neurosphenoma cell line (ATCC HTB-10) derived from a 14-year-old Caucasian female with neurosphenoma, with metastasis to the supra-orbital area.
BRAUNOR01	pNCY	This random primed library was constructed using RNA isolated from stratum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were foci of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolephrosclerosis, nodular colloid goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease.
BRAIDN03	pNCY	This normalized library was constructed from 6.7 million independent clones from a brain tissue library. Starting RNA was made from RNA isolated from diseased hypothalamus tissue removed from a 37-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema. The library was normalized in 2 rounds using conditions adapted from Sorese et al., PNAS (1994) 91:2228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48-hour/round) reannealing hybridization was used. The library was linearized and recircularized to select for insert containing clones.

WO 02/086069

PCT/US02/12464

Table 6

Library	Vector	Library Description
BRKSTNGT23	pINCY	Library was constructed using RNA isolated from diseased breast tissue removed from a 35-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated nonproliferative fibrocystic disease. Family history included type II diabetes, atherosclerotic coronary artery disease, acute myocardial infarction, hyperlipidemia, and coronary artery bypass.
BSCNNGT03	pINCY	Library was constructed using RNA isolated from cadaver nucleus tissue removed from the brain of a 92-year-old male. Pathology indicated several small cerebral infarcts, but no senile plaques or neurofibrillary degeneration. Patient history included throat cancer which was treated with radiation.
DENDNT01	pINCY	Library was constructed using RNA isolated from treated dendritic cells from peripheral blood.
FIBRUNT02	pINCY	Library was constructed using RNA isolated from an untreated MG-63 cell line derived from an osteosarcoma removed from a 14-year-old Caucasian male.
KIDCTMT01	pINCY	Library was constructed using RNA isolated from kidney cortex tissue removed from a 65-year-old male during nephroureterectomy. Pathology for the associated tumor tissue indicated grade 3 renal cell carcinoma within the mid-portion of the kidney and the renal capsule.
LUNGHEW03	pINCY	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
LUNGNON07	pINCY	This normalized lung tissue library was constructed from 5.1 million independent clones from a lung tissue library. Starting RNA was made from RNA isolated from lung tissue. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldi et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
LUNGNOT02	PBLUESCRIPT	Library was constructed using RNA isolated from the lung tissue of a 47-year-old Caucasian male, who died of a subarachnoid hemorrhage.
LUNGNOT18	pINCY	Library was constructed using RNA isolated from left upper lobe lung tissue removed from a 66-year-old Caucasian female. Pathology for the associated tumor tissue indicated a grade 2 adenocarcinoma. Patient history included cerebrovascular disease, atherosclerotic coronary artery disease, and pulmonary insufficiency. Family history included a myocardial infarction and atherosclerotic coronary artery disease.

WO 02/086069

PCT/US02/12464

Table 6

Library	Vector	Library Description
NEREDIN03	pINCY	This normalized dorsal root ganglion tissue library was constructed from 1.05 million independent clones from a dorsal root ganglion tissue library. Starting RNA was made from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema, acute bronchopneumonia, bilateral pleural effusions, pericardial effusion, and malignant lymphoma (natural killer cell type). The patient presented with pyrexia of unknown origin, malaise, fatigue, and gastrointestinal bleeding. Patient history included probable cytomegalovirus infection, liver congestion, and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, respiratory failure, pneumonia of the left lung, natural killer cell lymphoma of the pharynx, Bell's palsy, and tobacco and alcohol abuse. Previous surgeries included colonoscopy, closed colon biopsy, adenotonsillectomy, and nasopharyngeal endoscopy and biopsy. Patient medications included Diflucan (fluconazole), Deliasone (prednisone), hydrocodone, Lorabid, Alprazolam,
		Reaxzone, ProMace-Cytabom, Etoposide, Cisplatin, Cycarabine, and dexamethasone. The patient received radiation therapy and multiple blood transfusions. The library was normalized in 2 rounds using conditions adapted from Soares et al., FNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
OVARTUT10	pINCY	Library was constructed using RNA isolated from ovarian tumor tissue removed from the left ovary of a 58-year-old Caucasian female during a total abdominal hysterectomy, removal of a solitary ovary, and repair of inguinal hernia. Pathology indicated a metastatic grade 3 adenocarcinoma of colonic origin, forming a partially cystic and necrotic tumor mass in the left ovary, and an adenocarcinoma of colonic origin, forming a nodule in the left mesovarium. A single intramural leiomyoma was identified in the myometrium. The cervix showed mild chronic cystic cervicitis. Patient history included benign hypertension, follicular cyst of the ovary, colon cancer, benign colon neoplasm, and osteoarthritis. Family history included emphysema, myocardial infarction, atherosclerotic coronary artery disease, benign hypertension, and hyperlipidemia.
PLACNOB01	PBLUESCRIPT	Library was constructed using RNA isolated from placenta.
PLACNOT05	pINCY	Library was constructed using RNA isolated from placental tissue removed from a Caucasian male fetus, who died after 18 weeks' gestation from fetal demise.
THEPINO103	pINCY	Library was constructed using RNA isolated from untreated THP-1 cells. THP-1 is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (cf. Int. J. Cancer (1980) 26:171).
UTREDT07	pINCY	Library was constructed using RNA isolated from diseased endometrial tissue removed from a female during endometrial biopsy. Pathology indicated in phase endometrium with missing beta 3, Type I defects.

WO 02/086069

PCT/US02/12464

Table 6

Library	Vector	Library Description
UTRSTWROZ	pCDNA2.1	This random primed library was constructed using pooled cDNA from two different donors. cDNA was generated using mRNA isolated from endometrial tissue removed from a 32-year-old female (donor A) and using mRNA isolated from myometrium removed from a 45-year-old female (donor B) during vaginal hysterectomy and bilateral salpingo-oophorectomy. In donor A, pathology indicated the endometrium was secretory phase. The cervix showed severe dysplasia (CIN III) focally involving the squamocolumnar junction at the 1, 6 and 7 o'clock positions. Mild koilocytic dysplasia was also identified within the cervix. In donor B, pathology for the matched tumor tissue indicated multiple (23) subserosal, intramural, and submucosal leiomyomata. Patient history included stress incontinence, extrinsic asthma without status asthmaticus and normal delivery in donor B. Family history included cerebrovascular diseases, depression, and atherosclerotic coronary artery disease in donor B.

WO 02/086069

PCT/US02/12464

Table 7

Program	Description	References	Parameter Threshold
ABFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABIPARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, tblastx, and tblastn.	Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:405-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25:3389-3402.	ESTs: Probability value = 1.0E-8 or less; Full Length sequences: Probability value = 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, fasta, ffastx, ffastn, and fblastx.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci. USA</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183:63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	ESTs: fasta E value = 1.00E-6; Assembled ESTs: fasta Identify = 95% or greater and Match length = 200 bases or greater; ffastx E value = 1.0E-8 or less; Full Length sequences: fasta score = 100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PPAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) <i>Nucleic Acids Res.</i> 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) <i>Methods Enzymol.</i> 206:388-405; and Attwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37:417-424.	Probability value = 1.0E-3 or less

WO 02/086069

PCT/US02/12464

Table 7

Program	Description	Reference	Parameter/Threshold
Procrim HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as Pfam, SMART and TIGRFAM.	Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> 235:1501-1531; Sommeringer, E.L.L. et al. (1998) <i>Nucleic Acids Res.</i> 26:520-522; Durbin, R. et al. (1998) <i>Our World View</i> , in a Nussliel, Cambridge Univ. Press, pp. 1-350.	Pfam, SMART or TIGRFAM hits: Probability value = 1.0E-3 or less; Signal peptide hits: Score = 0 or greater
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Grishkov, M. et al. (1988) <i>CABIOS</i> 4:61-66; Grishkov, M. et al. (1989) <i>Methods Enzymol.</i> 183:146-159; Bairoch, A. et al. (1997) <i>Nucleic Acids Res.</i> 25:217-221.	Normalized quality score \geq GCG specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) <i>Genome Res.</i> 8:175-185; Ewing, B. and P. Green (1998) <i>Genome Res.</i> 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489; Smith, T.F. and M.S. Waterman (1981) <i>J. Mol. Biol.</i> 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) <i>Genome Res.</i> 8:195-202.	
SFSscan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) <i>Protein Engineering</i> 10:1-6; Claverie, J.M. and S. Audic (1997) <i>CABIOS</i> 12:431-439.	Score = 3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determines orientation.	Persson, B. and P. Argos (1994) <i>J. Mol. Biol.</i> 237:182-192; Persson, B. and P. Argos (1996) <i>Protein Sci.</i> 5:363-371.	

WO 02/086069

PCT/US02/12464

Table 7

Program	Description	Reference	Perimeter Threshold
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L., et al. (1998) Proc. Sixth Int. Conf. On Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence (AAAI) Press, Menlo Park, CA, and MIT Press, Cambridge, MA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 02/086069

PCT/US02/12464

Table 8

SEQ ID NO.	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
57	7500308	6824818H1	SNP00058395	172	952	A	A	G	noncoding	n/a	n/a	n/a	n/a
57	7500308	6824818H1	SNP00142506	233	791	A	A	G	noncoding	n/a	n/a	n/a	n/a
57	7500308	6824818H1	SNP00058395	580	952	A	A	G	noncoding	n/a	n/a	n/a	n/a
57	7500308	6824818H1	SNP00142506	419	791	A	A	G	noncoding	n/a	n/a	n/a	n/a
59	7503839	1233381H1	SNP00105660	163	3212	A	A	G	noncoding	n/a	n/a	n/a	n/a
59	7503839	1903833H1	SNP00051538	159	3011	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7503839	2396122H2	SNP00051538	111	3011	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7503839	2439251H1	SNP00051538	184	3011	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7503839	3525859H1	SNP00051538	114	3011	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7503839	5676172H1	SNP00021725	41	1071	C	C	A	noncoding	n/a	n/a	n/a	n/a
59	7503839	6367830H1	SNP0003533	249	3361	C	C	A	noncoding	n/a	n/a	n/a	n/a
59	7503839	7579662H1	SNP00021725	113	1071	C	C	A	noncoding	n/a	n/a	n/a	n/a
59	7503839	7683343H1	SNP00021725	485	1071	C	C	A	noncoding	n/a	n/a	n/a	n/a
59	7503839	7763355J1	SNP00051538	85	3011	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7503839	8450847J1	SNP00003533	257	3361	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7503839	8454612J1	SNP00003533	255	3361	C	C	T	noncoding	n/a	n/a	n/a	n/a
59	7503839	8454612J1	SNP00003533	255	3361	C	C	T	noncoding	n/a	n/a	n/a	n/a
60	7503698	1336111H1	SNP00151238	61	1312	A	A	G	E297	n/a	n/a	n/a	n/a
60	7503698	1904303H1	SNP00151238	115	1312	A	A	G	E297	n/a	n/a	n/a	n/a
60	7503698	3417040H1	SNP00151239	194	1563	C	C	T	F381	n/a	n/a	n/a	n/a
60	7503698	5943105H1	SNP00151239	45	1561	C	C	T	S380	n/a	n/a	n/a	n/a

WO 02/086069

PCT/US02/12464

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30,
 - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4-5, SEQ ID NO:7-12, SEQ ID NO:14-22, and SEQ ID NO:24-30,
 - c) a polypeptide comprising a naturally occurring amino acid sequence at least 94% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:3,
 - d) a polypeptide comprising a naturally occurring amino acid sequence at least 98% identical to an amino acid sequence of SEQ ID NO:6,
 - e) a polypeptide comprising a naturally occurring amino acid sequence at least 92% identical to an amino acid sequence of SEQ ID NO:13,
 - g) a polypeptide comprising a naturally occurring amino acid sequence at least 93% identical to an amino acid sequence of SEQ ID NO:23,
 - h) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, and
 - i) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-30.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

WO 02/086069

PCT/US02/12464

7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
- 5 9. A method of producing a polypeptide of claim 1, the method comprising:
- a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
- 10 b) recovering the polypeptide so expressed: --
10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-30.
- 15 11. An isolated antibody which specifically binds to a polypeptide of claim 1.
12. An isolated polynucleotide selected from the group consisting of:
- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:31-60,
- 20 b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:32, SEQ ID NO:34-35, and SEQ ID NO:37-60,
- c) a polynucleotide comprising naturally occurring a polynucleotide sequence at least 94% identical to the polynucleotide sequence selected from the group consisting of
- 25 SEQ ID NO:31 and SEQ ID NO:33,
- d) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 98% identical to the polynucleotide sequence of SEQ ID NO:36,
- e) a polynucleotide complementary to a polynucleotide of a),
- f) a polynucleotide complementary to a polynucleotide of b),
- 30 g) a polynucleotide complementary to a polynucleotide of c),
- h) a polynucleotide complementary to a polynucleotide of d), and
- i) an RNA equivalent of a)-h).
13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a
- 35 polynucleotide of claim 12.

WO 02/086069

PCT/US02/12464

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
 - b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.
16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
 - b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-30.
19. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition of claim 17.
20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
 - b) detecting agonist activity in the sample.
21. A composition comprising an agonist compound identified by a method of claim 20 and a

WO 02/086069

PCT/US02/12464

pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

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

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

WO 02/086069

PCT/US02/12464

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

- 5
- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
 - b) detecting altered expression of the target polynucleotide, and
 - c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

10 29. A method of assessing toxicity of a test compound, the method comprising:

- 15
- a) treating a biological sample containing nucleic acids with the test compound,
 - b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,
 - c) quantifying the amount of hybridization complex, and
 - d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.
- 20

30. A diagnostic test for a condition or disease associated with the expression of SECP in a biological sample, the method comprising:

- 25
- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
 - b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
- 30

31. The antibody of claim 11, wherein the antibody is:

- 35
- a) a chimeric antibody,
 - b) a single chain antibody,
 - c) a Fab fragment,
 - d) a F(ab')₂ fragment, or

WO 02/086069

PCT/US02/12464

- e) a humanized antibody.
32. A composition comprising an antibody of claim 11 and an acceptable excipient.
- 5 33. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 32.
34. A composition of claim 32, wherein the antibody is labeled.
- 10 35. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 34.
36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:
- 15 a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-30, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibodies from said animal, and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence
- 20 selected from the group consisting of SEQ ID NO:1-30.
37. A polyclonal antibody produced by a method of claim 36.
- 25 38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.
39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:
- a) immunizing an animal with a polypeptide consisting of an amino acid sequence
- 30 selected from the group consisting of SEQ ID NO:1-30, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- 35 d) culturing the hybridoma cells, and

WO 02/086069

PCT/US02/12464

- e) isolating from the culture monoclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30.
- 5 40. A monoclonal antibody produced by a method of claim 39.
41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.
42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression
10 library.
43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.
- 15 44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30 in a sample, the method comprising:
- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) detecting specific binding, wherein specific binding indicates the presence of a
20 polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30 in the sample.
45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-30 from a sample, the method comprising:
- 25 a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID
30 NO:1-30.
46. A microarray wherein at least one element of the microarray is a polynucleotide of claim
13.
47. A method of generating an expression profile of a sample which contains polynucleotides,
35 the method comprising:

WO 02/086069

PCT/US02/12464

- a) labeling the polynucleotides of the sample,
b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
5 c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target
10 polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

15 50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

20

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

25

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains
30 multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

35 56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

WO 02/086069

PCT/US02/12464

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.
58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.
59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.
60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.
61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.
62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.
64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.
65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

WO 02/086069

PCT/US02/12464

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.
76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.
- 5 77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.
78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.
79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.
- 10 80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.
81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.
- 15 82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.
83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.
84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.
- 20 85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.
86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:31.
- 25 87. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:32.
88. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:33.
89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:34.
- 30 90. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:35.
91. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:36.
- 35 92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:37.

WO 02/086069

PCT/US02/12464

93. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:38.
94. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:39.
- 5 95. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:40.
96. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:41.
97. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:42.
- 10 98. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:43.
99. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:44.
- 15 100. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:45.
101. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:46.
- 20 102. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:47.
103. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:48.
- 25 104. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:49.
- 30 105. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:50.
106. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:51.
- 35

WO 02/086069

PCT/US02/12464

107. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:52.

108. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
5 NO:53.

109. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:54.

10 110. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:55.

111. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:56.

15 112. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:57.

113. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
20 NO:58.

114. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:59.

25 115. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:60.

WO 02/086069

PCT/US02/12464

<110> INCYTE GENOMICS, INC.
KLAMMER, Aaron A.
HAFALIA, April J.A.
DUGGAN, Brendan M.
WARREN, Bridget A.
EMERLING, Brooke M.
TRIBOULEY, Catherine M.
ARVIZU, Chandra S.
HONCHELL, Cynthia D.
NGUYEN, Damien B.
KALLICK, Deborah A.
YUE, Henry
AU-YOUNG, Janice K.
RAMKUMAR, Jayalaxmi
LI, Joana X. -----
THANGAVELU, Kavitha
GIETZEN, Kimberly J.
DING, Li
BAUGHN, Mariah R.
YAO, Monique G.
WALIA, Narinder K.
MASON, Patricia M.
LAL, Preeti G.
GRAUL, Richard C.
REDEY, Roopa
BECHA, Shanya D.
SAPPERSTEIN, Stephanie K.
RICHARDSON, Thomas W.
FRAN, Uyen K.
ELLIOTT, Vicki S.
TANG, Y. Tom
AZIMZAI, Yalda
LU, Yan
XU, Yuming

<120> SECRETED PROTEINS

<130> PF-0949 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/285,207; 60/287,114; 60/288,640; 60/290,516; 60/292,184;
60/343,553; 60/358,279; 60/366,041; 60/357,002

<151> 2001-04-20; 2001-04-27; 2001-05-03; 2001-05-11; 2001-05-18;
2001-21-12; 2002-02-20; 2002-03-19; 2002-02-13

<160> 60

<170> PERL Program

<210> 1
<211> 289
<212> PRT

WO 02/086069

PCT/US02/12464

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1895273CD1

<400> 1

```

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

```

<210> 2

<211> 159

<212> PRT

<213> Homo sapiens

WO 02/086069

PCT/US02/12464

<220>

<221> misc_feature

<223> Incyte ID No: 70072222CD1

<400> 2

```

Met Ser Gln Ala Trp Val Pro Gly Leu Ala Pro Thr Leu Leu Phe
 1      5      10      15
Ser Leu Leu Ala Gly Pro Gln Lys Ile Ala Ala Lys Cys Gly Leu
 20      25      30
Ile Leu Ala Cys Pro Lys Gly Phe Lys Cys Cys Gly Asp Ser Cys
 35      40      45
Cys Gln Glu Asn Glu Leu Phe Pro Gly Pro Val Arg Ile Phe Val
 50      55      60
Ile Ile Phe Leu Val Ile Leu Ser Val Phe Cys Ile Cys Gly Leu
 65      70      75
Ala Lys Cys Phe Cys Arg Asn Cys Arg Glu Pro Glu Pro Asp Thr
 80      85      90
Pro Val Asp Cys Arg Gly Pro Leu Glu Leu Pro Ser Ile Ile Pro
 95      100     105
Pro Glu Arg Val Arg Val Ser Leu Ser Ala Pro Pro Pro Pro Tyr
110     115     120
Ser Glu Val Ile Leu Lys Pro Ser Leu Gly Pro Thr Pro Thr Glu
125     130     135
Pro Pro Pro Pro Tyr Ser Phe Arg Pro Glu Glu Tyr Thr Gly Asp
140     145     150
Gln Arg Gly Ile Asp Asn Pro Ala Phe
155

```

<210> 3

<211> 559

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3559223CD1

<400> 3

```

Met Ala Ala Ser Glu Asp Gly Ser Gly Cys Leu Val Ser Arg Gly
 1      5      10      15
Arg Ser Gln Ser Asp Pro Ser Val Leu Thr Asp Ser Ser Ala Thr
 20      25      30
Ser Ser Ala Asp Ala Gly Glu Asn Pro Asp Glu Met Asp Gln Thr
 35      40      45
Pro Pro Ala Arg Pro Glu Tyr Leu Val Ser Gly Ile Arg Thr Pro
 50      55      60
Pro Val Arg Arg Asn Ser Lys Leu Ala Thr Leu Gly Arg Ile Phe
 65      70      75
Lys Pro Trp Lys Trp Arg Lys Lys Lys Asn Glu Lys Leu Lys Gln
 80      85      90
Thr Thr Ser Ala Leu Glu Lys Lys Met Ala Gly Arg Gln Gly Arg
 95      100     105
Glu Glu Leu Ile Lys Lys Gly Leu Leu Glu Met Met Glu Gln Asp
110     115     120
Ala Glu Ser Lys Thr Cys Asn Pro Asp Gly Gly Pro Arg Ser Val
125     130     135

```

3/61

WO 02/086069

PCT/US02/12464

Gln Ser Glu Pro Pro Thr Pro Lys Ser Glu Thr Leu Thr Ser Glu	140	145	150
Asp Ala Gln Pro Gly Ser Pro Leu Ala Thr Gly Thr Asp Gln Val	155	160	165
Ser Leu Asp Lys Pro Leu Ser Ser Ala Ala His Leu Asp Asp Ala	170	175	180
Ala Lys Met Pro Ser Ala Ser Ser Gly Glu Glu Ala Asp Ala Gly	185	190	195
Ser Leu Leu Pro Thr Thr Asn Glu Leu Ser Gln Ala Leu Ala Gly	200	205	210
Ala Asp Ser Leu Asp Ser Pro Pro Arg Pro Leu Glu Arg Ser Val	215	220	225
Gly Gln Leu Pro Ser Pro Pro Leu Leu Pro Thr Pro Pro Lys	230	235	240
Ala Ser Ser Lys Thr Thr Lys Asn Val Thr Gly Gln Ala Thr Leu	245	250	255
Phe Gln Ala Ser Ser Met Lys Ser Ala Asp Pro Ser Leu Arg Gly	260	265	270
Gln Leu Ser Thr Pro Thr Gly Ser Pro His Leu Thr Thr Val His	275	280	285
Arg Pro Leu Pro Pro Ser Arg Val Ile Glu Glu Leu His Arg Ala	290	295	300
Leu Ala Thr Lys His Arg Gln Asp Ser Phe Gln Gly Arg Glu Ser	305	310	315
Lys Gly Ser Pro Lys Lys Arg Leu Asp Val Arg Leu Ser Arg Thr	320	325	330
Ser Ser Val Glu Arg Gly Lys Glu Arg Glu Glu Ala Trp Ser Phe	335	340	345
Asp Gly Ala Leu Glu Asn Lys Arg Thr Ala Ala Lys Glu Ser Glu	350	355	360
Glu Asn Lys Glu Asn Leu Ile Ile Asn Ser Glu Leu Lys Asp Asp	365	370	375
Leu Leu Leu Tyr Gln Asp Glu Glu Ala Leu Asn Asp Ser Ile Ile	380	385	390
Ser Gly Thr Leu Pro Arg Lys Cys Lys Lys Glu Leu Leu Ala Val	395	400	405
Lys Leu Arg Asn Arg Pro Ser Lys Gln Glu Leu Glu Asp Arg Asn	410	415	420
Ile Phe Pro Arg Arg Thr Asp Glu Glu Arg Gln Glu Ile Arg Gln	425	430	435
Gln Ile Glu Met Lys Leu Ser Lys Arg Leu Ser Gln Arg Pro Ala	440	445	450
Val Glu Glu Leu Glu Arg Arg Asn Ile Leu Lys Gln Arg Asn Asp	455	460	465
Gln Thr Glu Gln Glu Glu Arg Arg Glu Ile Lys Gln Arg Leu Thr	470	475	480
Arg Lys Leu Asn Gln Arg Pro Thr Val Asp Glu Leu Arg Asp Arg	485	490	495
Lys Ile Leu Ile Arg Phe Ser Asp Tyr Val Glu Val Ala Lys Ala	500	505	510
Gln Asp Tyr Asp Arg Arg Ala Asp Lys Pro Trp Thr Arg Leu Ser	515	520	525
Ala Ala Asp Lys Ala Ala Ile Arg Lys Glu Leu Asn Glu Tyr Lys	530	535	540
Ser Asn Glu Met Glu Val His Ala Ser Ser Lys His Leu Thr Arg	545	550	555

WO 02/086069

PCT/US02/12464

Phe His Arg Pro

<210> 4
 <211> 1222
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3441255CD1

```

<400> 4
Met Pro Lys Gly Gly Ala Pro Pro Trp Ile Met Ala Leu Met Phe
 1      5      10      15
Thr Gly His Leu Leu Phe Leu Ala Leu Leu Met Phe Ala Phe Ser
 20     25     30
Thr Phe Glu Glu Ser Val Ser Asn Tyr Ser Glu Trp Ala Val Phe
 35     40     45
Thr Asp Asp Ile Asp Gln Phe Lys Thr Gln Lys Val Gln Asp Phe
 50     55     60
Arg Pro Asn Gln Lys Leu Lys Lys Ser Met Leu His Pro Ser Leu
 65     70     75
Tyr Phe Asp Ala Gly Glu Ile Gln Ala Met Arg Gln Lys Ser Arg
 80     85     90
Ala Ser His Leu His Leu Phe Arg Ala Ile Arg Ser Ala Val Thr
 95    100   105
Val Met Leu Ser Asn Pro Thr Tyr Tyr Leu Pro Pro Pro Lys His
110   115   120
Ala Asp Phe Ala Ala Lys Trp Asn Glu Ile Tyr Gly Asn Asn Leu
125   130   135
Pro Pro Leu Ala Leu Tyr Cys Leu Leu Cys Pro Glu Asp Lys Val
140   145   150
Ala Phe Glu Phe Val Leu Glu Tyr Met Asp Arg Met Val Gly Tyr
155   160   165
Lys Asp Trp Leu Val Glu Asn Ala Pro Gly Asp Glu Val Pro Ile
170   175   180
Val His Ser Leu Thr Gly Phe Ala Thr Ala Phe Asp Phe Leu Tyr
185   190   195
Asn Leu Leu Asp Asn His Arg Arg Gln Lys Tyr Leu Glu Lys Ile
200   205   210
Trp Val Ile Thr Glu Glu Met Tyr Glu Tyr Ser Lys Val Arg Ser
215   220   225
Trp Gly Lys Gln Leu Leu His Asn His Gln Ala Thr Asn Met Ile
230   235   240
Ala Leu Leu Thr Gly Ala Leu Val Thr Gly Val Asp Lys Gly Ser
245   250   255
Lys Ala Asn Ile Trp Lys Gln Ala Val Val Asp Val Met Glu Lys
260   265   270
Thr Met Phe Leu Leu Asn His Ile Val Asp Gly Ser Leu Tyr Glu
275   280   285
Gly Val Ala Tyr Gly Ser Tyr Thr Ala Lys Ser Val Thr Gln Tyr
290   295   300

```

5/61

WO 02/086069

PCT/US02/12464

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

WO 02/086069

PCT/US02/12464

```

Thr Glu His Val Val Ser Ile Val Thr Asp Tyr His Asn Leu Lys
725 730 735
Thr Arg Phe Asn Tyr Leu Gly Phe Gly Gly Phe Ala Ser Val Ala
740 745 750
Asp Gln Gly Gln Ile Thr Arg Phe Gly Leu Gly Thr Gln Ala Ile
755 760 765
Val Lys Pro Val Arg His Asp Arg Ile Ile Phe Pro Phe Gly Phe
770 775 780
Lys Phe Asn Ile Ala Val Gly Leu Ile Leu Cys Ile Ser Leu Val
785 790 795
Ile Leu Thr Phe Gln Trp Arg Phe Tyr Leu Ser Phe Arg Lys Leu
800 805 810
Met Arg Trp Ile Leu Ile Leu Val Ile Ala Leu Trp Phe Ile Glu
815 820 825
Leu Leu Asp Val Trp Ser Thr Cys Ser Gln Pro Ile Cys Ala Lys
830 835 840
Trp Thr Arg Thr Glu Ala Glu Gly Ser Lys Lys Ser Leu Ser Ser
845 850 855
Glu Gly His His Met Asp Leu Pro Asp Val Val Ile Thr Ser Leu
860 865 870
Pro Gly Ser Gly Ala Glu Ile Leu Lys Gln Leu Phe Phe Asn Ser
875 880 885
Ser Asp Phe Leu Tyr Ile Arg Val Pro Thr Ala Tyr Ile Asp Ile
890 895 900
Pro Glu Thr Glu Leu Glu Ile Asp Ser Phe Val Asp Ala Cys Glu
905 910 915
Trp Lys Val Ser Asp Ile Arg Ser Gly His Phe Arg Leu Leu Arg
920 925 930
Gly Trp Leu Gln Ser Leu Val Gln Asp Thr Lys Leu His Leu Gln
935 940 945
Asn Ile His Leu His Glu Pro Asn Arg Gly Lys Leu Ala Gln Tyr
950 955 960
Phe Ala Met Asn Lys Asp Lys Lys Arg Lys Phe Lys Arg Arg Glu
965 970 975
Ser Leu Pro Glu Gln Arg Ser Gln Met Lys Gly Ala Phe Asp Arg
980 985 990
Asp Ala Glu Tyr Ile Arg Ala Leu Arg Arg His Leu Val Tyr Tyr
995 1000 1005
Pro Ser Ala Arg Pro Val Leu Ser Leu Ser Ser Gly Ser Trp Thr
1010 1015 1020
Leu Lys Leu His Phe Phe Gln Glu Val Leu Gly Ala Ser Met Arg
1025 1030 1035
Ala Leu Tyr Ile Val Arg Asp Pro Arg Ala Trp Ile Tyr Ser Met
1040 1045 1050
Leu Tyr Asn Ser Lys Pro Ser Leu Tyr Ser Leu Lys Asn Val Pro
1055 1060 1065
Glu His Leu Ala Lys Leu Phe Lys Ile Glu Gly Gly Lys Gly Lys
1070 1075 1080
Cys Asn Leu Asn Ser Gly Tyr Ala Phe Glu Tyr Glu Pro Leu Arg
1085 1090 1095
Lys Glu Leu Ser Lys Ser Lys Ser Asn Ala Val Ser Leu Leu Ser
1100 1105 1110
His Leu Trp Leu Ala Asn Thr Ala Ala Ala Leu Arg Ile Asn Thr
1115 1120 1125
Asp Leu Leu Pro Thr Ser Tyr Gln Leu Val Lys Phe Glu Asp Ile
1130 1135 1140

```

WO 02/086069

PCT/US02/12464

Val His Phe Pro Gln Lys Thr Thr Glu Arg Ile Phe Ala Phe Leu
 1145 1150 1155
 Gly Ile Pro Leu Ser Pro Ala Ser Leu Asn Gln Ile Leu Phe Ala
 1160 1165 1170
 Thr Ser Thr Asn Leu Phe Tyr Leu Pro Tyr Glu Gly Glu Ile Ser
 1175 1180 1185
 Pro Thr Asn Thr Asn Val Trp Lys Gln Asn Leu Pro Arg Asp Glu
 1190 1195 1200
 Ile Lys Leu Ile Glu Asn Ile Cys Trp Thr Leu Met Asp Arg Leu
 1205 1210 1215
 Gly Tyr Pro Lys Phe Met Asp
 1220

<210> 5
 <211> 696
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1958917CD1

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

WO 02/086069

PCT/US02/12464

Ser Val Val Leu Ile Thr Lys Pro Thr Thr Val Lys Glu Ala Gln
230 235 240
Asn Leu Gln Cys Ala Ala Ser Thr Ala Gln Glu Ser Cys Pro Pro
245 250 255
Lys Pro Pro Arg Ala His Glu Leu Lys Leu Leu Val Arg Asn Ile
260 265 270
His Val Leu Leu Leu Ser Glu Pro Gly Ala Ser Gly His Ile Asn
275 280 285
Ala Val Cys Val Val Gln Leu Asn Asp Pro Val Gln Arg Phe Ser
290 295 300
Ser Thr Leu Thr Lys Asn Thr Pro Asp Leu Met Trp Glu Glu Glu
305 310 315
Phe Thr Phe Glu Leu Asn Ala Lys Ser Lys Glu Leu His Leu Gln
320 325 330
Ile Ser Glu Ala Gly Arg Ser Ser Glu Gly Leu Leu Ala Thr Ala
335 340 345
Thr Val Pro Leu Asp Leu Phe Lys Lys Gln Pro Ser Gly Pro Gln
350 355 360
Ser Phe Thr Leu Thr Ser Gly Ser Ala Cys Gly Ser Ser Val Leu
365 370 375
Gly Ser Val Thr Ala Glu Phe Ser Tyr Met Glu Pro Gly Glu Leu
380 385 390
Lys Ser Trp Pro Ile Pro Pro Pro Val Pro Ala Ala Lys Ile Glu
395 400 405
Lys Asp Arg Thr Val Met Pro Cys Gly Thr Val Val Thr Thr Val
410 415 420
Thr Ala Val Lys Thr Lys Pro Arg Val Asp Val Gly Arg Ala Ser
425 430 435
Pro Leu Ser Ser Asp Ser Pro Val Lys Thr Pro Ile Lys Val Lys
440 445 450
Val Ile Glu Lys Asp Ile Ser Val Gln Ala Ile Ala Cys Arg Ser
455 460 465
Ala Pro Val Ser Lys Thr Leu Ser Ser Ser Asp Thr Glu Leu Leu
470 475 480
Val Leu Asn Gly Ser Asp Pro Val Ala Glu Val Ala Ile Arg Gln
485 490 495
Leu Ser Glu Ser Ser Lys Leu Lys Leu Lys Ser Pro Arg Lys Lys
500 505 510
Ser Thr Ile Ile Ile Ser Gly Ile Ser Lys Thr Ser Leu Ser Gln
515 520 525
Asp His Asp Ala Ala Leu Met Gln Gly Tyr Thr Ala Ser Val Asp
530 535 540
Ser Thr His Gln Glu Asp Ala Pro Ser His Pro Glu Arg Ala Ala
545 550 555
Ala Ser Ala Pro Pro Glu Glu Ala Glu Ser Ala Gln Ala Ser Leu
560 565 570
Ala Pro Lys Pro Gln Glu Asp Glu Leu Asp Ser Trp Asp Leu Glu
575 580 585
Lys Glu Pro Gln Ala Ala Ala Trp Ser Ser Gln Val Leu Leu Asp
590 595 600
Pro Asp Gly Asp Glu Leu Ser Glu Ser Ser Met Ser Val Leu Glu
605 610 615
Pro Gly Thr Ala Lys Lys His Lys Gly Gly Ile Leu Arg Lys Gly
620 625 630
Ala Lys Leu Phe Phe Arg Arg Arg His Gln Gln Lys Asp Pro Gly
635 640 645

WO 02/086069

PCT/US02/12464

Met Ser Gln Ser His Asn Asp Leu Val Phe Leu Glu Gln Pro Glu
 650 655 660
 Gly Ser Arg Arg Lys Gly Ile Thr Leu Thr Arg Ile Leu Asn Lys
 665 670 675
 Lys Leu Leu Ser Arg His Arg Asn Lys Asn Thr Met Asn Gly Ala
 680 685 690
 Pro Val Glu Pro Cys Thr
 695

<210> 6
 <211> 436
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6219465CD1

<400> 6
 Met Gly Arg Ala Arg Pro Gly Gln Arg Gly Pro Pro Ser Pro Gly
 1 5 10 15
 Pro Ala Ala Gln Pro Pro Ala Pro Pro Arg Arg Arg Ala Arg Ser
 20 25 30
 Leu Ala Leu Leu Gly Ala Leu Leu Ala Ala Ala Ala Ala Ala
 35 40 45
 Val Arg Val Cys Ala Arg His Ala Glu Ala Gln Ala Ala Ala Arg
 50 55 60
 Gln Glu Leu Ala Leu Lys Thr Leu Gly Thr Asp Gly Leu Phe Leu
 65 70 75
 Phe Ser Ser Leu Asp Thr Asp Gly Asp Met Tyr Ile Ser Pro Glu
 80 85 90
 Glu Phe Lys Pro Ile Ala Glu Lys Leu Thr Gly Ser Thr Pro Ala
 95 100 105
 Ala Ser Tyr Glu Glu Glu Glu Leu Pro Pro Asp Pro Ser Glu Glu
 110 115 120
 Thr Leu Thr Ile Glu Ala Arg Phe Gln Pro Leu Leu Pro Glu Thr
 125 130 135
 Met Thr Lys Ser Lys Asp Gly Phe Leu Gly Val Ser Arg Leu Ala
 140 145 150
 Leu Ser Gly Leu Arg Asn Trp Thr Ala Ala Ala Ser Pro Ser Ala
 155 160 165
 Val Phe Ala Thr Arg His Phe Gln Pro Phe Leu Pro Pro Pro Gly
 170 175 180
 Gln Glu Leu Gly Glu Pro Trp Trp Ile Ile Pro Ser Glu Leu Ser
 185 190 195
 Met Phe Thr Gly Tyr Leu Ser Asn Asn Arg Phe Tyr Pro Pro Pro
 200 205 210
 Pro Lys Gly Lys Glu Val Ile Ile His Arg Leu Leu Ser Met Phe
 215 220 225
 His Pro Arg Pro Phe Val Lys Thr Arg Phe Ala Pro Gln Gly Ala
 230 235 240
 Val Ala Cys Leu Thr Ala Ile Ser Asp Phe Tyr Tyr Thr Val Met
 245 250 255

WO 02/086069

PCT/US02/12464

```

Phe Arg Ile His Ala Glu Phe Gln Leu Ser Glu Pro Pro Asp Phe
260 265 270
Pro Phe Trp Phe Ser Pro Ala Gln Phe Thr Gly His Ile Ile Leu
275 280 285
Ser Lys Asp Ala Thr His Val Arg Asp Phe Arg Leu Phe Val Pro
290 295 300
Asn His Arg Ser Leu Asn Val Asp Met Glu Trp Leu Tyr Gly Ala
305 310 315
Ser Glu Ser Ser Asn Met Glu Val Asp Ile Gly Tyr Ile Pro Gln
320 325 330
Val Ser Ala Gln Glu Ala Pro Ile Gln Met Glu Leu Glu Ala Thr
335 340 345
Gly Pro Ser Val Pro Ser Val Ile Leu Asp Glu Asp Gly Ser Met
350 355 360
Ile Asp Ser His Leu Pro Ser Gly Glu Pro Leu Gln Phe Val Phe
365 370 375
Glu Glu Ile Lys Trp Gln Gln Glu Leu Ser Trp Glu Glu Ala Ala
380 385 390
Arg Arg Leu Glu Val Ala Met Tyr Pro Phe Lys Lys Val Ser Tyr
395 400 405
Leu Pro Phe Thr Glu Ala Phe Asp Arg Ala Lys Ala Glu Asn Lys
410 415 420
Leu Val His Ser Ile Leu Leu Trp Gly Ala Leu Asp Asp Gln Ser
425 430 435
Cys

```

```

<210> 7
<211> 652
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3576625CD1

```

```

<400> 7
Met Ala Ala Ala Ala Leu Pro Pro Arg Pro Leu Leu Leu Leu Pro
1 5 10 15
Leu Val Leu Leu Leu Ser Gly Arg Pro Thr Arg Ala Asp Ser Lys
20 25 30
Val Phe Gly Asp Leu Asp Gln Val Arg Met Thr Ser Glu Gly Ser
35 40 45
Asp Cys Arg Cys Lys Cys Ile Met Arg Pro Leu Ser Lys Asp Ala
50 55 60
Cys Ser Arg Val Arg Ser Gly Arg Ala Arg Val Glu Asp Phe Tyr
65 70 75
Thr Val Glu Thr Val Ser Ser Gly Thr Asp Cys Arg Cys Ser Cys
80 85 90
Thr Ala Pro Pro Ser Ser Leu Asn Pro Cys Glu Asn Glu Trp Lys
95 100 105
Met Glu Lys Leu Lys Lys Gln Ala Pro Glu Leu Leu Lys Leu Gln
110 115 120

```

WO 02/086069

PCT/US02/12464

Ser Met Val Asp Leu Leu Glu Gly Thr Leu Tyr Ser Met Asp Leu
 125 130 135
 Met Lys Val His Ala Tyr Val His Lys Val Ala Ser Gln Met Asn
 140 145 150
 Thr Leu Glu Glu Ser Ile Lys Ala Asn Leu Ser Arg Glu Asn Glu
 155 160 165
 Val Val Lys Asp Ser Val Arg His Leu Ser Glu Gln Leu Arg His
 170 175 180
 Tyr Glu Asn His Ser Ala Ile Met Leu Gly Ile Lys Lys Glu Leu
 185 190 195
 Ser Arg Leu Gly Leu Gln Leu Leu Gln Lys Asp Ala Ala Ala Ala
 200 205 210
 Pro Ala Thr Pro Ala Thr Gly Thr Gly Ser Lys Ala Gln Asp Thr
 215 220 225
 Ala Arg Gly Lys Gly Lys Asp Ile Ser Lys Tyr Gly Ser Val Gln
 230 235 240
 Lys Ser Phe Ala Asp Arg Gly Leu Pro Lys Pro Pro Lys Glu Lys
 245 250 255
 Leu Leu Gln Val Glu Lys Leu Arg Lys Glu Ser Gly Lys Gly Ser
 260 265 270
 Phe Leu Gln Pro Thr Ala Lys Pro Arg Ala Leu Ala Gln Gln Gln
 275 280 285
 Ala Val Ile Arg Gly Phe Thr Tyr Tyr Lys Ala Gly Lys Gln Glu
 290 295 300
 Val Thr Glu Ala Val Ala Asp Asn Ala Leu Gln Gly Thr Ser Trp
 305 310 315
 Leu Glu Gln Leu Pro Pro Lys Val Glu Gly Arg Ser Asn Ser Ala
 320 325 330
 Glu Pro Asn Ser Ala Glu Gln Asp Glu Ala Glu Pro Arg Ser Ser
 335 340 345
 Glu Arg Val Asp Leu Ala Ser Gly Thr Pro Thr Ser Ile Pro Ala
 350 355 360
 Thr Thr Thr Thr Ala Thr Thr Thr Pro Thr Pro Thr Thr Ser Leu
 365 370 375
 Leu Pro Thr Glu Pro Pro Ser Gly Pro Glu Val Ser Ser Gln Gly
 380 385 390
 Arg Glu Ala Ser Cys Glu Gly Thr Leu Arg Ala Val Asp Pro Pro
 395 400 405
 Val Arg His His Ser Tyr Gly Arg His Glu Gly Ala Trp Met Lys
 410 415 420
 Asp Pro Ala Ala Arg Asp Asp Arg Ile Tyr Val Thr Asn Tyr Tyr
 425 430 435
 Tyr Gly Asn Ser Leu Val Glu Phe Arg Asn Leu Glu Asn Phe Lys
 440 445 450
 Gln Gly Arg Trp Ser Asn Met Tyr Lys Leu Pro Tyr Asn Trp Ile
 455 460 465
 Gly Thr Gly His Val Val Tyr Gln Gly Ala Phe Tyr Tyr Asn Arg
 470 475 480
 Ala Phe Thr Lys Asn Ile Ile Lys Tyr Asp Leu Arg Gln Arg Phe
 485 490 495
 Val Ala Ser Trp Ala Leu Leu Pro Asp Val Val Tyr Glu Asp Thr
 500 505 510
 Thr Pro Trp Lys Trp Arg Gly His Ser Asp Ile Asp Phe Ala Val
 515 520 525
 Asp Glu Ser Gly Leu Trp Val Ile Tyr Pro Ala Val Asp Asp Arg
 530 535 540

WO 02/086069

PCT/US02/12464

```

Asp Glu Ala Gln Pro Glu Val Ile Val Leu Ser Arg Leu Asp Pro
545 550 555
Gly Asp Leu Ser Val His Arg Glu Thr Thr Trp Lys Thr Arg Leu
560 565 570
Arg Arg Asn Ser Tyr Gly Asn Cys Phe Leu Val Cys Gly Ile Leu
575 580 585
Tyr Ala Val Asp Thr Tyr Asn Gln Gln Glu Gly Gln Val Ala Tyr
590 595 600
Ala Phe Asp Thr His Thr Gly Thr Asp Ala Arg Pro Gln Leu Pro
605 610 615
Phe Leu Asn Glu His Ala Tyr Thr Thr Gln Ile Asp Tyr Asn Pro
620 625 630
Lys Glu Arg Val Leu Tyr Ala Trp Asp Asn Gly His Gln Leu Thr
635 640 645
-----
Tyr Thr Leu His Phe Val Val -----
650

```

```

<210> 8
<211> 91
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4765758CD1

```

```

<400> 8
Met Trp Lys Leu Arg Glu Arg Leu Ala Asp Leu Pro Lys Val Thr
1 5 10 15
Val Ile Tyr Gln Met Gln Thr Gly Phe Glu Pro Arg Thr Leu Val
20 25 30
Leu Leu Ser Val Val Tyr Ala His Pro Ser Trp Ala Arg Gly Asp
35 40 45
His Arg Ala Ser Val His Arg His Lys Thr Arg Ala Gln Phe Pro
50 55 60
Asp Leu Gly Tyr Lys Asp Asp Ala Tyr Thr Gly Lys Thr Phe Lys
65 70 75
Ile Ser Phe Lys Asn Lys Ser Pro Arg Asp Ser Ala Trp Leu Cys
80 85 90
Leu

```

```

<210> 9
<211> 155
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7236661CD1

```

WO 02/086069

PCT/US02/12464

<400> 9

```

Met Trp Ser Gly His Ile Trp Leu Cys Phe Leu Ser His Val Ser
 1           5           10           15
Trp Gly Ser Glu Ala Pro Val Ala Ser Phe Gly Gly Leu Cys Leu
 20           25           30
Cys Pro Leu Val Ile Gln Gly Ser Val Ser Glu Thr Ser Trp His
 35           40           45
Glu Thr Leu Trp Asp Pro His Trp Gly Ser Gly Ser Gly His Arg
 50           55           60
Arg Ile Gln Gln Ile Lys Val Arg Gln Pro Val Leu Leu Pro Leu
 65           70           75
Ser Lys Pro Phe Gln Trp Leu His Phe Ser Asp Leu Gly Leu Thr
 80           85           90
Ser Lys Arg Lys Ser Phe Asp Ala Gly Phe Gln Leu Pro His Pro
 95           100          105
Gly Glu Cys Ala Glu Glu Gly Glu Ala Thr Glu Val Asp Lys Arg
 110          115          120
His Leu Ser Glu Ala Lys Ala Ala Val Arg Leu Ala Arg Lys Ser
 125          130          135
Leu Lys Trp Leu Cys Ala Phe Arg Arg Val Ala Val Pro Trp Pro
 140          145          150
Leu Trp Leu Ala Pro
 155

```

<210> 10

<211> 765

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7714187CD1

<400> 10

```

Met Ile Trp Arg Ser Arg Ala Gly Ala Glu Leu Phe Ser Leu Met
 1           5           10           15
Ala Leu Trp Glu Trp Ile Ala Leu Ser Leu His Cys Trp Val Leu
 20           25           30
Ala Val Ala Ala Val Ser Asp Gln His Ala Thr Ser Pro Phe Asp
 35           40           45
Trp Leu Leu Ser Asp Lys Gly Pro Phe His Arg Ser Gln Glu Tyr
 50           55           60
Thr Asp Phe Val Asp Arg Ser Arg Gln Gly Phe Ser Thr Arg Tyr
 65           70           75
Lys Ile Tyr Arg Glu Phe Gly Arg Trp Lys Val Asn Asn Leu Ala
 80           85           90
Val Glu Arg Arg Asn Phe Leu Gly Ser Pro Leu Pro Leu Ala Pro
 95           100          105
Glu Phe Phe Arg Asn Ile Arg Leu Leu Gly Arg Arg Pro Thr Leu
 110          115          120
Gln Gln Ile Thr Glu Asn Leu Ile Lys Lys Tyr Gly Thr His Phe
 125          130          135
Leu Leu Ser Ala Thr Leu Gly Gly Glu Glu Ser Leu Thr Ile Phe

```

14/61

WO 02/086069

PCT/US02/12464

```

140                               145                               150
Val Asp Lys Arg Lys Leu Ser Lys Arg Ala Glu Gly Ser Asp Ser
155                               160                               165
Thr Thr Asn Ser Ser Ser Val Thr Leu Glu Thr Leu His Gln Leu
170                               175                               180
Ala Ala Ser Tyr Phe Ile Asp Arg Asp Ser Thr Leu Arg Arg Leu
185                               190                               195
His His Ile Gln Ile Ala Ser Thr Ala Ile Lys Val Thr Glu Thr
200                               205                               210
Arg Thr Gly Pro Leu Gly Cys Ser Asn Tyr Asp Asn Leu Asp Ser
215                               220                               225
Val Ser Ser Val Leu Val Gln Ser Pro Glu Asn Lys Ile Gln Leu
230                               235                               240
Gln Gly Leu Gln Val Leu Leu Pro Asp Tyr Leu Gln Glu Arg Phe
245                               250                               255
Val Gln Ala Ala Leu Ser Tyr Ile Ala Cys Asn Ser Glu Gly Glu
260                               265                               270
Phe Ile Cys Lys Glu Asn Asp Cys Trp Cys His Cys Gly Pro Lys
275                               280                               285
Phe Pro Glu Cys Asn Cys Pro Ser Met Asp Ile Gln Ala Met Glu
290                               295                               300
Glu Asn Leu Leu Arg Ile Thr Glu Thr Trp Lys Ala Tyr Asn Ser
305                               310                               315
Asp Phe Glu Glu Ser Asp Glu Phe Lys Leu Phe Met Lys Arg Leu
320                               325                               330
Pro Met Asn Tyr Phe Leu Asn Thr Ser Thr Ile Met His Leu Trp
335                               340                               345
Thr Met Asp Ser Asn Phe Gln Arg Arg Tyr Glu Gln Leu Glu Asn
350                               355                               360
Ser Met Lys Gln Leu Phe Leu Lys Ala Gln Lys Ile Val His Lys
365                               370                               375
Leu Phe Ser Leu Ser Lys Arg Cys His Lys Gln Pro Leu Ile Ser
380                               385                               390
Leu Pro Arg Gln Arg Thr Ser Thr Tyr Trp Leu Thr Arg Ile Gln
395                               400                               405
Ser Phe Leu Tyr Cys Asn Glu Asn Gly Leu Leu Gly Ser Phe Ser
410                               415                               420
Glu Glu Thr His Ser Cys Thr Cys Pro Asn Asp Gln Val Val Cys
425                               430                               435
Thr Ala Phe Leu Pro Cys Thr Val Gly Asp Ala Ser Ala Cys Leu
440                               445                               450
Thr Cys Ala Pro Asp Asn Arg Thr Arg Cys Gly Thr Cys Asn Thr
455                               460                               465
Gly Tyr Met Leu Ser Gln Gly Leu Cys Lys Pro Glu Val Ala Glu
470                               475                               480
Ser Thr Asp His Tyr Ile Gly Phe Glu Thr Asp Leu Gln Asp Leu
485                               490                               495
Glu Met Lys Tyr Leu Leu Gln Lys Thr Asp Arg Arg Ile Glu Val
500                               505                               510
His Ala Ile Phe Ile Ser Asn Asp Met Arg Leu Asn Ser Trp Phe
515                               520                               525
Asp Pro Ser Trp Arg Lys Arg Met Leu Leu Thr Leu Lys Ser Asn
530                               535                               540
Lys Tyr Lys Ser Ser Leu Val His Met Ile Leu Gly Leu Ser Leu
545                               550                               555
Gln Ile Cys Leu Thr Lys Asn Ser Thr Leu Glu Pro Val Leu Ala

```

WO 02/086069

PCT/US02/12464

```

560                               565                               570
Val Tyr Val Asn Pro Phe Gly Gly Ser His Ser Glu Ser Trp Phe
575                               580                               585
Met Pro Val Asn Glu Asn Ser Phe Pro Asp Trp Glu Arg Thr Lys
590                               595                               600
Leu Asp Leu Pro Leu Gln Cys Tyr Asn Trp Thr Leu Thr Leu Gly
605                               610                               615
Asn Lys Trp Lys Thr Phe Phe Glu Thr Val His Ile Tyr Leu Arg
620                               625                               630
Ser Arg Ile Lys Ser Asn Gly Pro Asn Gly Asn Glu Ser Ile Tyr
635                               640                               645
Tyr Glu Pro Leu Glu Phe Ile Asp Pro Ser Arg Asn Leu Gly Tyr
650                               655                               660
Met Lys Ile Asn Asn Ile Gln Val Phe Gly Tyr Ser Met His Phe
665                               670                               675
Asp Pro Glu Ala Ile Arg Asp Leu Ile Leu Gln Leu Asp Tyr Pro
680                               685                               690
Tyr Thr Gln Gly Ser Gln Asp Ser Ala Leu Leu Gln Leu Leu Glu
695                               700                               705
Ile Arg Asp Arg Val Asn Lys Leu Ser Pro Gly Gln Arg Arg
710                               715                               720
Leu Asp Leu Phe Ser Cys Leu Leu Arg His Arg Leu Lys Leu Ser
725                               730                               735
Thr Ser Glu Val Val Arg Ile Gln Ser Ala Leu Gln Ala Phe Asn
740                               745                               750
Ala Lys Leu Pro Asn Thr Met Asp Tyr Asp Thr Thr Lys Leu Cys
755                               760                               765

```

<210> 11

<211> 150

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5136540CD1

<400> 11

```

Met Phe Thr Leu Leu Val Leu Leu Ser Gln Leu Pro Thr Val Thr
1                               5                               10                               15
Leu Gly Phe Pro His Cys Ala Arg Gly Pro Lys Ala Ser Lys His
20                               25                               30
Ala Gly Glu Glu Val Phe Thr Ser Lys Glu Glu Ala Asn Phe Phe
35                               40                               45
Ile His Arg Arg Leu Leu Tyr Asn Arg Phe Asp Leu Glu Leu Phe
50                               55                               60
Thr Pro Gly Asn Leu Glu Arg Glu Cys Asn Glu Glu Leu Cys Asn
65                               70                               75
Tyr Glu Glu Ala Arg Glu Ile Phe Val Asp Glu Asp Lys Thr Ile
80                               85                               90
Ala Phe Trp Gln Glu Tyr Ser Ala Lys Gly Pro Thr Thr Lys Ser
95                               100                              105

```

16/61

WO 02/086069

PCT/US02/12464

Ala Leu Gln Pro Ser Met Lys Gly Gly Gly Thr Leu Pro Pro Ser
 110 115 120
 Phe Ser Glu Asp Leu Arg Arg Leu Pro Cys Leu His Cys Arg Leu
 125 130 135
 Leu Trp Arg Met Gln Asp Tyr Leu Leu Met Asn Arg Gln Trp Arg
 140 145 150

<210> 12
 <211> 685
 <212> PRT
 <213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3277403CD1

<400> 12

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

17/61

WO 02/086069

PCT/US02/12464

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

<210> 13

WO 02/086069

PCT/US02/12464

<211> 126
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1517569CD1

<400> 13
 Met Leu Leu Ser Ser Cys Leu Pro Pro Ala Asn Val Thr Thr Lys
 1 5 10 15
 Ala Ala Thr Pro Pro Leu Val Leu Ser Leu Thr Thr Ala Asp
 20 25 30
 Pro Ala Gly Lys Pro Ala Pro Cys Arg Val Thr Leu Thr Leu Leu
 35 40 45
 Arg Ala Ser Ile Pro Ala Thr Lys Arg Ala Ser Phe Leu Ser Ser
 50 55 60
 Phe Ile Lys Met Phe Phe Glu Glu Leu Glu Tyr Ile Leu Gly Phe
 65 70 75
 Leu Ser Leu Leu Lys Phe His Val His Val Ser Val Tyr Ser Ala
 80 85 90
 Ile Cys His Phe Gln Lys Glu Gly Thr Gly Asn Ser Arg Ser Phe
 95 100 105
 Thr Cys Thr Pro Glu Leu Phe Pro Arg Leu Gln Thr His Leu Arg
 110 115 120
 Ala Glu Gly Gly Ala Gln
 125

<210> 14
 <211> 149
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2415991CD1

<400> 14
 Met Asn Tyr Leu Phe Phe Phe Leu Thr Thr Ser Gly Leu Tyr Cys
 1 5 10 15
 Leu Ser Gly Ser His Gly Ser Asn Val Lys Tyr Ile Val Leu Thr
 20 25 30
 Tyr Phe Asn Cys Ser Trp Ser Leu Thr Ser Pro Gly Phe Arg Asp
 35 40 45
 Val Leu Lys Gly Ser Gln Leu Trp Gln Val Thr Asp Ser Trp Glu
 50 55 60
 Met Glu Arg Thr Lys Glu Tyr Ser Ser Cys Leu Thr Phe Leu Pro
 65 70 75
 Thr Ala Asp Ile Val Gln Ala Arg Val Met Glu Glu Leu Asn Leu
 80 85 90
 Leu Ala Ser Gln Ala Ala Pro Ile Pro Thr Ser Gln Cys Thr Ala
 95 100 105
 Pro Pro His Leu Phe Ser Pro Leu Ser Leu Thr Ser Pro Phe Ile
 110 115 120
 Met Ser His Lys Ser Gly Thr Val Gly Ser His Tyr Asn Leu Leu
 125 130 135

19/61

WO 02/086069

PCT/US02/12464

Cys His Arg Asp Ser Ile Phe Leu Ile Ser Asn His Val Ser
 140 145

<210> 15
 <211> 114
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2735742CD1

<400> 15
 Met Leu Ser Gly Ala Arg Cys Arg Leu Ala Ser Ala Leu Arg Gly
 1 5 10 15
 Thr Arg Ala Pro Pro Ser Ala Val Ala Arg Arg Cys Leu His Ala
 20 25 30
 Ser Gly Ser Arg Pro Leu Ala Asp Arg Gly Lys Lys Thr Glu Glu
 35 40 45
 Pro Pro Arg Asp Phe Asp Pro Ala Leu Leu Glu Phe Leu Val Cys
 50 55 60
 Pro Leu Ser Lys Lys Pro Leu Arg Tyr Glu Ala Ser Thr Asn Glu
 65 70 75
 Leu Ile Asn Glu Glu Leu Gly Ile Ala Tyr Pro Ile Ile Asp Gly
 80 85 90
 Ile Pro Asn Met Ile Pro Gln Ala Ala Arg Met Thr Arg Gln Ser
 95 100 105
 Lys Lys Gln Glu Glu Val Glu Gln Arg
 110

<210> 16
 <211> 519
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2768535CD1

<400> 16
 Met Thr Cys Pro Asp Lys Pro Gly Gln Leu Ile Asn Trp Phe Ile
 1 5 10 15
 Cys Ser Leu Cys Val Pro Arg Val Arg Lys Leu Trp Ser Ser Arg
 20 25 30
 Arg Pro Arg Thr Arg Arg Asn Leu Leu Leu Gly Thr Ala Cys Ala
 35 40 45
 Ile Tyr Leu Gly Phe Leu Val Ser Gln Val Gly Arg Ala Ser Leu
 50 55 60
 Gln His Gly Gln Ala Ala Glu Lys Gly Pro His Arg Ser Arg Asp
 65 70 75
 Thr Ala Glu Pro Ser Phe Pro Glu Ile Pro Leu Asp Gly Thr Leu
 80 85 90
 Ala Pro Pro Glu Ser Gln Gly Asn Gly Ser Thr Leu Gln Pro Asn
 95 100 105
 Val Val Tyr Ile Thr Leu Arg Ser Lys Arg Ser Lys Pro Ala Asn
 110 115 120

WO 02/086069

PCT/US02/12464

```

Ile Arg Gly Thr Val Lys Pro Lys Arg Arg Lys Lys His Ala Val
125 130 135
Ala Ser Ala Ala Pro Gly Gln Glu Ala Leu Val Gly Pro Ser Leu
140 145 150
Gln Pro Gln Glu Ala Ala Arg Glu Ala Asp Ala Val Ala Pro Gly
155 160 165
Tyr Ala Gln Gly Ala Asn Leu Val Lys Ile Gly Glu Arg Pro Trp
170 175 180
Arg Leu Val Arg Gly Pro Gly Val Arg Ala Gly Gly Pro Asp Phe
185 190 195
Leu Gln Pro Ser Ser Arg Glu Ser Asn Ile Arg Ile Tyr Ser Glu
200 205 210
Ser Ala Pro Ser Trp Leu Ser Lys Asp Asp Ile Arg Arg Met Arg
215 220 225
Leu Leu Ala Asp Ser Ala Val Ala Gly Leu Arg Pro Val Ser Ser
230 235 240
Arg Ser Gly Ala Arg Leu Leu Val Leu Glu Gly Gly Ala Pro Gly
245 250 255
Ala Val Leu Arg Cys Gly Pro Ser Pro Cys Gly Leu Leu Lys Gln
260 265 270
Pro Leu Asp Met Ser Glu Val Phe Ala Phe His Leu Asp Arg Ile
275 280 285
Leu Gly Leu Asn Arg Thr Leu Pro Ser Val Ser Arg Lys Ala Glu
290 295 300
Phe Ile Gln Asp Gly Arg Pro Cys Pro Ile Ile Leu Trp Asp Ala
305 310 315
Ser Leu Ser Ser Ala Ser Asn Asp Thr His Ser Ser Val Lys Leu
320 325 330
Thr Trp Gly Thr Tyr Gln Gln Leu Leu Lys Gln Lys Cys Trp Gln
335 340 345
Asn Gly Arg Val Pro Lys Pro Glu Ser Gly Cys Thr Glu Ile His
350 355 360
His His Glu Trp Ser Lys Met Ala Leu Phe Asp Phe Leu Leu Gln
365 370 375
Ile Tyr Asn Arg Leu Asp Thr Asn Cys Cys Gly Phe Arg Pro Arg
380 385 390
Lys Glu Asp Ala Cys Val Gln Asn Gly Leu Arg Pro Lys Cys Asp
395 400 405
Asp Gln Gly Ser Ala Ala Leu Ala His Ile Ile Gln Arg Lys His
410 415 420
Asp Pro Arg His Leu Val Phe Ile Asp Asn Lys Gly Phe Phe Asp
425 430 435
Arg Ser Glu Asp Asn Leu Asn Phe Lys Leu Leu Glu Gly Ile Lys
440 445 450
Glu Phe Pro Ala Ser Ala Val Ser Val Leu Lys Ser Gln His Leu
455 460 465
Arg Gln Lys Leu Leu Gln Ser Leu Phe Leu Asp Lys Val Tyr Trp
470 475 480
Glu Ser Gln Gly Gly Arg Gln Gly Ile Glu Lys Leu Ile Asp Val
485 490 495
Ile Glu His Arg Ala Lys Ile Leu Ile Thr Tyr Ile Asn Ala His
500 505 510
Gly Val Lys Val Leu Pro Met Asn Glu
515

```

<210> 17

21/61

WO 02/086069

PCT/US02/12464

<211> 1164
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6948851CD1

```

<400> 17
Met Ala Leu Phe Pro Ala Phe Ala Gly Leu Ser Glu Ala Pro Asp
1      5      10      15
Gly Gly Ser Ser Arg Lys Glu Leu Asp Trp Leu Ser Asn Pro Ser
20     25     30
Phe Cys Val Gly Ser Ile Thr Ser Leu Ser Gln Gln Thr Glu Ala
35     40     45
-----
Ala Pro Ala His Val Ser Glu Gly Leu Pro Leu Thr Arg Ser His
50     55     60
Leu Lys Ser Glu Ser Ser Asp Glu Ser Asp Thr Asn Lys Lys Leu
65     70     75
Lys Gln Thr Ser Arg Lys Lys Lys Lys Glu Lys Lys Lys Lys Arg
80     85     90
Lys His Gln His His Lys Lys Thr Lys Arg Lys His Gly Pro Ser
95     100    105
Ser Ser Ser Arg Ser Glu Thr Asp Thr Asp Ser Glu Lys Asp Lys
110    115    120
Pro Ser Arg Gly Val Gly Gly Ser Lys Lys Glu Ser Glu Glu Pro
125    130    135
Asn Gln Gly Asn Asn Ala Ala Ala Asp Thr Gly His Arg Phe Val
140    145    150
Trp Leu Glu Asp Ile Gln Ala Val Thr Gly Glu Thr Phe Arg Thr
155    160    165
Asp Lys Lys Pro Asp Pro Ala Asn Trp Glu Tyr Lys Ser Leu Tyr
170    175    180
Arg Gly Asp Ile Ala Arg Tyr Lys Arg Lys Gly Asp Ser Cys Leu
185    190    195
Gly Ile Asn Pro Lys Lys Gln Cys Ile Ser Trp Glu Gly Thr Ser
200    205    210
Thr Glu Lys Lys His Ser Arg Lys Gln Val Glu Arg Tyr Phe Thr
215    220    225
Lys Lys Ser Val Gly Leu Met Asn Ile Asp Gly Val Ala Ile Ser
230    235    240
Ser Lys Thr Glu Pro Pro Ser Ser Glu Pro Ile Ser Phe Ile Pro
245    250    255
Val Lys Asp Leu Glu Asp Ala Ala Pro Val Thr Thr Trp Leu Asn
260    265    270
Pro Leu Gly Ile Tyr Asp Gln Ser Thr Thr His Trp Leu Gln Gly
275    280    285
Gln Gly Pro Pro Glu Gln Glu Ser Lys Gln Pro Asp Ala Gln Pro
290    295    300
Asp Ser Glu Ser Ala Ala Leu Lys Ala Lys Val Glu Glu Phe Asn
305    310    315
Arg Arg Val Arg Glu Asn Pro Arg Asp Thr Gln Leu Trp Met Ala
320    325    330
Phe Val Ala Phe Gln Asp Glu Val Met Lys Ser Pro Gly Leu Tyr
335    340    345
Ala Ile Glu Glu Gly Glu Gln Glu Lys Arg Lys Arg Ser Leu Lys

```

WO 02/086069

PCT/US02/12464

```

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

```

WO 02/086069

PCT/US02/12464

```

770          775          780
Cys Leu Trp Lys Gln Tyr Ala His Leu Glu Trp Leu Leu Gly Asn
785          790          795
Thr Glu Asp Ala Arg Lys Val Phe Asp Thr Ala Leu Gly Met Ala
800          805          810
Gly Ser Arg Glu Leu Lys Asp Ser Asp Leu Cys Glu Leu Ser Leu
815          820          825
Leu Tyr Ala Glu Leu Glu Val Glu Leu Ser Pro Glu Val Arg Arg
830          835          840
Ala Ala Thr Ala Arg Ala Val His Ile Leu Thr Lys Leu Thr Glu
845          850          855
Ser Ser Pro Tyr Gly Pro Tyr Thr Gly Gln Val Leu Ala Val His
860          865          870
Ile Leu Lys Ala Arg Lys Ala Tyr Glu His Ala Leu Gln Asp Cys
875          880          885
Leu Gly Asp Ser Cys Val Ser Asn Pro Ala Pro Thr Asp Ser Cys
890          895          900
Ser Arg Leu Ile Ser Leu Ala Lys Cys Phe Met Leu Phe Gln Tyr
905          910          915
Leu Thr Ile Gly Ile Asp Ala Ala Val Gln Ile Tyr Glu Gln Val
920          925          930
Phe Ala Lys Leu Asn Ser Ser Val Phe Pro Glu Gly Ser Gly Glu
935          940          945
Gly Asp Ser Ala Ser Ser Gln Ser Trp Thr Ser Val Leu Glu Ala
950          955          960
Ile Thr Leu Met His Thr Ser Leu Leu Arg Phe His Met Lys Val
965          970          975
Ser Val Tyr Pro Leu Ala Pro Leu Arg Glu Ala Leu Ser Gln Ala
980          985          990
Leu Lys Leu Tyr Pro Gly Asn Gln Val Leu Trp Arg Ser Tyr Val
995          1000         1005
Gln Ile Gln Asn Lys Ser His Ser Ala Ser Lys Thr Arg Arg Phe
1010         1015         1020
Phe Asp Thr Ile Thr Arg Ser Ala Lys Pro Leu Glu Pro Trp Leu
1025         1030         1035
Phe Ala Ile Glu Ala Glu Lys Leu Arg Lys Arg Leu Val Glu Thr
1040         1045         1050
Val Gln Arg Leu Asp Gly Arg Glu Ile His Ala Thr Ile Pro Glu
1055         1060         1065
Thr Gly Leu Met His Arg Ile Gln Ala Leu Phe Glu Asn Ala Met
1070         1075         1080
Arg Ser Asp Ser Gly Ser Gln Cys Pro Leu Leu Trp Arg Met Tyr
1085         1090         1095
Leu Asn Phe Leu Val Ser Leu Gly Asn Lys Glu Arg Ser Lys Gly
1100         1105         1110
Val Phe Tyr Lys Ala Leu Gln Ser Cys Pro Trp Ala Lys Val Leu
1115         1120         1125
Tyr Leu Asp Ala Val Glu Tyr Phe Pro Asp Glu Met Gln Glu Ile
1130         1135         1140
Leu Asp Leu Met Thr Glu Lys Glu Leu Arg Val Arg Leu Pro Leu
1145         1150         1155
Glu Glu Leu Glu Leu Leu Leu Glu Asp
1160

```

<210> 18
<211> 112

WO 02/086069

PCT/US02/12464

<212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7040722CD1

<400> 18
 Met Met Glu Glu Gly Arg Phe Leu Val Leu Gly Leu Thr Leu Thr
 1 5 10 15
 Leu Ser Gly Glu Gln Thr Leu Gly Asn Leu Val His Tyr Ser Gln
 20 25 30
 Gln Arg Ala Val Pro Val Leu Val Gly Gly Glu Gly Leu Gly Pro
 35 40 45
 Gly Leu Ile Cys Ala Asn Ser Leu Glu Gly Gly Gly Arg Arg Cys
 50 55 60
 Arg Gly Gly Ala Glu Ala Thr Lys Glu Thr Ala Pro Pro Ala Pro
 65 70 75
 Gly Leu Ser Ala Asp Ile Pro Gly Glu Ala Thr Gly Gly Asp Phe
 80 85 90
 Gly Tyr Pro Glu Lys Gln Lys Gly Asn Gly Met Ser Asn Lys Thr
 95 100 105
 Leu Ser Thr Ser Ser Leu Phe
 110

<210> 19
 <211> 170
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6430290CD1

<400> 19
 Met Trp Leu Ser Ile Glu Gly Asn Ser Tyr Ala Asp Cys Gly Leu
 1 5 10 15
 Glu Asn Leu Gly Trp Leu Ser Cys Trp Leu Ser Arg Val Gly Trp
 20 25 30
 Phe Gly Ala Tyr Ser Gly Val Pro Cys Leu Pro Val Cys Leu Leu
 35 40 45
 Met Ile Leu Met Ser Ser Leu Arg Gly Gln Arg Leu Val Pro Gly
 50 55 60
 Ala Gly Pro Ser Cys Thr Leu Arg Phe Arg Trp Asp His Lys Gln
 65 70 75
 Arg Ser Thr Gly Gly Lys His Leu Cys Ala Ser Thr Leu His Ile
 80 85 90
 Asn Lys Leu Asp Gln Asn Met Glu Cys Thr Ser Leu Leu Arg Ile
 95 100 105
 Leu Glu Leu Ser Val Pro Pro Leu Gly Ala Pro Leu Trp Leu Lys
 110 115 120
 Thr Pro Pro Glu Ala Glu Gly Glu Glu Glu Ala Glu Arg Lys Glu
 125 130 135
 Asn Arg Arg Leu Gly Val Gly Ser Ser Arg Ala Glu Pro Ser Arg
 140 145 150
 Pro Ala Phe Leu Asp Ser Ser Gly Gly Ser Gly Asp Val Cys Cys

WO 02/086069

PCT/US02/12464

<212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6393813CD1

<400> 22
 Met Ala Val Val Leu Pro Ala Val Val Glu Glu Leu Leu Ser Glu
 1 5 10 15
 Met Ala Ala Ala Val Gln Glu Ser Ala Arg Ile Pro Asp Glu Tyr
 20 25 30
 Leu Leu Ser Leu Lys Phe Leu Phe Gly Ser Ser Ala Thr Gln Ala
 35 40 45
 Leu Asp Leu Val Asp Arg Gln Ser Ile Thr Leu Ile Ser Ser Pro
 50 55 60
 Ser Gly Arg Arg Val Tyr Gln Val Leu Gly Ser Ser Ser Lys Thr
 65 70 75
 Tyr Thr Cys Leu Ala Ser Cys His Tyr Cys Ser Cys Pro Ala Phe
 80 85 90
 Ala Phe Ser Val Leu Arg Lys Ser Asp Ser Ile Leu Cys Lys His
 95 100 105
 Leu Leu Ala Val Tyr Leu Ser Gln Val Met Arg Thr Cys Gln Gln
 110 115 120
 Leu Ser Val Ser Asp Lys Gln Leu Thr Asp Ile Leu Leu Met Glu
 125 130 135
 Lys Lys Gln Glu Ala
 140

<210> 23
 <211> 478
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5685755CD1

<400> 23
 Met Pro Arg Arg Gly Tyr Ser Lys Pro Gly Ser Trp Gly Ser Phe
 1 5 10 15
 Trp Ala Met Leu Thr Leu Val Gly Leu Val Thr His Ala Ala Gln
 20 25 30
 Arg Ala Asp Val Gly Gly Glu Ala Ala Gly Thr Ser Ile Asn His
 35 40 45
 Ser Gln Ala Val Leu Gln Arg Leu Gln Glu Leu Leu Arg Gln Gly
 50 55 60
 Asn Ala Ser Asp Val Val Leu Arg Val Gln Ala Ala Gly Thr Asp
 65 70 75
 Glu Val Arg Val Phe His Ala His Arg Leu Leu Leu Gly Leu His
 80 85 90
 Ser Glu Leu Phe Leu Glu Leu Leu Ser Asn Gln Ser Glu Ala Val
 95 100 105
 Leu Gln Glu Pro Gln Asp Cys Ala Ala Val Phe Asp Lys Phe Ile
 110 115 120
 Arg Tyr Leu Tyr Cys Gly Glu Leu Thr Val Leu Leu Thr Gln Ala

WO 02/086069

PCT/US02/12464

```

125          130          135
Ile Pro Leu His Arg Leu Ala Thr Lys Tyr Gly Val Ser Ser Leu
140          145          150
Gln Arg Gly Val Ala Asp Tyr Met Arg Ala His Leu Ala Gly Gly
155          160          165
Ala Gly Pro Ala Val Gly Trp Tyr His Tyr Ala Val Gly Thr Gly
170          175          180
Asp Glu Ala Leu Arg Glu Ser Cys Leu Gln Phe Leu Ala Trp Asn
185          190          195
Leu Ser Ala Val Ala Ala Ser Thr Glu Trp Gly Ala Val Ser Pro
200          205          210
Glu Leu Leu Trp Gln Leu Leu Gln Arg Ser Asp Leu Val Leu Gln
215          220          225
Asp Glu Leu Glu Leu Phe His Ala Leu Glu Ala Trp Leu Gly Arg
230          235          240
Ala Arg Pro Pro Pro Ala Val Ala Glu Arg Ala Leu Arg Ala Ile
245          250          255
Arg Tyr Pro Met Ile Pro Pro Ala Gln Leu Phe Gln Leu Gln Ala
260          265          270
Arg Ser Ala Ala Leu Ala Arg His Gly Pro Ala Val Ala Asp Leu
275          280          285
Leu Leu Gln Ala Tyr Gln Phe His Ala Ala Ser Pro Leu His Tyr
290          295          300
Ala Lys Phe Phe Asp Val Asn Gly Ser Ala Phe Leu Pro Arg Asn
305          310          315
Tyr Leu Ala Pro Ala Trp Gly Ala Pro Trp Val Ile Asn Asn Pro
320          325          330
Ala Arg Asp Asp Arg Ser Thr Ser Phe Gln Thr Gln Leu Gly Pro
335          340          345
Ser Gly His Asp Ala Gly Arg Arg Val Thr Trp Asn Val Leu Phe
350          355          360
Ser Pro Arg Trp Leu Pro Val Ser Leu Arg Pro Val Tyr Ala Asp
365          370          375
Ala Ala Gly Thr Ala Leu Pro Ala Ala Arg Pro Glu Asp Gly Arg
380          385          390
Pro Arg Leu Val Val Thr Pro Ala Ser Ser Gly Gly Asp Ala Ala
395          400          405
Gly Val Ser Phe Gln Lys Thr Val Leu Val Gly Ala Arg Gln Gln
410          415          420
Gly Arg Leu Leu Val Arg His Ala Tyr Ser Phe His Gln Ser Ser
425          430          435
Glu Glu Ala Gly Asp Phe Leu Ala His Ala Asp Leu Gln Arg Arg
440          445          450
Asn Ser Glu Tyr Leu Val Glu Asn Ala Leu His Leu His Leu Ile
455          460          465
Val Lys Pro Val Tyr His Thr Leu Ile Arg Thr Pro Lys
470          475

```

<210> 24

<211> 80

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 71728459CD1

WO 02/086069

PCT/US02/12464

<400> 24

```

Met Lys Val Thr Ala Leu Leu Gly Ala Leu Ser Pro Val Phe Ala
 1           5           10          15
Phe Val Ser Val Phe Ile Ser Trp Leu Ala Ser Phe Gly Asp Gln
 20          25          30
Lys Ser Ile Asp Ser Pro Pro Asp Glu Gln Gln Ser Asn Ser Tyr
 35          40          45
Thr Ser Gly Gln Ala Ala Ser Tyr Ser Gln Lys Ala Ile Gly Arg
 50          55          60
Lys Gly Asn Trp Leu Pro Tyr Ser Leu His Asp Glu Ala Ala Leu
 65          70          75
Gly Ser Gly Ser Trp
 80

```

<210> 25

<211> 505

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1904303CD1

<400> 25

```

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

```

29/61

WO 02/086069

PCT/US02/12464

Gly Ser Leu Gln Ile Pro Tyr Thr Thr Glu Asp Leu Ile Lys His
 245 250 255
 Tyr Asn Cys Gly Asp Leu Ser Ser Val Ile Leu Ser His Asp Ser
 260 265 270
 Ser Gln Val Pro Asn Phe Ile Asn Ala Thr Leu Pro Pro Gln Glu
 275 280 285
 Arg Ile Thr Ala Gln Glu Ile Asp Ser Tyr Leu Arg Arg Glu Leu
 290 295 300
 Ile Tyr Lys Arg Asn Glu Arg Ile Gly Lys Arg Val Lys Ala Leu
 305 310 315
 Leu Glu Glu Phe Pro Asp Lys Gly Phe Phe Ala Phe Gly Ala
 320 325 330
 Gly His Phe Met Gly Asn Asn Thr Val Leu Asp Val Leu Arg Arg
 335 340 345
 Glu Gly Tyr Glu Val Glu His Ala Pro Ala Gly Arg Pro Ile His
 350 355 360
 Lys Gly Lys Ser Lys Lys Thr Ser Thr Arg Pro Thr Leu Ser Thr
 365 370 375
 Ile Phe Ala Pro Lys Val Pro Thr Leu Glu Val Pro Ala Pro Glu
 380 385 390
 Ala Val Ser Ser Gly His Ser Thr Leu Pro Pro Leu Val Ser Arg
 395 400 405
 Pro Gly Ser Ala Asp Thr Pro Ser Glu Ala Glu Gln Arg Phe Arg
 410 415 420
 Lys Lys Arg Arg Arg Ser Gln Arg Arg Pro Arg Leu Arg Gln Phe
 425 430 435
 Ser Asp Leu Trp Val Arg Leu Glu Glu Ser Asp Ile Val Pro Gln
 440 445 450
 Leu Gln Val Pro Val Leu Asp Arg His Ile Ser Thr Glu Leu Arg
 455 460 465
 Leu Pro Arg Arg Gly His Ser His His Ser Gln Met Val Ala Ser
 470 475 480
 Ser Ala Cys Leu Ser Leu Trp Thr Pro Val Phe Trp Val Leu Val
 485 490 495
 Leu Ala Phe Gln Thr Glu Thr Pro Leu Leu
 500 505

<210> 26

<211> 321

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2911343CD1

<400> 26

Met Ser Gly Gly Lys Ser Ala Gln Gly Pro Glu Glu Gly Gly Val
 1 5 10 15
 Cys Ile Thr Glu Ala Leu Ile Thr Lys Arg Asn Leu Thr Phe Pro
 20 25 30
 Glu Asp Gly Glu Leu Ser Glu Lys Met Phe His Thr Leu Asp Glu
 35 40 45
 Leu Gln Thr Val Arg Leu Asp Arg Glu Gly Ile Thr Thr Ile Arg
 50 55 60
 Asn Leu Glu Gly Leu Gln Asn Leu His Ser Leu Tyr Leu Gln Gly

30/61

WO 02/086069

PCT/US02/12464

```

65          70          75
Asn Lys Ile Gln Gln Ile Glu Asn Leu Ala Cys Ile Pro Ser Leu
80          85          90
Arg Phe Leu Ser Leu Ala Gly Asn Gln Ile Arg Gln Val Glu Asn
95          100         105
Leu Leu Asp Leu Pro Cys Leu Gln Phe Leu Asp Leu Ser Glu Asn
110         115         120
Leu Ile Glu Thr Leu Lys Leu Asp Glu Phe Pro Gln Ser Leu Leu
125         130         135
Ile Leu Asn Leu Ser Gly Asn Ser Cys Thr Asn Gln Asp Gly Tyr
140         145         150
Arg Glu Leu Val Thr Glu Ala Leu Pro Leu Leu Leu Asp Leu Asp
155         160         165
Gly Gln Pro Val Val Glu Arg Trp Ile Ser Asp Glu Glu Asp Glu
170         175         180
Ala Ser Ser Asp Glu Glu Phe Pro Glu Leu Ser Gly Pro Phe Cys
185         190         195
Ser Glu Arg Gly Phe Leu Lys Glu Leu Glu Gln Glu Leu Ser Arg
200         205         210
His Arg Glu His Arg Gln Gln Thr Ala Leu Thr Glu His Leu Leu
215         220         225
Arg Met Glu Met Gln Pro Thr Leu Thr Asp Leu Pro Leu Leu Pro
230         235         240
Gly Val Pro Met Ala Gly Asp Ser Ser Pro Ser Ala Thr Pro Ala
245         250         255
Gln Gly Glu Glu Thr Val Pro Glu Ala Val Ser Ser Pro Gln Ala
260         265         270
Ser Ser Pro Thr Lys Lys Pro Cys Ser Leu Ile Pro Arg Gly His
275         280         285
Gln Ser Ser Phe Trp Gly Arg Lys Gly Ala Arg Ala Ala Thr Ala
290         295         300
Pro Lys Ala Ser Val Ala Glu Ala Pro Ser Thr Thr Lys Thr Thr
305         310         315
Ala Lys Arg Ser Lys Lys
320

```

```

<210> 27
<211> 60
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7500308CD1

```

```

<400> 27
Met Ser Gln Ala Trp Val Pro Gly Leu Ala Pro Thr Leu Leu Phe
1          5          10          15
Ser Leu Leu Ala Gly Pro Gln Lys Val Ile Leu Lys Pro Ser Leu
20         25         30
Gly Pro Thr Pro Thr Glu Pro Pro Pro Pro Tyr Ser Phe Arg Pro
35         40         45
Glu Glu Tyr Thr Gly Asp Gln Arg Gly Ile Asp Asn Pro Ala Phe
50         55         60

```

WO 02/086069

PCT/US02/12464

<210> 28
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7501098CD1

```

<400> 28
Met Phe Thr Leu Leu Val Leu Leu Ser Gln Leu Pro Thr Val Thr
  1           5           10          15
Leu Gly Phe Pro His Cys Ala Arg Gly Pro Lys Ala Ser Lys His
  20          25          30
----- Ala Gly Glu Glu Asp Leu Ile Trp Ser Ser Ser Leu Pro Ala Thr -----
                35                40                45

```

<210> 29
<211> 43
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503839CD1

```

<400> 29
Met Ala Leu Phe Pro Ala Phe Ala Gly Leu Ser Glu Ala Pro Asp
  1           5           10          15
Gly Gly Ser Ser Arg Lys Val Lys Glu Ile Met Leu Gln Leu Ile
  20          25          30
Leu Asp Ile Ala Leu Phe Gly Leu Arg Thr Phe Arg Leu
                35                40

```

<210> 30
<211> 456
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503698CD1

```

<400> 30
Met Ser Pro Trp Ser Trp Phe Leu Leu Gln Thr Leu Cys Leu Leu
  1           5           10          15
Pro Thr Gly Ala Ala Ser Arg Arg Gly Ala Pro Gly Thr Ala Asn
  20          25          30
Cys Glu Leu Lys Pro Gln Gln Ser Glu Leu Asn Ser Phe Leu Trp
  35          40          45
Thr Ile Lys Arg Asp Pro Pro Ser Tyr Phe Phe Gly Thr Ile His
  50          55          60
Val Pro Tyr Thr Arg Val Trp Asp Phe Ile Pro Asp Asn Ser Lys
  65          70          75
Glu Ala Phe Leu Gln Ser Ser Ile Val Tyr Phe Glu Leu Asp Leu

```

WO 02/086069

PCT/US02/12464

	80		85		90									
Thr	Asp	Pro	Tyr	Thr	Ile	Ser	Ala	Leu	Thr	Ser	Cys	Gln	Met	Leu
	95													105
Pro	Gln	Gly	Glu	Asn	Leu	Gln	Asp	Val	Leu	Pro	Arg	Asp	Ile	Tyr
	110													120
Cys	Arg	Leu	Lys	Arg	His	Leu	Glu	Tyr	Val	Lys	Leu	Met	Met	Pro
	125													135
Leu	Trp	Met	Thr	Pro	Asp	Gln	Arg	Gly	Lys	Gly	Leu	Tyr	Ala	Asp
	140													150
Tyr	Leu	Phe	Asn	Ala	Ile	Ala	Gly	Asn	Trp	Glu	Arg	Lys	Arg	Pro
	155													165
Val	Trp	Val	Met	Leu	Met	Val	Asn	Ser	Leu	Thr	Glu	Val	Asp	Ile
	170													180
Lys	Ser	Arg	Gly	Val	Pro	Val	Leu	Asp	Leu	Phe	Leu	Ala	Gln	Glu
	185													195
Ala	Glu	Arg	Leu	Arg	Lys	Gln	Thr	Gly	Ala	Val	Glu	Lys	Val	Glu
	200													210
Glu	Gln	Cys	His	Pro	Leu	Asn	Gly	Leu	Asn	Phe	Ser	Gln	Val	Pro
	215													225
Asn	Phe	Ile	Asn	Ala	Thr	Leu	Pro	Pro	Gln	Glu	Arg	Ile	Thr	Ala
	230													240
Gln	Glu	Ile	Asp	Ser	Tyr	Leu	Arg	Arg	Glu	Leu	Ile	Tyr	Lys	Arg
	245													255
Asn	Glu	Arg	Ile	Gly	Lys	Arg	Val	Lys	Ala	Leu	Leu	Glu	Glu	Phe
	260													270
Pro	Asp	Lys	Gly	Phe	Phe	Phe	Ala	Phe	Gly	Ala	Gly	His	Phe	Met
	275													285
Gly	Asn	Asn	Thr	Val	Leu	Asp	Val	Leu	Arg	Arg	Glu	Gly	Tyr	Glu
	290													300
Val	Glu	His	Ala	Pro	Ala	Gly	Arg	Pro	Ile	His	Lys	Gly	Lys	Ser
	305													315
Lys	Lys	Thr	Ser	Thr	Arg	Pro	Thr	Leu	Ser	Thr	Ile	Phe	Ala	Pro
	320													330
Lys	Val	Pro	Thr	Leu	Glu	Val	Pro	Ala	Pro	Glu	Ala	Val	Ser	Ser
	335													345
Gly	His	Ser	Thr	Leu	Pro	Pro	Leu	Val	Ser	Arg	Pro	Gly	Ser	Ala
	350													360
Asp	Thr	Pro	Ser	Glu	Ala	Glu	Gln	Arg	Phe	Arg	Lys	Lys	Arg	Arg
	365													375
Arg	Ser	Gln	Arg	Arg	Pro	Arg	Leu	Arg	Gln	Phe	Ser	Asp	Leu	Trp
	380													390
Val	Arg	Leu	Glu	Glu	Ser	Asp	Ile	Val	Pro	Gln	Leu	Gln	Val	Pro
	395													405
Val	Leu	Asp	Arg	His	Ile	Ser	Thr	Glu	Leu	Arg	Leu	Pro	Arg	Arg
	410													420
Gly	His	Ser	His	His	Ser	Gln	Met	Val	Ala	Ser	Ser	Ala	Cys	Leu
	425													435
Ser	Leu	Trp	Thr	Pro	Val	Phe	Trp	Val	Leu	Val	Leu	Ala	Phe	Gln
	440													450
Thr	Glu	Thr	Pro	Leu	Leu									
	455													

<210> 31
 <211> 1053
 <212> DNA
 <213> Homo sapiens

WO 02/086069

PCT/US02/12464

<220>

<221> misc_feature

<223> Incyte ID No: 1895273CB1

<400> 31

```

ccagcttggg gacaacatgt gggtttgtac aactetgctc ctttgggttc cagttgatgg 60
gcagtgggc tggctactac tgcaggtctc cagcagagtc ttcccggaag gagaacctct 120
ggccttgagg tgcctatgct ggaagatata gctgggtgac aatgtgcttt actatcgaaa 180
tggcaaaagc ttttaagtttt tccactggaa ttctaacctc accattctga aaaccaacat 240
aagtccaat ggcacctacc atgtctcagg catgggaaag catcgctaca catcagcagg 300
aatatctgtc actgtgaaag agctatttcc agctccagt ctgaatgcat ctgtgacatc 360
cccactcctg gagggaaalc tggtcacctc gagctgtgaa acaaaagttc tcttgagag 420
gctgtgtttg cagctttact tctcctctca catgggcage aagacctgc gaggcaggaa 480
cacatcctct gaataccaaa tactaactgc tagaagagaa gactctgggt tatactggtg 540
cgagctggc acagaggatg gaatgtctc taaccgacc cctgagtttg acctcaagt 600
gcttggctc cagttaccaa ctccctgtct gtttcaatgc cttttctatc tggcagtggt 660
aataatggtt ttagtgaaca ctgttctctg gttgacaata cgttaaaagc tgaagagaaa 720
gaaaagtggt aatttagaaa tctctttgga ttctggatc gagaagaagg taattlccag 780
ccttcaagaa gacagacatt tagaagaaga gctgaaatgt caggaacaaa aagaagaaca 840
gctgcaggaa ggggtgacc ggaaggagcc ccagggggcc acgtagcagc ggctcagttg 900
tggcccatg atctggaccg tcccctgcc acttgtctcc cgtgagcact ggttacaaac 960
atccaaaagt tcaacaaacac cagaactgtg tgtctcatgg tatataactc ttaaaacaaa 1020
taaatgaact gacttcaaaa aaaaaaaaa agg 1053

```

<210> 32

<211> 1579

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7007222CB1

<400> 32

```

agccccctga cgagggtggg agcattagga catgtccaca cgtcttcaag ggaaaatggc 60
taagggtgtg cagagagggc cagggaaagag gcttgggtag ggccgagcag ggaagcagaca 120
atcacttgtt gaagaaagat ccaagttccag gaagaaactt agggcagttt ggtgtcatgg 180
aaggaaactgg gggcccagcc gaggggctca agagcccttc agactctgcc tgagcagagg 240
tggctgtctg tcacccagcc tggagtgtag tgggtgtatc tcaactcact gcagccttga 300
actcctaggc tcaagtgaic ctccctacctc aacctccaga gtaactggga ctacagggaaa 360
gctcagttgg ccccagccca ggtgttccca agcttgggtc cccggctctg cgcaccactt 420
gctgttccgc ctgctggctg gcccccacaa gattgcagcc aaatgtggtc tcaactctgc 480
ctgccccaaa ggaattcaat gctgtggtga cagctgctgc cagggaagag agctcttccc 540
tggccccgtg aggatcttgg tcaatcatct cctggctatc ctgtccgtct tttgcatctg 600
tggcctggct aagtcttctt gtcgcaactg cagagagccg gagccagaca cccagtgga 660
ttgcccgggg cccctggaac tgccttccat calccccca gaggggtca gattatccct 720
ttctgccc ccacccccct acagtgaggt gattotgaag cccagcctgg gcccaactcc 780
cacagagcca cccctccct acagcttcag gcttgaagaa tataccgggg atcagagggg 840
cattgacaac ccggccttct gagtcaacct ctgctggaaa tottgcctc agcaacctcc 900
tcccagctgc ctccctggatc aagctagaga ctgctggcac cccaggaatg tccctgccc 960
tcttgcctg tctctgttca ttcttggatt taacttatta ctttttctgc ttctgtttcc 1020
accocagctg cctctcttct cctgaggttt aggttggagt gacagtttcc gccaccoccc 1080
cagcccaga aagagctgc cggaaagaaa atgctgacca ttggaggtgc ccaacagtag 1140
aatgggtctc tgtgaggggt agtaagagcc ccatttctgg aggtatgcca atcttgactg 1200
gacagccagc tctgagattt tatcagggca cttctatacc tgtgggacat tggactggat 1260
gagccctgag ccagcttcca ctccctacct aatagagaac tcaactgcacc caccacaac 1320

```

WO 02/086069

PCT/US02/12464

```

acatgataaa cacatgtcct cactgaatgt tactgattgc ggctgagggc ctgcoctctg 1380
ctgtgtgggg aggtgggtgg agaggtgagc ccagccactg ctgaggggtg cggtagatgg 1440
gtcgtgctgc ccaatccca ccaactgata gccacctggg agtctggga ggccagtcca 1500
tccatgggccc gccctgggag agagggctgt tctagatgta ttggctgtct gttttttgat 1560
gtctctgtgt gccaaacag 1579

```

<210> 33

<211> 2440

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3559223CB1

```

<400> 33
agcagtcctt gacgacgagg aatttttttt tttttttttt tttttttttt taattttctt 60
ttttaaaaag acgccccctc cagccccctc gccggtgacc ttggccgctt cggatgctct 120
gattccacgc ggctcgtctc aacttgcccc ccgcccggcc gggcccatgg ccgctcggga 180
ggacgggagc ggctcgtctc tctcgggggg ccgctcgcag agtgaccoca ggctcctcac 240
cgactctctg gccacctctc ccgcccggcc ccgggagaac ccagatgaga tggaccaaaac 300
gccccgggag cgtctcgaat atctggctctc agggattcga actccccctg tgaggagaaa 360
cagcaaacctg gccacctctg ccaagatctt caaacctctg aaatggagga aaaagaaaaa 420
cgaaaaactg aagcagacaa cgtcagcgtt ggagaagaag atggccggca ggcaaggccg 480
agaggagctc atcaagaagc ggctcgtgga gatgatggag caggatgctg aaagcaaac 540
ttgcaacccc gatggaggac cccgatctgt acagagtgaa ccaccactc ccaagtccga 600
gacgtgactc tcagaagatg cccagcccgg aagccccttg gccactggga cggaccaggt 660
ctccctggac aagccactgt cctcagctgc ccacttggac gatgcagcca agatgctctc 720
tgcatccagt ggtgaagaag cagacgttgg cagctctctc cccaccacca atgagctctc 780
ccaagcctta gctggggctg actcctctga cagctctctc agactcttg agagatcctg 840
gggcccagctc cccagccccc cactcctgac cactccgcca cccaaggcaa gctccaaaac 900
cacaaaaaat gtcacagccc aagccacact cttccaagcc tccagcatga agagtccga 960
ccctcctctc cggggccagc tctccacacc caccgggtct ccgcatctca ccaagctcca 1020
ccggctcttt ccccccaagc gogtcaattg ggagctgcac agggccttgg ccacgaagca 1080
ccgccaagac agttttcaag gaagagaagc taaaggtctc ccaagagagc ggctgagatg 1140
ccgtctgttg acaacttcca cgttgagagc ggccaaggag agggagagag cttggagctt 1200
tgacgggcca ttggaagaaca agcgaactgc cgttaaggaa tctgaggaga caaaggagaa 1260
cctgatcata aattctgaac tcaaacgca cttgcttttg tatcaggagc aggagggcgt 1320
gaacgactcc attattcttg gaacactgcc ccggaatgac aagaaggagc tccgtggcct 1380
gaagctaaag aaccggccca gcaaacagga actagaagac cggaaacatt tccccagaag 1440
gactgatgaa gaaagcagc agatccggca gacagatcag atgaagctt ccaaacggct 1500
gagcaaaaga cctgcccgtg aagagcttga gagaagaat atcttgaac aaaggaatga 1560
tcagacagag caggagaaga aaagagaat caagcaaga ttgacagaa agcttaatca 1620
gagaccactc gttgatgaat taagagacag aaaaattctg ataccattca gtgattactg 1680
ggaagtagna aaagccagc actatgacag gaggccagac aaaccttga cagactgtct 1740
agcagcagat aagccagcaa ttcytaaaga attaaatgag tacaagaatg atgaaatgga 1800
ggtacatgca tcaagcaagc acttgacaag attccacagg ccatagagat ttctctctga 1860
gaagaatttg tgtttaattt ttgatacca acactgaaca ttcatcagga aacttctctg 1920
aagttcagct caagactacc ctacctctg tgtttgtgag aagagttaga tcacacacac 1980
aggttcaact ttgaccacac ttacctgcaa gaggagtaac cagagacac acttctctcc 2040
ttctttgggt tctgaggagt gtgaactgtt ggggtcagtt aagaccacac ataactctat 2100
cagaagaaaa ctgtgttttg cttttcaacc ttgttttaca gttctgagt gtaatggagg 2160
acgggcaacg tgcattgtca ggctcaacc tcccaggcct ctgacatgag gacatgtga 2220
cagtgctcatt cagattatgt tcaaaaagac atttttatcc tgatcataat taatttgaaa 2280
actctttaag ttcattgtat acaagatgat ttaactgtatt atacttttcc ttttttatat 2340
aatgtctaac aaaaaataca cctgcaacat ttgatctct gtttaatttg tcttttaatt 2400

```

WO 02/086069

PCT/US02/12464

aaatgactac ttattgcagc aaaaaaaaaa aaaaaaaaaa 2440

<210> 34
<211> 4133
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3441255CB1

```

<400> 34
ctctcttggg caaacataca caaatgcac atacatgtgt ggtgagctta tcaccagta 60
tggttttctg tctagaatgt gactcttaac ttgaattttg gactgctttt tctctttttt 120
tacaatgtgt gttccaaactc tttgtgttaa atagatttaa gtaagggggg taatgtctaa 180
atccatagtg ttttttaacct gtaaccactc cctgtgtatt atgnaaaat taggattttt 240
aacgttatcc aaagtllttac tggaaagcaa acbtgtgccag ggacagagat atacaattta 300
agttttctct ttttggcaac tgcacttctt taaaatgtac tgsatgtcac ctggatttca 360
cagcatatca gatttacagt ctttgtctta tcaagccctt tactgtatgt tttatactaa 420
ccagatggga aacacattga gcatcatatc tgacatgtat gcctaagggg ggagctcccc 480
catggatcat ggcgttaaat tttacaggac atttactatt ctttagcatta ttgatgtttg 540
ctttctctac ttttgaggaa tctgtgagca attattccga atgggcagtt ttcacagatg 600
atatagatac gtttaaaaca cagaaagtgc aagatttcag acccaaccaa aagctgaaga 660
aaagtatgct tcatccaagt ttatatattg atgctggaga aatccaagca atgagacaaa 720
agctctgtgc aagccatttg catcttttta gagctatcac aagtgacag acagttatgc 780
tgtccaacc aacatactac ctacctccac caaagcatgc tgattttgct gccaaagtga 840
atgaaattta tggtaacaat ctgctctctt tagcatgtta ctgtttgtta tgcacagaag 900
acaaagtgc ctttgaattt gctctgggat atatggacag gatggttggc tacaagact 960
ggctagtaga gaatgaacca ggagatgagg ttccaaitgt ccattcctta acaggttttg 1020
coactgcttt tgacttttta tataacttat tagataatca tcaagacaa aaatacctg 1080
aaaaaatatg ggttattact gaggaaatgt acgagtatc caaggtccgc tcatggggca 1140
aacagcttct ccataaccac caagccacta atatgatagc attactcaca gggccttgg 1200
tgactggagt agataaagga tcaaaagcaa atatatggaa acaggtctga gtggatgca 1260
tggaaaagac aatgtttota ttgaatcata ttgttgatgg ttctttggat gaaggtgttg 1320
cctatggagc ctacacagct aaatccgtca cacagtatgt tttctggcc cagcccaatt 1380
ttaatatcaa caacttggat aataactggt taaagatgca cttttgttc tatltagca 1440
cccttttaac tggcttccaa agaactgttg gtaatagaca ttccaattat aattgtttt 1500
atgttccaga aagccagcta gttttcttgg abaaatctat cttaagaat ggagctgaa 1560
attgtttagc tcaagcaaat agaaagcacc gacctaaaga tggcccgatg tttcttcaa 1620
ctgcccnaag tggagtaact ctcccaactg aatacctctg gtatgatccc cagctcaac 1680
cacagccacc tgcgtatgat ggtactgcaa aatacacacc attcctaac tggggtgtg 1740
ttacttatgg gctgtgggtg ccaaacacac agaccaaac ctttctgtct ttaaatctg 1800
ggagctggg gggcagact gtgtatgaca tagttcattt tcaagcatat tctgtgattg 1860
atgggtggg aagttttaac ccaggacatg agcatccaga tcagaactca tttaactttg 1920
cccccaatgg acaaglattt gtttctgaag ctctctatgg acccaagtg agccacctta 1980
acaatgtatt ggtgtttgct ccatcaccct caagccagt taataagccc tgggaagtc 2040
aactgggaga atgtgcgcag tggcttaagt ggaactggca ggaggttgt gatgcagct 2100
gggaaataat cactgcctct caactggggg aaatggtatt tgbgagttgg gaagccgtgt 2160
ctgcttattc ttcagcaatg agactgaaaa gtgtatatcg tgctttgctt ctcttaatt 2220
ccaaacctct gctagttggt gatcatattg agagccaaga agattcccca ataaattctg 2280
tcagtgctct ctttcaaat ttggatattg attttaaata tatcccatat aagtttatga 2340
ataggtataa tgggtccatg atggatgtgt gggatgcaca ttacaatat ttttggttg 2400
atcatatgg caatagtccc atggccagta tacaggaagc agagcaagct gctgaattta 2460
aaaaacgatg gactcaattt gttaatgtta ctttcaagat ggaaccacca atcaaaagaa 2520
ttgcataatg cttttatggg ccatatatca atgtctccag ctgcagattt attgatagt 2580
ccaatcctgg acttcaagatt tctctcaatg tcaataatca tgaacatggt tttctatgt 2640

```

WO 02/086069

PCT/US02/12464

```

taactgatta ccataacctg aagacaagat tcaattatct gggattcggg ggccttgcca 2700
gtgtggctga tcaaggccaa ataaccggat ttggtttggg cactcaagca atagtaaagc 2760
ctgtaagaca tgataggatt attttccct ttggatttaa atttaatata gcagtggat 2820
taattttgtg cattagcttg gtgattttaa ctttccaatg gcgtttttac ctttctttta 2880
gaaaactaat gogatggata ttaatacttg ttattgcott gtggtttatt gagcttttg 2940
atgtgtggag cactttagt cagccattt gtgcaaatg gacaaggaca gaggctgagg 3000
gaagcaagaa gtcttctct tctgaaggc accacatga tcttctgat gttgtcatta 3060
ctcactctcc tggttcagga gctgaattc tcaacaact tttttcaac agtagtgatt 3120
ttctctacat cagggttctt acagcctaca ttgatattcc tgaactgag ttggaactg 3180
actcatttgt agatgctgt gaatggaagg tgcagabab ccgcagtgg ctttctggt 3240
tactccggag ctggttgag totttagtc aggacacaaa attacattg caaaaatcc 3300
atctgatga acccaatag ggtaaactg cccaatatt tgcgatgat aaggacaaa 3360
aaagaaaatt taaaaggaga gactctttgc cagaacaaag aagtcaaat aaaggcgct 3420
ttgatagaga tctggaatat attagggctt tgaagagaa cotggtttac tatccaagt 3480
cactctctgt gctcagttta agcagtgga gctggaactt aaagcttcat tttttcagg 3540
aagttttagg agcttctgat agggcattgt acatagtaag agaccctcg gcattgattt 3600
attcaatgtt gtacaatagt aaaccaagtc ttattctttt gaagaaatga ccagagcatt 3660
tagcaaaatt gtttaaaata gagggaggta aaggcaaatg taacttaaat tgggttatg 3720
ctttcgagta tgaaccattg aggaagaat tatcaaaact caaatcaaat gcagtctcc 3780
totgtctca cttgtggcta gcaaatcacg cagcagcctt gagaataaat acagatttg 3840
tgactactag ctaccagctg gtaagtttg aagatattgt goattttctt cagaaaacta 3900
ctgaaaggat ttttgccttt ctggaaatt cttgtctcc tctagttta aaccaaatat 3960
tgtttgccc ctctacaac cttttttacc ttccctatga aggggaaata tccaacata 4020
atactaagt ttggaacag aacttgccta gagatgaat taaactaatt gaaaacatc 4080
gctggactct gatgatcgc ctaggatata caaagtttat ggactaaat aaa 4133

```

```

<210> 35
<211> 4689
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1958917CB1

<220>
<221> unsure
<222> 4672
<223> a, t, c, g, or other

```

```

<400> 35
ggcgggtttc ctcccgttt gcccccag aggcocccg caacgctct cgcgggctc 60
cgaccccgcc ggcgccctc caccgccca ggcycctgt gccgcagca ggcggtgctc 120
tgaagtcgcc gttccggccc ggggctgcc ggaagttag gccgagcgg gccggggatg 180
cgcgatggcc atggcccggc tgggctcgt gctcggggg gccagtgcc tgcgctggt 240
gtcctctctc gtcggccccc tggcaccggt agcctctac ctgcgcagc ggcgctggtc 300
cagggcgcca ccccagccc agcggcgggc ggtggagcct ggaaggggg cgcgccggg 360
gtccgagcgc ctgctctctt ggtctctgac gctggcagc tggaggagcc agtggcaggc 420
ggcctgggtg accgcccga acgaggaggc cgagaggaaa gggggcccac ctttctgtc 480
ctttggggg gcccggcggc agcaggcact gtagctggt gtgcaggagg tctccagctc 540
gctcaggctg gcggaggaga aggtggtggt ctgtcacgtg gtgggcccagg ctattcagtt 600
cttggctcag gagacgctg cttgggtgct tggatccgg ctgtacgaca tgggctctc 660
ccctttccat ttacagctgg agttccacat gaaggagaag agagaggacc tccagattag 720
ctggtctctc atcagctgac cggaaatgac cgttaataac cagcccaaac cactggggga 780
ggaccagggt gctgagacaa gtcgcatgct tgacgttctc aaggacatct tgaagcattt 840
ggctggttct gctctccat cagtgttctt gattaccaag cccacagcty taaaggaagc 900

```

WO 02/08609

PCT/US02/12464

```

tcagaactta cagtgtgctg catctactgc tcaggaatcc tglcctccta aacctccaag 960
ggctcaacag ctgaagctac tggtagggaa catccacgtc ttgctgctca gggagccggg 1020
tgctccaggc cacattaatg cagtgtagct cgtgcagctg aacgatctcg ttcagaggtt 1080
ctccagcacc ctgacgaaaa acactccggg cctcatgttg gaagaggaat tcaacttcga 1140
gttagaatgc aagtcgaagg agttacaact gcagatttca gaggctgggc gatctcaga 1200
aggtctgctg ggcacggoga cagtctctct gacttattt aagaagcagc cttctggccc 1260
acagaccctc acgttgacca gccggctctg ctgcggcagc tgggtgctgg gctcgttcc 1320
ggcagagctc tcttacatgg aaactcgtga atlgaatcc tggcccatcc ctccccctgt 1380
tctgctgca aaaaatgaaa aggaccgcac ggtgatgccc tgtgggactg tggccaactc 1440
tgtcactgct gtagaagcca agcctcgcgt cgaagtgggg agggctgccc cgttgagctc 1500
tgattctccg gtagaagctc ccactcaagg gaaggtgctc gagaaggaca tctctgtcca 1560
ggccactgcc tgcgcgagcg ccccctcag aaaaacactc tcttctcag acacagaatt 1620
gttggtgttg aatggttcgg atccagtggt tgaagtggcc attgcacwgc tcagtgaatc 1680
ttcaaacgtg aaactcaagt cyccacggaa gaaaagcact alalacatct cagggaatc 1740
caaacactca ctactccagg accacagcgc tgcctgatg cagggtcaca gggcctctgt 1800
ggacagcacc caccagggag acgcccacac ccactccggg agggggcgag cctctgccc 1860
ggcagagggaa gccagtgtag cccagggcacc ccttgccccc aagccccagg aggaagagct 1920
agactccctgg gacttgagga aggagccaca ggcggcgga tgggagcagc aggtcctgct 1980
ggaccccagc ggtgatgagc tgcagagag cctccatgat gctctggagc cgggcaactc 2040
caaaaagcat aaaggaggaa tctcaaggaa aggtgcaaaag ctgttcttcc gccggcgcca 2100
tcaacagaaa gaccagggca tgaatcagtc acacaatgac cttgttcttc tggagcagcc 2160
agagggttcc cggaggaag gcactcaacct caccaggtac ctgaacaaag agctgctctc 2220
caggccaaga aacaagaaca ccactgaacgg tgcccccgtg gagccctgca cgtaggccct 2280
gaggtcaca cctccaagcc agaagacgtg caccatggt aactaccctc accaggaacc 2340
agccagtggt tcccgagat gtccagatgc cccgcttctc ttgctggggt tcttccaacc 2400
atctcgtcat ttaaaggaaa aacaaaaatc tagtctccag ccaggaggtc tctccagag 2460
agcaaaaaaa agcccactt gccaccagat gctaaatgag acttgacagc tgcagagctt 2520
ggctgtgctc catagctaag ggttaggggt caatattaga aggagattaa cattataagt 2580
gaataataat gctcctaag atgtggagg gcaggttga gggacttagt ttactaattc 2640
taccctaaca agtgttgttt tgggtccatg cctggaccac gtcaaaaaaa ggggtgccc 2700
ccctgtgctg tcactgtgaa tggaaaggat ggtgaccctc tcttcatctg ctgcttgaaa 2760
taaaaaatgc agctggccct gagtacaggg aagtggaaac taggcaggat ttggattaa 2820
tagcgaatct ttgataagaa tggagagcct acgacagatg taggaagctc tctactttg 2880
atattagcca tagaacttga acactaacta tatcctatgc atagtatgca gaaactttt 2940
ctaaagtctc tttagacctc ctgcaagtg gaagatctat atattctcac atgtgtttta 3000
cattttctcc tatcgtgta gaagctctaa taatgctagt gggcagttg acatccagg 3060
ttttttctcc tgcctgctat acttgctaaa caagagcaca gccggcctgt cagatgaagt 3120
caggagccat acgtgaccgc tctgtagaaca cagtaacca acacatacat ggattttgcc 3180
aagtgtgccc agtagccaaa caaagtctt tttagggcaa tagcgaatct tatttgtgt 3240
ctcaggtgct agtcttagga atggaaagtt aatacaaat agcccaact cgcaggcaat 3300
tccctctatg agcgttccag aatttgctg tgaacagctc tcttggacac aggttgggtt 3360
gccctgtttt gggtttgttt tggtaggaaa catcaaaaac ctggcaacc atttgaatat 3420
cccataatc attccagctg ctttctccat cagttgcctt tctatctcag ttcattaca 3480
gatctcaact ctgaagtgcc caactccagt agacatgctg gtcacaagag agtcaatct 3540
ggggtgaagt gttcttgaca gttaatatg atcaactttt ctccaaagac atgtaaaagg 3600
ctgttagcga agcttgctt ctgtcatgga gacggaaatg ggcaagcttc ctctcgtagc 3660
ctctgtgtaa tcttaaaaa ttaaatattt cgggggtaat agagccactg gtagtaaaaa 3720
acctatataa aaacaaaaat tataggattt tttctttttt agtaaaaaac tgaataaaaa 3780
ccaaaaatag aggatttttt tcttttttag taaaaaccta tataaaaaac aaaaattatg 3840
gattttttct tttttagtaa aaacctatct gaaaacaaa attataggat tttttctt 3900
tttagtaaaa acctgtataa aaacaaaaat tataggattt tttttctc ttttttagag 3960
agagagatta gaaaagcaca ttaggaaatt cacttlaaaa tgcgacttcc aaacttctta 4020
gggttccag gaattatcaa gtgactttaa aatgactttt ccaacctgct ttgttttaa 4080
aaattatatt ccagtttaa tcaattgtaa aaaagcactt ggagttcaa acaatgtgaa 4140
tactaccaag tttctgtccc caaagtcagg catcaactgt agtcttttgg gacagatggg 4200
acagatgttc actttaatgt tttactgaa gtttactgct tctttgcat gtggtaaaaa 4260

```

WO 02/08609

PCT/US02/12464

```

gaggctgaga catatttaag aattccaaga ggatattatg, tgtcagaatt tcagacactg 4320
atgagaagtt ttaattggtt cttttttatt tgattttgga attcaggtgc actctatcca 4380
agtgcaagga tatecagaagt ttttttttat ttaaaaaatt tttttttoga gatggagttt 4440
cactctgttg cccaggtctgg agtcaatgg cagcttactg caacctccac ctctctgttc 4500
aagcgattct cctgcctcag cctcccaagt agctgggatt acaggcacgc gccaccaac 4560
ctggtaatt ctatttagta gaaatggagt ctaccatgt tggtcaggct ggtctcgaac 4620
tctgaacctc aggtgatcca cccacctgg cacttacata aaccgcggtc gnccttgcg 4680
attttatt 4689

```

```

<210> 36
<211> 3294
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 6219465CB1

```

```

<400> 36
atgggcccggg cccggcccggg ccaaccccggg cgcgccagcc cgggcccgc cgcgcagcct 60
cccggcccac cgcgcgcgcg cgcgccttcc ctggcgcgtc tgggagccct gctggcccgc 120
gcccgtgcgc cgcgcgtccg ggtctgtcgc ccgccaagcc agcccagagc gcccgcggcg 180
caggaaactgg cgttgaagac cctggggaca gatggccctt tctcttttc ctctctggac 240
actgacgggg atagtacat cagccctgag gattccaac ccattgtga gaagctaaac 300
gggtcaactc cgcggccag ctacggagag gaggagtgc cccctgacc tagcggagag 360
acgtctacca tagaagccg attccagcct ctgctcccg agaccatgac caagagcaaa 420
gatggcttcc taggggtctc cgcctctgcc ctgctccgac tccgaaactg gacagccgac 480
gcctcaacaa gtgcagtgtt tgcaccccgc cacttccagc ccttctctc cccgcaaggc 540
caggagctgg gtgagccctg gtggatcac cccagttagc tgagcctgtt cactggtcac 600
ctgtccaaca accgcttcta tccaccgccc cccaagggca aggaggtcat catccaaccg 660
ctctgagaca tgttccacc tggcccttt gtgaagacc gctttgccc tcaggagact 720
gtgctctgcc tgaactccat cagcacttc tactacaatg tgatgttccg gatccatgcc 780
gagttccagc tcaagttagc gccgacttc cctctttggt tctccctctc tcaagtccac 840
ggccacatca tctctccaa agacgccacc cagctccgag acttccgctc cttcgtgcc 900
aaccacaggt ctctgaatgt ggacatggag tggccttac gggccagtga aagcagcaac 960
atggaggttg acatgggcta cataccocag gtgagcgcac aggaggtcc catccagatg 1020
gagctggagc ccaagggccc ctctgtgcc tccgtgatcc tggatgagga tggcagcatg 1080
atcgacagcc acctgcccctc aggggagccc ctgcagtttg tgtttgagga gatcaagtgg 1140
cagcaggagc tgaagctggg ggaagctgcc cggcgccttg aggtggccat gtacccttc 1200
aagaaggtct ctaactgccc gttcaactgag gocttcgacc gagccaaggc tgaagcaaac 1260
ctgtgcaact caactctgct gttgggggccc ctggatgacc agtctctgtc aggttcaggc 1320
cggactctcc gggagactgt cctggaaagt tggccatccc tcaacctgct caacgagagc 1380
tcaacagca cctggtccct gytgaaggag ctggaggaac tgcagaaca acaggagzac 1440
tcgtccacc agaagctggc tggcctgcac ctggagaagt acagcttccc cgtggagatg 1500
atgatctgcc tgcacaatgg caaccgtggtc calcaacatca atgccaacta ctctctggac 1560
atcaactccg tgaagccgga ggaatcagag agcaatctct tcaagctctc atccaccttt 1620
gaagaccctg ccacggccac ctacatgca tctctgaagg agggactccg cgttggcctg 1680
ccctctctcc agccctagag tgcctggagc ggcactctgt caccagcccc cagcctctac 1740
agccagagtg gtctctagcc caltctagac tgcagatgct gccactccc accccactcc 1800
taggctgctc tggagggtac aagatccact gagggtggcc accacagcct tggctccatg 1860
gtggcgggta gacaaggyat gcttgggctg actgggcaga ggaacctcta gctctgactg 1920
tcaactggct ctccctacc atttggctct ggaagctgct tggcccccac agatcagggc 1980
ctgggtgaac tccctggacc tttcttagcc agccgcacag tctaggccct tgtgggggta 2040
agaatggagc gaggagcagg ctaggaaagc ggggccacca cctctcctt gctttcagcc 2100
cttcccacag gaaacatcaa gaagcccag ccaggagggg ccaggctgcc aagcggctc 2160
ccctgtttat ctagagcctt cgttctctgc cataccocgg actgcccctc tgtgctctgat 2220

```

WO 02/086069

PCT/US02/12464

```

gtccccagct ggggtoagtc tcaacagag ccagtccttc ggagcctctg gccagaacce 2280
tccatccagag tggaaatcag acgggacccc ctgcagcttc cctgaccacg ccaactgacca 2340
gctatctggg gaagttaact gtgaaggggt ttctgccttt agcaatgggg ttactaagg 2400
gggttcccga gcccaggggc caaggcactc ccaccgccta ccttagcaca gggctctcgc 2460
aggactgggg gagccagcgc tctgcgcgcc cctcttgccc ctccagacct gcattccacag 2520
aagcacaacc cagccaacaa ccacagcctt ctccagagcc ggcactgtcc cggcaaccag 2580
gggtgccccg ggctagctct tctacctctg gggcaccacg gaetcccctt ggccactctt 2640
gggactttgg tccacgtcct gagccactga ccacggccag tctctctttt tatatgtcca 2700
gaaaagtgtt tttaacaaaa ctttctcatg gttttaggtt attttttat aaccoactg 2760
ctgaggagaa agggggggca gtggctctcc cggcagcagc cccatgatgg ctgaatccga 2820
aatctctgat ggttccagct tgaatctctt gcagctgcac ctatgggaag aagtgtctct 2880
ctctctcttc tctctctcag ctttttaaaa acagtctcca gaggatccat gatccccagc 2940
actgtcccac cctccacaaa gcccacaggg catgctgtta ctctctttaa ttaaggtctt 3000
gaagtcaggc tccccctccc ccagcccaca gttctctccc cccccctca ccccacggc 3060
ggctcactca gcttggcaga ggaagaagga aggcagacat ctccgcagcc actcctgggc 3120
cttttatgtg ccagtttacc caacttgcct tggggctgtc cactgagcct tcccagcca 3180
gtctgttctt caattttgtt ttgttttgtt ttgttttggg acgggactct getctgtcac 3240
ccaggctgga gtgatatggc tccatcttgg ctcactgcaa cctccacctc ccag 3294

```

```

<210> 37
<211> 3635
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3576625CB1

```

```

<400> 37
atggccgctg ccgcccctccc gcccccggccg ctgctccttc tgcgcctagt gctgctgctg 60
agcggccgcc ccacgcgcgc gcacagtaag gtgttttggg acctggacca ggtgaggatg 120
acctggaggg gctccgactg ccgtttgcaag tgcatoatgc ggcccctgag caaggacgcg 180
tgtagccgag tgcagctggt cggggccagc gtggaggact tctcaacggt ggagactgtg 240
agctcgggga ctgactgcgc ctgctcctgt accgcacctc cctcctctct caaccctgt 300
gagaacagat ggaagatgga gaaactcaaa aagcagcgc ccagctctct caagctgcag 360
tccatgttgg atctccctgga gggcaccctg tacagcatgg acttgatgaa ggtgcaagcc 420
taogtccaca aggtggcctc ccagatgaac aactggaag agagatcaa ggccaacctg 480
agcccggaga atgaggtggt gaaggcagc gtgcgcaccc tcaatgagca gttaggcaac 540
tatgagaatc actctgccc atgctgtggc atcaagaagg agctgtccgc cctggccctc 600
cagctgtctc agaaggatgc cgcgcgccgc cctgcccacc ctgcccaggg cactgtgtgc 660
aagcccagag acacagctag agnaaaagcc aaggacatca gcaagtatgg cagtgtgca 720
aaaagctttg cagacagagg cctcccacaaa cctcccaggg agaagctgct tcaaggtgag 780
aagctgagaa agggagcggg caagggcagt ttctctccag ccacagccaa gcccgcggcc 840
ctggcccagc agcaggtctg gatccggggc ttaectact acaagcagg caagcaggag 900
gtgaccgagg cgttggcaga caacgcctc cagggcaact cctggtgga gcaactgccc 960
cccaggtggt agggcaggtc caactccgca gagcccaact ccgcagagca ggtgaggtct 1020
gagcccaggt cctccagagc agtggacctg gcttctgga ccccacttc aatccctgcc 1080
accaccacca cgcaccacc caccacaacc cccaccaca gtctctctgc caccgagcca 1140
ccttcaggtc cagaagtctc cagcccaagg agagagggca gctgtgaggg caccctccgg 1200
gctgtgtgacc cccctgtgag gccaccacagc tatgggcgcc acgagggagc ctggatgaa 1260
gaocctgcag ctgagacga caggatctat gtcaccaact actactatgg aaacagcctg 1320
gtggagtccc gcaacctgga aaactcaag caagccctg ggagttaact gtacaagcta 1380
cctacaact ggatcggcac aggccactg gtgtaccagg gcgcttcta ctacaacgc 1440
gccttcacca agaaatcat caagtacgac ctacggcagc gcttctgtgc ctctgggccc 1500
ctgtgcccc acgtggtata tgaggacacc acaacttggg agtggcggg aactcggac 1560
attgactttg ccgtggatga gagcggcctg tgggtcactt accccgcctg ggacgacgc 1620

```

WO 02/086069

PCT/US02/12464

```

gatgaggccc agcccagggt gatcgtccct agtcgcttgg accccggcga tctctcctg 1680
caccggggaga ccaegtggaa gacacggctg cggcggaaact cctacgggaa ctgcttctctg 1740
gtgtgctggca tctgtatgc cgtggacacg tacaaccagc aggaaggcca ggtgcctac 1800
gctttcggaca cgcacacggg caccagcaca cgcceccagc tgccttctct caacgagcac 1860
gcctacacca cccagatcga ctacaacccc aaggagcggg tgcgttacgc ctgggacaat 1920
ggccaccagc tcaactacac cctccacttc gttgtctgag tggagacctg tctcccggg 1980
agaggggcag cagtgcggga ggggctttgc acagcagctc ctgcaactga cccagtcgc 2040
aataatltat tggggccag cccaggctg ggcctgggca tggagtgctt accaggattg 2100
agcttctctca gcccaccagtg ggtataactt gcttccactt gcagagcacc tggccaatga 2160
cttccaccac acttaccctg ttgattctcc tagcacctcc cttggaggta gagatcatga 2220
accattttaa cagacagga gacaggctca gagaggcacc gtccttgcc taacacctca 2280
gtgtgatca ggcaggctgt gctctcagga cagcccactt ttagggatga gggacttca 2340
ccagcccctc cctctcctgc cctcccctat ctcccctggt ctcccctggt ctctcaacce 2400
agtctccctt cccaggggca ctcaagaaca gaggcttcta gggccagtgt actggtgtgg 2460
ggggaggccc ctggctctgc ctggcactct agggccctgt tctggctgag ctgctgtgtg 2520
cctgggctt tgggcacctt agccaatgct ctgtctctt gtctttggcc agcccctca 2580
ggccagccca cccaaccgc tgtccggcca ccttccaaac ctctaccgct acctagctgc 2640
tgagcagaaa ccgcaaccgc agagaaaatc ccatctctct ttccaaggcc cctgtctctg 2700
ctatgctcat tttatlttc tcttattctt catcagtgcc gtcatttctt cctgcagagc 2760
cagcttgagaa ggcagccggc agctctgcca ggggggggag ctgagctgag gctccctctc 2820
caccagaagc acttggcgtt gttcacatag tcagccttgc ggtcccctcc ctggttctac 2880
ccagagcctt tgggctggga gtcgcgcttg tcttlttct ctgggcttcc aaaccaacaa 2940
cttttaccac ctacagggata cctccggctc tgccatgaat aagaccctag gcccaagtct 3000
gggtgggtgc caggatggca cagtttccct ctctcttgc agccctgacc tggctactgg 3060
gcagctggc ccagcagcct gggggctgca gaagacatgg tctgagtagt tgggtccagg 3120
ggaggcatca ggccttcttc tgggtgcag agagaaccag aggtgggaa cggggaggga 3180
agaactgagg gctctgagac tggacttttc ctggctcgac ccaggacttg ggttgaggat 3240
gcacgggggc caccttgcct ggggccaact gtcgctcccc agagcctcag acccaacag 3300
ccagagggac gaggccttca agcctgcccc tcttctctt ttttagacag tattttttag 3360
gctggaaaga aatlttctag cccaactccc tgtttcatga aagagaaaag aggtctagaa 3420
aagtttagag agcagctcag tctcaactg ggcagctggc acagccgat tctctcaga 3480
aggttctctg ctgtatcctt tattctcctc accattccca tcaaccgtga agagaagcct 3540
gctgtcaccg cagcccccaa gaatatagatg ccttctctt aattgcattt tttaaaacaa 3600
geaagtttcc ccaccagtga atgaaagctt tgtga 3635

```

```

<210> 38
<211> 635
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4765758C91

```

```

<400> 38
caacaatgta ggttttaatt ttcacatltt atagatgtgg aaattgagag agaggttagc 60
tgacttacc aaagtcacag taatatatca aatgcagaca ggattgaac ccaggactct 120
agttctggtg tccgtggttt atgctcatcc tctctgggct agaggagacc acagggcctc 180
cgttcacagg cataagacaa gagcccagtt tccagatctg ggttacaagg atgatgcata 240
cactgggaaa acttttaaga tttctttcaa gaacaagagc cctagagatt cggcctggct 300
gtgtctgtaa gtaagaatgt aatggctcta gactgggaaa ttaagttgc taagcaggtc 360
cttatccaca tttaatltca ccttctctca caatgcagac ttttttctca ctgtctagtg 420
attaagtca tggactttaa gtcatatgaa ctgggttggg aatgctgttt cataccctta 480
ctagctcgtt ggtcataggt aagtacttta acctctctga gccctagctt tctcatctgc 540
agggctggta tctcacagga ttgtgaagag tcaagtggga ctgtgttgac atgtttatcc 600
taggcacagg ttgtcaggtc aacttagttg ttcatt 635

```

WO 02/086069

PCT/US02/12464

<210> 39
<211> 793
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7236661CB1

<400> 39
ggcgccggtg gcttggtgtg gtggtttccc ttctggtgce ctactcggga ggttggtgct 60
ttgcccctatg acaggaagga tcaagcgtttc tctctttagc aaactgtgag gctggatgca 120
gagattccaga tactgctttt gtatggtggtg tcggggcata tctggtgtg tttctctcc 180
catgtttcat ggggaagtga agctcccggtg gctctctttg ggggctttg cctctgtccg 240
ctgggtgatac aggggtcggg atcagagacc tcttggcaag aaacctgtg ggaaccacat 300
tgggctctg gctctgggca cagaaggata caacaaatca aagtlggca gctctgctt 360
ctggcattaa gtaaaccttt tcaagtgctg cattttctg atctgggttt aacctcaaa 420
aggaagagct tgaacggcgg gtccacagct cctcatctg tggagtgtc cgaagagga 480
gaagcaacgg agtgggacaa accggaccctg tcaagagcga aagctgctg cagattggcg 540
aggaagctct taaaatggtt gtgtgctttc cggcgggtg ctgtgcatg gctctctg 600
ttggccctct gatggcagcc cgcgccacag taccatgtt ccaatggctc tccatgagg 660
agcacacctc tcatgaccac agctccggga agcagggtct ctgttcctg tcccgagaga 720
cccacggaaa aaattctctt tctctatggc aagaggact cctctccac gggggaggct 780
cacagagctc tgt 793

<210> 40
<211> 2741
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7714187CB1

<400> 40
ctctgatggc tgtttcctca gtcccacaaga gttgaacgg agaggctttt accttgcacc 60
aggtcgagaa cgtgatcagc cctttagaga aggaacctcc ttgtaacagg aattctgctg 120
ggaaacggcg tgtagcact accctccaag ataagatgca tgtttggagc tgtgtaaaat 180
gcccactgac tgtttgaaag aaggaaagg tgaactagggt gtcgaattaa ccttagatcc 240
acttaaaaat attctatcta gtcaaatgta cgttagcaaa gaagagagca ccaggabata 300
aaactgccac agcagcttgg aagacagatg attcagatgt ttgggacttt cttttgctg 360
ttacatagat ttgtttgtca tcatgcaqtt aagcagggtg tgagggaag ctggagaaat 420
gaaggtccta aatcccagc ggaagcatga talggcgaag cagagctggt gctgaattg 480
tctctctgat ggtctctatg gactggatag cactgactct tcaatgctg gttttagcgg 540
ttgctgctgt ttggatcag catgccacaa gccctctga ctggctctc tctgataagg 600
gacctctca tgcctcacag gaatacacag attttgtgga cagaagccgg cagggattta 660
gcacaagata caagatatac agggagtttg gccgctggaa agtaataac cttgcaqttg 720
agagaagaaa ttctcttggc tctctctgac ctcttgcccc tgaattcttc cgaacataa 780
gacttttggg acgtcgacct acccttcagc aaatcacaga aaaccttacc aagaatatg 840
ggaacacatt cttgctatct gctactctgg gaggagagga gtaacctaca atttttgtgg 900
acaagcggaa gttgagcaaa cgaagctgaag gaagtgattc caccaccaat agctcttgg 960
tcaactctga gacgctacat cagctagccg cttcttattt cattgacagg gacagacc 1020
ttcggagact tcaccaactt caaattgcat ccaactgcaat aaaggtaca gaaacocgga 1080
ctggctctct tggctgagc aactatgaca acctagattc tgtcagttct gttctggttc 1140
agagtcctga gaataagatt cagttgcaag ggcctcaagt actctccca gactatcttc 1200
aggaacgttt tgtacaagca gctttgagct acattgcttg caattcagag ggagagtta 1260

WO 02/086069

PCT/US02/12464

```

tctgcaagga aaatgactgc tgggtgctact gtgggtcccaa atttccagaa tgcaactgoc 1320
cctccatgga cattcaagcc atggaagaga atcttctctcg aataactgaa acctgggaaag 1380
cttacaacag tgactttgag gaatcagatg aattcaagtt atttatgaaa aggtcaccta 1440
tgaatttatt cctcaacaca tctactataa tgcatlltgt gacaatggat tctaattttc 1500
agcgcgctta tgaacaactg gagaacagca tgaacaact tttcctaag gcgcagaaaa 1560
ttgtacaaa gctttttagc cttagcaaga ggtgtcaata acaaccctc atcagcctgc 1620
caagacaag aacctcaacc tactggctta ctgcctcca gtctttctc tactgcaatg 1680
agaagcgctc cctaggcagc ttttcagaag agacgcactc gtgcacgtgt ccgaatgacc 1740
agggtgctcg caccgcttc ctgcccctga cagtgggaga cgcctctgc tgctgacat 1800
gpcacacaga caaccgoccc cgtcggcgca cctgcaaac ccgctacatg ctacgccaag 1860
ggctctgcaa gctggaagtc gcgagctcca ccgatoacta tattggcttt gaaactgacc 1920
tgaagatct cagatgaaa tatctgctgc agaaaaggga cagcaagata gaagtccatg 1980
ccatttttat cagcaatgac atgcgctcca atagctgggt tgatccctcc tggcgttaagc 2040
ggatgctctc caacttgaag agcaataagt acaagtcaag tctggctcat atgatttttg 2100
gtctctcttt acagatttc ttaactaaaa acagcaactt ggagccagtg ttgctgtttt 2160
atgcaatcc ctccggagcg agcaactctg agagctggtt tatgctcttg aatgaaaaa 2220
gctllccaga ctggggagcg actaagtctg acctaccctc gcagtgllal aactggacat 2280
taactctggg gaacaaatgg aagacatttt ttgagacagt acacatctac ctgagaagtc 2340
geatacaagtc caatggctcc aatggtaatg agagcattta ctatgaacct ctggagctta 2400
ttgacgcttc ccggaacctg ggtatataga aatcaataa cattedaagt ttggctaca 2460
gcattgactt tgacctgaa gcaattcggg acctgatttt gcagctggac taccctata 2520
ctcagggatc ccaggatcca gcacttttgc aactactaga gatcagagac cgtgtaata 2580
aactctcccc acctggctag cgtcgtctag atctttctc ttgctgtct cgtcalagac 2640
tcaagctctc tactagtgag gtggtgagga tccaactcgc tctgcaggcg tttaatgcca 2700
aattgccaaa cacaatggat tatgacacga ccaaatatg t 2741

```

```

<210> 41
<211> 1074
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5136540CB1

```

```

<400> 41
ggcaacccca ggcctctccc aggtttgccc gcggtggcca tccagaccct gggagagcgc 60
aggcccggag cgtgcgcgag gtttgagggc gccggagacc gagggcctgg cggccgaagg 120
aaccccccga agagagcctc ctggcccggg ggtcgtctga acatgcccgg ggggacacag 180
ttgttttgac agttgocaga ctatgtttac gctctgggtt ctactcagcc aactgcccac 240
agtaccctgt gggtttctct attgcgcaag aggtccaaag gcttctaaag atgcgggaga 300
agaagtgttt acatcaaaag aagaagcaaa ctitttcta catagaagcc ttctgtataa 360
tagatttgat ctggagctct tcactccogg caacctagaa agagagtgca atgagaact 420
ttgcaattat gaggagcca gagagatllt tgtggatgaa gatasaacga ttgcattttg 480
gcaggaalat tcagclaaag gaccaaccac aaaaatcagct ctccagcct clatgaaagg 540
gggaggcaca ctccctccat cattttcaga agacctgagg aggtgcctt gcttccattg 600
ccgcctctgt tggaggatgc agattaccct tcttatgaac aggcagtgcc gctgaccaga 660
aaacacagtg tttaccacc accaccatct cctgggcaca caaaaggatt tagggtatt 720
aaaaaatcta tctctctccc atctcactga ctacctgtc attttggtat aagaaatttg 780
tgttatttga taggcggggc atggtggctc atgcctgtaa tcccagcact ttggggagcc 840
aggagttcga gaccagcctg gccaacatgg tgaacccggg tctctactaa aaatcaaaa 900
atcacctagc cgtcatgggg catgcctgta gtcccaccta cttggaggcg tgaagcaga 960
gaattgctcg aacctgggag gcagaggttg cagtaagctg agatcacgcc actgcattcc 1020
agcctggggc acagagcaag actccatctc aaaaataaaa taaaaaaaaa aggg 1074

```

<210> 42

WO 02/086069

PCT/US02/12464

<211> 2550
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3277403CB1

```

<400> 42
gcggcagctg agcagaggcg gggcgccggg acctgcagtc gccaggattc cctccaggtg 60
acgatgctct ggtttccogg cgtcggggct ctgctgagc gttactccg ccgctcgct 120
gggattacgt gctgctctt gctgctactc aattgctcg ggtccccat gctctggct 180
tcctcttct tgacaggttc tgttgcaaaa tgtgaaaatg aaggtgaagt cctccagatt 240
ccatttatac cagacaacc ttgcataatg tctgtctgt tgaacaagg agtgaatgt 300
aagagagaga agtgcacctg cctgtccoga gactgtgcc tggccatcaa gcaggggga 360
gcctgtctg aacagtgcaa aggttgcacc latgaaaggaa ataccataaa cagctccctc 420
aaatggcaag gcccggtga gccctgtgtt ctacgccagt gccaggagg cgttgcaca 480
gagctctggg tgcctgtgt tgttctattt aaaaacctt tggagcatc ggaatgtgc 540
tgccccacat gctcagggtg tgtgtttgag ggtgtgagt atcagaagg ggaggaattt 600
cagccagaag gaagcaaatg taacaagtgt tctgtcactg gaggcagac acaatgtgtg 660
agagaagtct gtccattct ctctgtccc cagcaactta gtcacatac ccaggagacg 720
tgctgccca aalgtttggg tcagaggaaa gtgtttgacc tccctttgg gagctgcctc 780
tttcgaagtg atgtttatga caatggatcc tcatttctgt acgataactg cacagctgtg 840
acctcaggg actctactgt ggtttgcaag aggaagtgt cccaccctgg tggctgtgac 900
caaggccagg agggctgttg tgaagagtgc ctctacagag tgccccaga agacatcaaa 960
gtatgcaaat ttggcaacaa gatttccag gatggagaga tgtgtctctc tacaattgt 1020
accatctgtg ctgtgtgaa aggcaggacg gactgtccga ataagcagt cattcccacc 1080
agtagctgcc cacagggcaa aattctcaac agaaaaggat gctgtcctat ttgcactgaa 1140
aagcccgggc tttgcaaggt gtttggagat ccccactaca acacttttga cgttcggaca 1200
tttaacttcc aggggaagtg tcagtaagtt ttgacaaaag actgtctctc ccttgcctcg 1260
ccttccaggg tctgtgtgaa gaacgacgcc cgcggacaac gctcctctct gtggaccaag 1320
tcgggtggagc tgggtctggg cgagagcagg gtcagcctgc agcagcact caccgtgctc 1380
tggaaagctc cgcacatcgc gctccccctc cgcgcgccac acttccactc cgaactggat 1440
ggctacctct tgaagtgac caccaaaagc ggtttgaaa tatcttggga tggagacagt 1500
ttttagaaga tcatggctgc cgcgcatctc aaggccaagc tctgtgtctc ttgtggcaac 1560
tacaatggac ataaactga tgaactaatt ggtggagatg gaaactcaaa gttttagatg 1620
gatgactttg ctgaactctg gagggtggag tccaatgagt tctgcaacag acctcagaga 1680
aagccagtg ctgaactctg tcaagggaca gtcgaagtaa agctccggc ccatcagaga 1740
tgcaaaaagc tcaaatcctg gtagtttccg acctgccact cgaactggga ctacgccact 1800
ttctaccgtt cctgtgtgac agacatgtgt gaatgtccag tccataaaaa ctgttattgc 1860
gagtcatttt tggcatalac cgggctctgc cagagagagg gcatcaaat ccaactggag 1920
cctcagaca aattgtcagc caccacagtg aagcatggtg ctgtgtaccg tacctgtggt 1980
ccgggatgta tcaagacctg tgcaaacctg aatgaaattg gtccatgcaa caagccctgc 2040
gttctctggg gccactgtcc agcaaacctg gtccttcaca agggaaagtg calcaagcca 2100
gtcctttgtc cccagcgggt acctttgttt cgtacttcaa gactctgaaa tctgtgact 2160
ttgcaactga agcggaaagc ccaatgaagg actgcaatg ttgtgtccc gattctgtaa 2220
accacacac acagagata tctgtgata tatatataga tatattcaa aacattgcat 2280
catttatatg aactataggg gattattat atgtatattt ttgtctaaa gacatgtatt 2340
gtttctaga tccaaacctg taagccattg aacatgtgtt ataaatacac caggtgtttt 2400
taatttaata aggtggcatg cagatacatt ggaatagttt aacatcacat acattttgca 2460
tttttaagga agttttotaa gagccctcaa ttgcctgctt gtattaattt tagttttgat 2520
cagggttgg gggatccact atttaacgcc gcaccctgtt 2560

```

<210> 43
<211> 5833
<212> DNA

WO 02/086069

PCT/US02/12464

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1517569CB1

<400> 43

```

ctctctcag aaatcgatga gtgccggctc cagccgtgcc tgcattgggg ctcttctcag 60
gaccggctty ctgggtacct gtgccctctgc agcacaggct atgagggcgc ccactgtgag 120
ctggagaggg atgagtgccg agctccaccg tcagaaaatg gaggtcctg caggaaacct 180
caagggccct atgtctgccc gtgccctgca ggtctgtttg gatgccactg tgagacaggt 240
agggctcccc tccagtgggc cccacatgca gagccctggc ctctggaaga tcagagagga 300
ggtggggcaat cccacatctc cctgctgggc aggccactcg gggagaggga ccccagggc 360
tgcggccacc ttggggagga ggggtggagg tggaggtgct ggggagggct gagggcgccg 420
acctgccacc tgccctcttc tccctggcctc cgcctctaga ctcctcttt cctctctcc 480
cgtctctctc cctctctctc cccagactcc ccccttggag ctggggccca clctctgggt 540
gtctctcaaa ggtggaccgc tgggactcca gccctgcca gcatggaggc cgtctgaga 600
gcccgggggg ggcctaccctg tggctctgcc cagagagctt ctctggctac cactggcaga 660
cagtgagtag cccctgcttc tccagccctc gtggggcccg tggctattgc ctggccagca 720
accgtctcca cagtgcaacc tgcaaaagtg gctacacggg cggagactgc gccaaagagc 780
tctctccacc gacggccctc aagatggaga gagtggggga gagtggggtc tctatctct 840
ggaccgcccc caalgtctca gccgccaggg agatgcttga tggctaccgg gtcacctacg 900
tctctctcca cggctcttac cgcgcacagg acttctgtga caggaaccgc tctctgcacc 960
agctccaggg cctggcggcc ggcagggcct acaacatctc cgtctctca ctgaaagcaa 1020
acagttaaca caagaatgac atcagcaggg ctcgctgtct gctggccccc accgagcccc 1080
gcctgtggga aggtcttagg gtcaccaatg tgcaggttag caccatctca gtcagtgagg 1140
cctctccacc gatccggcat gccaccgtca gtcgggtccc tgtgtccact cgcaccctgt 1200
aggcctccag ggaccagggc accgatgtgg acaggagtgt gacagggttc accttaggg 1260
cctctctccc tggaaagagg tacaccatcc agctgaccac cctcagtggg ctccagggag 1320
aggatcacc caccagagag ctggccaccg cgcgcaccga cgtgtggacc cggccctgc 1380
ctccagcaaa cctgaccgcc gcccgagtca ctgccacctc tgcaccctgt gctctgggatg 1440
cccagctacc aggcagcttg ctggaggctt atgtcatcaa tgtgaccacc agccagagca 1500
ccaagagccc ctatgtccc aacgggaagc tggctctcta caccgtgccc gacctgtctc 1560
cgggacggcg gtaaccagctc tctgtgatag cagtgcagag caccggctcc gggccagagc 1620
acagcagggc cgcaccctc tacatcatca cctcccag ggatggctct gacagagct 1680
ggcaccaggg aggaaccac cctcgggtgc tcaagaacag accccccccg cgcgcctgc 1740
cggagctggc cctgctcaat gaccacagcg cccccagac cccccccagc cccccaggt 1800
tctcggagct tgtggacggc agaggaagag tgagcgcag gttcgggtgc taaccagca 1860
aagcagccc cgtgagatca caaccacag cctcggcga gctcagagac atggaggaag 1920
cccccaagg gctcagctcg gccctccagc tccctgaaca cggcagcaag gacctcgaa 1980
acgtccctcg caactgttca gaaaaccctc gtcagaacgg aggcactgtg gtgcccggcg 2040
cagacgccca cagctgtgac tggggcccag ggttcaagg cagacgctgc gagctcgct 2100
gtataaaggt gtcctccccc tgcacaaggg tgttctcga gacaaggccc tttccagct 2160
gggagggagg cgtctgtcac caggtgtata aagagtota ccgagltcac caagacatct 2220
gcttcaaaag gagctgtgaa agcaaacgcc tcaagaagac ccccaacagg aacaagaat 2280
aggtcagac actgggaaaa tcttaaggat ttaagaagct ctgtttacac tcccaacc 2340
tcagcgttt ctaaccacca ggaagatgag gttcaaaac tggatgaaa aggacacct 2400
gagaaaaggt cctagctgga gtcagtcctc tctgtgacct ctctctcag gctctagag 2460
gacagatggc cagcctctg cacacaccag cccaccctga gagaccctc tgggaccac 2520
caoctgtgag cctctgagtg cgtttaaaga cctctgtccc taccacaagc tgcagttct 2580
gaaggtgtag tctgtgtctc tgcggatgag atgacagctc gccatctccc ggaatcagtg 2640
aggtgtctag tcagccaccg tctctcagta tgcagaacc tgttcttaga ctccaaagcc 2700
agagaaagaa ttctcccttc gaggcccaac aaattgagaa ggaactgtga tggaccaact 2760
ccaaaaagaa gacgggggca gggctgaaag ggcagagacc agatgatgtc agaaggaag 2820
cgggttgca gacacagccg cccctgtctc ggtctccag cgtgttatg acgctctgc 2880
aggtcagaca gccatctat ggactagtta acactaaggt ggagttcaga cttttttaga 2940

```

WO 02/086069

PCT/US02/12464

```

caacggcgcc actggcagcc tttctctatc aagggtcaga cggtaaacgt tttcagcttt 3000
gcagaccaga ggtccctgtg getacagtag cgcagacaca gccacaggca tgcattgaa 3060
tggctggggc tatgttccaa taaaaactta tttacaataa cagggtgggg ccaaatggc 3120
ccatgggcoct taittgggta accctgttct atgagatcac ctaggcttca gcctaaaca 3180
gtggaagcca tcccctgaat gacaagtca cagggtatca aagaagacc cctgaatttt 3240
catggaaaaa gctaltcaga cccctgtctg gaaagtaag gcacactgcc acgaagcagc 3300
aaggacgctc tacaagtctc agtgcacac agatggacac ctgggtggg ctggacaatg 3360
tttaaggctc cttttagtcc atgactcaag tgatactgtt ttaggctatc aggtagtaaa 3420
cacgatetta gacatcccca tctttgtaag cagaacagta cggcacttca ccacatctgc 3480
ttcccaccat gcttctaagc agctgttctc cccctgtcaa tggttacaac aaagcagcca 3540
ccccaccacc tctctgttgg agcttcacga ccctgagccc agctggaaag ccagcgcctc 3600
gcgcgctcac cctgactctg ctccagagcca gcaatccagc cacaaagagg gcctccttcc 3660
tttctctctt cataaaaatg ttttttgaag agtttagata ttttttagc tttttatctt 3720
tattaaaatt ccatgtgcat gttctgtgtt attctgcaat ttgtctattt cagaagaag 3780
gaacggcaca ttcgggaagt ttcacactga ctcccagctc gttcccagg ctccagaacc 3840
acttgaagcc aagaagtgga gctcaatgaa gggctctggg ctccagcgaag ggtcactggg 3900
tgaactgaga gaaaccacgt tcacaaacgc gtaactgcga ctctctgccc cctggggacc 3960
tgtctactgt tgcctatgta agctacagca ttaactagcag atgctaaat cagagtatatt 4020
cactggaaaa gtagtgaaat cctactagga aactttctac tccctactag gacctcaagc 4080
ccctagacca cacagcaaat gctaatatgc tccagtgtta gcttagaagc ctgtgttcaa 4140
caagaactgg ctctgtgagtc ccaagcttgg tgccacaag cnaatgctaa tatgctccag 4200
tgttagctta gaagccttgt gtaacaacga actggctcct ggttcccaag ctgtgttcaa 4260
cacagcaaat gctaatatgc tccagtgta gcttagaagc ctgtgttcaa caagaactgg 4320
ctcctgagtc ccaagtgtg tcacaggaact tgcctaltgg gatgtttccc acattaaata 4380
tcagtaaaa agacttctctg gtgctcagga attacagttc gttcttgaaa cattocaaag 4440
agggcccacc agcttttccc atgtggcttc ttttaaaaac tcaaatggct tcttgaaaa 4500
tactcaaatg ccaccaaaag aatttagtaa taatagaact agaaaaactg caggagcata 4560
aagattctg tatcaaaaatg aaagaagcaa tctgtgagtg ctgaattacc catctgctaa 4620
tgaacccggg atggactgat catctaacca agtgagact gaggattcta cttagtctcc 4680
cgactgggta caacaacagc ctaggttcta gggagggtgg cagtgaccgg gatgocacaa 4740
tgggaagaga aatgaaacac ctggcacagt gaaatgtctc atttccaaac agtlttgcct 4800
atggccaagc aaggcaataa agacaaactc tccctttccc ccattgctgt tgggtgtcca 4860
ggtaacaagta agggaattct tctgtgtccc actgtcctcc agtagagaac ccaacagcca 4920
agggcccacc tcaggtatca gctcaactcc tgcacctccc cttagcagga actccttcca 4980
ctggcaaaag actgccactg ccatctgact caactgtggc gatgtggagc gagtcaaccg 5040
gctgttttcc ttttcaaaa caaaagtctt tttcttggca gtcacgctgt aagcagagcc 5100
tgctgggaga acaaaaagca cctagatttc agtctgaaat ccccaaat gcactgagct 5160
cacactacc aggggtatcc cagtgcatag gggaaaggaa cccggctgaa aaacagctc 5220
cttatttttc ttttaataa aataatacga tcctaagctc atttccacc tgaagtgttc 5280
acgagtgaac agtcaacata ctgtgtgga ccaggctta gatgagttc tcaggctcag 5340
cactgacatt ctggcccgga tcaactcttc ttgtgggacc atcctgtgca ctgcaaggtg 5400
tttcacagca cccctggcct ctaccacta gaatctaca atctaccaga tctaggatc 5460
tagttgattc tagaatgcta ccaagagcac ttgaattttg gttccatcat caaaaatgcc 5520
ttctgcatca atcctatgcc agtttccctc aaagagggc taactggaat gttctaggat 5580
gtaccacac tccaggacc tcttagagct catgccaaca ggcagagcc tctatctcag 5640
ggatcacccc ggctgaacac aaatctctct tcttcttccc caacatttca aattgttctt 5700
cggactcaat gagtccctca aggtgacacc ccccccccc cccccaccc acctggggg 5760
gtcagcagat gctccctgtg gccctggctt cagcccaact gctccatgac cccatgctc 5820
tctctctgg ttc 5833

```

```

<210> 44
<211> 2264
<212> DNA
<213> Homo sapiens
<220>

```

WO 02/086069

PCT/US02/12464

<221> misc_feature
<223> Incyte ID No: 2415991CB1

```

<400> 44
gtaagcccca atatacacta cataaaaaacg ttagggtctgc ctgtaatccc agcactttgg 60
gaagctgagg tgggtagatc acttgaggtc aggagttcga gaaccagctg gccaacatgg 120
cgaaacctcg tctctactaa aaatacaaaa ttttagctgg gtgttatggg gggcacctgt 180
aatcccagct actttgaaga tgggttagca gactcacttg aaccagggag ggggaggttg 240
cagtgaggcca agattttgccc actgcactcc agcctgggtg acagagcaag actctgtctc 300
acaaaaaaaa aaaaaaaaaa gggctgtaca taggcagcaa actaagctgc agtgatgttg 360
cctatattta aatthttctca aatggccaag cctctgatgg ctactttatt tgagcaatag 420
ttgagactta taattgccta taataaaca acaaatgaa ctatttgttt tttttctca 480
caacatctgg cctatattgt ctgtcaggaa gccatggctc caatgtaag tacatagttc 540
taacatactt caactcagc tggctccctga cctccaccgg ttocagagat gttcttaaag 600
gaagccagct gtggcaggtc acagattcat gggaaatgga aagaaccaag gaatatagct 660
cttgctccac ctttcaaccd actgctgata tagttcaagc cagagttaat gaagaactta 720
acttaactgc ctctcaggct gctcctatcc ctacctcca gtgtacagcc cctcccctc 780
tctttagctc cctttccctc acttcccctt ttataatgtc acacaatatc gggacagtag 840
gatcacatta taacctactt tgbctatagg attcgalitt tcttatatca aatcagtttt 900
cotgaaaccg agctggggcca tatgcactca atgtctaata catacttatt aatgtaccgg 960
atatggcctt tgcctctgga tatcagcaat ahattataaa aggttccagt agatgagagc 1020
attgtgctgt aatacaattg cagttaattg tgccaataaa gatattgtac tghtacagtc 1080
tttagagttaa agccgcttga atgcagcatg caacttcaat taacagaca atcagggtag 1140
gctagagata accacaaaaa tctatctggc ctactgcag ccaactatg gttagcaaat 1200
ggagagata gtttglygtc cattaattga cctctttca ccaactaga ctagaacttc 1260
tctttcaggt catttattaa atatgattga aatgttataa agttccgaa catgattcat 1320
gatgataaaa atatacaca actgataaaa gaacttaaga actttatata tttcctgttg 1380
cctcaaaaag taacagaaat tattctttaga gctttgattt tagctatcct aattactgca 1440
aataaatttt tgttcttata gttttaaact aaaaagaaaa gtcttgttat aaaacctaa 1500
gcttgaatc atattaaata aatgtattgt acatagtgga aaatttccag tagctaattt 1560
aaaatttcag aaaatgcctat aaagaattt tgattcaagt atttaactg tttagtattg 1620
catgcttctt attaaccgaa aatgataata ccatttagtt tagtgatcag tatgagaagc 1680
aatacctaact cctatgttgc tattgtattt tttcctagtt ggtgtgctg ctcagaaaaa 1740
catatactgt atgtgtatc atacctgtgt atataaaaa ggtcaattta tatatttttc 1800
tatagaaaaa tggagtaca agttccctat ctcccatatt tattttgcca tagtaaaatg 1860
gccacattga tgataatttc tagaactagt ttctgagatt gtcagccctt tgtctaaaat 1920
aatggcagta ttaatgattg acttctgtca ctgccatagt tacctggatt gtcagccttg 1980
gtagcctttg tctaaagtcc taaagagttc caaaaaaat gtgttgaat ttaattgcta 2040
aatagtgctt ggtgattctt bacagtagga atgtataaa ttttcttgca aataagttat 2100
ttactgctat tgatattgaa taatttctct tttattcaga tatatttcaa aaagcaatgaa 2160
tatatgataa ttcataaatt gtatacttta ccagtaagtt ttcagaggaa ataaagactt 2220
ttaaatcctt aaaaaaaaaa aaaaaaaaaa aaaaaaaa aaaa 2264

```

<210> 45
<211> 1638
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2735742CB1

```

<400> 45
tcgagctggc cactattacg gcgcagtggt ctggaaaggg tctgaatcta aagcagaggc 60
agcataacta tcttccagac tectgtttgt tgttttggc ttaaatctag aatccctgt 120
catgatggaa attgactata atacaactgt aatttttaant agcacattaa attattatgt 180

```

WO 02/086069

PCT/US02/12464

```

acctatcact tttgtgctg ctagtgtatc aaaaaggat tgttatacat atactaata 240
tatgacattt gagaacaact gttggatbtt tacttctgta gaataattca gatgagacta 300
aaggccttat aagtctactt attttgcctg ggtgaggcct ggtctccggc tgccagacca 360
tgctgagtg agcagctgc aggcctgcct cagcctgcg ggaacgcgc ggcggcgt 420
ccgctgctgc cgtaggtgc ctgcacgct cgggtctgc gctttggcc gaccgggca 480
agaagactga gtagccgcc cgcgacttc atccggcct gctggatcc ctgggtgcc 540
cgctctccaa gaagccgctc agatattgag catcaacaaa tgaattgatt aatgaagat 600
tgggaatagc ttatccaatc attgatggga tcctataat gataccacag gcagctagga 660
tgacactaag aagtaagaag caagaagaag tggagcagc ctagtccata atttaaaaa 720
attaaaaaaa cgcacagcc aacttttctt aataccata acctttaaa acacagtgc 780
aggtataaag tggaaagaaa gaattgttct gtctctctc acgttgactg ttcttattcc 840
actggtttct ttacaggaac tttctctctc agcctctgt gaagaaaact tcccacagg 900
ctgcactagc acagccagcc tttgctttta cagcctgctc ttgcttatta ccatcaccgt 960
gtatgtattc ttccaccttt ggaattggat ggtattaaa ctctcaggc ataattgatg 1020
caactcggct caatagctg tatatartaa tggatagctc tggcctcga tctctgaaag 1080
ctcaaatgga tggatattag ttggcgggaa agaggctttg ctttcggcat atcaggctta 1140
ggactgtggg aggtttaagt tgcagatgct tctttattg tactctgtt ctgacctgt 1200
ttttgaaag cctgactta taactgctgt atcagaagaa acattttgac agtctctgg 1260
ttggagatga acatccclaa tgcacatgt atgactatt ctattccat tcatctaga 1320
gtcaatgaaa tttgttttg ctgtttgtt tagctcaag gtcttggta aagtcacatg 1380
ttagagatga cgaataaat tcaaaaggag tgatgtgga atagtcctc taaggagag 1440
aaatgacttt gaacgaatgt gatataaac cacataatca aatgaaaact tcatgtact 1500
acaaaaactg agtttgtaaa attaccctca tttcttgac attaaatgct tatattgca 1560
ataaacatgt tgacactttc ctataaaaaa taaccagtt tgcagtmeta aaaaaaaaa 1620
agggggcgg cgcactag 1638

```

- <210> 46
- <211> 4790
- <212> DNA
- <213> Homo sapiens
- <220>
- <221> misc_feature
- <223> Incyte ID No: 2768535CB1
- <220>
- <221> unsure
- <222> 4768
- <223> a, t, c, g, or other

```

<400> 46
gacgtcttcc ttaccttttt ttctgaactt ctaggccttc totttccaga actggtggaa 60
gaacnaatgaa acggccaaga tggtaagaaa caagccgat ttctcttgg ggagactgat 120
aatttaaaag gttgtttgtg tcagaaacat tcccagcttc atcaacaaac ctttctctcc 180
acctctgccc actggagacc acttacatcc cgaagcggac gcggcagctg aagtcaaggaa 240
acatcgcacc acattagcag gagccaatc cagactttaa actccgttca acatgtggat 300
ggccagagaga aatgacctgt ccagacaagc cggggcagct catabaactgg ttcatctgct 360
coctgtgctc ccgcgggtg cgttaagctct ggagcagcgc gcgtccaagg acccggagaa 420
acctctgctc gggcactgag tgtccatctc acttgggctt cctggtgagc caggtgggga 480
gggcctctct ccagcatgga caggcggctg agaaggggoc acatcgagc cgcgacaccg 540
ccgagccatc cttccctgag ataaccctgg atggtaccct gcccctcca gagtcccagg 600
gcaatgggtc cactctgcag cccaatgtg tgtacaltac cctacgctcc aagcgcagca 660
agccggccaa tatccgtggc accgtgaagc ccaagcgcag gaaaaagcat gcaytggcat 720
cgctgcccc agggcaggag gctttgtctg gaccatccct tcagccgagc gaagcgcaa 780
gggaagctga tgcgtagca cctgggtacg ctcaggagac aaacctggtt aagattggag 840
agcgacctg gaggttgggt cggggtccgg gagtgcgagc cgggggccca gacttctgc 900

```

WO 02/086069

PCT/US02/12464

```

agcccagctc cagggagagc aacattagga tctacagcga gagcgcctcc tcttggctga 960
gcaaaagatga catccgaaga atgcgactct tggcggacag cgcagtgcca gggctccggc 1020
ctgtgtcttc taggagcggg gcccgcttgc tggctctgga gggggggcca cctggcgctg 1080
tgctccgctg tggccctagc cctctgtggc tctcaagca gcccttggag atgagtggag 1140
tgtttgcctt ccacctagac aggatcctgg ggctcaacag gcccttggag tctgtgagca 1200
ggaagcaga gttcatccaa gatggccgcc catgccccat cattcttggg gatgcattct 1260
tatcttcagc aagtaatgac acccattctt ctgttaagct caactgggga acttatcagc 1320
agttgtcga acagaaatgc tggcagaatg gccgagtacc caagcctgaa ctggttctga 1380
ctgaaataca tcatcatgag tggccaaga tggcaactct tgattttttg ttacagattt 1440
ataatgcctt agatcaaaat tgctgtggat tcagacctcg caaaggagat gccctgtgac 1500
agaaatgatt gagcccaaaa tgtgatgacc aagttctgc gccctagca cacatbatcc 1560
agcgaagcca tgaccsaagg catttggctt tlatagacaa caaggttttc ttgacagga 1620
gtgaagataa cttaacttc aaattgttag aaggcatcaa agattttcca gcttctgag 1680
ttctgtttt gaagagccag caattagggc agaaactctc tcagctcttg tttctgata 1740
aagtgrattg ggaagtcaa ggaggtagac naggaattga aaagcttatc gatgtaatg 1800
aacaacagagc caaattctt atcacctata tcaatgcaca cggggtcaaa gtattacctc 1860
tgaatgaatg acaaaagaat ctctctggta ggggtttaga tataatttag catttttgg 1920
ttgttttta aatcaagcac atcaacctca agcccgctta gcaatggagg agtcagatg 1980
aatacgtcaa ataaatgact ttaaccaagt agcttaaat ggacttagca ctgtatgcat 2040
acttaaaaag gttttgaaa acaactact tgagaaatat ttgtttat tttctctaa 2100
catcatgcta tgttcatgc tgaacatctg acaacagaaa ttctgattt tattctagct 2160
aagtttgaa aacatttgtc atgctgttta atagaaaact gcaaacagca gatactgact 2220
ccatataaa accatatttt gtgcccgttt gactgtctcg accaaatact aatgggaaca 2280
attcttgagc ttttctgttt gctgattgtt aacatagagc agtctctaca ctacctgag 2340
gcaactctac attggaacac tgaggcttac agctgcaag agctcagag ctgaccatc 2400
atttaaacag aaatgctggt ttatttgcaa aatcaccagt atattttcta ttgtctctat 2460
aaaaaatcag tcaatlaagt acaagaatca tattttccat tctcttttag aaatttatt 2520
tgttgcctct atggaatca ttcacatctg acaattata tgttaaagag ttttaacttc 2580
tctatttttg tcaatttgt atctagtggc tgagaaatta aataattcta agtatgag 2640
ttacctatct gaaatgtac ttacagagta tcaatttaaa atggatgct cttaaaaaat 2700
ttgttactt ttaccaacaa tgaatataaa tttatgata ttttataat aatgtgaat 2760
tcttataaat ttgtctatg tacttatatt taatttgatt taatggttac tgccagata 2820
ttgagaaatg gtccaatat tgagtgtgtt tcaatatatt atctgctta tttcaacatg 2880
agtaatatga gcaaaataag ttaaaacctg cgtctgatca atttctctca tgactagaac 2940
taaaacagta aatttgaca atattaagcc tcaataatc atctccaaac tcttcttaac 3000
actttttaa tcagattgga agcatggac aaatcaggtt catgtgttc atctttatgt 3060
cctttgcaa tatccaagat catccatatt ggtagatatt cacatggagt ttcaaatca 3120
gaatagatta caattacott cctgccccta cacatcctac tcttattta aaagtctct 3180
ttgtgacttt caattctctg aaagttaaaa aatcacatgt gagaatgatt ataataact 3240
ctctcctgic tttcaccggt tacgtctggtt atgtctgaaa tacaccacat tttcttgtt 3300
ctgtccaagg ttaactcaat atctgtctga aagagaacta ctaaccacgt tacaatagag 3360
gctagtcttg aaaaaaaaaa tctatagatc taattgatc aatgtgaaa caaatgca 3420
aaataatggt taaagtataa gagaagtgg accaaggaga gagaatcat ttgaaatct 3480
aattgtagct tttctaggct cacattcatg tactactttt agcaccctta tgggtctgc 3540
tcgcccctcg gacagltgag ctttggatta tcttctctt caattttccc tctattgac 3600
cgagtgctc cctctgctc tacagattta tagtactcct tggctctttt gagtctccac 3660
ttttactcac tctctctggg atttttaaga tctctttctt ctcttataaa tcatctctc 3720
aatgaaaatt agcctaccaa aagtttggag actggaatcc tactttgag cactgacttg 3780
aaataactct tttggcaagt tgctgacat cctgtcttac caagtgga tatttgoat 3840
tttactgctt aaaaactttt tttttttta ccatcttat ccaaatttat catattgat 3900
gtaggactaa caggcttttt agaagctggc ttaactttg agtctcaagc tacaatgct 3960
ttggcgagcc tggctctccc acgtgaggtt ttaactttg ttatttgcct ccagtattc 4020
caaaatgctt attaaatgaa aggccagga acatgtttat tttagtccc ttgtctttt 4080
aacaattttg ttttgaatc aatgagtaat tcatgatgaa ttatttttga ctaatgala 4140
gccgaaggcc aagcttttaa tctaatagg taatgttctt cttttgtctt atgaaacaa 4200
tgagaatact ctgtgcattt caaatgcact ccgattatgc tgtgtttta ttcaataag 4260

```

WO 02/086069

PCT/US02/12464

```

cacaatatgt gttttattha taacttcata acaaaccttat aatataataa ttttaocttag 4320
cagacatgca aaagcttatt cttgtgtgac ttactttctt taagctaata atataaaaaat 4380
aaatatgtat cttaaaaatc tataataaaa cattagaaat taaagabatg tgctttttat 4440
tttgagatg agttoaattg cttctgtaga tgtgttttca gagctaggtg cagaggaatg 4500
tttgcacct ttagcgggtg aaaaagaaag agagtcaaga attttgttg atgtgtttg 4560
tgtgtgcata tatttgatat catcattata tttgtaact ttggactbtg aatcatagcc 4620
tgtttattct actgtgocat taaatatact ttaocttata cataacgaat aaataacota 4680
gatgtgatatt tatttacaca ctgtcttatg aaattcttat gtttatgca gcacagaaca 4740
agcacacatt tagagcattc atctgaaant tatatgaaa ttgaccatgg 4790

```

```

<210> 47
<211> 3916
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 6848851CB1

```

```

<400> 47
tgcctccgct cgcctcagtc ccttgacgac cctcctctct gggcccccgc cctcccgctt 60
cgggggtcaag ccccagagag cgcctcgaaa accacatttc ccagagtga cgcgacggc 120
aggggtcctc agaccggcgc tgcctcgccg gcgccatccc tatagagaag aacggaggtg 180
cggcctcttg tcatggcgct gttcccgacc ttltcggggc ttagttaggc tcccgatggc 240
gggagctcca ggaagagttt agactggctg agcaaccocaa gcttttctgt tggatccata 300
acgtccctga gccaaacaac tgaagcagct ccagcccatg tttctgaagg gttaccgctg 360
acaaggagtc atctgaaatc agagtcttca gatgaaagt acactaaca aaagctcaaa 420
caacaagta gaaaaaagaa gaaagagaaa aagaaaaaaa ggaagcatca gcatcataag 480
aaaaacaaga ggaagcatgg gccctcgagt agcagcaggt ctgagacaga caccgattct 540
gaaaaggaca aaccttccag aggcgttggg gccagtaaaa aggaatctga ggaaccgaa 600
caaggaaata atgctgcagc tgatactgga catcgtcttg ttggcttga ggaacttcag 660
gctgtgacgg gagaaccctt cagaacagat aagaaccag atctcgcaaa ctgggagtac 720
aagtctctct accgagggga tatagcaaga tacaagagga aaggagactc ctgctctggc 780
attaacctta agaagcagtg catatcttgg gaagggactt ccacagagaa gaagcattca 840
cgcacagcag ttgaacgcta ttttactaag aagagtgctg gattaatgaa catcagtgga 900
gttgacatta cgagtaaaac tgaacctccc tcaatctgagc ccatctcttt tataccagtg 960
aaggacttgg aagatggcgc tctcttaca acctgttga atctctggg gatattgat 1020
cagtcaccca cacattggct acaaggacag ggtcctccag agcaggaatc aaagcagcca 1080
gagccacagc cagacagcga gagtggcgct ctcaaggcca agtgtaggga gtttaacagg 1140
aggttgccgg agaactctcg ggatacgcag ctgtggatgg cattgtlct ttttcaggac 1200
gaggtcatga aaagctctgg cctgtatgcc atcagggaa gggagcagga aaagcgaag 1260
aggtccctga agctcattct ggaagaagag ctggccatcc tggagcgggc cattgagagc 1320
aacagagca gtgtggatct gaaactggcc aagctgaagc tctgacagaa gttctgggag 1380
cctccactc tggtaaaga tggcagaaa ctgatattt tgeatcccaa taatacagcc 1440
ctttggcaga aataaccttt attttccag agccagttta gtaoctttc gatatacaaa 1500
atccacagtc tttatggaaa atgcttgagc actttgtctg ctgttaagga cggcagcabc 1560
ttalctcacc ctgcgttgc tggcaaggaa gaggccatgt ttgcactctt tottcagcag 1620
tgcactttc tggcggcagg tggcactctt gagaaggcca tctcattgtt ccaggccatg 1680
gtggacttoa cctttctcaa acccgacagc gtgaaagatc tgcctaccaa aggacaggtg 1740
gaattctttg aacctcttgg gacagtgga gacccccggc ctggggagaa gggagcccca 1800
ggctggaagg cgtggatgca ccagcaggaa cgaggtggct ggggtgtcat caaccagat 1860
gaggatgac atgaaccaga agagatgac caggaaataa aagataagac tctgccagg 1920
tggcagatct ggcttctctc tgagcgttcc cgtgaccaga ggcactggcg gccctggcgc 1980
cctgataaga ccaagaagca aaccgaggaa gactgtgagg atccccagag acaggtgttg 2040
tttgatgata ttgggcaatc tttgatcaga ctttccagcc atgatcttca gttccagctg 2100
gtggggccct tcttcagatt cttgggtgtg ccttctggct ttactctccc agctcctctg 2160

```

WO 02/086069

PCT/US02/12464

```

ctttatctgg ccatggatga gaacagcabc tttgataatg gactttatga tgaaaagccc 2220
ttgacttttt tcaacccttt gttttctggg gctagctgtg ttggccgcat ggataggttg 2280
gggtatctct gctggaccag ggttcagAAC cgagagggcg aggagttcat ccgcaatgtc 2340
ttccaccctg tcatgccttt attttcaggg aaagagaagt cccagctctg cttctcctgg 2400
ttacagtatg agattgcaaa ggtcatttgg tgcctgcaca ctaaaaaca gaagagatta 2460
aagtctcaag ggaagaactg caaaaaacta gccagaatc tccttaagga gccagaaac 2520
tgcacaact tttgctgtg gaagcagtat gcaacatcgg agtggttgct tggcaacacg 2580
gggatgcca gaaaagtttt tgacacagca cttgcatcgg cagggaagcg agaactgaaa 2640
gactctgacc tctgtgagct cagctgcctc tatgctgagc tggaggttga gctgtcgcca 2700
gaagtggaga gggctgccc agctcgagct gttccatbat taaccaagt gactgagagc 2760
agccctatg ggcctacac tggacaggtg ttggtgttcc acattttgaa agcgcgaaag 2820
gcttatgagc agcaactgca ggaactgttg ggtgacagct gtgtctccaa tccagctccc 2880
accgattcct gtagecgcct aattagcctg gctaaatgct tcatgctctt ccagtatttg 2940
accataggga ttgatgctgc tgtgcagata taagaacagg tgtttgcaaa actgacagt 3000
tctgttttcc cagaaggctc tggcgagggg gacagtgcga gctcccagag ttggaccagt 3060
-----
gttctcgaag ccatcaccct gatgcacag agcctgctga gatccacat gaaagtgagt 3120
gtttaccctg tggccctctt gcygagggca ctctcaccag ctttaaaagt gtatccaggc 3180
aaocaggctc ttggaggctc ctatgtacag attcagaata agtcccacag tgcagcaaa 3240
accaggagat tttttgacac aatcaccagg tctgccaaac ccttggagcc ttggttgbtt 3300
gcaattgagc ctgagaaact gaggagagaga ctggtgaaa ctgtccagag gttagacggt 3360
agagatgacc agccacaact tectgagacc gcttlaatgc atcgcatcca agccctggtt 3420
gaaaatgcca tggcagcaga cagtggcagc cagtccccc tgcctgtgag gatgtatttg 3480
aaccttctg tttccttagg aaataaagaa agaagcaaa ggtatttcta caaagcaact 3540
cagagtggcc ctgggcaaa ggtgtgttac ctggacccg tggagtattt ccccgatgag 3600
atgcagaga tcttgacct gatgactgag aaggagctcc ggtgtgccc gccctggag 3660
gagctggagc tgcctgtgga ggtattgaga gcagtgggaa aacgggctgt gctgcgag 3720
ccagttgcc cccctgcgg agctaggagg ccgagcaga gaacgtgtgt gttaggagaa 3780
ctcggctttt gaaatgttct tctcgtatg taataatgtg gctgtccagc ctctcacc 3840
ttgcacactt ttgggtgtg taaatgacac aaaagttatt tacatattat atatgataat 3900
acgcgtgtat atgaca 3916

```

```

<210> 48
<211> 1702
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7040722CB1

```

```

<400> 48
ctccctgagc ccacgggagc ctgggaacca gccagccatg ctgggagggg tcccacctgc 60
gttggggtag gaggagaca cagtaccagt tcccttcaaa caactcaccg tgcctgggtt 120
gtaaacatcc ccgtcgttca ccagcacttt gttaaagagg acaacgcccc catcactggg 180
gaaaggcttc tgggtgagcc ccgcagaaaa agacaccaga gaaggcaccg gagctcctgg 240
agcagagaaa ggatggaatg gattgaccgg ggtgcttctg tctctcaggg ggagggcctg 300
gaaaactgag gctcactgac cagcccaggg tgcagggccc attttatttc acccaaatgt 360
gtgaagcaaa ctctcactgc caaacagcca gagggttccc tgtctgaggg atcagcagcg 420
aggacagctg aggaatctga gaaacccgac agactcttgt taggaggggt tgtgtgaggc 480
tgggcatgta agcatatgag atgagatgt gtatgtcgtg ctcatcctc tagggagctg 540
gtgagcacac atattaggag gccagggccag gcagtgggtg catcatcctc tagggagctg 600
cttcttact aataaatcag taaaacctgg agcccagga agctcatcag gaccacacaa 660
gggcagagge ccaccocagt atccccttgg gaaactcaat gagctgggat gatggagag 720
ggaagatccc ttgtcttggg cctgactctg accctctctg gggagcaaac ctctgggaaat 780
ttgctccact atagtcagca gaggcctgtc ccagtgttgg tgggtgggga agggctgggc 840
ccaggcttga tttgtgctaa tagcctggag ggaggagga ggagatgcag aggggggca 900

```

WO 02/086069

PCT/US02/12464

```

gaggccacca agggagaccgc tctctctgct cctgggtctct cagcagacat acctggggaa 960
gtctacaggag gtgacttcgg gtatctctgaa aaacagaaag gaaatggcat gtcaaacaaag 1020
acctcaagca ccagctcctt gttctgagcc ccccttttcc tctgctctc acaagttcaa 1080
atogagtctt cactactata gaaacagaca tagaacaagc gcaccactg tcactgccc 1140
cctlcaagca gagggcaatt tgtaccagcc cgtccctgct tctcactag ccaattatc 1200
ctggcgggga agagggcaact gagggcaagc tggctgcagg aaagaaagag atgggcactg 1260
gagcccttga gacttctctg tgaactgagag agtcacacaa gcccaactaac atttgcctc 1320
tagcagtgtt atctgttcaa gactctacc ttoactaga ctgaatgag tacttataa 1380
atggaggcat ggatgggtgg gtgcagaagt ggagggacaa atagatgat gattggatgc 1440
acgcaggaaa gactgcacag cttgatgtgt ggaagatgt gcacatgat aggtgggtgg 1500
acagatatgt gcattgatac atgggtgggt ggttccttgg ggtgcccag cctctgata 1560
tgaggatgca tatgtgggtg gatggatgta tcatgggtg galacacgga tacatgtgtg 1620
gtgggtggc tggatggaag gatggatgga agggatacac ggatagatgc atgggtgagt 1680
gctgatgca cggagggtg cg 1702

```

```

<210> 49
<211> 1462
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Inyte ID No: 6430290CB1

```

```

<400> 49
aagaccctgt ctcaaaaaaa agaaaaaaat cagttgcaga tgtccatacc ttatgatgat 60
tagccaagga ctgggtgacac acctcatggc agattgcacg tcagggtcga gagccagtgg 120
ctgaggacag ctctatgctc aggcctgtgt ttcccacagc tttgtgccc tctgcagccc 180
tottgtgccc agacaagag tagttaaaaa cagcagtaga gagggcaaga catggaactg 240
cotacttttt tctcaagtgg gaaaagtgct tgggtgagac atgtggctgt ccatgaaag 300
gaacagctat gctgactgtg ggttggaaaa tctaggctgg ctgtcttctg ggtgtccc 360
tgttggctgg tttggtgect acagtgagat ccttctctg cctgtgtgtc tcttaatgat 420
tctgatgca tctctgcgag gtccagagtt ggttcccggc gcaggcccat ctgtcaacct 480
gaggttcaga tgggaccaca agcagcgtag cacaggggga aaacacttgt gcgctgac 540
cttgacatc aacaagttag accagaacat ggagtgcacc agcctctca gaattctga 600
gtctctgta ccaccactag gggctctct gtggtcaag acctctctg aggcagaggy 660
tgaggaggag gcagaagaa aagagaacag gagacttga gtgggcaga gtagagcaga 720
accagcagc cccgctttt tggatctct tggggctct ggggatgtgt gttgtggag 780
tggagagtg tgacgagcag aggcagaca gctgctgaaa gcgatcact ctctccttgt 840
ggggatcct gggagaaaca tgggtgatct ctgcagtaag gcagcagat cctctgtggg 900
ccttatggg cggctctggg tctctgact ctaactcctg gcaacaaat gggctgcaag 960
gtgggtctcc tggtagatgc tcaggaaagg ttgltcaact cctcacccc tectgagctg 1020
cttctagcc tgcagcccc gtctttggal cctttttgg cactctgct ggtcagagc 1080
ttccgtgga gcagcagag ctgctctctc ctgggtctct ctgggagca ccatgagct 1140
tggcagttc ctctatattg tgcattctg gttcaaac caccgccc tcagacagct 1200
gtacttccc ggcaggeagt aggatgggg tgggggtgga ggcgtggg cctgcaact 1260
tcctgctct gctctgctc ttgttagcca gtgagttgga ggcgctggg ctacagctat 1320
gtctccaag gaggacttgg aatgaacccc agacacaaag ggtgaggtca gtggttctg 1380
cagcaagaag gctgagttag tgaacacctg tcacagctcc caggatcctt ggtcagctc 1440
cctggctccg aatctcccc ct 1462

```

```

<210> 50
<211> 3958
<212> DNA
<213> Homo sapiens

```

WO 02/086069

PCT/US02/12464

<220>

<221> misc_feature

<223> Incyte ID No: 2640251CB1

<400> 50

```

tttttttttt tttttttteta aagtttcagt cagtttttatt acatccagtc cagtaagctg 60
tgcgtgtotag tggatccac acatacttta ttagttttat ctttgttcag cccccaaca 120
aaaatttgggaa aagattgcta cacatgacag agcgacagta cccatggggtg acatgacagc 180
ccataataca gttctctggtt toacattaaa aagttaccat tcaatccgtc caaacatcca 240
ccccatccac aegttctctc aaagactagc atggttctgt ccattaccctt gctctaatc 300
caattccatc gttctcaata cttatttgaa aaggcacatt tggtcagctc gatgtcattc 360
ttaggggatc tttgctctct cctggccatt actgcttgc tttatgtact gaaaagtcat 420
ccattgaaag gttctgttcc gaattgtcgc acaaaagcag ctgtaaaacc cactccattt 480
ccctccctaa atlacctccc tccaggtctg ttccctcaaa gatcgtccct tagtgggtca 540
cagtaactcal gaacctctgg tacaattgat acaggtctgag aagtaagaat tcaactttca 600
cttagggcaa ctcaattccct gacctgaacc ggggtgttct cacccttggc catggttaga 660
tctctttcct tgatagaaca agcaaaaagag gggctgacct atctccctgt aacatctggt 720
gactgtgggc ggcctgggccc agccttgctg tgtttgcctc agatgcagct aaggaagctc 780
cagttgtagc cttggcacag ctgagggccat gccattgcta cccgggaagct ggctctgacg 840
cgttgtgact tagaacacttt acaagcttga gccattccct actagttcat caactttcct 900
tccctatgat ctgcaaaagat acagagagaaa tggcattcca aaatccatg gatgactttt 960
ctgaatgtcc tagaatgttt ggcaaaagatc tggcctaaaac tagtactatg gaatagaaaa 1020
ggaaaagtga catttgtgta atttatttttg gaaatgcaaa ggaattatta tgcagcaatc 1080
ttaactttgt gtacatacca ctacacctccc ttaaaattca gcagaaccag aggtgcca 1140
aatgtcaaaa aatattctagt cattccaata tgcraaacaa gactgccttg ttttaaaaaa 1200
acaaaataaa tgtaaaaatt gtatgactct taccagtgta aaagctttac taacacagcc 1260
ctgggtgcta tttccagata ctgaaaagat ctaagtaggc gcttaagaac tgccttttaa 1320
cccattccaa acctccatat taaactagga tgaggtgca ggaagctgag aaggaagccc 1380
tgggtgggaa cctggatgct ctttccctgga gaacagcttt gctatttccag tggtaacctc 1440
gaaacttcaa aatattgact gtacaaatta ctcaagttta ttttggccta attattctta 1500
agaaaggtgg tcttttttta gaaaagggaa gcagaagcta attgaaatc cattactcag 1560
gcctctagta agtccccctc tcaigtatgc actattcact gagggccatg aggytgagcc 1620
cctgctgggg gggcagggca gggctgtgag ctcaaggacc catccctgtg ctcaagagca 1680
tgggaactcca atgagagac acactatgta ggtgcaattt ctgatttctt agttttttgac 1740
caaaatagca gtaatttcaa acagttgaat gctctgatcg atactatgct agttgtgat 1800
gaatgcttg aatggtatg aatgctctgg accaaagaga acagcaacat gtccttagtc 1860
ctcagctcaa gaagctccatc gactccctct ctcaaatgtt acgtctgtgt ttttcaaaag 1920
coattgtgac aggtcaaaaac cctggagtat ttggtttgtg tttaaatgta ctgcaactc 1980
tggaaaacat caacctaaag cctggatgca ggaagcaggg atgtcgagga tgtgaattag 2040
agtgtctggg tbtgctaabg cctagggagc catoagccac tbtgtctggt tgggggatcc 2100
tgtatgact ggaacctggg ccttccattac gggctccggt tttggagtc agagggattg 2160
gcattttaa agactgctg cttacccttt gcttggcatg cacattcagg atgtgacttc 2220
agcttccatc atcctggcca acaaaaaaaa gttcaaaaaa ttgttttcaa ttaattttct 2280
gaagatcttt gtagggagc accctctcga gatgaaacca atttctgtc cacacagcca 2340
ccatctctt atatagatgg ggaaaaaatcc agaaaagggg gattttcaag agctggcttg 2400
aagccccctg atgtccctt ggcagagcca aagaatgagc ccagagcctg agattcaaga 2460
tgtgtctc tctgaaacac accgtggagc aatgtgtgct acctgcagat gccgcactca 2520
gcagctgga atgcagctc tcaatgcoctc acctagaaac gcttctctg ctgaatggct 2580
gtttgctctc agcagcaact gtcagcatag ctccggccctc tctgccccgt gccaaaggct 2640
gactaggacc agaagaatca gaaagcatgg tggagtcgca aagcattggc aaagccagca 2700
catctgtgag catgttacac atgggcttgc ctaagtattt cctctctgt gaacatcact 2760
taggagctac gtcagttaaag gcatctgtgag acacaggtgt ggttggagga tcatgagggg 2820
actactccac aggccttgg agagttgcat tttgggggct tctctagaaa gctctggggg 2880
aeggaagcag tggcctcctg gtcagctaca gacacagggc ggttagggcc cgtcttgcct 2940
tcatgtctt ttaaacacaa atggcagctc acaaatattc ctctaatat tcttccatc 3000
agcctagggc cctcaagag acgtgcaaac caaagtattt gtttttctca ctttattctt 3060

```

WO 02/086069

PCT/US02/12464

```

aacctagat gtggcaagtc tggaccactg ctttagagaa agctggatta aatctcttga 3120
gcagaagtag agaccctaga gaatcttccac tggataacct gttccacact ttggccaata 3180
tttatgaata agttctaaag cctatgcgaa gcccaattctg ttttgcttgg ggaccagcca 3240
gaccactcct caactccagt ggggcctttt ttgtgtcttt cacatcactg ttacaagtgg 3300
tggctggaca ataatgaca tllaatcac tgtccocacat atggatata tagaaaaacc 3360
ctaatgtttt aaaatltttt taatcactag ctgcttttao aaagcttttg atgcaatttg 3420
ttagaggata tgcagtcga gatattccag tactactgtg tgatcatatg aaaaagtgg 3480
gtttctctc tgcacaagtc aatagggtgg tagactaac caccagtgct ctactctct 3540
ctctccgatt aatcgtgttc actatattgt ggcagaatca gtcacgtccc ccagacctoa 3600
agctgcatct tccaaatggt tgggtctctt taaagagtcc ccgtggctgg gotttctctg 3660
aagaagtcca ttaaatccoc tgcctctata agctgcagga gtgccttagg atatggcccg 3720
tcagctctgg ggalgacct gcactgtctc ccagalttca ctoaggttc gtattttct 3780
ctgccacctt gaactacaga acaggtctct agtagcttga taagctttaa gctagaatga 3840
gatggaaagg tgggaaagac agacaggaaag tgtagctctt agtctatggg gctctcctta 3900
ggggacaaaa agatggctta ggctccgtgg gttttcaact cgtagaaaaa ataacctg 3958

```

```

<210> 51
<211> 826
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3839350CB1

```

```

<400> 51
gattctttcc actgaggagg caaggatgga gttgctgctg acccgtgcag atttgetgct 60
ggtaacatac tcatgggeaa agtctgaaaa acctcttctc gccctgggca aggcagatgg 120
ggcagccocag caggctggct ttgccacccg tctccgatat catccaacac agtcagccag 180
gtttgttggc cctctcactc acattctcat ttgccaacaa gatgggttgg gctatgctgt 240
ctttgcacag ccatctgac caggggacat ttgtgtgggt gtgtcccccg gagctgcatg 300
tctgcaggtt cagattgggg tcagctttgc tgggtagtyt ggttccctgc catggcacca 360
aagaagaggg ctggcggatt agtctcttgg tctctcactg ctttctagg acacagcggg 420
ccatgtggac tgacagctgg agacttcata aagatggcct aaaaaattag gacactgtcc 480
tgttgacaaa gcccttaggc ctgcccccat caggcctcat ggttcttttt caggcctaag 540
ccagggttct gaaatcactc gtaatgtgaa tttatatttc tagggcccat tgcctgcatc 600
ctttccagca agatgtoatt ctgagatltt tgtcaltctc gggcgttatg aggaactgcca 660
agggacacaa ttccatggga gaaaagtltt ctltcaagct gcaagccagc tctggctctc 720
ggagtttctc cttaggattg caagacagta gttccctgtc ctttaggtct tctccagaaa 780
tccttaaaa tagcaataa aataagatat aagcaaaaa aaaaaa 826

```

```

<210> 52
<211> 729
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 6393813CB1

```

```

<400> 52
aagaattcgg aacgagggcg ggggggggtc aggatcctcc acaggtaggc gcagtcagct 60
ggagctccgc ggcggtccgc cggctcgtgga gggcgtgtcc tgcggccgca tggcogtagt 120
gttccggcgc gttgtggagg agtctctgag cgagatggcg gcggcgtgac agggagcgc 180
gcgaattcct gatgaatc tcgttatcgtc gaagtttctc tttggctcat cagccaacca 240
ggccttggac ctagtgtgac gacagtccat caocttaatc tcatcaccca gtggaaggcg 300

```

WO 02/086069

PCT/US02/12464

```

tgtttaaccag gtccttggaa gttccagtaa aacatacaca tgtttggctt cttgtcatta 360
ctgttcacatg cctgcatttg cattctcagt gctacggaag agtgacagca tccctgtgcaa 420
gcatactcttg gcagtttaacc tgaatcaggt tatgaggacc tgtcagcagc taagtgtctc 480
tgacaagcag ttgactgaca tattattgat ggagaagaaa caagaagcat aaaaggactg 540
caggaggtgc tgtgggtlgg agccgtgggc tgtggagggg ttgtgtatga tgagaagccc 600
tgtacagtcct tgtcaagaaa taccttgagc cagtcctctga gacgcttcgg taaaaaatgt 660
ccctggatgg aatcaagatt ttaaatcaca ataaagccta atatcatggt gtgtccacaa 720
aaaaaaaaa                                     729

```

```

<210> 53
<211> 1610
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5685755CB1

```

```

<400> 53
tgccaacaga aaggggggtcc cttccctgtg ccaggatcct atataaggaa gagggatgtg 60
agcagagtggt tctggcttca ccgtccctcc agtggacttg ggggctgagc ataaagatgc 120
ctaggagaggg ctactccaag cctgggtcct ggggcagctt ctgggcatg ctgaccttgg 180
tgggctctgt caccatgca gcacagagag ccgatgttgg cggggaggca gctggcaact 240
ccatacaacca ctcccagggcg gtgctccagc gcttgcaaga gctgctcgg cagggcaacg 300
ccagcagatgt ggttctggcg gtgcaggctg cgggcaacga tgaggtccgg gtattccaacg 360
cccaaccgctt gctgctggga ctgcaagctg agctgttccg ggagctgcta agtaaccaga 420
gcagagcggt gctgcaggag ccaacaggact gcgcgctgt cttcgacaag ttcaatcaggt 480
acctgtactg cggggagctg accgtgctgc tgaccaagcc catccccctg cacagactgg 540
ccccaagta cggcgtgtcc tccctgcagc gcggcgtggc cgactacatg cggcgcaacc 600
tggcggggagg cggggcccg gcggtgggct ggtaccacta cggggtgggc accggggacg 660
aggccctggc cgagagctgc ctgcagttcc tggcctggaa cctgtcggcc gtggggcca 720
gcaccagatg gggcgcctg agccccgagc tgctctggca gctcctgcaa cgtcgggacc 780
tgggtctgca ggatgaactg gactgttcc acgcctgga ggcctggctg ggtcgcgcgc 840
ggcgcctccc tgccgtggcc gagcggggcc tgccgcccac acgctaccac atgatccca 900
cggcacagct gttccagctg caggcggcct cggcagccct gcgcgccac ggcccgggg 960
tggcagacct cctgtgcag gcctaccagt tccacgcgc gtgcgcctg cactacgcca 1020
agttcttcca cgtcaacggc agcctctcc tgccccgcaa ctacctcggc cccgctgggg 1080
ggccccctg ggtcatcaac aaccggccc gcgacgacc cagcaacagc ttccagacgc 1140
agctgggccc gagtggccc gagcggggcc gccgggtcac ctggacaact cctctctcgc 1200
cgcgctggct gcccgtaagc ctggcgcccc ttlacgcgga cgcgcgggc actgctctgc 1260
cgcgcgcgcg cccggaggac ggcgcagccc ggttggtggt gacccggcc agcagcggcg 1320
gcacgcggcg gggcgtgagc ttccagaaga cagtgtggt gggggcccgc cagcaagacc 1380
gcctgtggt cgcacaagcc tacagcttcc accagagcag cggagggccc ggcgaacttc 1440
tggcgcacgc cgaactgcag cggcgcaact ccgagtacct ggttgagaac gccctgcaac 1500
tgcaactcat cgtcaagccc gtatacaca cccttatccg gaccccaag tagcctcggg 1560
gtctgggaat aaaggcctgg ctaggtggt ccctgtgggt ctgggaaggg 1610

```

```

<210> 54
<211> 869
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 71728459CB1

```

WO 02/086069

PCT/US02/12464

```

<400> 54
ctctaggaca tccctgcaga aaggagagaa ccaaggagaa ggagtgcag gtgttttgc 60
aggcctgtat ctcagccag atacaaagcc cccaatgta caoctgctct cattatgtct 120
cattctgtag cagttaatca agttaattat gagcagttga gcctggggag cttttactga 180
cagtgcaacc tataattct ctgaaagtta cagagccaga gatgaaggtt actgctcttt 240
taggagccct gctccagtt ttgtctttt ttctgtttt taittcatgg tttagctct 300
ttggagacca gaagagcatt gacagccccc cagatgagca gcagagcaat tcttacta 360
gtggtcaagc agcttcttac agccaaaggg ctataggaag aaagggtaac tggctgccc 420
acagtttaca cgatgaagct gccttgggct ctgggtcttg gtgatttgc cagaagattg 480
ttctcagag ctacagccca cacattctcc taaaggcctt ctctctggca taggtcaag 540
acagaaccgc agagagggaa atgagccaga acaaatatgt catctgatga cctcaccccc 600
agcaaatggg cagggagag agctagaatg aagaagccag aacatggagc accatgaaa 660
gaaactgccc aagcaagaga agtggcatgt gggctttaa gaccaggtt agaaaacagc 720
agccctgacc ttgtgttga agtagcgggg cgaggtatct tcatagaagt gttaatctg 780
tactctctcc ccgattctga gtaggctgct tttaacataga gatgatcaat gtgtacttgc 840
ctaatttaatt acaactgcc cccagtgta 869

```

```

<210> 55
<211> 2209
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1904303CB1

```

```

<400> 55
gcctgtgccc tgtgctccg cttgtctgct ctccccgcc ttggccttct ctgcgcccc 60
cctttcgctt cctctgtctt tgtgtctctt ctctccctgc tegtctctgt gtccttggac 120
tgctccttct gctctgcgcg ctctctggtc gctccctccg tgttgcctcg ttgtgtggtt 180
agctgtcttc ctttctctct gtctcctcgc tctttgtctt ttttcttttg gttctgtgtg 240
ctctgtctcg tctctggctt tcttggtttt tgtgcccaag cccccaagtt ggcctgtgct 300
gccggctaca cccagccccc gcccgaggcc gccccggcgc cccggctagc caggcggggc 360
ggcccaactc tgccctaact ctctcctccg cgttccgggc ctgggagccg ctttggagag 420
gatgagctcc tggagctggt tctctgtgca gaccctctgc ctctcccca cgggcccagc 480
ttcggggcgc ggggcccgcg gcaccgccaa ctgcgagctc aagccccaac aaagcagact 540
gaattccttc ttgtggacca ttaagcgaga cccaccatct tacttctttg gcaacaatca 600
tgtcccctac acccgagttt gggacttcat ccccgacaac tctaaggagg ctttctctca 660
gagcagcatt gtgtactttg agttgatctt cacagacccc tatacctact cagctctctac 720
cagctgtcag atgctgccc acggcgagaa cctccaagat gtgctcccca gggacatcta 780
ctgcccctcc aagcgcaccc tggagtatgt caagctcatg atgccccttg ggtgacccc 840
agaccagcgc ggcgaagggc tctacgcaga ctactcttct atgctcttg ccggaacctg 900
ggagcgaagc aggcctgtct ggtgtatgct catggtcaac tcccggactg aagtggacat 960
taagtcocct ggagtgcctg tcttagaact gttccttggc caggagcctg agcggctgag 1020
gaaacagact gggcagtggt aaaaagtggg agagcagtcg catccattga atgggtttaa 1080
cttttcacag gtcactcttg ctttgaacca gaccctcttg cagcaggaaa gcttgcgagc 1140
aggcagctct cagatccctt acaagcggga ggatctctat aaacactata actcggggga 1200
cctcagctcc gtcactctca gccatgacag ctccccagtt cccaatttta ttaatgccac 1260
gctaccacct caggagcgca tcaactgtca ggagattgac agctacttac gccggagact 1320
tgacaaggcc ttctctcttg cctttggagc tggctatttc atgggcaaca acacagtgct 1380
ggaatgtttg cggcgtgaag gctatgaggt agaacaagcc ctctgtggac gaccacatca 1440
caaaagggag agtaaaaaga cctccacagc gcccaactctg tccacatct ttgtccaaa 1500
agtccctacc ctggaagtac cggcaccaga agccgtatcc tcaaggcact caacgctgac 1560
tccccttctg tcccggcctg gaagtgcga cacgccagtt gaggccaac agaggttccg 1620
gaagaagcgc agcgggtcac agcggggccc gcgactccgg caattcagcg atctgtgggt 1740

```

WO 02/086069

PCT/US02/12464

```

ccgcctggag gagagtgaca tagtcccga actccaggtc cctgtccctgg acaggccat 1800
ctccactgaa ctggggctcc ctccgccctgg gcattcccac cacagccaga tgggtggccag 1860
cagtgctctgc ctgtctctct ggactcctgt gttctgggtg ctgggtctgg ctttccaaac 1920
agagacacc cctctgtaac gactggaagc accaggctaa gaacctgacc cctggacctt 1980
gaagaatggc cttccctgta ctcccattc tggcttagcc ttgttggccc caatccagaa 2040
gagactgctc tgaaaaagc gggcccagtg tgattctctc ctttccaagg aatgtgactt 2100
tgggttatcc aactttgggg gcagggtgtac agttttgtaa catagtgagt tgtgtgaaaa 2160
taattataa atgagttgta tgaataataa aaaaaaaaa aaaaaaaa 2209

```

```

<210> 56
<211> 1520
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2911343CB1

```

```

<400> 56
ggccaaaggga gcccgcttt acctcccac tggggtcca gtctcttagc aacgactccg 60
gcttcttagg aactgtctct ttctcaacca ttcttgccea caacaccca gcttgetgce 120
agcaaaagccc ctccaacccc ctcaaaactc agaccttca cctcaattta ctgtttttct 180
cgacctcaac ccagcaattt tctctgaatc aacctcttt cctctctccc taagtctctg 240
agtttctctc tctctctgcc atcctgagtt ctgccatct tagggggccc caagacctct 300
ctttctgttc ctctcccgc tcagaaccagc agccttttat tttagatcat gctcggaggg 360
aagtccagccc aggtcccaga ggaagggggc gtctgcata ctgaagccct tatcactaag 420
cggaacttga ctttccctga agatggggaa ctgtcagaga agatgttcca cactctttag 480
gaactgcaga ctgtccgctt ggaacggggg gggattacta ctatcaggaa cttagaaggg 540
ctccagaatc ttcacagtct ctatctgcaa gggataaga tccagcaaat tgagaacctg 600
gcttgcctcc cctccttggc dtctctgtct ctggcaggaa accaaatcag gcagggtgaa 660
aacctctctg acctcccctg cctccagttt ctggaccttt ctgagaacct gatagaaca 720
ttgaagctgg atgagttccc ccagagcctt ctcatctca acctgtctgg aaacagctgc 780
accaaacagg atggtaccgc cagctgggtg acagaagccc tgcacctct cctggacctg 840
gaccggcagc ctgtgtggga ggcctggatt tcggaagagg aggatgaagc ctcaagcag 900
gaggagttcc cagagctgag tggcccattc tgctcagaga gaggcttccc caaggacctg 960
gagcaggagc tgagcaggca cagggagcac cggcaacaga cggccctgac agagcactg 1020
ctgaggatgg agatgcagcc caccctcacc gacctgccc tgctacctgg ggtgccatg 1080
gctgggggaca ccagcccttc tggcactctc gcgcaagggg agggacagct cctcagggcc 1140
gtctctcacc ccagggcttc ctctcccacc aagaaccat gcagctctgat tcccagggcc 1200
caccaaagct ctttctgggg aaggaaaggg gcacgagcag ccacagcccc caaggctctc 1260
gtggctgagg ccccagcacc aaccaaaact acggccaaga gaagcaagaa atgattctct 1320
gtcaaccttt ctctactagt gggcaggagt gggcctgccc cctcttctca gacctctgac 1380
ctgtgcagaa agcccactcc cagtaaaagt tctctagggc ctgagatgac tttctatgct 1440
acttggggta tctcagggga aaaaaccagt gaaagctcca ggaacaaaa taccagagctc 1500
agtgagccac ctgaaaaaaa

```

```

<210> 57
<211> 1282
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7500308CB1

```

<400> 57

WO 02/086069

PCT/US02/12464

```

agccctgca gcagggtggg agcattagga catgtccaca cgtcttcaag ggaaaaatggc 60
taagggctgt cagagaggcc cagggaagag gcctgggtag gggcgagcag ggagcagaca 120
atcaactgtt gaaggaagat ccaagtccag gaaggaactt agggcagttt ggtgtcatgg 180
aaggaactgg gggccagacc gaggggtca agagcccttc agactctgcc tgagcgaggg 240
tgggtctctg tcaccagggc tggagtgtag tgggtgtgac tcagctcact gcagccttga 300
actcctagcc tcaagtgatc ctccctaccc aacctccaga gtaactggga ctacagggaa 360
gctcagtgcc ccccaagcca ggaatgcccc agcttgggtc cccggcctcg cggccacctt 420
gctgttcagc ctgctggctg gcccccacaa ggtgattctg aagcccagcc tggcccacac 480
tcccacagag ccaccctctc cctacagctt caggcctgaa gaatataccg gggatcagag 540
gggcattgac aaccggcctc tctgagtcac ctctgctctg gaatcttggc atcagcaacc 600
tctctccccc tgctctctgg atcaagctag agactgtgag caccocagga atgtccctgc 660
ccatctggcc gtgtctctgt tcaattcttg atttaactta ttaacttttc tctctctgtt 720
tccaccaccg ctgctctctt tgtctctgag gttaggctgg agtgacagtt tcccccacc 780
ccccagccca agaagagggc tggcggaaag naaatgctga ccattggagg tgcacaacag 840
tagaatgggc tactgtgagg ggtagtaaga gcccaatttc tgggggtatg cgaacttga 900
ttgagacacc agctctgaga ttttatcagg gcaactctat acctgtggga catgtgactg 960
gatgagccct gagccagctt ccaactctac ctgaatagag aactcactgc acccaccacc 1020
aacacatgat aaacacatgt cctcaactgaa tgttactgat tggcgctgag ggcctgcctc 1080
tggctgtgtg gggagggtgg tggagaggtg agcccagcca ctgctgagag gtgctgtgat 1140
gggtctctcg cgcgcacatc ccaccactga tggaccactc gggaggctcg ggagggcagt 1200
ccatccatgg gccgcctcag gagagagggc tgttctagat gtaattggctg tctgtttttt 1260
gatgtctctg tctgccaacc ag 1282

```

```

<210> 58
<211> 1228
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7501098CB1

```

```

<400> 58
gcgaacccca ggcccttccc aggtttgccc gcgggggcca tccagaccct gcggagagcg 60
agcccgggag cgtgcgcyag gtttgagggc gccggagacc gagggcctgg cggccgaagg 120
aaccgcccca agaagagcct ctggcccggg ggctgctgga acatgtgccc ggggacacag 180
tttgtttgac agttgcagag ctatgtttac gcttctggtt ctactcagcc aactgcccac 240
agttaccctg ggtttccttc attgcccaag aggtccaaag gcttctaagc atgcccggaga 300
agaagatttg atctggagct ctctcactccc ggcacactag aaagagagtg caatgaagaa 360
ctttgcaatt atgaggagac cagagagatt tttgtggatg aagataaaac gattgcattt 420
tggcaggcaat atcagctaaa agcacaacc ccaaatcag ctcttcagcc gtctatgaaa 480
gggggagggc cactcctccc atcattttca gaagacctga ggaagctgcc ttgtctccat 540
tgcgcctctc tgtggaggat gcaggattac ctctttatga acagggcagtg gcctgtacca 600
gaaaacacag tgtttcacca ccaaccacat atcctgggca cacaacagga tttagggtat 660
ttaaaaaatc tatgtctctc ccatctcaact gactaacctg tcaattttgt atagaanaat 720
tgtgtttatt gataggccgg gcatggtggc tcatgcctgt aatcccagca ctltgggagg 780
ccaggagttc gagaccagcc tggccaacat ggtgaaaccc ggtctctact aaaaattcaa 840
aaataactca ggcgtcatgg ggcctgcctg tagtcccacc tacttgggag gctgagagcg 900
gagaattgct cgaacctggg aggcagaggt tgcagtaagc tgagatcacg ccaactgcat 960
ccagcctggg cgacagagca agactccatc tcaaaaataa aataaaaaaa gaaaaaaga 1020
aaagagagag aaaagagag agggagagag agatgaagga ggaagagag gagaaagaga 1080
agaagagaga gaagagagacc acaaaagaca tggactatcc aacttttatg acaactgacg 1140
gaataaggag aatagtcatg tactgtacac gaagtctgtc tgcatctgga ctgaaactgat 1200
catcatcagt gatagagact ttgatcta 1228

```

<210> 59

WO 02/086069

PCT/US02/12464

<211> 3582

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7503839CE1

<400> 59

```

gcctctccct gccccgccc gccccagtc cttgacgacc ctctctctgt ggccccccc 60
ctcccgcttc ggggtcaagc cccagagagc gcgcgaaaa ccacatttcc cagagtgcac 120
cgcgacggaa ggggtcctca gaccggcgct cgtcgcggg gcgcatccct atagagaaga 180
acggaggtac ggcctgtggt catggcgctg ttccagcct ttgcggggct tagtgaagct 240
cccagtgccg ggagctccag gaaagtcagc gaaataatgc tgcagctgat actggacalc 300
gctttgtttg gcttgaggac atbcagcctg tgacgggaga aaccttcaga acagataaga 360
aacctgattc tgcgaactgg gactacaagt ctctctaccg aggggatata gcaagataca 420
agaggaagag agactcctgc cttggcattt accctaagaa gcagtgcata tcttgggaag 480
ggactccac agagaagaag caticacgca agcaggttga acgctatctt actaagaaga 540
gtgtgggatt aatgaacalc gatggagtty ccattagcag taaaaactga cctccctcat 600
ctgagcccat ctctttata ccagtgaaag acttggaaag tgcggctcct gttacaacct 660
gggtgaatcc tctgggattc tatgatcagt caaccacaca ttggctaca ggacagggct 720
ctccagagca ggaatcaaaag cagccagacg cacagccaga cagcgagagt cgggctctca 780
agggcaaggt ggagaggttt aacagggggg tgcgggagaa tctctgggat acgcagctgt 840
ggatggcatt tttgtctttt cagggcaggg tcatgaaaaa tctctggctg tatgcaatg 900
aggaagagga cagggaaaag cgaagaggtt ccttgaagct cattctggag aagaagctgt 960
ccattctgga cggggccatt gagagcaacc agagcagttg ggaatctgaa cctggccaagc 1020
tgaagctcgc cacagagttc tgggagccct ccaactctgt caaagagttg cagaactga 1080
tatttttga tcccaataat acagcccttt ggcaagaata ccttttatt ttgccagacc 1140
agttttagtac cttttcgata tcaaaaattc acagtcttta tggaaaatgc ttgagcactt 1200
tctctgtgt taaggacggc agcattttat ctcaacctgc gttgctggc acggaagagg 1260
ccatgtttgc actctttctt cagcagtgcc actttctgc gtaggtggc cactctgaga 1320
agggcattcc atgtttccag gccatggtgg acttcaactt cttcaaaccc gacagcgtaa 1380
aagatctgcc taccaaagga caggtggaat tctttgaacc cttttgggac agtggagagc 1440
cccggctcgt ggagaaggga gcccgaggct ggaagcgtg gatgcaacc caggaacagag 1500
gtggtggtgt ggtcatcacc ccagatgagg atgacgatga accagaagag gatgacacag 1560
aaataaaga taagactctg cccaggtggc agatctggct tgetgctgag cgttccctgt 1620
accagagcca ctggcgcccc tggcgccctg atagaccaa gaagcaaac gaggagagct 1680
gtgaggtacc cgagagacag gtgttgtttg atgatattgg gcaatctttg atcagacttt 1740
ccagccatga tcttcaagtc cagctgttgg aggccttctt cgaatctttg ggtgtgcttt 1800
ctggctttac tctccagcc tctgtctttt atctggccat ggaatgaaac agcatctttg 1860
ataatgactc tbatgatgaa aagcccttga ctttttcaa cctttgtttt tctggggcta 1920
gtgtgtcttg ccgcatggat aggttgggct atctctgctg gaccaggggt cagaaccagag 1980
agggcgagga gttcactccg aatgttctcc acctgtctat gcttttattt tcaggcaanag 2040
agaagtcoca gctctgcttc tctgggttac agtatgagat tgcnaaggtc atttgtgtgc 2100
tgcacactaa aaacaagaag agattaaagt ctcaagggaa gaaactgaaa aaactagcca 2160
agaatctcct taaggagcca gaaaactgca acaactttg cctgtggaag cagtatgca 2220
atctggagly gttccttggc aacaaggagg atgcaagaaa agttttgac acagcacttg 2280
gcagctcagg aagcagagaa ctgaaagact ctgacctctg ttagctcagt ctgctctatg 2340
ctgagctgga ggtggagctg tgcacagaag tgagaagggc tgcacagct cgagctgttc 2400
acataattac caagctgact gagagcagcc cctatgggac ctacactgga caggtgtttg 2460
ctgttcacat tttgaaagcy cgaaggcctt atgagcagc actgcaaggac tgtttgggtg 2520
acagctgtgt ctccaaatca gctccaccg attcctgtag ccgctcaat agcctgacta 2580
aatgtctcat gctctccag tatttgacca tagggattga tgetgctgtg cagatatacg 2640
aacaggtgtt tgcaaaactg aacagttctg ttttcccaga aggtctctgg cagggggaca 2700
gtgccagctc ccagagtgg accagtttcc togaagccat cacaccgat caacagagcc 2760
tgetgagatt ccaatgaaa gtgagtgttt acccgctggc ccctctgca gaggcactct 2820

```

WO 02/086069

PCT/US02/12464

```

cacaggcttt aaagtgtgat ccaggcaacc aggttccttg gaggtccat gtacagattc 2880
agataaagtc ccacagtgc acgaaaacca ggagatttt tgacacaatc accaggtctg 2940
ccaaaccctt ggagccttgg ttgtttgcaa ttgaagctga gaaactgagg aagagactgg 3000
tggaaactgt ccagaggtta gacggtagag agatccacgc cacaattcct gagaccgctg 3060
taatgcacgt gatccaagcc ctgtttgaaa atgccatcgc cagcgcacgt ggcagccagt 3120
gccctctgct gtggaggatg tatttgaact ttctggtttc cttaggaatc aaagaaagaa 3180
gcaaaggtgt attctacaaa gcacttcaga gttgcccttg ggcaaaagtg ttgtacctgg 3240
acgccgtgga gtatttcccc gatgagatgc aggagatcct ggcactgatg actgagaagg 3300
agctccgggt gcgctgccc ctggaggagc tggagctgct gctggaggat tagagagcag 3360
tggaaaaagc gctgtgacct gcgaggccaa gttgcccaac ctgcccagct agggggcgcg 3420
agcagagaac gtgtgtgcta ggagaactcg gcttttgaaa tgtctctctc cgatagtaat 3480
aatgtgggct gccagcctct cacatcttgc acactttttg ggtgtgtaaa tgacacaaaa 3540
gttatttaca tattatataT gataatcgc gtgtatata ca 3582

```

<210> 60

<211> 2053

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7503698CB1

<400> 60

```

gectgtgccc tgtgctccg cttgtctgct ctccccgcc ttggcctctc ctcccccgcc 60
cctttccgct cctctgtcct tgtgtctctc ctctcccgte tgcctctcgt gtccttggac 120
tgcctctctc gtctccgccc ctctcttggtc gctccctccg ttgtgtctcg ttgtgtggtt 180
agctgtcttc cttttctctt gttccgtcgc tctttgcttc ttttctcttg gttctgtgtg 240
tctgtgctct tctctgctct tcttggtttt tgtcccacgg cccccacttt ggcctgtctc 300
gccgggtcca cccagcccc gcccgaggcc gcccccggcc cccggctagc cagggcgggc 360
ggccacactc tgcctactc ctctccctccg cgttccgggc ctccggagcc ccttgaggag 420
gatgagtccc tggagctggt tctctctgca gacctctgc ctctgcccac cgggcgcagc 480
ttcggggcgc ggggcgccc gcaccgcca ctccgagctc aagccccaac aaagcagact 540
gaattctctc ttgtggacca ttaagcgaga cccaccatct tacttctttg gcacaatcca 600
tgcctccgtc acccaggttt gggacttcat ccccgacaac tctaaaggag ctttctctca 660
gagcagcatt gtgtaacttg agttggatct cacagacccc tataccatct cagctctcac 720
cagctgtcag atgtgcccac agggcgagaa cctccaagat gtgtcccaca gggacatcta 780
ctgcgcctc aagcgcacc tggagtatgt caagctcatg atgccttgt ggalgacccc 840
agaccagcgc ggcaaggggc tctacgcaga ctacctcttc aatgctatg ccggaactgt 900
ggagcgcag aggcctgtct ggggtatgct catggtcaac tccctgacty aagtggacat 960
taagtccctg gtagtgcctg tcttagacct gttccttgc caggagctgt agcggctgag 1020
gaaacagact ggggcagtyg aaaagtgga agagcagtgc ctccattga atgggtgaa 1080
cttttcacag gttcccaatt ttaataatgc cactgaccca cctcaggagc gcatcactgc 1140
tcaggagatt gacagctact tacgcccggg gctgatctac aagcgaatg agagatagg 1200
gaagcgggtg aagcctcttt tggaggagtt cctgacaaa ggcctctctt ttgctcttgg 1260
agctgtctat ttoatgggca acaacacagt gctggatgt ttgcggcgtg aagcctatga 1320
ggtagaacac gccctgtctg gaccgacctc ccccaagggy aagagtaaaa agactctcac 1380
acygcccact ctgtccacca tctttgctcc aaaagtccct accctggaag taaccggccc 1440
agaagccgta tctcagggc actcaacgct gctctccctt gtgtcccggc ctggaagctg 1500
cgacacgccc agtgaggcgc aacagaggtt ccggaagaa cggagcgggt cacagcggag 1560
gcgcgactc cggcaatcca gcatctgtg ggtccgctg gaggagagty acatagtccc 1620
gcaactccag gtccctgtcc tggacagcca catctccact gaactcgggc tccctcggcg 1680
tgggcattcc caccacagcc agatggtggc cagcagtgcc tgcctgtctc tctggactcc 1740
tgtgtctctg gtgtgtgtgc tggctttcca aacagagaca cccctctctg aacagctgga 1800
agcaccaggc taagaaacctg acccctcggg cttgaagaaT gggcattcct gtactccaca 1860
ttctgtctca gcttgttgg gcccaatcca gaagagacty ctcttgaaaa agcggcccag 1920

```

WO 02/086069

PCT/US02/12464

```
fgttgattct tototttcca aggaatgga ctttgggta tccaactttg ggggcaggtg 1980
tacegttttg taacatagtg eggtgtgga aaataaatta taaatgagtt gbatgaaat 2040
aaataccttt ttttgataa aaa                               2063
```

【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
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 02/086069 A3

(51) International Patent Classification: C07H 15/11,
C12P 21/06, C12N 15/85, 15/65

(21) International Application Number: PCT/US02/12464

(22) International Filing Date: 19 April 2002 (19.04.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/285,207 20 April 2001 (20.04.2001) US
60/287,114 27 April 2001 (27.04.2001) US
60/288,640 3 May 2001 (03.05.2001) US
60/290,516 11 May 2001 (11.05.2001) US
60/292,184 18 May 2001 (18.05.2001) US
60/343,553 21 December 2001 (21.12.2001) US
60/357,002 13 February 2002 (13.02.2002) US
60/358,279 20 February 2002 (20.02.2002) US
60/366,041 19 March 2002 (19.03.2002) US

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

(75) Inventors; and
Inventors/Applicants (for US only): KLAMMER,
Aaron, A. [US/US]; 565 Ortega Avenue #1, Mountain
View, CA 94040 (US). HAFALIA, April, J., A. [US/US];
2227 Calle de Primavera, Santa Clara, CA 95054 (US).
DUGGAN, Brendan, M. [AU/US]; 243 Baena Vista
Avenue #306, Sunnyvale, CA 94086 (US). WARREN,
Bridget, A. [US/US]; 2250 Homestead Court #2, Las
Altos, CA 94024 (US). EMERLING, Brooke, M. [US/US];
1735 Woodland Avenue #71, Palo Alto, CA 94303 (US).
TRIBOULEY, Catherine, M. [FR/US]; 1121 Tennessee
Street #5, San Francisco, CA 94107 (US). ARVIZU,
Chandra, S. [US/US]; 1706 Morocco Drive, San Jose,
CA 95125 (US). HONCHELL, Cynthia, D. [US/US];
400 Laurel Street #203, San Carlos, CA 94070 (US).
NGUYEN, Dannie, B. [US/US]; 1403 Ridgewood Drive,
San Jose, CA 95118 (US). KALLICK, Deborah, A.
[US/US]; 58 Linda Avenue, Atherton, CA 94027 (US).
YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale,
CA 94087 (US). AU-YOUNG, Janice, K. [US/US];
233 Golden Eagle Lane, Brisbane, CA 94005 (US).
RAMKUMAR, Jayalaxmi [IN/US]; 34359 Maybird
Circle, Fremont, CA 94555 (US). LI, Joana, X. [US/US];
1264 Geneva Avenue, San Francisco, CA 94112 (US).

THANGAVELU, Kavitha [IN/US]; 1950 Montecito
Avenue #23, Mountain View, CA 94304 (US). GIETZEN,
Kimberly, J. [US/US]; 691 Los Hueces Drive, San Jose,
CA 95123 (US). DING, Li [CN/US]; 3553 Alma Street,
#146, Palo Alto, CA 94306 (US). BAUGHN, Mariah,
R. [US/US]; 14244 Santiago Road, San Leandro, CA
94577 (US). YAO, Monique, G. [US/US]; 1189 Woodgate
Drive, Carmel, IN 46033 (US). WALLA, Narinder, K.
[US/US]; 890 Davis Street A #205, San Leandro, CA
94577 (US). MASON, Patricia, M. [US/US]; 360 Clarke
Lane, Morgan Hill, CA 95037 (US). LAL, Preeti, G.
[US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US).
GRAUL, Richard, C. [US/US]; 682-29th Avenue, San
Francisco, CA 94121 (US). REDDY, Roopa [IN/US];
1233 W. McKinley Drive #3, Sunnyvale, CA 94086 (US).
BECHA, Shantya, D. [US/US]; 21062 Gary Drive #
117, Castro Valley, CA 94546 (US). SAPPERSTEIN,
Stephanie, K. [US/US]; 3531 Highland Avenue, Redwood
City, CA 94062 (US). RICHARDSON, Thomas, W.
[US/US]; 616 Canyon Road #107, Redwood City, CA
94062 (US). TRAN, Uyen, K. [US/US]; 2638 Mabury
Square, San Jose, CA 95133 (US). ELLIOTT, Vicki, S.
[US/US]; 3770 Polton Place Way, San Jose, CA 95121
(US). TANG, Y., Tom [US/US]; 4230 Ranwick Court, San
Jose, CA 95118 (US). AZIMZAI, Valda [US/US]; 5518
Boulder Canyon Drive, Castro Valley, CA 94552 (US).
YAN, Lu [CA/US]; 3885 Corrina Way, Palo Alto, CA
94305 (US). XU, Yuning [US/US]; 1739 Walnut Drive,
Mountain View, CA 94040 (US).

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

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

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

[Continued on next page]



WO 02/086069 A3

(54) Title: SECRETED PROTEINS

(57) Abstract: The invention provides human secreted proteins (SECP) and polynucleotides which identify and encode SECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of SECP.

WO 02/086069 A3 

Published:
— with international search 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.

(88) Date of publication of the international search report:
6 March 2003

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US02/12464
A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : C07H 15/11; C12P 21/06; C12N 15/85, 15/63 US CL : 536/23.1; 435/69.1, 325, 320.1 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) U.S. : 536/23.1; 435/69.1, 325, 320.1 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) MPSRCH		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database Genbank, Accession No. AAW00839, ARUFFO et al. See entire document.	3
---		-----
Y		6, 7, 9
X	Database Genbank, Accession No. X14356 M21091, ALLEN et al. Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcR). Gene sequence. Nucleic acid res. 1988, Vol 16, No. 24, page 11824. See entire document.	12, 13
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents:		
"A"	document defining the general state of the art which is not considered to be of particular relevance	"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.
"E"	earlier application or patent published on or after the international filing date	"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.
"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)	"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.
"O"	document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family.
"P"	document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search 08 September 2002 (08.09.2002)	Date of mailing of the international search report 06 DEC 2002	
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer ANTHONY CAPUTA Telephone No. 703-308-0916 <i>Kevin Callan for</i>	

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 K 48/00	A 6 1 K 48/00	4 B 0 6 5
A 6 1 P 1/04	A 6 1 P 1/04	4 C 0 8 4
A 6 1 P 1/16	A 6 1 P 1/16	4 C 0 8 5
A 6 1 P 1/18	A 6 1 P 1/18	4 C 0 8 6
A 6 1 P 3/10	A 6 1 P 3/10	4 H 0 4 5
A 6 1 P 5/14	A 6 1 P 5/14	
A 6 1 P 5/40	A 6 1 P 5/40	
A 6 1 P 7/00	A 6 1 P 7/00	
A 6 1 P 7/02	A 6 1 P 7/02	
A 6 1 P 7/04	A 6 1 P 7/04	
A 6 1 P 7/06	A 6 1 P 7/06	
A 6 1 P 7/08	A 6 1 P 7/08	
A 6 1 P 9/00	A 6 1 P 9/00	
A 6 1 P 9/04	A 6 1 P 9/04	
A 6 1 P 9/10	A 6 1 P 9/10	
A 6 1 P 11/00	A 6 1 P 9/10	1 0 1
A 6 1 P 11/06	A 6 1 P 9/10	1 0 3
A 6 1 P 13/08	A 6 1 P 11/00	
A 6 1 P 13/10	A 6 1 P 11/06	
A 6 1 P 13/12	A 6 1 P 13/08	
A 6 1 P 17/00	A 6 1 P 13/10	
A 6 1 P 17/02	A 6 1 P 13/12	
A 6 1 P 17/06	A 6 1 P 17/00	
A 6 1 P 19/02	A 6 1 P 17/02	
A 6 1 P 19/06	A 6 1 P 17/06	
A 6 1 P 19/08	A 6 1 P 19/02	
A 6 1 P 19/10	A 6 1 P 19/06	
A 6 1 P 21/00	A 6 1 P 19/08	
A 6 1 P 21/04	A 6 1 P 19/10	
A 6 1 P 25/00	A 6 1 P 21/00	
A 6 1 P 25/02	A 6 1 P 21/04	
A 6 1 P 25/08	A 6 1 P 25/00	
A 6 1 P 25/12	A 6 1 P 25/02	
A 6 1 P 25/14	A 6 1 P 25/08	
A 6 1 P 25/16	A 6 1 P 25/12	
A 6 1 P 25/20	A 6 1 P 25/14	
A 6 1 P 25/28	A 6 1 P 25/16	
A 6 1 P 27/02	A 6 1 P 25/20	
A 6 1 P 27/06	A 6 1 P 25/28	
A 6 1 P 27/16	A 6 1 P 27/02	
A 6 1 P 29/00	A 6 1 P 27/06	
A 6 1 P 31/04	A 6 1 P 27/16	
A 6 1 P 31/10	A 6 1 P 29/00	
A 6 1 P 31/12	A 6 1 P 29/00	1 0 1
A 6 1 P 31/18	A 6 1 P 31/04	
A 6 1 P 33/00	A 6 1 P 31/10	
A 6 1 P 33/02	A 6 1 P 31/12	

A 6 1 P	33/10	A 6 1 P	31/18	
A 6 1 P	35/00	A 6 1 P	33/00	
A 6 1 P	35/02	A 6 1 P	33/02	
A 6 1 P	37/02	A 6 1 P	33/10	
A 6 1 P	37/04	A 6 1 P	35/00	
A 6 1 P	37/06	A 6 1 P	35/02	
A 6 1 P	37/08	A 6 1 P	37/02	
A 6 1 P	43/00	A 6 1 P	37/04	
C 0 7 K	14/47	A 6 1 P	37/06	
C 0 7 K	16/18	A 6 1 P	37/08	
C 1 2 M	1/00	A 6 1 P	43/00	1 1 1
C 1 2 N	1/15	C 0 7 K	14/47	
C 1 2 N	1/19	C 0 7 K	16/18	
C 1 2 N	1/21	C 1 2 M	1/00	A
C 1 2 N	5/10	C 1 2 N	1/15	
C 1 2 P	21/02	C 1 2 N	1/19	
C 1 2 P	21/08	C 1 2 N	1/21	
C 1 2 Q	1/68	C 1 2 P	21/02	C
G 0 1 N	33/15	C 1 2 P	21/08	
G 0 1 N	33/50	C 1 2 Q	1/68	A
G 0 1 N	33/53	C 1 2 Q	1/68	Z
G 0 1 N	33/566	G 0 1 N	33/15	Z
		G 0 1 N	33/50	Z
		G 0 1 N	33/53	D
		G 0 1 N	33/53	M
		G 0 1 N	33/566	
		C 1 2 N	5/00	A
		A 6 1 K	37/02	
		C 1 2 N	15/00	F

(31)優先権主張番号 60/290,516
(32)優先日 平成13年5月11日(2001.5.11)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/292,184
(32)優先日 平成13年5月18日(2001.5.18)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/343,553
(32)優先日 平成13年12月21日(2001.12.21)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/357,002
(32)優先日 平成14年2月13日(2002.2.13)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/358,279
(32)優先日 平成14年2月20日(2002.2.20)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/366,041
(32)優先日 平成14年3月19日(2002.3.19)
(33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,

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

- (72)発明者 ハファリア、エープリル・ジェイ・エイ
アメリカ合衆国カリフォルニア州9 5 0 5 4・サンタクララ・コーレデプリマベータ 2 2 2 7
- (72)発明者 ダガン、ブレンダン・エム
アメリカ合衆国カリフォルニア州9 4 0 8 6・サニーバイル・# 3 0 6・ブエナビスタアベニュー 2 4 3
- (72)発明者 ワレン、ブリジット・エイ
アメリカ合衆国カリフォルニア州9 4 0 2 4・ロスアルトス・# 2・ホームステッドコート 2 2 5 0
- (72)発明者 エマーリング、ブルック・エム
アメリカ合衆国カリフォルニア州9 4 3 0 3・パロアルト・# 7 1・ウッドランドアベニュー 1 7 3 5
- (72)発明者 トリボレー、キャサリン・エム
アメリカ合衆国カリフォルニア州9 4 1 0 7・サンフランシスコ・# 5・テネシーストリート 1 1 2 1
- (72)発明者 アービズ、チャンドラ・エス
アメリカ合衆国カリフォルニア州9 5 1 2 5・サンノゼ・モロッコドライブ 1 7 0 6
- (72)発明者 ホンチェル、シンシア・ディー
アメリカ合衆国カリフォルニア州9 4 0 7 0・サンカルロス・# 2 0 3・ローレルストリート 4 0 0
- (72)発明者 ニュエン、ダニエル・ピー
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・リッジウッドドライブ 1 4 0 3
- (72)発明者 カリック、デボラー・エイ
アメリカ合衆国カリフォルニア州9 4 0 2 7・アサートン・リンダアベニュー 5 8
- (72)発明者 ユエ、ヘンリー
アメリカ合衆国カリフォルニア州9 4 0 8 7・サニーバイル・ルイスアベニュー 8 2 6
- (72)発明者 オウ・ヤング、ジャニス・ケイ
アメリカ合衆国カリフォルニア州9 4 0 0 5・プリズベーン・ゴールドデンイーグルレーン 2 3 3
- (72)発明者 ランクマール、ジャヤラクシミ
アメリカ合衆国カリフォルニア州9 4 5 5 5・フレモント・メイバードサークル 3 4 3 5 9
- (72)発明者 リー、ジョアナ・エックス
アメリカ合衆国カリフォルニア州9 4 1 1 2・サンフランシスコ・ジュネーブアベニュー 1 2 6 4
- (72)発明者 サンガベル、カピサ
アメリカ合衆国カリフォルニア州9 4 3 0 4・マウンテンビュー・# 2 3・モンテシトアベニュー 1 9 5 0
- (72)発明者 ギーツェン、キンバリー・ジェイ
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・ロスウエコドライブ 6 9 1
- (72)発明者 ディング、リー
アメリカ合衆国カリフォルニア州9 4 3 0 6・パロアルト・# 1 4 6・アルマストリート 3 3 5 3
- (72)発明者 ボーゲン、マライア・アール
アメリカ合衆国カリフォルニア州9 4 5 7 7・サンレアンドロ・サンティアゴロード 1 4 2 4 4
- (72)発明者 ヤオ、モニーク・ジー
アメリカ合衆国インディアナ州4 6 0 3 3・カーメル・ウッドゲートドライブ 1 1 8 9
- (72)発明者 チョーラ、ナリンダー・ケイ

アメリカ合衆国カリフォルニア州 9 4 5 8 7 ・ ユニオンシティ ・ # 7 1 2 ・ ユニオンスクエア 3
3

(72)発明者 メイソン、パトリシア・エム

アメリカ合衆国カリフォルニア州 9 5 0 3 7 ・ モーガンヒル・クラークレーン 3 6 0

(72)発明者 ラル、ブリーティ・ジー

アメリカ合衆国カリフォルニア州 9 5 0 5 6 ・ サンタクララ・ピーオーボックス 5 1 4 2

(72)発明者 グラール、リチャード・シー

アメリカ合衆国カリフォルニア州 9 4 1 2 1 ・ サンフランシスコ・トゥエンティナインズアベニュー 6 8 2

(72)発明者 レディ、ルーパ

アメリカ合衆国カリフォルニア州 9 4 0 8 6 ・ サニーベイル・# 3 ・ ウェストマッキンレーアベニュー 1 2 3 3

(72)発明者 ベチャ、シャニア・ディー

アメリカ合衆国カリフォルニア州 9 4 5 4 6 ・ カストロバレー・# 1 1 7 ・ ゲイリードライブ 2 1 0 6 2

(72)発明者 カレート、ステファニー・ケイ

アメリカ合衆国カリフォルニア州 9 4 0 6 2 ・ レッドウッドシティ・ハイランドアベニュー 3 5 3 1

(72)発明者 リチャードソン、トマス・ダブリュ

アメリカ合衆国カリフォルニア州 9 4 0 6 2 ・ レッドウッドシティ・# 1 0 7 ・ キャニオンロード 6 1 6

(72)発明者 トラン、ユエン・ケイ

アメリカ合衆国カリフォルニア州 9 5 1 3 3 ・ サンノゼ・メイブリースクエア 2 6 3 8

(72)発明者 エリオット、ビッキー・エス

アメリカ合衆国カリフォルニア州 9 5 1 2 1 ・ サンノゼ・ポルトンプレイスウェイ 3 7 7 0

(72)発明者 タング、ワイ・トム

アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ・ランウィックコート 4 2 3 0

(72)発明者 アジムザイ、ヤルダ

アメリカ合衆国カリフォルニア州 9 4 5 5 2 ・ カストロバレー・ボールダーキャニオンドライブ 5 5 1 8

(72)発明者 ヤン、リュ

アメリカ合衆国カリフォルニア州 9 4 3 0 3 ・ パロアルト・コリーナウェイ 3 8 8 5

(72)発明者 スー、ユーミング

アメリカ合衆国カリフォルニア州 9 4 0 4 0 ・ マウンテンビュー・ウォルナットドライブ 1 7 3 9

F ターム(参考)

2G045	AA35	AA40	BA11	BB50	DA13	DA36	FB02	FB03		
4B024	AA01	AA11	BA80	CA04	CA09	DA02	EA02	HA14	HA17	
4B029	AA07	FA12								
4B063	QA08	QA13	QQ42	QQ53	QR32	QR36	QR55	QR62	QS25	QS34
4B064	AG01	AG27	CA10	CA19	CA20	CC24	DA01	DA13		
4B065	AA90X	AA93Y	AB01	CA24	CA44	CA46				
4C084	AA01	AA02	AA07	AA13	AA17	BA01	BA08	BA22	BA23	CA18
	DC50	NA14	ZA01	ZA02	ZA05	ZA06	ZA15	ZA16	ZA20	ZA33
	ZA34	ZA36	ZA40	ZA42	ZA45	ZA51	ZA54	ZA55	ZA59	ZA66
	ZA68	ZA75	ZA81	ZA89	ZA94	ZA96	ZA97	ZB05	ZB08	ZB09
	ZB11	ZB13	ZB15	ZB26	ZB27	ZB33	ZB35	ZB37	ZB38	ZB39
	ZC06	ZC08	ZC31	ZC35	ZC42	ZC54	ZC55			
4C085	AA13	AA14	CC03	DD01						
4C086	AA01	AA02	AA03	BC42	BC43	CB07	DA38	EA17	EA18	MA01
	MA04	NA14	ZA01	ZA02	ZA05	ZA06	ZA15	ZA16	ZA20	ZA33

ZA34	ZA36	ZA40	ZA42	ZA45	ZA51	ZA54	ZA55	ZA59	ZA66	
ZA68	ZA75	ZA81	ZA89	ZA94	ZA96	ZA97	ZB05	ZB08	ZB09	
ZB11	ZB13	ZB15	ZB26	ZB27	ZB33	ZB35	ZB37	ZB38	ZB39	
ZC06	ZC08	ZC31	ZC35	ZC42	ZC54	ZC55				
4H045	AA10	AA11	AA20	AA30	BA10	CA40	DA75	EA20	EA50	FA74

專利名称(译)	分泌蛋白質		
公开(公告)号	JP2004535174A	公开(公告)日	2004-11-25
申請号	JP2002583584	申請日	2002-04-19
[标]申請(專利权)人(译)	洞察Genomics公司		
申請(專利权)人(译)	洞察基因组公司		
[标]發明人	クラマーアーロンエイ ハファリアエープリルジェイエイ ダガンブレンダンエム ワレンブリジットエイ エマーリングブルックエム トリボレーキャサリーンエム アービズチャンドラエス ホンチエルシンシアディー ニュエンダニエルビー カリックデボラーエイ ユエヘンリー オウヤングジャンスケイ ランクマールジャヤラクシミ リージョアナエックス サンガベルカピサ ギーツエンキンバリージェイ ディングリー ボーグンマライアアール ヤオモニークジー チョーラナリンダーケイ メイソンパトリシアエム ラルプリーティジー グラールリチャードシー レディルーパ ベチャシャニアディー カレートステファニーケイ リチャードソントマスダブリュ トランユエンケイ エリオットビッキーエス タングワイトム アジムザイヤルダ ヤンリュ スーユーミング		
發明人	クラマー、アーロン・エイ ハファリア、エープリル・ジェイ・エイ ダガン、ブレンダン・エム ワレン、ブリジット・エイ エマーリング、ブルック・エム トリボレー、キャサリーン・エム アービズ、チャンドラ・エス ホンチエル、シンシア・ディー ニュエン、ダニエル・ビー カリック、デボラー・エイ ユエ、ヘンリー		

オウ-ヤング、ジャニス・ケイ
 ランクマール、ジャヤラクシミ
 リー、ジョアナ・エックス
 サンガベル、カピサ
 ギーツェン、キンバリー・ジェイ
 デイニング、リー
 ボーグン、マライア・アール
 ヤオ、モニーク・ジー
 チョーラ、ナリンダー・ケイ
 メイソン、パトリシア・エム
 ラル、プリーティ・ジー
 グラール、リチャード・シー
 レディ、ルーパ
 ベチャ、シャニア・ディー
 カレート、ステファニー・ケイ
 リチャードソン、トマス・ダブリュ
 トラン、ユエン・ケイ
 エリオット、ビッキー・エス
 タング、ワイ・トム
 アジムザイ、ヤルダ
 ヤン、リュ
 スー、ユー・ミング

IPC分类号 G01N33/50 A61K31/7088 A61K38/00 A61K39/395 A61K45/00 A61K48/00 A61P1/04 A61P1/16 A61P1/18 A61P3/10 A61P5/14 A61P5/40 A61P7/00 A61P7/02 A61P7/04 A61P7/06 A61P7/08 A61P9/00 A61P9/04 A61P9/10 A61P11/00 A61P11/06 A61P13/08 A61P13/10 A61P13/12 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/08 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/08 A61P25/12 A61P25/14 A61P25/16 A61P25/20 A61P25/28 A61P27/02 A61P27/06 A61P27/16 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/02 A61P37/04 A61P37/06 A61P37/08 A61P43/00 C07K14/47 C07K16/18 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12P21/02 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566

CPC分类号 A01K2217/05 A61K38/00 A61P1/04 A61P1/16 A61P1/18 A61P11/00 A61P11/06 A61P13/08 A61P13/10 A61P13/12 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/08 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/08 A61P25/12 A61P25/14 A61P25/16 A61P25/20 A61P25/28 A61P27/02 A61P27/06 A61P27/16 A61P29/00 C07K14/47

FI分类号 C12N15/00.ZNA.A A61K31/7088 A61K39/395.D A61K39/395.N A61K45/00 A61K48/00 A61P1/04 A61P1/16 A61P1/18 A61P3/10 A61P5/14 A61P5/40 A61P7/00 A61P7/02 A61P7/04 A61P7/06 A61P7/08 A61P9/00 A61P9/04 A61P9/10 A61P9/10.101 A61P9/10.103 A61P11/00 A61P11/06 A61P13/08 A61P13/10 A61P13/12 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/08 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/08 A61P25/12 A61P25/14 A61P25/16 A61P25/20 A61P25/28 A61P27/02 A61P27/06 A61P27/16 A61P29/00 A61P29/00.101 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/02 A61P37/04 A61P37/06 A61P37/08 A61P43/00.111 C07K14/47 C07K16/18 C12M1/00.A C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/68.A C12Q1/68.Z G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K37/02 C12N15/00.F

F-TERM分类号 2G045/AA35 2G045/AA40 2G045/BA11 2G045/BB50 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA80 4B024/CA04 4B024/CA09 4B024/DA02 4B024/EA02 4B024/HA14 4B024/HA17 4B029/AA07 4B029/FA12 4B063/QA08 4B063/QA13 4B063/QQ42 4B063/QQ53 4B063/QR32 4B063/QR36 4B063/QR55 4B063/QR62 4B063/QS25 4B063/QS34 4B064/AG01 4B064/AG27 4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA01 4C084/AA02 4C084/AA07 4C084/AA13 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/BA23 4C084/CA18 4C084/DC50 4C084/NA14 4C084/ZA01 4C084/ZA02 4C084/ZA05 4C084/ZA06 4C084/ZA15 4C084/ZA16 4C084/ZA20 4C084/ZA33 4C084/ZA34 4C084/ZA36 4C084/ZA40 4C084/ZA42 4C084

/ZA45 4C084/ZA51 4C084/ZA54 4C084/ZA55 4C084/ZA59 4C084/ZA66 4C084/ZA68 4C084/ZA75
 4C084/ZA81 4C084/ZA89 4C084/ZA94 4C084/ZA96 4C084/ZA97 4C084/ZB05 4C084/ZB08 4C084
 /ZB09 4C084/ZB11 4C084/ZB13 4C084/ZB15 4C084/ZB26 4C084/ZB27 4C084/ZB33 4C084/ZB35
 4C084/ZB37 4C084/ZB38 4C084/ZB39 4C084/ZC06 4C084/ZC08 4C084/ZC31 4C084/ZC35 4C084
 /ZC42 4C084/ZC54 4C084/ZC55 4C085/AA13 4C085/AA14 4C085/CC03 4C085/DD01 4C086/AA01
 4C086/AA02 4C086/AA03 4C086/BC42 4C086/BC43 4C086/CB07 4C086/DA38 4C086/EA17 4C086
 /EA18 4C086/MA01 4C086/MA04 4C086/NA14 4C086/ZA01 4C086/ZA02 4C086/ZA05 4C086/ZA06
 4C086/ZA15 4C086/ZA16 4C086/ZA20 4C086/ZA33 4C086/ZA34 4C086/ZA36 4C086/ZA40 4C086
 /ZA42 4C086/ZA45 4C086/ZA51 4C086/ZA54 4C086/ZA55 4C086/ZA59 4C086/ZA66 4C086/ZA68
 4C086/ZA75 4C086/ZA81 4C086/ZA89 4C086/ZA94 4C086/ZA96 4C086/ZA97 4C086/ZB05 4C086
 /ZB08 4C086/ZB09 4C086/ZB11 4C086/ZB13 4C086/ZB15 4C086/ZB26 4C086/ZB27 4C086/ZB33
 4C086/ZB35 4C086/ZB37 4C086/ZB38 4C086/ZB39 4C086/ZC06 4C086/ZC08 4C086/ZC31 4C086
 /ZC35 4C086/ZC42 4C086/ZC54 4C086/ZC55 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30
 4H045/BA10 4H045/CA40 4H045/DA75 4H045/EA20 4H045/EA50 4H045/FA74

優先権
 60/285207 2001-04-20 US
 60/287114 2001-04-27 US
 60/288640 2001-05-03 US
 60/290516 2001-05-11 US
 60/292184 2001-05-18 US
 60/343553 2001-12-21 US
 60/357002 2002-02-13 US
 60/358279 2002-02-20 US
 60/366041 2002-03-19 US

外部リンク Espacenet

摘要(译)

本発明提供了鉴定和编码SECP的人分泌蛋白 (SECP) 和多核苷酸。本
 发明还提供表达载体，宿主细胞，抗体，激动剂和拮抗剂。本发还提
 供了用于诊断，治疗或预防与SECP异常表达有关的疾病的方法。

と5を参照)。

接頭コード	解析タイプやプログラムの例
GNN 、 GFG 、 ENST	例えば、GENSCAN (Stanford University, CA, USA) または FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)を用いた、ゲノム配列群からのエキソン予測
GBI	手動で編集された、ゲノム配列群の解析
FL	ステッチまたはストレッチゲノム配列(実施例5参照)
INCY	EST 配列群の、ゲノムへのマッピングからの、完全長転写物と エキソンとの予想。エキソンと転写物を予想するために、ゲノ ム位置と EST 組成のデータが組み合わされる。